

Wei Lu · Qiaoyan Wen ·  
Yuqing Zhang · Bo Lang ·  
Weiping Wen · Hanbing Yan ·  
Chao Li · Li Ding · Ruiguang Li ·  
Yu Zhou (Eds.)

Communications in Computer and Information Science

1299

# Cyber Security

17th China Annual Conference, CNCERT 2020  
Beijing, China, August 12, 2020  
Revised Selected Papers

Editorial Board Members

Joaquim Filipe 

*Polytechnic Institute of Setúbal, Setúbal, Portugal*

Ashish Ghosh

*Indian Statistical Institute, Kolkata, India*

Raquel Oliveira Prates 

*Federal University of Minas Gerais (UFMG), Belo Horizonte, Brazil*

Lizhu Zhou

*Tsinghua University, Beijing, China*

More information about this series at <http://www.springer.com/series/7899>

Wei Lu · Qiaoyan Wen ·  
Yuqing Zhang · Bo Lang ·  
Weiping Wen · Hanbing Yan ·  
Chao Li · Li Ding · Ruiguang Li ·  
Yu Zhou (Eds.)

# Cyber Security

17th China Annual Conference, CNCERT 2020  
Beijing, China, August 12, 2020  
Revised Selected Papers



*Editors*

Wei Lu  
CNCERT/CC  
Beijing, China

Yuqing Zhang  
University of Chinese Academy of Sciences  
Beijing, China

Weiping Wen  
Peking University  
Beijing, China

Chao Li  
CNCERT/CC  
Beijing, China

Rui Guang Li  
CNCERT/CC  
Beijing, China

Qiaoyan Wen  
Beijing University of Posts  
and Telecommunications  
Beijing, China

Bo Lang  
Beihang University  
Beijing, China

Hanbing Yan  
CNCERT/CC  
Beijing, China

Li Ding  
CNCERT/CC  
Beijing, China

Yu Zhou  
CNCERT/CC  
Beijing, China



ISSN 1865-0929

ISSN 1865-0937 (electronic)

Communications in Computer and Information Science

ISBN 978-981-33-4921-6

ISBN 978-981-33-4922-3 (eBook)

<https://doi.org/10.1007/978-981-33-4922-3>

© The Editor(s) (if applicable) and The Author(s) 2020. This book is an open access publication.

**Open Access** This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd. The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

## Preface

The China Cyber Security Annual Conference is the annual event of the National Computer Network Emergency Response Technical Team/Coordination Center of China (hereinafter referred to as CNCERT/CC). Since 2004, CNCERT/CC has successfully held 16 China Cyber Security Annual Conferences. As an important bridge for technical and service exchange on cyber security affairs among the industry, academics, research, and application, the conference has played an active role in safeguarding cyber security and raising social awareness.

Founded in August 2001, CNCERT/CC is a non-governmental non-profit cyber security technical center and the key coordination team for China's cyber security emergency response community. As the national CERT of China, CNCERT/CC strives to improve the nation's cyber security posture and safeguard the security of critical information infrastructure. CNCERT/CC leads efforts to prevent, detect, alert, coordinate, and handle cyber security threats and incidents, in line with the guiding principle of "proactive prevention, timely detection, prompt response and maximized recovery."

This year, due to the COVID-19 pandemic, the China Cyber Security Annual Conference was held online in China on August 12, 2020, on the theme of "Jointly Combating against Threats and Challenges" as the 17th event in the series. The conference featured one main session and six sub-sessions. The mission was not only to provide a platform for sharing new emerging trends and concerns on cyber security and discussing countermeasures or approaches to deal with them, but also for finding ways to join hands in managing threats and challenges posed by this year's COVID-19 pandemic. There were over 2.95 million visits to our online event and over 1,500 comments received live. Please refer to the following URL for more information about the event: <http://conf.cert.org.cn>.

We announced our CFP (in Chinese) on the conference website, after which 58 submissions were received by the deadline from authors in a wide range of affiliations, including governments, NGOs, research institutions, universities, financial institutions, telecom operators, and companies. After receiving all submissions, we randomly assigned every reviewer with five papers, and every paper was reviewed by three reviewers. All submissions were assessed on their credibility of innovations, contributions, reference value, significance of research, language quality, and originality. We adopted a thorough and competitive reviewing and selection process which went in two rounds. We first invited the reviewers to have an initial review. Based on the comments received, 31 papers passed and the authors of these 31 pre-accepted papers made modifications accordingly. Moreover, 3 papers among those 31 pre-accepted ones were invited as keynote papers in sub-sessions of our conference. In the second round modified papers were reviewed again. Finally, 17 out of the total 58 submissions stood out and were accepted. The acceptance rate was around 29.3%.

The 17 papers contained in this proceedings cover a wide range of cyber-related topics, including cryptography, intrusion/anomaly detection, malware mitigation, systems security, social network security and privacy, access control, denial-of-service attacks and hardware security implementation, etc.

We hereby would like to sincerely thank all the authors for their participation, and our thanks also go to the Program Committee chair and members for their considerable efforts and dedication in helping us solicit and select the papers of quality and creativity.

At last, we humbly hope the proceedings of CNCERT 2020 will shed some light for forthcoming researchers in the research and exploration of their respective fields.

September 2020

Wei Lu  
Hanbing Yan

# Organization

## Program Committee

### Committee Chairs

Wei Lu CNCERT/CC, China  
Hanbing Yan CNCERT/CC, China

### Committee Members

Yang Zhang CISPA Helmholtz Center for Information Security,  
Germany  
Zhenkai Liang National University of Singapore, Singapore  
Guoai Xu Beijing University of Posts and Telecommunications,  
China  
Bo Lang Beihang University, China  
Purui Su Institute of Software, Chinese Academy of Sciences,  
China  
Weiping Wen School of Software and Microelectronics,  
Peking University, China  
Xinhui Han Institute of Computer Science and Technology,  
Peking University, China  
Haixin Duan Tsinghua University, China  
Chao Zhang Tsinghua University, China  
Senlin Luo School of Information and Electronics, Beijing Institute  
of Technology, China  
Hua Zhang Beijing University of Posts and Telecommunications,  
China  
Jiang Ming The University of Texas at Arlington, USA  
Min Yang Fudan University, China  
Baoxu Liu Institute of Information Engineering, Chinese Academy  
of Sciences, China  
Meng Xu Georgia Institute of Technology, USA  
Yongzheng Zhang Institute of Information Engineering, Chinese Academy  
of Sciences, China  
Huaxiong Wang Nanyang Technological University, Singapore  
Guojun Peng Wuhan University, China  
Qiang Wang Carleton University, Canada  
Xinguang Xiao Antiy Cooperation, China  
Chunhua Su University of Aizu, Japan  
Xueying Li Topsec Cooperation, China  
Kui Ren Zhejiang University, China

Yuanzhuo Wang	Institute of Computing Technology, Chinese Academy of Sciences, China
Wenling Wu	Institute of Software, Chinese Academy of Sciences, China
Feifei Li	Stanford University, USA
Stevens Le Blond	Max Planck Institute for Software Systems, Germany
Yaniv David	Technion, Israel
Siri Bromander	University of Oslo, Norway
Zoubin Ghahramani	University of Cambridge, UK
Li Ding	CNCERT/CC, China
Zhihui Li	CNCERT/CC, China
Tian Zhu	CNCERT/CC, China

# Contents

## Access Control

PassEye: Sniffing Your Password from HTTP Sessions by Deep Neural Network . . . . .	3
<i>Zhiqing Rui, Jingzheng Wu, Yanjie Shao, Tianyue Luo, Mutian Yang, YanJun Wu, and Bin Wu</i>	
Research on the Development Route of International Communication Accesses . . . . .	16
<i>Tianpu Yang, Junshi Gao, Xiaoming Chen, Yanchun Guo, and Shuo Sun</i>	

## Cryptography

A Secure Ranked Search Model Over Encrypted Data in Hybrid Cloud Computing . . . . .	29
<i>Jiuling Zhang, Shijun Shen, and Daochao Huang</i>	
Based on GAN Generating Chaotic Sequence . . . . .	37
<i>Xuguang Chen, Hongbin Ma, Pujun Ji, Haiting Liu, and Yan Liu</i>	
MinerGate: A Novel Generic and Accurate Defense Solution Against Web Based Cryptocurrency Mining Attacks . . . . .	50
<i>Guorui Yu, Guangliang Yang, Tongxin Li, Xinhui Han, Shijie Guan, Jialong Zhang, and Guofei Gu</i>	
Research on Industrial Internet Security Emergency Management Framework Based on Blockchain: Take China as an Example . . . . .	71
<i>Haibo Huang, Yuxi Gao, Min Yan, and Xiaofan Zhang</i>	
Research Status and Prospect of Blockchain Technology in Agriculture Field . . . . .	86
<i>Dawei Xu, Weiqi Wang, Liehuang Zhu, and Ruiguang Li</i>	

## Denial-of-Service Attacks

Practical DDoS Attack Group Discovery and Tracking with Complex Graph-Based Network . . . . .	97
<i>Yu Rao, Weixin Liu, Tian Zhu, Hanbin Yan, Hao Zhou, and Jinghua Bai</i>	

**Hardware Security Implementation**

Research on the Remote Deployment Design of OTN Electrical Racks . . . . . 117  
*Tianpu Yang, Junshi Gao, Haitao Wang, Guangchong Dai, and Rui Zhai*

**Intrusion/Anomaly Detection and Malware Mitigation**

An Effective Intrusion Detection Model Based on Pls-Logistic Regression  
with Feature Augmentation. . . . . 133  
*Jie Gu*

DeepHTTP: Anomalous HTTP Traffic Detection and Malicious Pattern  
Mining Based on Deep Learning. . . . . 141  
*Yuqi Yu, Hanbing Yan, Yuan Ma, Hao Zhou, and Hongchao Guan*

**Social Network Security and Privacy**

A Label Propagation Based User Locations Prediction Algorithm  
in Social Network . . . . . 165  
*Huan Ma and Wei Wang*

Personalized Differentially Private Location Collection Method  
with Adaptive GPS Discretization . . . . . 175  
*Huichuan Liu, Yong Zeng, Jiale Liu, Zhihong Liu, Jianfeng Ma,  
and Xiaoyan Zhu*

**Systems Security**

Analysis on the Security of Satellite Internet . . . . . 193  
*Huan Cao, Lili Wu, Yue Chen, Yongtao Su, Zhengchao Lei,  
and Chunping Zhao*

A Survey on Cyberspace Search Engines . . . . . 206  
*Ruiguang Li, Meng Shen, Hao Yu, Chao Li, Pengyu Duan,  
and Lihuang Zhu*

Brief Introduction of Network Security Asset Management for Banks . . . . . 215  
*Yumo Wang and Qinghua Zhang*

Embedded Security-Critical Device Resource Isolation. . . . . 222  
*Xuguo Wang, Shengzhe Kan, and Yeli Xu*



**Author Index . . . . . 235**

# **Access Control**





# PassEye: Sniffing Your Password from HTTP Sessions by Deep Neural Network

Zhiqing Rui<sup>1</sup> , Jingzheng Wu<sup>1</sup> , Yanjie Shao<sup>1</sup>, Tianyue Luo<sup>1</sup>, Mutian Yang<sup>1,2</sup>, Yanjun Wu<sup>1</sup>, and Bin Wu<sup>1</sup>

<sup>1</sup> Institute of Software, Chinese Academy of Sciences, Beijing, China  
{zhiqing, jingzheng08, yanjie, tianyue, mutian, yanjun, wubin}@iscas.ac.cn

<sup>2</sup> Beijing ZhongKeWeiLan Technology, Beijing, China

**Abstract.** Passwords are the most widely used method for user authentication in HTTP websites. Password sniffing attacks are considered a common way to steal password. However, most existing methods have many deficiencies in versatility and automation, such as manual analysis, keyword matching, regular expression and SniffPass. In this paper, to better describe the problem, we propose a HTTP Sessions Password Sniffing (HSPS) attack model which is more suitable in HTTP environment. Furthermore, we propose PassEye, a novel deep neural networkbased implementation of HSPS attack. PassEye is a binary neural network classifier that learns features from the HTTP sessions and identifies Password Authentication Session (PAS). We collected 979,681 HTTP sessions from the HTTP and HTTPS websites for training the binary classifier. The results show that PassEye is effective in sniffing the passwords with an accuracy of 99.38%. In addition, several measures are provided to prevent HSPS attacks in the end.

**Keywords:** Password sniffing attack · Deep neural network · Website security · Network traffic analysis

## 1 Introduction

Password is a traditional identity authentication method [1]. However, this authentication method has many security problems, which has been criticized for a long time. Some more secure methods have been proposed for the same purpose, such as fingerprint, asymmetric key, 2-step verification, one-time password, but password is still the most widely used one due to its convenience, simplicity, and user habits. This gives attackers the opportunity to perform brute force attacks, password sniffing attacks and password reuse attacks. The widespread use of plain text password transmission and weakly encrypted password transmission in HTTP websites makes password sniffing attacks more easily.

This work was supported by National Key Research and Development Program of China (2017YFB0801900), National Natural Science Foundation of China (61772507) and the Key Research Program of Frontier Sciences, CAS (ZDBS-LY-JSC038).

© The Author(s) 2020

W. Lu et al. (Eds.): CNCERT 2020, CCIS 1299, pp. 3–15, 2020.

[https://doi.org/10.1007/978-981-33-4922-3\\_1](https://doi.org/10.1007/978-981-33-4922-3_1)

Traditional methods of password sniffing attacks include manual analysis, keyword matching, regular expression and automatic tool [2, 3], which can attack some HTTP websites. Session is the basic unit of communication between the client and the server in the HTTP protocol [4] including request and response messages. HTTP websites usually perform password authentication through sessions. Because of the diversity of websites, manual analysis is probably the most common and effective measure. For examples, attackers listen to the network traffic packets and search for PAS, Keyword matching is also fast and effective, but experiments show that it has a high false-positive rate, and regular expression is an upgraded version of the former two. Attackers can write several regular expressions to match the PAS of some websites. However, writing regular expressions for all websites is an impossible task. Therefore, some automatic password sniffing tools have been proposed, e.g., SniffPass [5] and Password Sniffer Spy [6]. These tools support some protocols, such as POP3, IMAP4, SMTP, FTP, and HTTP Basic authentication, and do not support password authentication in HTTP webpage form, resulting in low availability in HTTP website password sniffing attacks. Overall, the current methods have many deficiencies in terms of versatility and automation.

Currently, more and more websites use HTTPS protocol to protect the security of data transmission and prevent man-in-the-middle attacks, thereby greatly enhancing the security of websites. However, since the user may try to install the root certificate in the web browser due to the temptation of the attacker or the request of the network administrator, the attacker can track the user's web browsing request by setting a transparent proxy server. In this paper, we propose an HSPS attack model if an attacker can obtain unencrypted traffic logs of users browsing the web. We define PAS as a session containing a password authentication request message. And the attacker intends to sort out PAS for users, so that any website can be accessed from numerous of traffic logs.

To overcome the shortcomings of previous methods, we have developed a password sniffing attack tool based on deep neural networks, called PassEye. Firstly, PassEye takes the HTTP session as input and uses designed rules to extract the features from the HTTP session. Preprocessing feature data is required: getting the invalid items removed, and the feature data normalized and one-hot encoded. The correlation rate between each feature and the plaintext password feature can be calculated by XGBoost algorithm [7], and the features with high rates can then be selected. Secondly, the basic architecture of PassEye is a neural network. The loss function, the number of layers and neurons, and the activation function are elaborately designed to build the network. The selected feature data is used to train the neural network model. Finally, PassEye can perform password sniffing attacks on network traffic.

In the experiments, an approach was first designed to collect labeled training data. 979,681 HTTP sessions were collected as our raw dataset and 7,697 were labeled as PAS. Secondly, the designed feature extraction and selection methods of PassEye were used to collect features from the raw data. 58 features were extracted and the top 20 were selected for the subsequent training. Thirdly, python and TensorFlow are used to build a deep learning neural network for binary classification, and it was trained by using the selected data and features. Experimental results show that the accuracy, f1-score, precision and recall of PassEye reach 0.9931, 0.9931, 0.9932 and 0.9931 respectively, which successfully proves the superiority of PassEye.

In summary, our **contributions** are as follows:

- A new HSPS attack model is proposed for website traffic password sniffing in HTTP and HTTPS protocols.
- We design and implement PassEye, a practical HSPS attack tool based on deep neural networks.
- We also show that PassEye is effective deep neural networks in HSPS attack.

**Outline.** The rest of this paper is organized as follows. In Sect. 2, we provide background on password attacks, password sniffing attacks, and the application of neural network to network traffic classification. In Sect. 3, we define the HSPS attack model. In Sect. 4, we show the design of PassEye. In Sect. 5, we present an evaluation of PassEye. Finally, we provide conclusions and future work in Sect. 6.

## 2 Background

### 2.1 Password Attack

Due to the vulnerability of password authorization, password attacks have been the focus of many scholars. Current research on password authentication is mainly focus on the evaluation of password security [8] and the optimization of password guessing methods [8–12]. Traditional password guessing methods are based on dictionary, Markov model or probabilistic context-free grammar (PCFG) [11]. Melicher et al. [8] use a neural network for password guessing attacks for the first time, and the evaluation results show outstanding performance. Following Melicher’s work, neural network methods for password guessing have developed rapidly in recent years.

### 2.2 Password Sniffing Attack

Compared with password guessing attacks, there is little research on password sniffing attacks, since it does not have good versatility currently. In fact, it can directly capture plain text passwords from network traffic without guessing, which is more time-saving and of greater practical value. This is an motivation for our research.

There are four traditional methods of password sniffing attacks, such as manual analysis, keyword matching, regular expression, and automatic tools. These methods are also applicable to HTTP website attacks. Manual analysis is based on traffic dump tools (e.g. Wireshark, TcpDump) or man-in-the-middle (MITM) proxy tools (e.g. fiddle, Burp Suite, mitmproxy). Attackers manually search and filter the raw network traffic logs and find which packet contain plain passwords. This can be the most common and effective method due to the complexity of websites. Keyword matching is fast, which uses password’s keywords (e.g. ‘password’, ‘pwd’, ‘passwd’) to match the content of the network traffic. However, experiments show that it has a high false positive rate. Compared with these methods, regular expression can bring more accurate results. According to the patterns of the site’s PAS, attackers can write regular expressions to match the usernames and passwords. However, since regular expressions are usually specifically

designed and do not support a wider range of websites, attackers need to learn the pattern of PAS for each website. Therefore, it is indeed a time-consuming method for attackers. Since SniffPass [5] and Password Sniffer Spy [6] are two automatic tools that support only a few patterns, such as POP3, IMAP4, SMTP, FTP, and HTTP basic authentication, and do not support password authentication in HTTP webpage form, their availability in HTTP website password sniffing attacks is quite low. In summary, current methods have many deficiencies in terms of versatility and automation.

### 2.3 Neural Network in Network Traffic Classification

The deep neural network has shown superior performance in software developing and analysis [13, 14] and has been widely used for classification and detection of network traffic logs in recent years [15, 16], such as attack detection, traffic content classification. Liu et al. [16] use a two-step neural network, Payload Locating Network and Payload Classification Network, for web attack detection, and the precision of the evaluation results reaches 100%.

Traffic content type identification is also an important application of deep neural networks [17]. Lotfollahi et al. [18] take the extracted first 20 bytes of the IP header, first 20 bytes of the TCP/UDP header and first 1460 bytes of payload as the input to the CNN/SAE deep neural network classification model, and the classification precision for Tor, webpage, audio and video reaches 95%.

## 3 Attack Model

The ultimate goal of a passive attacker in a traditional password sniffing attack is to intercept the user’s password. The attack model is shown in Fig. 1.

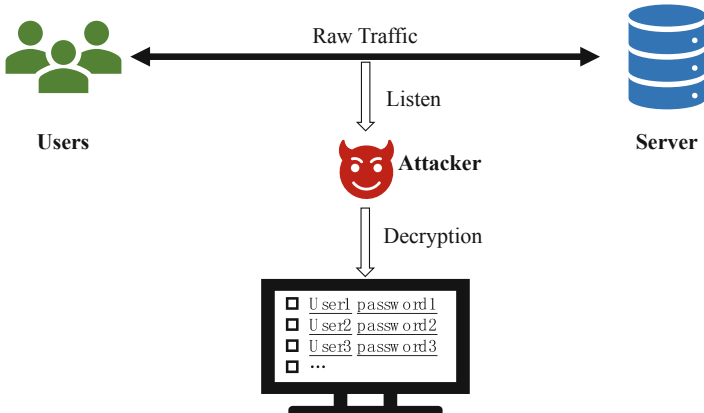


Fig. 1. Traditional password sniffing attack model.

HTTPS is a common and effective method to prevent MITM attacks on websites. Mi et al. [19] analyze the insecurity of IP proxy and Krombholz et al. [20] reveal the

problems encountered by HTTPS in practice. Their research shows that HTTPS is not absolutely secure. In addition to the above work, there are many attack methods for MITM attacks in HTTPS. Users may be tempted to install a root certificate in a web browser, and the attackers can set a transparent proxy server to track users' web browsing requests. DNS hijacking is also effective in HTTPS MITM attack.

In this paper, we propose an HSPS attack model, focusing on the classification of the HTTP sessions, and assuming that HTTPS traffic has been perfectly decrypted into HTTP sessions. Figure 2 shows the HSPS model. Users use browsers to surf the internet, and the browsers send HTTP/HTTPS requests to the webserver and receive the response.

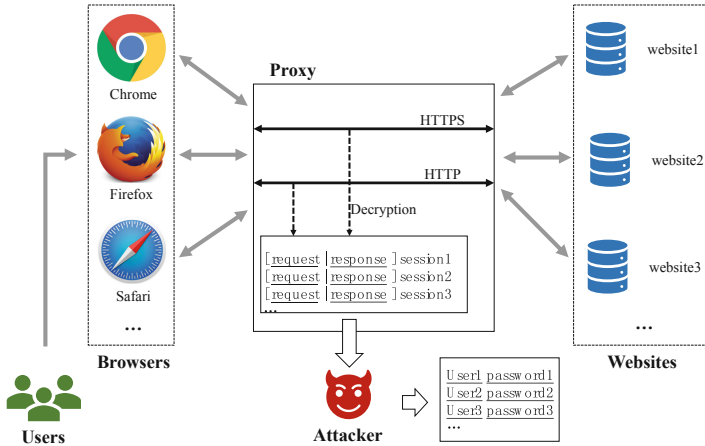


Fig. 2. HTTP session password sniffing (HSPS) attack model.

In the process of messages transposition, there is a transparent proxy that can perfectly decrypt HTTPS traffic and parse the traffic into request and response. For some reason, attackers can monitor the decrypted HTTP sessions. The goal of the attackers is to filter out PAS, and then parse the user's password information as much and as accurately as possible from a large amount of HTTP sessions.

## 4 PassEye Design

Due to the lack of versatility of previous methods, this paper proposes PassEye, a password sniffing attack tool based on deep neural networks, which can steal password in numerous HTTP traffic logs.

### 4.1 Overview

Figure 3 shows an overview of the PassEye design. The input of the PassEye was the HTTP sessions containing request and response messages. Then we designed a feature extraction method that could extract feature data from the HTTP sessions, which was

helpful for PAS. The invalid items in the feature data were removed, and the feature data was normalized and one-hot encoded. The correlation rate between each feature and the plaintext password feature was calculated using the XGBoost [7] features with high correlation were selected features with high correlation. After these steps, the HTTP sessions was transformed into feature vectors, which could be used to train the deep neural network model, PassEye, designed in this paper. The results show that this method can perform password sniffing attacks in the HTTP sessions.



**Fig. 3.** An overview of the PassEye.

### 4.2 Feature Extraction and Selection

Feature extraction is very important for neural network models. The more important the features can represent the PAS, the more accurate and generalized the machine learning model can be.

In this paper, a total of 21 plain passwords related features extracted from the HTTP sessions are listed in Table 1.

**Table 1.** Features extracted from the http session.

Name	Meaning	Type	Name	Meaning	Type
Session number	The session number in each record	Int	Response set cookie	Whether the response header has the 'Set Cookie'Field	Bool
Count pk	The occurrences of password keywords in the request message	Int	Response cookie len	The length of 'Set Cookie' field in the response header	Int
Count uk	The occurrences of username keywords in the request message	Int	Response code	Response status code	Enum

(continued)

**Table 1.** (continued)

Name	Meaning	Type	Name	Meaning	Type
Request content len	The length of the request content	Int	Time request	Time taken for the browser sending the request to the server	Float
Response content len	The length of the response content	Int	Time response	Time taken for the server sending the response to the browser	Float
Request header len	The length of the request header	Int	Time all	Time taken from the beginning of the request to the end of the response	Float
Response header len	The length of the response header	Int	Content type request	The list of content types in the request header 'Content-Type' field	List
Request header count	The number of the request header fields	Int	Content type Response	The list of content types in the response header 'Content-Type' field	List
Response header count	The number of the response header fields	Int	Content type accept	The list of content types in the request header 'Accept' field	List
Request cookie len	The length of the request header cookie field	Int	Is https	Whether this session uses HTTPS protocol	Bool
Request cookie count	The number of key-value pairs in the request header cookie field	Int			

It is worth mentioning that the “count pk” feature counts the number of times that the password keywords appear in the request messages. Password keywords are words with high frequency around passwords in statistics. In other words, we take the keyword matching method as a feature in PassEye method.

After the step of feature extraction, a HTTP session is abstracted into a list of features. To better show the correlation between discrete features and plain passwords, PassEye

uses one-hot encoding to convert discrete feature variables into multiple Boolean features. Z-score standardization is used to keep the features within a similar numerical range, which can be described as follows:

$$z = \frac{x - \mu}{\sigma}$$

where  $x$  denotes the eigenvalue to be normalized,  $\mu$  denotes the arithmetic mean,  $\sigma$  denotes the standard deviation of the feature, and  $z$  denotes the target value.

To quantify the impact of each feature on the plain password, PassEye calculates its correlation using the XGBoost algorithm. The top  $k$  of the above features are selected.

Through the above steps, we can obtain a  $1 * k$  feature vector, which can be used as the input of the neural network. The feature vector can well keep the information of the session itself and its correlation with the PAS, so that the performance of the neural network can be improved.

### 4.3 Neural Network Model

Our model consists of 5 layers, including an input layer, 3 hidden layers, and an output layer. The input layer has  $k$  nodes, which correspond to the feature vectors on a one-to-one basis. The three hidden layers contain 5 nodes, 5 nodes, and 1 node, respectively. The activation function in hidden layers 1 and 2 is ReLU, while that in hidden layer 3 is Sigmoid. The output layer has 2 nodes, corresponds to the two classification results: PAS and non-PAS. The optimizer is Adam, the learning rate is 0.001, and the loss function is Binary Cross Entropy.

During the training process, the random value of the initialization of the model weights ranges from  $-1$  to  $1$ . The batch size is 32, the number of steps per epoch is 100, and the maximum epoch is 10,000. An early stop condition is set to prevent over-fitting. The training will stop if the model does not show any improvement in 20 consecutive epochs.

## 5 Evaluation

We demonstrate the effectiveness of PassEye by answering the following questions:

- Q1.** Does PassEye perform better than keyword matching and regular expression methods?
- Q2.** What are the characteristics of PassEye compared to traditional methods in HSPS attacks?

### 5.1 Environment Setup

The hardware environment and main softwares are as followed.

Hardware: (1) CPU: Intel E7 4809v4 \* 2; (2) Memory: 128G; (3) Disk: 8T SSD.  
 Software: (1) OS: Ubuntu Linux 18.04 LTS; (2) Python 3.6.9; (3) TensorFlow 1.0; (4) XGBoost 0.9.0; (5) Docker 19.03.2; (6) mitmproxy 4.0.4; (7) Selenium 141.0; (8) Chrome 78.0.3904.108.



## 5.2 Dataset

We designed a new approach to collect labeled training data: using selenium and chrome to simulate browsing and logging into a website, and then using mitmproxy as a middle-man proxy to collect HTTP traffic logs. The experiment target site was Alexa China’s top 500 website [21]. 224 of the sites use HTTPS while 276 do not. We wrote a script for each of these websites, and several browsing and login records could be captured by executing each script. Each record generated a set of usernames and passwords randomly as well as several HTTP sessions. As a result, a total of 43,619 records (with corresponding usernames and passwords) and 979,681 HTTP sessions were collected as our raw dataset. Text search was used to see if the plaintext password corresponding to the HTTP sessions exists in the request message of the sessions, and a PAS was labeled when the answer was yes. In the end, 7,697 PAS were obtained. sample set: Due to the disparity in proportion between PAS and non-PAS samples, we used weighted random sampling to select a subset from the raw dataset as the sample set for training, which contains 5,875 PAS and 11,058 non-PAS.

We then divided the sample set into a training set, a validation set, and a test set at a ratio of 0.64:0.16:0.20.

## 5.3 Effectiveness of PassEye

We then extracted features from the training set described in PassEye Design. After one-hot encoding, 58 features were collected. XGBoost was used to calculate the correlation of the features, and the top 20 were selected to train the deep neural network, as shown in Fig. 4.

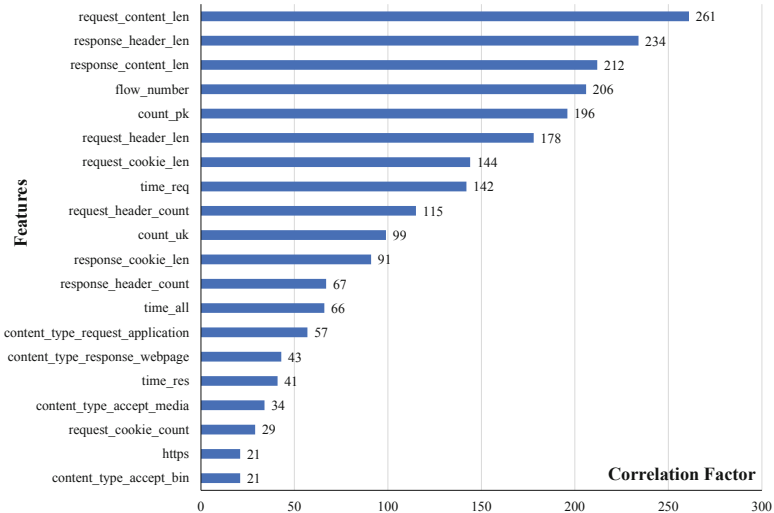


Fig. 4. Correlation of the top 20 features.

The training and validation sets were used to train the machine learning model described in PassEye Design. The test set was used to test the performance of the trained model.

For comparison, we also implemented keyword matching and regular expression methods as our baselines, and the test set was the same one.

### Performance of PassEye

Table 2 shows the accuracy, precision, recall, and f1-score results for the three methods. As can be seen from the table, all the performance metrics of PassEye are over 99.2%. Furthermore, all the metrics of PassEye are the highest, followed by the regular expression, and the performance of the keyword matching is the worst. It can be concluded that PassEye significantly surpasses these traditional methods in terms of the performance.

**Table 2.** The performance of the three methods.

Method	Accuracy	Precision	Recall	F1-score
Keyword matching	81.87%	85.67%	82.92%	81.65%
Regular expression	97.40%	96.50%	97.89%	97.13%
PassEye	<b>99.38%</b>	<b>99.46%</b>	<b>99.20%</b>	<b>99.33%</b>

### Characteristics of PassEye

Table 3 shows the characteristics of different password sniffing attack methods. Manual analysis, keyword matching, regular expression, and SniffPass, which can be seen as the representative of automatic tool, are presented in it for comparison, along with PassEye. The evaluation metrics include automaticity, versatility, scalability, independence, fastness, and robustness. Automaticity refers to whether it can run completely automatically without human intervention. Versatility refers to whether it can be used on any website. Scalability refers to whether the method supports extensions for use on new websites. Independence refers to whether this method can perform password sniffing attacks independently. Fastness refers to whether the method can run fast enough. Robustness refers to whether the method is effective enough in the face of unknown situations.

As can be seen from Table 3, PassEye has the characteristics of automaticity, versatility, scalability, fastness and robustness, except for independence. All other methods are not robust. Despite that SniffPass owns the independence, PassEye is still the best choice after the comprehensive consideration of all characteristics.

Therefore, it can be summarized that PassEye has the best characteristics among all these methods.

**Table 3.** The characteristics of different password sniffing attack methods.

	Manual analysis	Keyword matching	Regular expression	SniffPass	PassEye
Automaticity	×	✓	✓	✓	✓
Versatility	✓	×	×	×	✓
Scalability	×	✓	✓	×	✓
Independence	×	×	×	✓	×
Fastness	×	✓	✓	✓	✓
Robustness	×	×	×	×	✓

## 5.4 Discussion

It can be seen from the experiment results that PassEye has brilliant performance and best characteristics compared with some other traditional methods.

In the experiments, we also calculated the correlation between each feature and whether it is PAS. The correlation is shown in Fig. 4. The figure shows that the five features that have the most influence on the classifier are request\_content\_len, response\_header\_len, response\_content\_len, session\_number and count\_pk. This has given us some implications for preventing against HSPS attacks.

To prevent HSPS attacks, websites can make the following changes to the above features:

- Randomly change the length of the request content, the length of the response header, and the length of the response content by padding arbitrary characters.
- Have several unrelated random sessions between the browser and the server before the former sending the password authentication request messages to the latter. The goal is to change the session number. – Obfuscate and encrypt the fields of PAS.

In addition, there are some conventional ways to prevent password sniffing attacks. Websites can asymmetrically encrypt or hash passwords before sending login requests. Using self-built algorithms to obfuscate the content of requests is also an effective way. Changing the way of password authentication can be a solution as well, such as using one-time password, 2-step verification, etc.

## 6 Conclusion and Future Work

This paper proposed an HSPS attack model, which is a perfect expression for the problem of website password sniffing. PassEye, a tool based on deep neural networks, was proposed to implement the attack. We also designed a feature extraction method for HSPS attacks. In Evaluation, the experiment results verified the effectiveness of PassEye and deep neural networks in HSPS attacks. Some prevention strategies for websites were provided as well.

In the future, we will explore methods to make PassEye more robust, such as CNN, RNN or other machine learning models. The classification of obfuscated and hashed passwords can be considered and added to make PassEye more practical.

## References

1. Wang, D., Wang, P., He, D., Tian, Y.: Birthday, name and bifacial-security: understanding passwords of Chinese web users. In: 28th USENIX Security Symposium (USENIX Security 19), pp. 1537–1555. USENIX Association, Santa Clara (2019)
2. Jammalamadaka, R.C., Van Der Horst, T.W., Mehrotra, S., Seamons, K.E., Venkasubramanian, N.: Delegate: a proxy based architecture for secure website access from an untrusted machine. In: 2006 22nd Annual Computer Security Applications Conference (ACSAC 2006), pp. 57–66. IEEE, Miami Beach (2006)
3. Password Sniffing Attack. In: SSH.COM (2020). <https://www.ssh.com/attack/password-sniffing>. Accessed 3 Dec 2019
4. Mozilla: a typical HTTP session. In: MDN Web Docs (2019). <https://developer.mozilla.org/en-US/docs/Web/HTTP/Session>. Accessed 20 Oct 2019
5. SniffPass Password Sniffer - Capture POP3/IMAP/SMTP/FTP/HTTP passwords. In: NirSoft. [https://www.nirsoft.net/utills/password\\_sniffer.html](https://www.nirsoft.net/utills/password_sniffer.html). Accessed 22 Oct 2019
6. SecurityXploded: Password Sniffer Spy : Free Tool to Sniff and Capture HTTP/FTP/POP3/SMTP/IMAP Passwords (2020). <https://www.SecurityXploded.com>. Accessed 1 Jan 2020
7. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD 2016, pp. 785–794. ACM Press, San Francisco (2016)
8. Melicher, W., et al.: Fast, lean, and accurate: modeling password guessability using neural networks. In: 25th USENIX Security Symposium (USENIX Security 16), pp. 175–191. USENIX Association, Austin (2016)
9. Hitaj, B., Gasti, P., Ateniese, G., Perez-Cruz, F.: PassGAN: A Deep Learning Approach for Password Guessing. [arXiv:170900440](https://arxiv.org/abs/1709.00440) [cs, stat] (2017)
10. Pal, B., Daniel, T., Chatterjee, R., Ristenpart, T.: Beyond credential stuffing: password similarity models using neural networks. In: 2019 IEEE Symposium on Security and Privacy (SP), pp. 417–434. IEEE, San Francisco (2019)
11. Liu, Y., et al.: GENPass: a general deep learning model for password guessing with PCFG rules and adversarial generation. In: 2018 IEEE International Conference on Communications, ICC 2018, May 20, 2018–May 24, 2018. Institute of Electrical and Electronics Engineers Inc., p Cisco; et al.; Huawei; National Instruments; Qualcomm; Sprint (2018)
12. Muliono, Y., Ham, H., Darmawan, D.: Keystroke dynamic classification using machine learning for password authorization. *Proc. Comput. Sci.* **135**, 564–569 (2018). <https://doi.org/10.1016/j.procs.2018.08.209>
13. Duan, X., et al.: VulSniper: focus your attention to shoot fine-grained vulnerabilities. In: Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, Macao, China, pp. 4665–4671 (2019)
14. Yang, M., Wu, J., Ji, S., Luo, T., Wu, Y.: Pre-Patch: find hidden threats in open software based on machine learning method. In: Yang, A., et al. (eds.) SERVICES 2018. LNCS, vol. 10975, pp. 48–65. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-94472-2\\_4](https://doi.org/10.1007/978-3-319-94472-2_4)
15. Prasse, P., Machlica, L., Pevny, T., Havelka, J., Scheffer, T.: Malware detection by analysing network traffic with neural networks. 2017 IEEE Security and Privacy Workshops (SPW), pp. 205–210. IEEE, San Jose (2017)

16. Liu, T., Qi, Y., Shi, L., Yan, J.: Locate-then-detect: real-time web attack detection via attention-based deep neural networks. In: Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, Macao, China, pp. 4725–4731 (2019)
17. Yao, Z., et al.: Research review on traffic obfuscation and its corresponding identification and tracking technologies. *Ruan Jian Xue Bao/J. Softw.* **29**(10), 3205–3222 (2018). (in Chinese). <http://www.jos.org.cn/1000-9825/5620.htm>
18. Lotfollahi, M., Zade, R.S.H., Siavoshani, M.J., Saberian, M.: Deep packet: a novel approach for encrypted traffic classification using deep learning. [arXiv:170902656](https://arxiv.org/abs/170902656) [cs] (2018)
19. Mi, X., et al.: Resident evil: understanding residential IP proxy as a dark service. In: 2019 IEEE Symposium on Security and Privacy (SP), pp. 1185–1201. IEEE, San Francisco (2019)
20. Krombholz, K., Busse, K., Pfeffer, K., Smith, M., von Zezschwitz, E.: “If HTTPS were secure, i wouldn’t need 2FA” - end user and administrator mental models of HTTPS. In: 2019 IEEE Symposium on Security and Privacy (SP), pp. 246–263. IEEE, San Francisco (2019)
21. Alexa China Siterank. <http://www.alexa.cn/siterank>. Accessed 28 Nov 2019

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





# Research on the Development Route of International Communication Accesses

Tianpu Yang<sup>(✉)</sup>, Junshi Gao, Xiaoming Chen, Yanchun Guo, and Shuo Sun

China Mobile Group Design Institute Co., Ltd., Beijing 10080, China  
yangtianpu@cmdi.chinamobile.com

**Abstract.** With the implementation of China's Belt and Road Initiative, a new wave of globalization is taking shape, promoting the growth of international service requirements, which requires pre-deployment of international infrastructure. The construction of international communications infrastructure is an important guarantee for China's major international activities, external communication activities, and the normal operation of global and regional economies. International Communication Accesses is an important part of international infrastructure. The development and construction of international accesses is not an intrinsic mode, which involves many factors. It needs long-term planning and local adaptation; it relies on both the policy environment and basic network resources; it should consider both return on investment and convenience services. This document puts forward the future construction route of international communication accesses based on the analysis of factors including macro policies, geographical environments, service requirements, circuit quality improvement, transmission resources, fund support, and security assurance.

**Keywords:** International communication access · Channel access · International submarine cable · Cross-border terrestrial cable

## 1 Background

With the implementation of the Belt and Road Initiative, the new wave of China's globalization is developing continuously, accelerating the interaction and integration between China and other countries. In terms of personnel mobility, the number of Chinese outbound personnel in 2018 soared to a new height. The number of outbound personnel from the Chinese mainland reached 150 million, an increase of 14.5% over the previous year. By the end of 2018, more than 27,000 Chinese investors have established approximately 43,000 foreign direct investment enterprises in 188 countries (regions), and the investment from China covers more than 80% of countries (regions). China has set up more than 10,000 international enterprises in countries (regions) along the Belt and Road.

The construction of international communications infrastructure is an important guarantee for China's major international activities, external communication activities, and the normal operation of global and regional economies. Therefore, it is an indispensable prerequisite for responding to China's Belt and Road Initiative, serving Chinese enterprises, and supporting China's globalization.

## 2 Current Situation of International Communication Networks

International communication networks mainly consist of international communication infrastructure, which includes some points and lines. The points include the international communication accesses inside China and the international nodes outside China. The lines include international submarine cables and cross-border terrestrial cables.

### 2.1 International Communication Access

International communication accesses shall include international communication channel accesses (channel access for short), international communication service accesses (international access for short), and border international communication accesses. International accesses are service transfer points between national communication service networks and international communication service networks. They are mainly used to implement service interconnection and data exchange between the communications networks of operators from the Chinese mainland and the communications networks of foreign operators and operators in Hong Kong, Macao, and Taiwan. The international accesses can effectively shorten optical cable routes, thereby reducing the international circuit delay and improving circuit security. Since international accesses transmit cross-border information, they need to be supervised by government departments. Currently, international accesses in China are mainly constructed by China Telecom, China Mobile, and China Unicom. Up to now, China has set up 11 international accesses distributed in Beijing, Shanghai, Guangzhou, Kunming, Nanning, Urumqi, Hohhot, Fuzhou, Xiamen, Harbin, and Shenzhen.

The services transferred by international accesses include voice, public Internet, data private line, and international Internet transfer services. Since voice and public Internet services are strictly supervised, and the government approval procedure is complex, only Beijing, Shanghai, and Guangzhou are full-service international accesses, and others are data private line or Internet transfer accesses.

Channel accesses are transfer points between national communications transmission channels and international communications transmission channels. Therefore, they are mainly located at international submarine cable landing stations or cross-border terrestrial cable access equipment rooms.

### 2.2 International Submarine Cable Construction

The ocean covers 71% of the earth's surface, and there is no land between the Oceania, the American continent, and the Eurasia-Africa continent. Only 44 of the nearly 200 countries around the world do not have coastlines. More than 95% of the global international communication traffic is transmitted through submarine optical cables. After years of construction, submarine optical cables routed from China can be directly connected to North America, Asia, Europe, and Africa, and can be transferred to South America, Africa, and Oceania. China has implemented direct network interconnection with key countries including the United States, Japan, Singapore, and UK. By the end of 2018, five international submarine cable landing stations have been established in the Chinese

mainland, including Qingdao, Shanghai Nanhui, Shanghai Chongming, Shanghai Lingang, and Shantou, and two submarine cable landing stations connecting to Taiwan have been established in Fuzhou and Xiamen. In addition, Chinese operating enterprises have established international submarine cable landing stations in Tseung Kwan O and Chung Hom Kok of Hong Kong. There are nine international submarine cables landed on the Chinese mainland. China's telecommunications operating enterprises have a bandwidth of over 40 Tbit/s on submarine cables, and are constructing and planning a batch of projects. In the direction to the US, there are TPE and NCP. In the direction to Southeast Asia, there are APG, SJC, APCN2, EAC, and C2C. In the direction to Europe there are SMW3 and FLAG.

### 2.3 Cross-Border Terrestrial Cable Construction

China borders 14 countries. In addition to international submarine cable construction, cross-border terrestrial cable construction is also indispensable, like the Silk Road Economic Belt and the 21st-century Maritime Silk Road. Cross-border terrestrial cables function as the Silk Road in international communications to connect neighboring countries and lead to Europe and the African continent through neighboring countries. Currently, China has 17 international terrestrial cable border stations, including Horgos, Alashankou, Manzhouli, Pingxiang, and Ruili. It has established cross-border terrestrial cable systems with 12 neighboring countries except Bhutan and Afghanistan, and the system bandwidth exceeds 70 Tbit/s.

## 3 Factors Affecting International Access Establishment

To establish international accesses, the following factors need to be considered: policy environment, geographical environment, necessity of the establishment, and whether the conditions for the establishment are present.

### 3.1 Policy Environment

The policy environment is the major factor to be considered in the establishment of international accesses, because the establishment should be supported by relevant laws and regulations and policies, including country-level macro-policies and local policies.

Country-level macro-policies are divided into two types: relatively broad strategic policies, including the Belt and Road Initiative, continuous reform of international free trade areas, and establishment of the Guangdong-Hong Kong-Macao Greater Bay Area and Xiong'an New Area; closely-related industry policies, including the *Outline of the National Information Development Strategy* issued by the General Office of the State Council and the *Information and Communication Industry Development Plan (2016–2020)* issued by the Ministry of Industry and Information Technology.

Local policies are more specific measures formulated by local governments based on national strategies and related policies, such as the *13th Five-Year Development Plan for the Information and Communications Industry in XX Province*.



### 3.2 Geographical Environment

The geographical environment is an important reference condition for establishing international accesses. The purpose of establishing international accesses is to transfer and supervise international communication services, for which the most important thing is stable communication and easy construction and maintenance. Therefore, when establishing an international access, you need to consider both the geographical location and natural conditions of the selected city, including the risks of natural disasters such as earthquakes and floods. The requirements for geographically selecting the city vary from international access to international access. For example, to establish a global international access, select a location far away from the existing international accesses in Beijing, Shanghai, and Guangzhou, unless the existing international accesses are no longer capable of sustainable development. Regional international accesses should be deployed in regional centers or border provincial cities. In this way, regional international accesses can effectively work with channel accesses to reduce the circuit transfer delay. Therefore, regional international accesses should be deployed based on existing or future international submarine cables and cross-border terrestrial cables.

### 3.3 Necessity of International Access Construction

Establishing international accesses aims to meet international service requirements and improve the quality of international circuits. Service requirements drive the establishment of international accesses. Improving circuit quality is an important way to improve the competitiveness of international communications operators and ensure customer resources.

**Service Requirements.** International service requirements are the important prerequisite for establishing an international access. In other words, it is necessary to establish an international access only when international service requirements are sufficient. The measurement of international services mainly includes the significance and volume of international services. As to significance, mainly consider whether the regions where the international services pass through have great strategic significance to China. As to service volume, check whether the service volume of the existing international accesses reaches the upper limit, whether the current international services will continue to grow in the next few years, and how the business revenue is.

**Quality Improvement.** The improvement of international circuit quality includes the reduction of circuit delay and the improvement of circuit security. Under the current technical conditions, the only way to reduce the circuit delay is to reduce the length of optical cable routes. To improve circuit security, optical cables and circuit transmission protection technologies should be used. Setting proper international accesses can effectively reduce international circuit route diversion, thereby reducing the delay. In addition, the reduction of the optical cable route length can also reduce the probability of optical cable interruption and improve circuit security.

### 3.4 Feasibility of International Access Construction

**Transmission Resources.** International accesses are mainly used for international circuit transfer and switching. They need to connect to national circuits and outbound circuits. Therefore, the selected international accesses must have abundant national transmission resources and international transmission resources. For national transmission resources, inter-province backbone transmission networks, intra-province backbone transmission networks, and networks with multiple outgoing optical cable routes are preferred. For international transmission resources, regions that have international submarine cables or cross-border terrestrial cables and international transmission resources reaching the Europe, Asia, and North America are preferred. National transmission resources are important carriers of international circuits (consisting of pass-through circuits and cross-border circuits) inside a country. International transmission resources are important guarantee for overseas transmission of international circuits.

**Fund Support.** The construction of an international access has high requirements on equipment room conditions. Generally, the equipment room must be owned by an enterprise. In addition, the power supply and subsequent communication assurance must meet the highest standards. Therefore, the investment is high. On the other hand, due to the restrictions of national supervision conditions, the international access site needs to reserve the corresponding equipment room location and power supply for the supervision equipment, which increases the enterprise cost. Therefore, fund guarantee is also an important condition for the selection and construction of the international access.

### 3.5 Security Assurance Measures

Generally, the international access provides security assurance based on the highest level in the communications industry. Therefore, during international access construction, a reasonable and effective network and information security assurance system needs to be formulated, dedicated personnel need to be specified for ensuring network and information security, and a complete and feasible network and information security assurance technical solution needs to be formulated.

## 4 International Access Establishment Case

Based on the preceding factors and process, China Mobile chose the Urumqi international access to analyze the process of establishing a regional international access.

### 4.1 Environment Factor

**National Policy.** In March 2015, the National Development and Reform Commission, Ministry of Foreign Affairs, and Ministry of Commerce of the People's Republic of China jointly released the Vision and Actions on Jointly Building Silk Road Economic Belt and 21st-Century Maritime Silk Road: We should make good use of Xinjiang's

geographic advantages and its role as a window of westward opening-up to deepen communication and cooperation with Central, South and West Asian countries, make it a key transportation, trade, logistics, culture, science and education center, and a core area on the Silk Road Economic Belt. On land, the Initiative will focus on jointly building a new Eurasian Land Bridge and developing China-Mongolia-Russia, China-Central Asia-West Asia and China-Indochina Peninsula economic corridors by taking advantage of international transport routes, relying on core cities along the Belt and Road and using key economic industrial parks as cooperation platforms. The China-Pakistan Economic Corridor and the Bangladesh-China-India-Myanmar Economic Corridor are closely related to the Belt and Road Initiative, and therefore require closer cooperation and greater progress. The Ministry of Industry and Information Technology (MIIT) has formulated the Information and Communications Industry Development Plan (2016–2020). The Plan specifies that the China-Russia, China-ASEAN, China-South Asia, and China-Central Asia cross-border terrestrial cable construction should be particularly considered, China will continue the cross-border optical cable construction with neighboring countries, establish direct cross-border optical cables with countries and regions where conditions permit based on business development, expand and optimize the existing cross-border systems if possible, and explore cross-border transfer.

**Geographical Conditions.** Xinjiang borders Russia, Kazakhstan, Kyrgyzstan, Tajikistan, Pakistan, Mongolia, India, and Afghanistan. The capital city Urumqi has a relatively good economic, humanistic, and natural environment, and can effectively ensure the construction and operation of the international access.

## 4.2 Necessity Analysis

**Service Requirements.** According to China Mobile's prediction on the service development in Central Asia and Southwest Asia, by 2021, China Mobile will have 2,750 Gbit/s international services passing from Urumqi to Central Asia and Southwest Asia, including 2,490 Gbit/s private line services (including pass-through services) and 260 Gbit/s Internet transfer services. It mainly provides data private lines for Chinese enterprises' go-global, enterprises that invest in China, and operators and enterprises in Central Asia and Southwest Asia.

**Circuit Quality Improvement.** Currently, the international data private lines to Central Asia and Southwest Asia are transferred through the Beijing or Guangzhou international access, and then to Central Asia and Southwest Asia through Urumqi/Korla. For services from Xinjiang and surrounding provinces to Central Asia and Southwest Asia, 15 hops are required for one round trip of a circuit, the transmission distance increases by about 6000 km, and the delay increases by about 60 ms, making the cost of circuit construction, operation, and maintenance high (Fig. 1).

After the Urumqi international access is set up, services in the western region can be transmitted to Central Asia and Southwest Asia directly through Urumqi, reducing hops by 15 and the delay by 60 ms, and saving the transmission investment. Services from Central Asia and South Asia to Europe can be transferred directly from Urumqi.



Fig. 1. International circuit route diagram for Central Asia and Southwest Asia

Urumqi has become an important transfer point for international services. This changes the layout of international service accesses in eastern China, southwest China, Beijing, Shanghai, Guangzhou, and Kunming, optimize networks, and improve network security.

### 4.3 Construction Feasibility Analysis

**Transmission Resources.** Urumqi is the backbone node of China Mobile’s international and government/enterprise private transport networks, the backbone node of China Mobile’s inter-province backbone transport network, and the backbone node of the provincial backbone network in Xinjiang Uygur Autonomous Region. The city has more than three outgoing optical cable routes. In addition, China Mobile has set up four channel accesses, including Alashankou, Horgos, Atushi, and Tashikuergan in Xinjiang, to connect Kazakhstan, Kyrgyzstan, and Pakistan. In a word, Urumqi has abundant national and international transmission resources and is suitable to set up an international access.

**Fund Support.** China Mobile is one of the world’s top 500 enterprises with strong capital strength and has sufficient funds to set up an international access in Urumqi.

Based on the preceding analysis, the construction of the Urumqi regional international access in meets the policy requirements. The geographical advantages are prominent. The service requirements are urgent. The circuit quality is improved obviously. The existing transmission resources are sufficient. The fund and security can be guaranteed. Therefore, it is feasible to set up the Urumqi international access.

## 5 International Access Development Recommendations

Based on China’s international communications development and current international communications infrastructure construction, international accesses should pay attention

to the balance and capability improvement in future development to promote the overall development of China's economy and society, maintain national network and information security, and serve the Belt and Road Initiative. Many other factors also need to be considered in the development of international accesses. The following are some suggestions proposed based on the current national economy, policy environment, and construction process.

**National Economic Belts Provide Prerequisites for International Access Development.** With the continuous development of economic globalization, China has continuously launched new economic circles to promote the development of regional economy and foreign trade. Since the Shanghai Free-Trade Zone (Shanghai FTZ) was listed in 2013, the number of China FTZs has reached 18 in six years. In the past six years, China's FTZs have developed from the eastern coast to the western inland and formed a Wild Goose Queue in China's opening-up. The purpose of establishing n FTZs is to build an open economy, further improve trade facilitation and even liberalization, and build core hubs for international trade and logistics, especially in coastal and border provinces of China. For example, Shandong mainly promotes the continuous conversion of old and new kinetic energy, promotes the development of marine economy with high quality, and deepens regional economic cooperation between China, Japan, and South Korea, and promotes the construction of new bases for the opening-up. Guangxi will, by deepening its open cooperation with ASEAN, promoting the construction of a new international land-sea trade channel, and exploring the development and opening-up of border areas, to form an important gateway for the economic integration of the Silk Road Economic Belt and the 21st Century Maritime Silk Road. Yunnan will cooperate with neighboring countries such as Vietnam, Laos, and Myanmar to build an important node that connects South Asia-Southeast Asia channels, and promote the formation of China's radiation center and opening-up frontier oriented to the South Asia and Southeast Asia.

China's establishment of domestic FTZs and continuous construction of new economic zones provide necessary preconditions and future planning direction for the development of international accesses. For example, since the Hainan FTZ was set up, based on Hainan's unique geographical location, submarine cables can be routed from Hainan to the Asia Pacific region, and the Hainan international access can be established to provide a new channel for submarine cables from the western China to the Asia Pacific region, reducing delay and providing communication assurance for foreign enterprises to invest in Hainan. Similarly, Guangxi, Shandong, Zhejiang, Jiangsu, and even the Northeast China are likely to become international accesses.

**International Terrestrial and Submarine Cables Promote Balanced Development of Regional International Accesses.** Regional international accesses are supplements to the global full-service international accesses such as Beijing, Shanghai, and Guangzhou. Developing regional international accesses is an effective means to reduce the international circuit delay and improve the circuit security, and can achieve the balanced development of domestic international accesses.

In addition to national macroeconomic factors, the development of regional international accesses needs to be based on the current construction of international submarine cables and cross-border terrestrial cables. The construction of regional international

accesses is significant only when national transmission resources and cross-border transmission system resources such as terrestrial and submarine cables are available. With the development of China, international submarine cables and cross-border terrestrial cables will be gradually constructed, and new channel accesses will also be gradually set up. At that time, regional international accesses will be set up based on actual service requirements and new network layout.

Take Zhuhai as an example. When the Hong Kong–Zhuhai–Macau Bridge is completed, the cross-bridge optical cables between Zhuhai, Hong Kong, and Macao are also deployed. Therefore, a regional international access can be set up in Zhuhai to cover the western cities of Guangdong. In this way, services from the western cities of Guangdong to Hong Kong and Macao will no longer need to be transferred to Guangzhou or Shenzhen, reducing the average optical cable routes by 200 km and the optical cable routes on Hong Kong–Macao circuits by 300 km, which improves the quality of international circuits (Fig. 2).



**Fig. 2.** Directions of the international private lines between western Guangdong, Hong Kong, and Macao

**The Capabilities of International Accesses Need to be Improved.** Currently, only Beijing, Shanghai, and Guangzhou are full-service international accesses, covering voice, public Internet, and data private line services. Other regional international accesses do not cover voice and public Internet services. With the deepening of reform and opening-up, the demand for international voice and Internet services will increase. Currently, Beijing, Shanghai, and Guangzhou international accesses as a whole have not improved their capability of carrying voice services, and their capability of carrying Internet services lags far behind the growth level of global Internet capacity. This will create a contradiction between demand growth and capability improvement. To solve this problem, on the one hand, the capacity of the existing international accesses can be expanded to improve the transfer and exchange capabilities; on the other hand, the service pressure of the existing international accesses can be shared by adding international accesses or expanding the service coverage of the existing regional international accesses.

As the capital of China, Beijing is the political center. With the migration of other functionalities of Beijing and the existing equipment room conditions of the international access, it is recommended that an international access be set up in the Xiong'an New Area

as the backup of the Beijing international access to gradually share or carry all services of the Beijing international access. With the continuous increase of Internet traffic, the regional international accesses of coastal or border provinces can be upgraded to increase the public Internet egress capability and reduce the network load in China.

## 6 Conclusion

The development and construction of international accesses is not an intrinsic mode, which involves many factors. It needs long-term planning and local adaptation; it relies on both the policy environment and basic network resources; it should consider both return on investment and convenience services. Therefore, we can put forward a more reasonable and accurate route in the future development of international accesses only by constantly tracking the new situation, new technologies, and new resources at home and abroad.

## References

1. Jie, Z.: Analysis of China's Current International Communications Construction and Development Suggestions, Telecommunications Technology, August 2013
2. China Academy of Information and Communications Technology, White Paper on China International Optical Cable Interconnection (2018), August 2018

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# **Cryptography**





# A Secure Ranked Search Model Over Encrypted Data in Hybrid Cloud Computing

Jiuling Zhang<sup>(✉)</sup>, Shijun Shen, and Daochao Huang

CNCERT/CC, Beijing 100029, China  
zhangjl@cert.org.cn

**Abstract.** The security issue is becoming more and more prominent since user's private information being outsourced to the somewhat untrustworthy cloud. Encrypting the information before uploading them to the cloud is one of ultimate solutions. Secure searchable encryption schemes and secure ranking schemes have been proposed to help retrieving the most relevant documents over the cloud. However the present methods are encumbered by the huge computing and communicating occupation of the cipher text. In this paper, a fully homomorphic encryption based secure ranked search model over the hybrid cloud is proposed. By introducing hybrid cloud, which typically composed by private cloud and public cloud, the high cost of computing and communicating of the cipher text is transferred to the trustworthy private cloud, in which the decrypting are performed. The client does not need to perform any heavy computations, thence making the secure ranking practical from the client's point of view.

**Keywords:** Secure ranked search · Fully homomorphic encryption · Hybrid cloud computing

## 1 Introduction

With the unprecedented growing of information, as well as the limited storage and computation power of their terminal, or the limited capacity of battery, the users are outsourcing more and more individual information to remote servers or clouds. However, since the public cloud are not fully trustworthy, the security of the information stored on the cloud could not be guaranteed. The security issue has attracted a variety of attentions in both the area of engineering and research. One solution to meet the needs of data outsourcing while preserving privacy is to encrypt them before storing them on the cloud, and this is one of the generally accepted ultimate methods. After the sensitive data are encrypted with some scheme, the cipher text of sensitive private information may be deemed as secure and could be outsourced to public cloud. However, once the data are encrypted into cipher text, the processing and utilizing of the cipher text information will be the subsequent problem that needs to be taken into consideration. With the accumulation of the information outsourced over the cloud, the collection will be so large that the retrieving of the encrypted form of information is the subsequent conundrum.

Several kinds of schemes have been proposed since the pioneering work of Song et al. [1], in which a cryptography scheme for the problem of searching on encrypted data is proposed. The proposed scheme is practical and provably secure, as from the query the server cannot learn anything more about the plain text. Although the scheme performs well over the linear search, it is almost impractical over the huge information retrieval scenario. In another work, Goh introduced a Bloom filter based searchable scheme with a complexity of the number of documents in the collection over the cloud [2], which is also not applicable in huge information retrieval scenario. Other work includes a secure conjunctive keyword search over the encrypted information with a linear communication cost [6], privacy-preserving multi-keyword ranked searches over encrypted cloud data [3, 10], and fuzzy keyword search over encrypted cloud [4].

In the general huge plaintext retrieval scenario, different pieces of information are organized in the form of documents. The information size stored in the cloud is very large, and the acquisition of the requested information should be implemented with the help of retrieval methods. A number of documents may contain a given query and if the query is searched, many documents may be retrieved. After the more or less relevant documents are retrieved, the ranking of them over the cloud computing is necessary. Actually in the large information retrieval scenario, the retrieved information should be ranked by the relevance scores between the document and the queries. This is due to that the number of documents contains a keyword or a multiple of keywords is so large that it is hard to obtain the most relevant documents from the client's point of view. The most relevant documents should be retrieved and given to the users. In plaintext retrieval, a similarity based ranking schemes named the locality sensitive hashing [7] and an inner product similarity to value the relevance between the query and the document are separately presented [4].

In huge collection cipher text retrieval, a one-to-many order-preserving mapping technique is employed to rank sensitive score values [13]. A secure and efficient similarity search over outsourced cloud data is proposed in [14]. There are also work exploring semantic search based on conceptual graphs over encrypted outsourced data [8]. Though the one-to-many order-preserving mapping design facilitates efficient cloud side ranking without revealing keyword privacy and there is no cost on the client's terminal, the precision of this model is lower than that over the plaintext scenario. There is a counterbalance between the accuracy of the results and the security as the statistical information is provided. There are also efforts utilizing the fully homomorphic encryption [5] to calculate the relevance scores between the documents and queries. However, since the encryption and decryption are all performed on the client's terminal, and they are also resource consuming, the time cost is also intolerable.

In order to solve the problem that enormous computation and communication emerged in the fully homomorphic encryption based ranking, the hybrid cloud [12] is introduced to employ. Hybrid cloud generally consists public cloud and private cloud. The public cloud is provided by the entrepreneur, and not fully trust worthy, while the private cloud belongs to the organization, and thus trustable. Hybrid cloud also described the architecture and cooperation among different cloud vendors, and gave solution on the communication, storage, and computation among different cloud [11]. Here, we make the assumption that there is at least one secure private cloud in the hybrid cloud. The

plain text information is handled over the trustworthy or private cloud. With the cooperation with other public clouds on which encrypted information are stored and processed, the secure ranking model is proposed.

This paper is organized as follows, the related work is reviewed in Sect. 2, and then a secure ranked search model over the hybrid cloud is introduced in Sect. 3. Some experiments are carried out in Sect. 4. Finally, a conclusion is drawn in Sect. 5.

## 2 Related Work

### 2.1 The Okapi BM25 Model Over Plain Text

In information retrieval, a document  $D$  is generally processed into a bag of words. The collection of documents is denoted by  $C$ . The documents and queries are generally preprocessed and stemmed, the index and inverted index are also built to facilitate further retrieval [9], the details are omitted here.

There are a variety of mature information retrieval models, which varies from the linear search to Boolean model to ranked vector space model (VSM) models. Different retrieval model applies in different scenarios. The ranking models are used most frequently for general purposes.

Okapi BM25 model [15] is one of the most popular ranking model for obtaining the relevance scores of documents and queries. In the Okapi BM25 model, the term frequency is defined by Eq. 1.

$$TF(q_i) = f(q_i, D) \quad (1)$$

While the inverse document frequency is given by Eq. 2.

$$IDF(q_i) = \log \frac{N}{n(q_i)} \quad (2)$$

In which  $f(q_i, D)$  means the occurrence frequency of  $n(q_i)$  in  $D$ .  $n(q_i)$  means the number of documents which contain  $q_i$ .

The Okapi relevance scores between a query and a document is given by Eq. 3.

$$\text{Score}(D, Q) = \sum_{q_i \in Q} TF(q_i) \times IDF(q_i) \quad (3)$$

The relevance between a document and a query is quantified by the Okapi relevance scores.

### 2.2 Fully Homomorphic Encryption

Homomorphism is a very valuable property of encryption algorithms, which means that the computation results over cipher texts corresponds to that of the computation over plaintext. Fully homomorphic encryption (FHE) is both additive homomorphic and multiplicative homomorphic, satisfying both the Eqs. 4 and 5.

$$D(E(a) \oplus E(b)) = a + b \quad (4)$$

$$D(E(a) \otimes E(b)) = a \times b \quad (5)$$

Where  $\oplus$  means the “addition” over the cipher text, while  $\otimes$  denotes the “multiplication” over the cipher text.

In this work, the term frequency TF and inverse document frequency IDF values are encrypted by FHE separately. The documents which contain the terms are encrypted by some other encryption scheme, such as AES, only to protect the information stored on the public cloud.

All the information is thence uploaded to the public cloud after encrypted by a certain encryption scheme. The cipher text of Score (D, Q) could also be obtained.

### 2.3 The Applicability of Hybrid Cloud Computing

We assume that the hybrid cloud is simply constructed by one private cloud and one public cloud. The private cloud stores the client’s sensitive information and the public cloud performs computation over cipher text information.

A new scheme based on the private and the public cloud platform is proposed here. The public cloud in this hybrid cloud scenario is assumed to have the following characteristics: the computing resource is very enormous, and the resource allocated to a client can be elastically provided in order to meet the client’s computation demands.

The private cloud actually acts as an agent for the client in the scenario. Since the computation and storage resources are relatively abundant over private cloud, it has enough computing power to just encrypt a user’s plaintext information. The bandwidth between the private cloud and the public cloud is also large enough to transfer the cipher text of the relevance scores.

## 3 Secure Ranked Search Model Over Hybrid Cloud

A new encryption based secure and efficient retrieval scheme over hybrid cloud is proposed in this section.

### 3.1 The Architecture of the Secure Ranked Search Model Over Hybrid Cloud

There are three parties in this architecture, the client, the private cloud, and the public cloud. As shown in Fig. 1.

In the building process, the client uploads original sensitive information to the private cloud, as shown by step (1) in Fig. 1. The private cloud preprocesses the documents, and encrypts the TF, IDF values and the document itself. The encrypted information are then uploaded to the public cloud, as shown by step (2). Over the public cloud, an inverted index is built, and a variety of corresponding computations are performed.

In the retrieval process, the client gives a certain keyword to the private cloud, as shown by step (3). The private cloud encrypts the word, and search over the public cloud, as shown by step (4). On the public cloud, the calculation over the cipher text is carried out. The cipher text of evaluation scores are downloaded by the private cloud, as shown

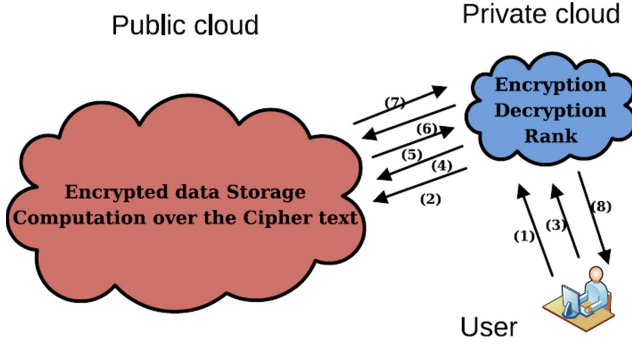


Fig. 1. Schematic diagram of secure ranked search model

by step (5). After the decryption, the scores are ranked, thence the top N document IDs are sent to the public cloud, as shown by step (6). Then the private cloud downloads the encrypted document, as shown by step (7). After decryption, the plaintext documents are given back to the clients, as shown by step (8).

### 3.2 The Implementation of Fully Homomorphic Encryption Based Secure Ranking Scheme

In the inverted index building process, the computation of encryption of the plain text are performed over the private cloud.

The encrypted form of term frequency is expressed as Eq. 6.

$$v_{tf} = (FHE(tf_1), FHE(tf_2), \dots, FHE(tf_N)) \tag{6}$$

The encrypted form of inverse document frequency is given as Eq. 7.

$$v_{idf} = (FHE(idf_1), FHE(idf_2), \dots, FHE(idf_N)) \tag{7}$$

In the ranking process, the computation such as the addition and multiplication over the cipher text are performed over the public cloud.

The full process can be described as the following, Firstly the TF and in decimal form are transformed into binary, then each of them is encrypted, the relevance is obtained after addition and multiplication. The process is shown in Fig. 2.

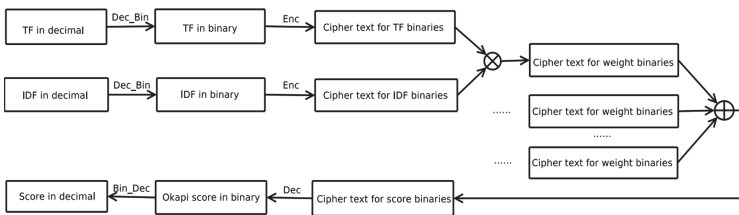


Fig. 2. Implementation of the FHE based ranking

The process of calculating relevance scores between the query and the document is given as Eq. 8.

$$\text{FHE}(\text{score}) = \sum_{q_i} \text{FHE}(\text{tf}_i) \times \text{FHE}(\text{idf}_i) \quad (8)$$

Thence the relevance scores in FHE form are obtained over the hybrid cloud. By decrypting them, the documents could be subsequently ranked.

## 4 Experiment Result and Future Work

### 4.1 Preliminary Experimental Result

Based on the proposed retrieval and ranking model over hybrid cloud, some preliminary experiments are carried out. The experiment utilized a small-sized Cranfield collection. The experimental result is compared with the order preserving scheme (OPE), which is employed in [13].

The precision of top  $N$  retrieved documents and the MAP [9] are used to evaluate different ranking schemes. The experimental result is shown in the following table (Table 1).

**Table 1.** The comparison result of different methods.

Metric	OPE	Okapi BM25
Map	0.283	0.416
P@5	0.310	0.440
P@10	0.223	0.310
P@20	0.150	0.199
P@30	0.117	0.150
P@50	0.082	0.103
P@100	0.050	0.059

The tentative experimental result demonstrates that the order preserving encryption based retrieval result is dramatically lower than that of the Okapi BM25 ranking models for the crucial P@N criteria.

### 4.2 Future Work

While retrieving, the proposed scheme needs the private cloud to download all cipher text of the relevance scores of possibly relevant documents, which also would be enormous. In order to make it more practicable, the future work may incorporate both the OPE and the FHE over the hybrid cloud. By OPE, a pre-rank could be performed over the public cloud, and give a top  $M$  relevance scores to private cloud. Here,  $M$  should be a

large enough number, say 10000. Then the private cloud then decrypts the top  $M$  scores and ranks them. By this way, both the computation and communication cost over the private cloud would be limited, the efficiency of retrieving and ranking will be greatly enhanced.

## 5 Conclusion

A fully homomorphic encryption based secure ranked search model over the hybrid cloud is proposed, the implementation of the retrieval and ranking process are described in detail. Experimental result shows its precedence over the existing purely OPE based ranking. In the future, we would incorporate both OPE and the FHE to implement industrial model while preserving user's privacy over the hybrid cloud.

**Acknowledgments.** This work is supported by The National Key Research and Development Program of China (2016YFB1000105).

## References

1. Song, D.X., Wagner, D., Perrig, A.: Practical techniques for searches on encrypted data. In: Proceeding 2000 IEEE Symposium on Security and Privacy, pp. 44–55 (2000)
2. Goh, E.-J.: Secure indexes. IACR Cryptology ePrint Archive, p. 216 (2003)
3. Curtmola, R., Garay, J., Kamara, S., Ostrovsky, R.: Searchable symmetric encryption: improved definitions and efficient constructions. *J. Comput. Secur.* **19**, 895–934 (2011)
4. Li, J., Wang, Q., Wang, C., Cao, N., Ren, K., Lou, W.: Fuzzy keyword search over encrypted data in cloud computing. In: 2010 Proceedings IEEE INFOCOM, pp. 1–5 (2010)
5. Gentry, C.: Fully homomorphic encryption using ideal lattices. In: STOC 2009
6. Golle, P., Staddon, J., Waters, B.: Secure conjunctive keyword search over encrypted data. In: ACNS (2004)
7. Kuzu, M., Islam, M.S., Kantarcioglu, M.: Efficient similarity search over encrypted data. In: 2012 IEEE 28th International Conference on Data Engineering, pp. 1156–1167 (2012)
8. Fu, Z., Huang, F., Sun, X., Vasilakos, A., Yang, C.: Enabling semantic search based on conceptual graphs over encrypted outsourced data. *IEEE Trans. Serv. Comput.* **12**, 813–823 (2019)
9. Manning, C.D., Raghavan, P., Schütze, H.: Introduction to Information Retrieval, 1st edn. Cambridge University Press, Cambridge (2005)
10. Vishvapathi, P., Reddy, M.J.: Privacy-preserving multi-keyword ranked search over encrypted cloud data (2016)
11. Wang, H., Ding, B.: Growing construction and adaptive evolution of complex software systems. *Sci. China Inf. Sci.* **59**, 1–3 (2016)
12. Wang, H., Shi, P., Zhang, Y.: JointCloud: a cross-cloud cooperation architecture for integrated internet service customization. In: IEEE 37th International Conference on Distributed Computing Systems, pp. 1846–1855 (2017)
13. Wang, C., Cao, N., Ren, K., Lou, W.: Enabling secure and efficient ranked keyword search over outsourced cloud data. *IEEE Trans. Parallel Distrib. Syst.* **23**, 1467–1479 (2012)
14. Wang, C., Ren, K., Yu, S., Urs, K.M.: Achieving usable and privacy-assured similarity search over outsourced cloud data. In: Proceedings IEEE INFOCOM, pp. 451–459 (2012)
15. Whissell, J.S., Clarke, C.L.: Improving document clustering using Okapi BM25 feature weighting. *Inf. Retr.* **14**, 466–487 (2011)

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.







# Based on GAN Generating Chaotic Sequence

Xuguang Chen<sup>1</sup>, Hongbin Ma<sup>1</sup>(✉), Pujun Ji<sup>1</sup>, Haiting Liu<sup>1</sup>, and Yan Liu<sup>2</sup>

<sup>1</sup> Electronic Engineering College, Heilongjiang University, No. 74 Xuefu Road, Harbin, China  
mahongbin@hlju.edu.cn

<sup>2</sup> National Computer Network Emergency Response Technical Team/Coordination Center of  
China, Beijing, China  
liuyan@cert.org.cn

**Abstract.** In this paper, an adversarial encryption algorithm based on generating chaotic sequence by GAN is proposed. Starting from the poor leakage resistance of the basic adversarial encryption communication model based on GAN, the network structure was improved. Secondly, this paper used the generated adversarial network to generate chaotic-like sequences as the key  $K$  and entered the improved adversarial encryption model. The addition of the chaotic model further improved the security of the key. In the subsequent training process, the encryption and decryption party and the attacker confront each other and optimize, and then obtain a more secure encryption model. Finally, this paper analyzes the security of the proposed encryption scheme through the key and overall model security. After subsequent experimental tests, this encryption method can eliminate the chaotic periodicity to a certain extent and the model's anti-attack ability has also been greatly improved. After leaking part of the key to the attacker, the secure communication can still be maintained.

**Keywords:** GAN · Chaos model · Key generation · Data protection

## 1 Introduction

With the development of science and technology, the importance of data is becoming more and more obvious, and data protection is also highly valued. Information security includes a wide range of encryption technology is one of the important technologies to ensure information security.

In 2016, Abadi et al. proposed an adversarial encryption algorithm based on neural network, which consists of the communication party Alice and Bob and the attacker Eve. Eve tried to decipher the communication model between Alice and Bob. Alice and Bob tried to learn how to prevent Eve's attack. The three used this confrontation training to increase their performance. However, this model has no way to show what the two communication parties and the attacker learned in the end, nor can they judge whether the password structure is safe. Therefore, this paper first analyzed the security of the scheme in detail and statistics the security of the leaked part of the key and then improved the structure of the existing network according to the remaining problems in the system. In addition, a key generation model based on GAN was constructed. It takes Logistic

mapping as input, and generated chaotic sequence as encryption key with the help of confrontation training between networks, and inputs into the subsequent encryption model to obtain a more secure encryption algorithm. To get a more secure encryption algorithm. Finally, this paper analyzed the security of the proposed encryption scheme through the key and the overall model security.

## 2 Basic Adversarial Encryption Algorithm and Chaotic Map

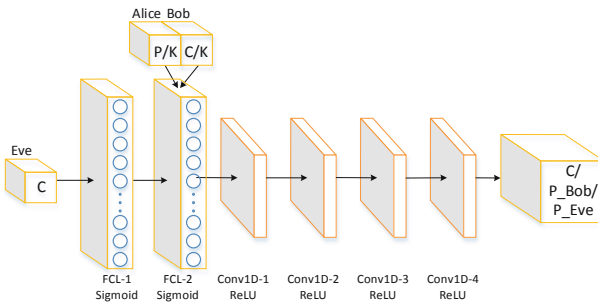
### 2.1 Basic Adversarial Encryption Communication Model

In 2016, Abadi M, etc. first proposed the use of GAN to implement encrypted communication [1]. Through this technology, encrypted and secure communication during enemy monitoring is realized. Its work is based on traditional cryptography scenarios, and its workflow is shown in Table 1.

**Table 1.** Training of basic adversarial encryption communication model based on GAN.

The training steps of the basic adversarial encryption communication model based on GAN
1. Alice encrypts plaintext $P$ with key $K$ to generate ciphertext $C$
2. Bob gets ciphertext $C$ and decrypts it when he knows the key $K$ to get message $P_{Bob}$
3. Eve obtains ciphertext $C$ and decrypts $P_{Eve}$ without key $K$
4. The communication model and attack model are optimized in the training
5. Finally, $P_{Bob}$ is the same as $P$ , and the gap between $P_{Eve}$ and $P$ is as wide as possible

In terms of internal model construction, Alice and Bob have the same model. The Eve network adds a fully connected layer to simulate the key generation process. Eve is trained to improve his decryption ability to make  $P_{Eve} = P$ . Figure 1 shows the network structure model of Alice, Bob, and Eve.



**Fig. 1.** The network structure of Alice, Bob and Eve.

The L1 distance is used to calculate the distance between the plaintext and its estimated value. The L1 distance is defined as:

$$d(P, P') = \frac{1}{N} \sum |P_i - P'_i| \quad (1)$$

Eve's loss function is defined by an expected value:

$$L_E(\theta_A, \theta_E) = E_{P,K}[d(P, D_E(\theta_E, E_A(\theta_A, P, K)))] \quad (2)$$

Where  $\theta_A$  and  $\theta_E$  are the neural network parameters of Alice and Bob respectively.  $P$  is the encryption output  $E_A(\theta_A, P, K)$  when the plaintext  $K$  is the key and  $D_E(\theta_E, C)$  is the decryption output when Eve inputs ciphertext  $C$ . Similarly, Bob's loss function is defined as follows:

$$L_B(\theta_A, \theta_B) = E_{P,K}[d(P, D_B(\theta_B, E_A(\theta_A, P, K), K))] \quad (3)$$

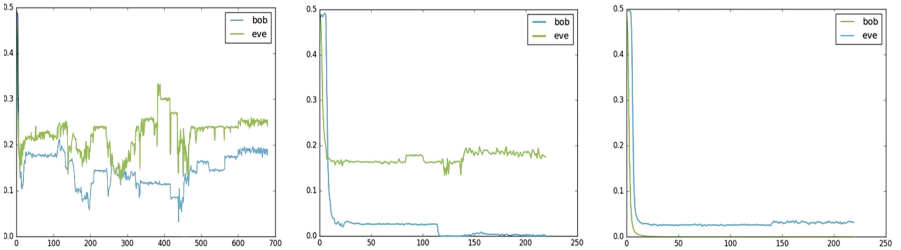
Where  $D_B(\theta_B, C, K)$  is the decryption output when Bob inputs ciphertext  $C$  and key  $K$ .

Alice and Bob should exchange data accurately while defending against attacks. Therefore, the joint loss function of communication parties is defined by combining  $L_B$  and  $L_E$ , which is defined as follows:

$$L_{A,B}(\theta_A, \theta_B) = L_B(\theta_A, \theta_B) - L_E(\theta_A, O_E(\theta_A)) \quad (4)$$

Where  $O_E(\theta_A)$  is the optimal Eve found by minimizing the loss.

However, according to continuous research, people gradually found the problems exposed by this basic model. The main problem is that when part of the key and plaintext are leaked, it is no longer an acceptable encrypted communication scheme [2]. This will cause losses that are difficult to assess for the data protection of both parties. Experiments showed that neither Bob nor Eve could easily converge when the amount of plaintext leaked to Eve was small. When the number of leaked bits in plaintext exceeded 9 bits, the two could gradually converge, but at this time Eve's decryption ability also increased. When the number of leaks is 16 bits, both could get good performance. The experimental results are shown in Fig. 2, where the ordinate indicates the decryption error rate, and the abscissa indicates the number of training rounds.



**Fig. 2.** The decryption of Bob and Eve when the amount of leaked plaintext is 8, 10, and 16 bits respectively.

## 2.2 Logistic Chaotic Map

Up to now, many classic chaotic models have been widely used, such as Logistic mapping, Tent mapping, Lorenz chaotic model, etc. In this paper, Logistic mapping is selected as the encryption method, which has a simple mathematical form and its complex dynamic behavior [3]. Its system equation is as follows:

$$x_{n+1} = \mu x_n (1 - x_n) \quad (5)$$

The range of  $\mu$  in the formula is 0 to 4, which is called the logistic parameter. Studies have shown that when  $x$  is between 0 and 1, the Logistic map is in a chaotic state. Outside this range, the sequence generated by the model must converge to a certain value [4]. And when the value of  $\mu$  is 3.5699456 to 4, the value generated by the iteration presents a pseudo-random state. In other ranges, convergence will occur after a certain number of iterations, which is unacceptable to us.

In this paper, GAN can automatically learn the characteristics of data distribution of the real sample set, to automatically learn the data distribution of Logistic mapping and generate variable sequences as keys for subsequent encryption models. Besides, the mapping equations of  $\mu = 3.5699456$  and  $\mu = 4$  are selected as the input of the GAN model. The powerful learning ability of the model is used to generate the distribution between the two input data distributions and try to fit the original data distribution.

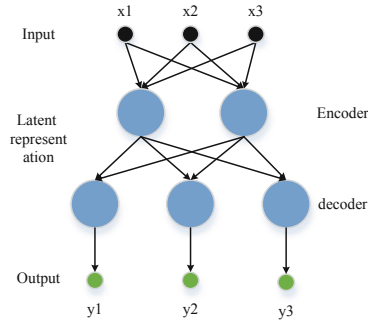
## 3 Adversarial Encryption Algorithm Based on Logistic Mapping

### 3.1 Key Generator Based on GAN

In this section, the GAN was used to simulate the chaotic model, and then a random sequence similar to the chaos was generated as an encryption key. The GAN model contains two systems, namely the discriminating system and the generating system. The results of the two systems will be opposed to each other and will be updated in turn [5]. The principle of the binary game is used to achieve the optimal state.

**Improvement of Key Generator Network Structure.** GAN has high requirements for hardware devices. If the network structure is too complex, GAN may not only lead to the collapse of the platform but also reduce the efficiency of the key generation to some extent. Because of this situation, this paper proposed a form of self-encoder and GAN fusion to generate data.

The self-encoder can traverse the input information, convert it into efficient potential representations, and then output something that it wants to look very close to the input. The model includes generative networks and discriminant networks, except that part of the generative network uses self-encoders as the basic structure. Besides, in the overall framework, the information source should not only be used as the input source for the generator to extract the generation factor, but also need to be mixed with the real samples from the training set, because the generated fake samples contain key information features. The self-encoder consists of two parts: the encoder extracts potential features



**Fig. 3.** Basic structure of self-encoder.

from the input source, and the decoder replaces these potential features with the output source. Its structure is shown in Fig. 3.

The key generation model is combined with the GAN network and the self-encoder, in which the self-encoder is applied to the generating network part and the rest is composed of the GAN model. The experimental results showed that this method had a great improvement in the generation effect, and it could almost coincide with the original model distribution in the later stage of training.

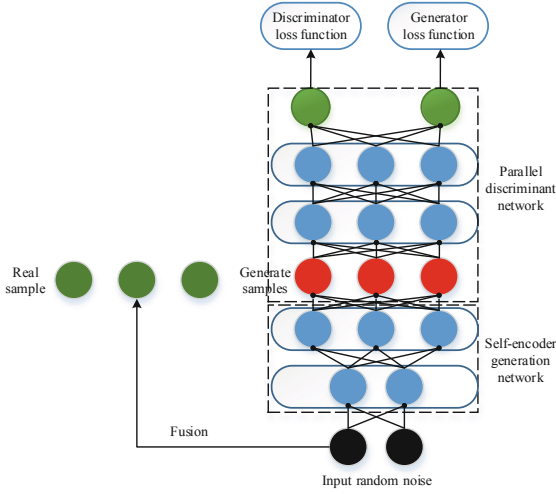
In addition to the improvement of the above structure, this paper also introduced the concept of parallel training. The original GAN model uses the idea of serial training, in which network A is used to distinguish between true and false tags, and network B is used to generate false tags. This method has many disadvantages in the training process, for example, the number of rounds of discriminating network training cannot be determined. In this paper, the idea of parallel training was used to solve the two problems. The parallel structure made the two networks compete at the same time, and the training was no longer sequential, which greatly improved the training speed. The improved network structure is shown in Fig. 4.

The self-encoder network consists of convolutional networks and deconvolutional networks, which correspond to encoder and decoder respectively. The convolution layer number of both is 4, and the activation function of the convolution layer is ReLU. The structure of the auto-encoder generation network is shown in Fig. 5.

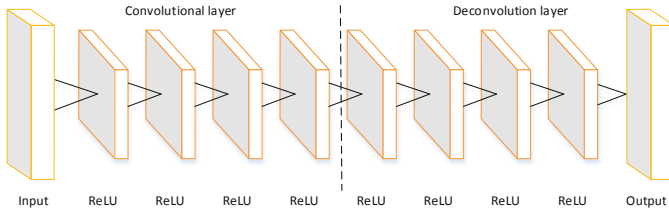
The number of layers in the discriminant network is also 4, as shown in Fig. 6.

**Simulation to Test the Improved Effect.** In this paper, the improvement of training speed is verified through simulation. With the help of Python, the CPU training data collected for 200 times by serial and parallel models were counted, excluding the time consumption at the start of training. The results show that the parallel GAN cancels the cumbersome operations such as discriminating network parameter transmission and feedforward of the B network. It and can better adapt to the popular platform with moderate computing power and the time consumption is also greatly reduced. Statistics are shown in Table 2.

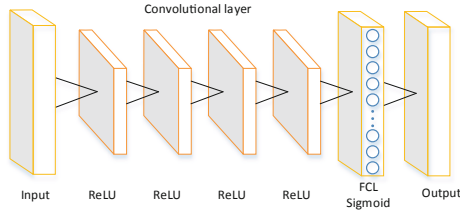
According to the study of the chaos model, when the parameter is greater than 3.5699456 and less than or equal to 4, Logistic mapping enters into the chaos state.



**Fig. 4.** Parallel training network structure.



**Fig. 5.** Generate network structure.



**Fig. 6.** Discriminate network structure.

**Table 2.** Time-consuming comparison of serial/parallel GAN network training.

Network type	Minimum value	Maximum value	Average value
Serial GAN	1.7241	1.9857	1.8325
Parallel GAN	0.8754	1.4732	1.1473

The purpose of this section is to train a network. The mapping equation  $\mu = 3.5699456$  and  $\mu = 4$  is selected as the input of the GAN generation model. During the training process, the networks competed with each other. When the accuracy d generated by the discriminator is 0.5, the key generated at this time is extracted and used for the subsequent encryption system. The key generation algorithm is as follows (Table 3):

**Table 3.** Key generation algorithm based on GAN.

---

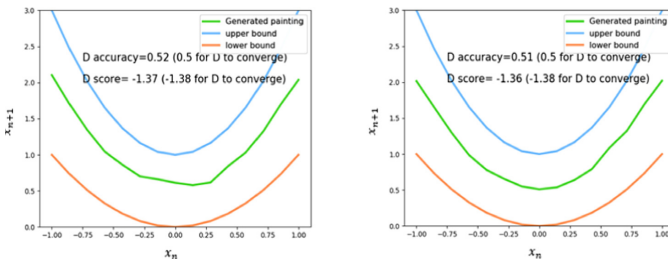
The key generation algorithm based on GAN:

---

1. Initialize network parameters
2. While  $i < N$ , do:
  - a. for k-step, do:
    - (1) Select  $m$  samples  $\{z_1, z_2, \dots, z_m\}$  from  $P_z$
    - (2) Select  $m$  samples  $\{x_1, x_2, \dots, x_m\}$  from  $P_{data}$
    - (3) Fixed generator, update discriminator network parameters by the gradient ascent algorithm:  $\nabla_{\theta_D} \frac{1}{m} \sum_{i=1}^m [\log D(x^{(i)}) + \log(1 - D(G(z^i)))]$
  - b. A block of  $m$  samples  $\{z_1, z_2, \dots, z_m\}$  is sampled from the initial distribution  $P_x$
  - c. Fixed discriminator, updates generator network parameters by gradient descent algorithm:  $\nabla_{\theta_G} \frac{1}{m} \sum_{i=1}^m 1 - \log(1 - D(G(z^i)))$

---

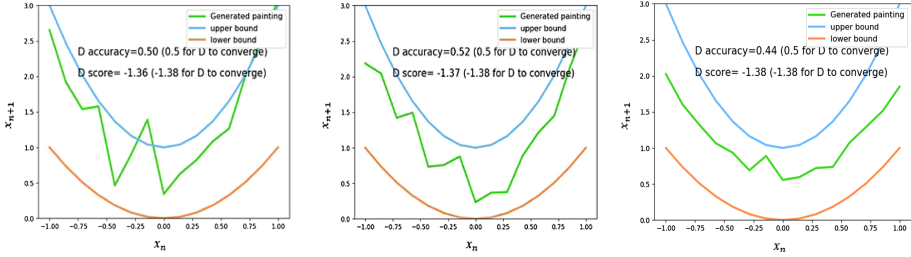
The results of the training are shown below. Figure 7 shows the results when the number of iterations is 5000 and the discriminator output accuracy is 0.5. Figure 8 shows the unexpected results generated by the model when the number of iterations is less than 500, or the discriminator output accuracy is not selected properly.



**Fig. 7.** The number of iterations is 5000 and the discriminator output accuracy is 0.5.

### 3.2 The Overall Encryption Algorithm Design

The adversarial encryption algorithm described in this paper mainly includes sufficient security of the encryption system and encryption key. 3.1 Section introduced the process



**Fig. 8.** The number of iterations is less than 500, or the discriminator output accuracy is not selected properly.

**Table 4.** Improved adversarial encryption algorithm based on GAN.

---

Process of the encryption algorithm

---

1. Plaintext  $[P_0, P_1, \dots, P_n]$  and key  $[k_0, k_1, \dots, k_n]$  are sent to the Alice network
  2. From the bit-angle mapping formula  $f(b) = \arccos(1 - 2b)$ , the bit-angle is converted to Angle  $[a_0, a_1, \dots, a_n]$
  3. The full connection layer uses the hidden variable  $[h_0, h_1, \dots, h_n]$  as the initial ciphertext with the Angle information
  4. Using formula  $f^{-1}(a) = \frac{1 - \cos(a)}{2}$ , the initial ciphertext in 3 is transformed into the final ciphertext  $[C_0, C_1, \dots, C_n]$
  5. Alice, the encryptor, sends ciphertext  $[C_0, C_1, \dots, C_n]$  to Bob
  6. Bob and Eve receive ciphertext and output plaintext P
  7. Start alternate training of Alice, Bob and Eve networks
  8. When Eve's decryption accuracy is high, Alice selects the encryption key again and circulates 5, 6, 7 and 8
  9. Stop training when Eve decrypts plaintext like random guesses
- 

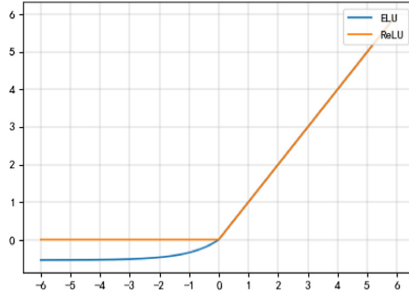
of key generation of simulated chaos model. This method can well hide information about the chaotic map and increase the difficulty of decoding. Then the key and plaintext are entered into the GAN counter communication model. By means of confrontation training, an efficient encryption method that can eliminate chaotic cycles is obtained. In the test session, this article only showed the performance research of the random key generation by the adversarial algorithm and the partial leakage of the key.

In view of the problems in the model in Sect. 2.1, this section made some improvements, mainly through the replacement of the activation function and the enhancement of the neural network structure to strengthen the encrypted communication model.

Figure 1 shows that the activation function used in the basic model is ReLU. According to the property of it, the result is always 0 when the input is negative, so when the neuron is negative, this property will affect the weight update [6]. To solve this problem, the ELU activation function is selected in this paper to replace ReLU, which alleviates the phenomenon of large area neuron death during weight updating. In addition, it is



negative at  $x$  and has a small absolute value, which has good anti-interference ability. And the ELU is better suited for subsequent normalization operations. The comparison of the two is shown in Fig. 9.



**Fig. 9.** Comparison of ReLU function and ELU function.

The model described in Sect. 2.1 cannot communicate normally after the increase in the amount of information leaked, that is, neither Bob nor Eve can decrypt normally, and the decryption ability reaches its limit. Therefore, this paper enhanced the network model of Bob and Eve to improve their decryption ability. By adding the full connection layer, the decryption ability of Bob and Eve was increased synchronously. The activation function took the tanh function, and the structure of Alice's network remained unchanged.

In addition, to prevent the model from falling into the local optimal mode due to the rising stability, and thus the performance cannot be improved, normalization processing was added to the full connection layer in this paper to improve the network's ability to learn the optimal state. Through experimental verification, data normalization maps the value range of data to different regions, eliminating the problem that the accuracy of classification results is affected by the inconsistency of data size range. The improved Alice, Bob, and Eve network models are shown in Fig. 10.

The improved adversarial encryption algorithm based on GAN is shown in the following (Table 4).

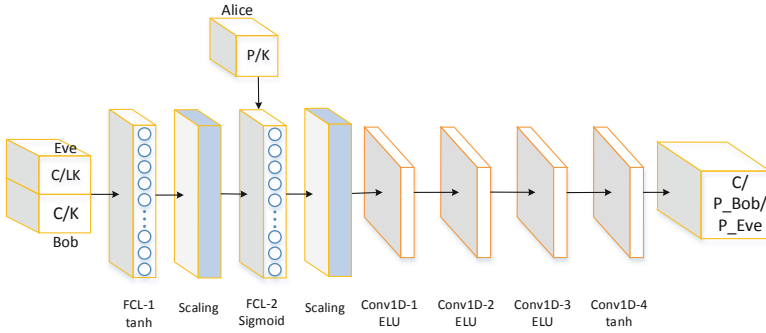


Fig. 10. Improved network model of Alice, Bob and Eve.

## 4 Experimental Procedure

### 4.1 Experimental Environment

This experiment was conducted on the Window System, using TensorFlow as the network learning framework, and demonstrated the performance research of the random key generation by the adversarial algorithm and the partial leakage of the key. In the key generation stage, this article chose Adam optimizer, and the learning rate is 0.0008. A Mini-batch with M of 4096 was used in the encryption-model.

### 4.2 Key Security Analysis

FID can well evaluate the quality of the generated data. It can give the same judgment results as human vision, and the computational complexity of FID is not high. Therefore, FID is selected as the performance evaluation index of GAN in this paper. The formula of FID is as follows:

$$FID = \|\mu_r - \mu_g\|^2 + T_r \left( \sum r + \sum g - 2 \left( \sum r \sum g \right)^{1/2} \right) \quad (6)$$

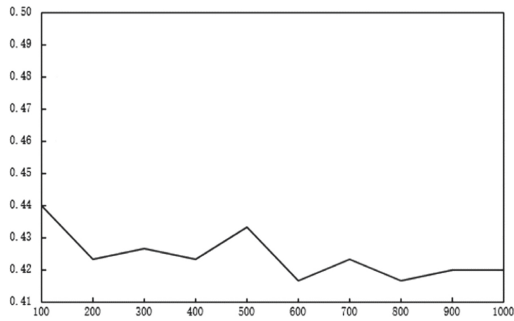
The evaluation process is shown in Table 5.

In this paper, the distribution of chaotic map data generated by GAN is within the upper and lower bounds. Figure 11 statistics the FID distance between the generated samples and the real samples in the early training range, where the abscissa represents the number of training rounds, and the ordinate represents the FID value. It can be seen from Fig. 11 that as the training progresses, the FID value gradually becomes smaller, indicating that the performance of GAN through training is also continuously increasing.

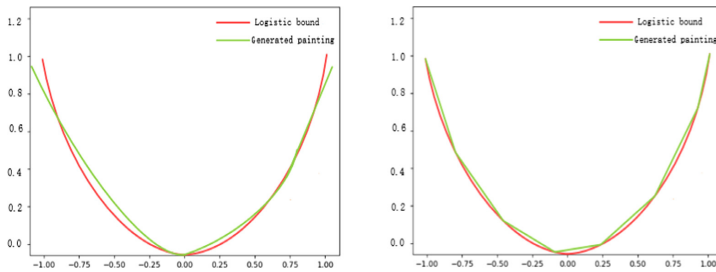
To prevent attackers from attacking through fuzzy keys, the sensitivity of keys must be high enough, and chaotic systems can satisfy this characteristic well [7]. To verify the high sensitivity of the key generated by the key generator,  $\mu = 4$  was set in the experiment. The initial value of  $x_{01} = 0.256$  and  $x_{02} = 0.264$  were selected and input into the model to observe the sensitivity of the period. The results show that, when the initial selection gap is small, even when the distribution generated after training is almost

**Table 5.** FID evaluation process.

FID evaluation process
1. Send the samples generated by the generation network and the samples generated by the discriminant network to the classifier
2. Abstract features of the middle layer of the classifier
3. Assuming that the abstract feature matches the Gaussian distribution, estimate the mean and variance of the generated and training samples
5. Calculate the distance between two Gaussian distributions and use this to evaluate the performance of GAN

**Fig. 11.** FID distance between generated sample and real sample.

fitted to the original distribution, the difference between the two distributions generated is still obvious. This indicates that the generated chaotic key has high sensitivity. Besides, a different key sequence is used for each encryption process, which further improves the security of the generated key. The test results are shown in Fig. 12.

**Fig. 12.** Training results of the key generator based on GAN.

### 4.3 Analysis of Encryption Model Performance

The security of the encrypted communication system is inseparable from the ability of the attacker. When the attacker has a strong ability, the model still needs to ensure the security of the communication as far as possible. Therefore, in order to verify the security of the model in this paper, we let Eve know the ciphertext and some keys simultaneously to test the model in this paper. The results are shown in Fig. 13.

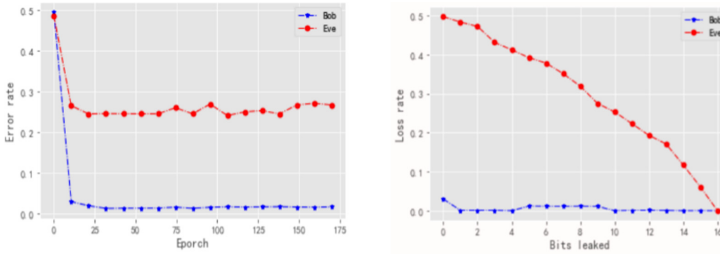


Fig. 13. Results of the security evaluation of the encryption model.

As can be seen from the figure above, as the training progresses, both Bob and Eve can achieve convergence rapidly, and Eve's error during convergence is very high. At this time, it can be considered that the communication between Alice and Bob is safe. In addition, the encryption performance of the communication model is much better than that of the basic model when the amount of key leakage increases gradually. According to the test and analysis, the adversarial encryption algorithm based on generating the chaotic sequence by GAN is secure. After a certain number of rounds of training, the model tended to be stable, and the performance of the model in resisting attacks was also improved.

## 5 Conclusion

Based on referring to a lot of literature and based on the anti-encryption communication model, this paper introduced a chaos model to optimize the generation mode of key and proposes a counter-encryption model based on Logistic mapping. And analyze the security of the entire system through model analysis, key analysis, and other methods. Finally, it is concluded that the key to the encryption algorithm can be changed from time to time, and the periodic problems in chaotic encryption can be eliminated to a certain extent. In addition, compared with the basic anti-encryption communication model, the security of the encryption model is greatly improved.

## References

1. Abadi, M., Andersen, D.G.: Learning to Protect Communications with Adversarial Neural Cryptography. ICLR (2017)

2. Raghunathan, A., Segev, G., Vadhan, S.: Deterministic public-key encryption for adaptively-chosen plaintext distributions. *J. Cryptol.* **31**(4), 1012–1063 (2018). <https://doi.org/10.1007/s00145-018-9287-y>
3. Lin, Z., Yu, S., Li, J.: Chosen ciphertext attack on a chaotic stream cipher. In: Chinese Control and Decision Conference (CCDC), pp. 5390–5394 (2018)
4. Ashish, Cao, J.: A novel fixed point feedback approach studying the dynamical behaviors of standard logistic map. *Int. J. Bifurcat. Chaos.* **29**(01) (2019)
5. Tramer, F., Kurakin, A., Papernot, N., Goodfellow, I., Boneh, D., Mc Daniel, P.: Ensemble adversarial training: attacks and defenses. *EprintArxiv* (2017)
6. Jain, A., Mishra, G.: Analysis of lightweight block cipher FeW on the basis of neural network. In: Yadav, N., Yadav, A., Bansal, J.C., Deep, K., Kim, J.H. (eds.) *Harmony Search and Nature Inspired Optimization Algorithms*. AISC, vol. 741, pp. 1041–1047. Springer, Singapore (2019). [https://doi.org/10.1007/978-981-13-0761-4\\_97](https://doi.org/10.1007/978-981-13-0761-4_97)
7. Purswani, J., Rajagopal, R., Khandelwal, R., Singh, A.: Chaos theory on generative adversarial networks for encryption and decryption of data. In: Jain, L.C., Virvou, M., Piuri, V., Balas, V.E. (eds.) *Advances in Bioinformatics, Multimedia, and Electronics Circuits and Signals*. AISC, vol. 1064, pp. 251–260. Springer, Singapore (2020). [https://doi.org/10.1007/978-981-15-0339-9\\_20](https://doi.org/10.1007/978-981-15-0339-9_20)

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





# MinerGate: A Novel Generic and Accurate Defense Solution Against Web Based Cryptocurrency Mining Attacks

Guorui Yu<sup>1</sup>, Guangliang Yang<sup>2</sup>, Tongxin Li<sup>1</sup>, Xinhui Han<sup>1</sup> (✉), Shijie Guan<sup>1</sup>, Jialong Zhang<sup>3</sup>, and Guofei Gu<sup>2</sup>

<sup>1</sup> Peking University, Beijing 100871, China  
{yuguorui, litongxin, hanxinhui, 1600012835}@pku.edu.cn

<sup>2</sup> Texas A&M University, Texas, TX 77843, USA  
ygl@tamu.edu, guofei@cse.tamu.edu

<sup>3</sup> ByteDance AI Lab, Beijing 100098, China  
zjl.xjtu@gmail.com

**Abstract.** Web-based cryptocurrency mining attacks, also known as cryptojacking, become increasingly popular. A large number of diverse platforms (e.g., Windows, Linux, Android, and iOS) and devices (e.g., PC, smartphones, tablets, and even critical infrastructures) are widely impacted. Although a variety of detection approaches were recently proposed, it is challenging to apply these approaches to attack prevention directly.

Instead, in this paper, we present a novel generic and accurate defense solution, called “MinerGate”, against cryptojacking attacks. To achieve the goal, MinerGate is designed as an extension of network gateways or proxies to protect all devices behind it. When attacks are identified, MinerGate can enforce security rules on victim devices, such as stopping the execution of related JavaScript code and alerting victims. Compared to prior approaches, MinerGate does not require any modification of browsers or apps to collect the runtime features. Instead, MinerGate focuses on the semantics of mining payloads (usually written in WebAssembly/asm.js), and semantic-based features.

In our evaluation, we first verify the correctness of MinerGate by testing MinerGate in a real environment. Then, we check MinerGate’s performance and confirm MinerGate introduces relatively low overhead. Last, we verify the accuracy of MinerGate. For this purpose, we collect the largest WebAssembly/asm.js related code with ground truth to build our experiment dataset. By comparing prior approaches and MinerGate on the dataset, we find MinerGate achieves better accuracy and coverage (i.e., 99% accuracy and 98% recall). Our dataset will be available online, which should be helpful for more solid understanding of cryptojacking attacks.

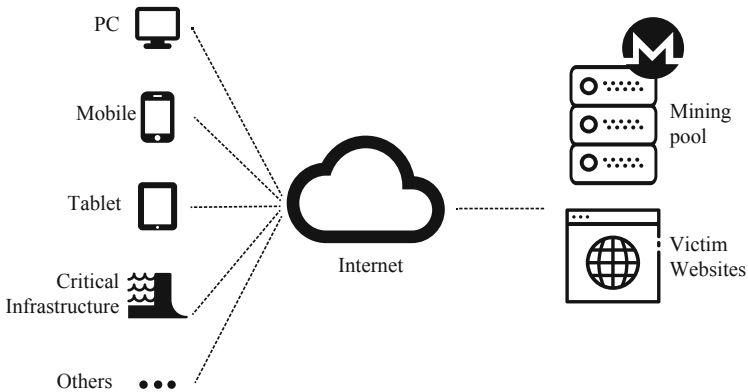
**Keywords:** Cryptojacking · WebAssembly · asm.js

## 1 Introduction

Recently, cryptocurrency mining attacks, also known as cryptojacking attacks, are becoming increasingly popular. Different from regular attacks, which usually aim at the access or destruction of private data or sensitive functionalities, this attack mainly focuses on stealing the computing resources (e.g., CPU) of victim Internet-connected devices for covertly mining cryptocurrencies and accumulating wealth.

Although the mining attack does not make malicious and notorious actions, it can still cause serious consequences. For example, the mining code usually occupies the most (or even the whole) of physical resources (e.g., CPU, Memory, and network), which results in all services and apps in the victim devices become inactive, unresponsive, or even crashed. Furthermore, this attack also significantly reduces the life cycle of hardware, such as the battery of laptops and smartphones.

With the significant development of web techniques (e.g., WebSocket [31], Web Worker [30], WebAssembly [9], asm.js [6]), more and more mining attacks are moved to the web platform, which means they can be simply launched by embedding JavaScript snippets. The attack scenario is shown in Fig. 1. First, in the victim websites, attackers include mining script code [15, 16], which is used to initialize the environment, and download and execute mining payloads. Please note that in general the mining payloads are written in WebAssembly/asm.js, which are intermediate languages and allow web browsers to run low-level languages (e.g., C/C++) for near-native performance.



**Fig. 1.** Example attack scenarios

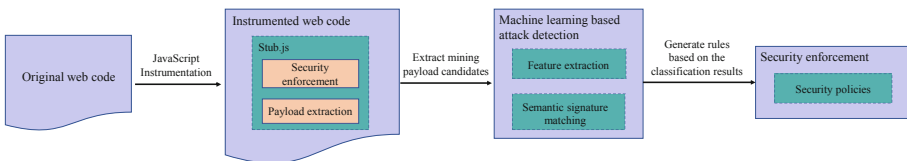
Up to now, a large number of diverse platforms (e.g., Windows, Linux, Android, and iOS) and devices (e.g., PC, smartphones, tablets, and even critical infrastructures) have been widely impacted. For example, recent studies [1, 5] showed popular applications running in smartphones or PC might silently launch the attacks by executing web mining payloads in the background. Furthermore, critical infrastructures (e.g., industrial control systems) may also be threatened by the mining attack. A recent report showed that a water utility [20] was attacked which might cause its industrial control application to be paused and even crashed.

Even worse, the attack may be hardly stopped once the related web code is executed in the background. A recent report [4] showed attackers could even continue mining with the help of service worker after closing the infected web page.

Therefore, a defense solution that can provide protection on all various devices and eliminate the threats of mining attacks is expected. Recently, a variety of detection solutions [10, 12, 13, 17, 32] have been proposed. However, these approaches do not meet the requirements. First, they are not scalable. Most of them [13, 17, 32] require the modification of web browser engines to collect runtime features, such as the usage of CPU, memory, and network activities. The above solutions not only bring considerable additional overhead to the browser, but also make it difficult to deploy the defense. Second, in case users access infected websites, the mining code should be immediately stopped. However, prior approaches [10, 12] does not meet the requirements. Third, the user experience should not be significantly influenced. However, prior tools may introduce high overhead. For example, [32] introduced almost 100% overhead.

Furthermore, prior approaches may face high false positives and negatives. To identify mining code, they either use a blacklist to block the access of infected websites, or leverage heuristic features to detect mining code. For the blacklist-based tools (e.g., Firefox [2]), it is difficult to keep up with the rapid iteration of mining websites, and thus may cause high false negatives. For the heuristic features, these features mainly include 1) the usage of CPU, memory, and network, 2) CPU cache events, and 3) cryptographic instructions. In our test, we find it is challenging for existing approaches to distinguish between benign CPU-intensive code and mining code.

Instead, in this paper, we propose a novel, general, and accurate protection solution, called MinerGate, that can automatically and effectively provide security enforcement on end devices against mining attacks. To achieve the goal, MinerGate is deployed as an extension of network gateways or proxies. As shown in Fig. 2, our approach is three-fold. First, MinerGate monitors network traffic to catch cryptocurrency mining payloads. For this purpose, MinerGate instruments all network traffic by injecting pre-defined JavaScript code “stub.js”, which will be executed in local devices. The injected stub code is responsible for extracting WebAssembly/asm.js code and enforcing security rules. When stub.js uncovers web code written in WebAssembly/asm.js in a victim device, it will send the content or the reference of related code to MinerGate for further analysis.



**Fig. 2.** MinerGate’s workflow

Second, different from prior approaches, which rely on the analysis on collected runtime features, MinerGate mainly focuses on understanding the semantics (e.g., CG



and CFG) of WebAssembly/asm.js code. Through data-driven feature selection, MinerGate determines and extracts semantic-related features and forwards these features to a machine learning engine for determining the existence of mining code. Last, once mining code is found, MinerGate notifies the victim device (i.e., stub.js) to apply security rules, such as stopping the execution of web code and alerting the victim user and the network administrator.

In our evaluation, we first verify the correctness of MinerGate by testing MinerGate in a real environment. Then, we check MinerGate’s performance and confirm MinerGate introduces relatively low overhead. Last, we verify the accuracy of MinerGate. For this purpose, we first address the challenge there is still not a reliable labeled dataset of cryptojacking mining payloads. To create such a dataset with ground truth, we systematically collect WebAssembly/asm.js code from the 10 million web pages, and NPM [11]. As a consequence, our dataset includes not only mining code from 4659 pages, but also 243 projects related to benign WebAssembly/asm.js. We will open up this dataset for the follow-up research. This dataset should be helpful for a better understanding of mining attacks.

Based on the dataset, we compare MinerGate and prior tools. We find MinerGate achieves better accuracy and coverage (i.e., 99% accuracy and 98% recall).

To sum up, we make the following contributions:

- We propose the novel, generic and accurate defense solution “MinerGate” against mining attacks.
- MinerGate obtains high accuracy by extracting and applying semantic-based features with help of call graph (CG) and control flow graph (CFG).
- We build the largest ground truth dataset.
- We compare MinerGate and existing related approaches, and show MinerGate is scalable, effective and accurate.

## 2 Background

### 2.1 Cryptocurrency Mining and Cryptojacking Attacks

Cryptocurrencies are digital assets designed to work as a medium of exchange that uses strong cryptography to secure financial transactions, control the creation of additional units, and verify the transfer of assets [34]. The cryptocurrency uses a distributed database, blockchain, to store the transactions in units of blocks. Each block mainly includes a unique ID, the ID of the preceding block, the timestamp, the nonce, the difficulty, and transaction records. A valid block contains a solution to a cryptographic puzzle involving the hash of the previous block, the hash of the transactions in the current block, and a cryptocurrency address which is to be credited with a reward for solving the cryptographic puzzle. The specific cryptographic puzzle is to find a block of data whose hash value is smaller than a set value which is decided by the difficulty. Most data of the block are known, and the miner should find the unknown part in a limited time. Once the pronomeral, typically is the nonce, is found, the miner will submit it to get profit. This process is called cryptocurrency mining [7].

Cryptojacking, the unauthorized use of hardware of others to mine cryptocurrency, has become the biggest cyber threat in many parts of the world. Cryptojacking was a burgeoning industry in 2018, there have been 13 million cryptojacking attempts in the case of a 40% increase in 2018 [19]. Using crypto-mining malware, criminals have mined earning up to 56 million USD in 2018. There are many reasons why cryptojacking is overgrowing. One of the most important reasons is the simplicity of deployment. Cryptojacking can be easily deployed by inserting a statement in the HTML, such as `<script src="attacker.com/mining.js"></script>`. This allows the attackers to deploy mining payloads to victim websites without actual control because of XSS or other vulnerabilities. The simplicity of cryptojacking leads to the threat of cryptojacking attacks as long as the cryptocurrency exists. There is no correlation between the existence of such an attack and whether or not a service is alive.

## 2.2 Related Web Techniques

In past years, web techniques made tremendous progress, which makes it feasible to launch mining attacks using web code. For example, the worker mechanisms provide the possibility of running web code in parallel and the background. WebAssembly/Asm.js provide chances to run mining code in machine-instruction level.

Asm.js is an embedded domain specific language that serves as a statically typed assembly-like language. It is a JavaScript subset that allows web code written in low-level languages, such as C/C++. In order to apply asm.js in runtime, the function body of asm.js code must define with a directive prologue “use asm” or “most asm”. WebAssembly [26] is an abstraction over modern hardware, making it language-, hardware-, and platform-independent, with use cases beyond just the Web. WebAssembly is a binary instruction format (bytecode) for a stack-based virtual machine which is different from a text form of asm.js. WebAssembly is designed as a portable target for compilation of high-level languages like C/C++/Rust. Moreover, WebAssembly is committed to getting the speed closer to the native code, and it is faster than asm.js. Currently, WebAssembly can be only be loaded and executed by JavaScript, JavaScript calls WebAssembly in three steps: 1) loading WebAssembly bytecode, 2) compiling bytecode, and 3) instantiating and executing compiled code.

Asm.js and WebAssembly have similarities in many respects. For example, they are both statically typed assembly-like languages, and they have similar instruction sets, which makes it possible for them to convert between each other.

The earnings of cryptojacking attackers are strongly related to the mining speed, so the attackers implement the core logic of mining with WebAssembly and asm.js. We suggest it is more effective and robust to analyze the WebAssembly/asm.js code instead of other scaffolding code. Previous works related to WebAssembly/asm.js malware analysis only concentrate on instruction features, which makes it challenging to classify mining applications.

### 3 System Overview

#### 3.1 Challenges and Our Solutions

In order to design and implement a generic defense solution against web-based mining attacks, several challenges are raised. More details about these challenges and our corresponding solutions are discussed below.

- *Diverse platforms and devices.* Nowadays, many different devices, such as PC, mobile devices and infrastructure devices, are connected to the Internet. They all are potentially affected by mining attacks. Considering these devices usually have their own operating systems, it is challenging to offer general protection.

To address it, we design and implement MinerGate as an extension of a network proxy (e.g., network firewall or gateway). MinerGate can protect all devices behind it. In practice, once a mining attack occurs, MinerGate can enforcedly stop the attack code and alert network administrators.

Please also note that considering HTTPS are frequently used, we assume that MinerGate can monitor all network traffic, including HTTPS-based communication. This can be achieved by installing MinerGate's certificate in all devices under the protection.

- *Obfuscated web code.* Web code, especially the code injected by adversaries, is frequently obfuscated in practice. This poses challenges to extracting adversaries' essential mining code. To address the problem, MinerGate instruments the web code and hijack crucial JavaScript APIs, which are helpful to extract the parameters related to mining code.

However, due to the natural flexibility of JavaScript, adversaries may still bypass the above solution. To deal with this issue, we introduce a self-calling anonymous function to protect instrumented web code, and carefully handle the creation of new JavaScript contexts.

- *Unknown mining semantics.* As introduced in Sect. 2, WebAssembly/asm.js have been widely deployed in web mining code. However, up to known, their inside semantics are still unclear, especially considering there are already many variants of the existing mining code. This may significantly reduce the detection accuracy.

To address this problem, we do program analysis on WebAssembly/asm.js code and extract all call graph (CG) and control flow graph (CFG). Although CG and CFG are basic things for program analysis, automatically generating CG and CFG is still not an easy task, especially considering indirect-call instructions are frequently used.

- *Difficulty of mining code determination.* WebAssembly/asm.js is frequently used not only in mining but also in another area, such as gaming and data processing. It is difficult to distinguish between them accurately. In this work, we address this issue by applying machine learning. However, although existing work discovered a variety of features available for machine learning, they may cause high false positives.

Instead, we extract features from mining semantics (e.g., CG and CFG) and obtain high accuracy. However, it is challenging to apply graph-based features in machine learning, which cause performance issues and affect scalability. To handle it, we analyze the code in units of semantic modules instead of functions or files to break the solid lines in the analysis.

- *Difficulty of stopping mining code.* Once mining attacks occur, hardware resources (e.g., CPU and memory) may be immediately occupied by adversaries. This poses challenges to stopping the corresponding malicious code in time.

To deal with this problem, we stop the execution of the mining thread through the function hijacking beforehand and cut off the source of malicious code.

As shown in Fig. 2, MinerGate contains three major modules: 1) JavaScript Instrumentation, which is used to instrument network traffic to inject `stub.js` for extracting WebAssembly/asm.js code, and enforcing security rules; 2) Machine Learning Based Detection, which can do program analysis on payloads to extract semantic-related features; 3) Security Enforcement, which defines and enforces security rules. For each module, more details are presented in the following sections.

### 3.2 JavaScript Instrumentation

As introduced in Sect. 3, MinerGate injects the JavaScript file “`stub.js`” into all web code to extract WebAssembly/asm.js code and apply security enforcement. This is achieved by hijacking and instrumenting several crucial JavaScript APIs. Please also keep in mind that the `stub.js` file is always placed at the beginning of web code, which can ensure all target JavaScript APIs are already instrumented before they are actually used by mining code.

In the next subsections, we explain how `stub.js` works. Furthermore, we also present our protection, which prevents adversaries bypass or destroy `stub.js` and our instrumented JavaScript APIs.

**WebAssembly/asm.js Code Extraction:** Our JavaScript API hijacking solution is designed based on the key observation: no matter where adversaries save the mining code, such as a URL or encrypted string, the key JavaScript APIs, such as `WebAssembly.instantiate` for WebAssembly must be called. Hence, in `stub.js`, we hijack all crucial JavaScript APIs to extract and collect all required parameters, which are sent back to MinerGate for further analysis. These hijacked APIs are listed in Table 1.

**Table 1.** Hooked APIs for WebAssembly mining payload extraction.

WebAssembly API	Description
<code>instantiate()</code>	Compiles and instantiates WebAssembly code
<code>instantiateStreaming()</code>	Compiles and instantiates a module from a streamed source
<code>compile()</code>	Compiles a Module from WebAssembly binary code
<code>compileStreaming()</code>	Compiles a Module from a streamed source
<code>Module()</code>	Synchronously compiles WebAssembly binary code to a Module

Let us use `WebAssembly.instantiate` as an example to describe how these APIs are hijacked and instrumented. Because JavaScript is a dynamic language, all objects can be replaced so that we can forge a `WebAssembly.instantiate` function

object and replace the original one. In this fake function, we first use a WebSocket connection to send the function parameter (the WebAssembly payload) asynchronously to the gateway and continue to execute the original code. No matter how the mining code is saved and how the code is obfuscated, the mining code will be identified and sent to MinerGate. In addition, the payload is sent asynchronously, without blocking code execution and increasing overhead.

For `asm.js`, we need some extra effort to extract them. Since attacker can dynamically invoke the `asm.js` compiler by APIs like `eval`, `Function`, etc. We need to hijack any API that will trigger code compilation. As introduced in Sect. 2, before the `asm.js` code is parsed and compiled, it must be defined with the prologue directive “use asm” or “most asm” [6]. This principle offers hints to extract `asm.js` code from APIs. More specifically, we first do syntax analysis on the parameter of `eval` to build the AST. Next, we scan the AST to identify all functions. Then, we check each function to determine the existence of “use asm”. Finally, in addition to the `asm.js` that appear directly in the HTTP traffic, the payloads found in the API are also sent to the gateway for analysis.

In addition to extracting WebAssembly/`asm.js` code, `stub.js` are also used to enforce security rules. More details are discussed in Sect. 3.5.

```
(function () {
  var ori_api=WebAssembly.instantiate;
  WebAssembly.instantiate = function (buf, importObj){
    if (isMalicious(buf)) {
      // Refuse to load malicious modules.
      return null;
    } else {
      return ori_api(buf, importObj);
    }
  };
})();
// Variable "ori_api" will not able to be accessed out of the scope.
```

**Protections on `stub.js`:** The `stub.js` solution can effectively extract the mining code and apply security enforcement. However, there are still several ways that adversaries may bypass and destroy the solution. To mitigate the problem, we provide the following protections:

- *Locating original APIs.* Considering if adversaries can find and access that variable, adversaries may still normally and freely use the hijacked APIs. To address this issue, we place `stub.js` inside a self-calling anonymous function. As a result, even though adversaries may find the local variables where the original APIs are saved in, such as calling `Function.toString()` to check the source code of the hijacked APIs, adversaries cannot still access them.

Furthermore, to improve the security of `toString()` and hide our defenses roughly, we can also hijack the function `toString()` to confuse the attackers.

- Starting a new web context. Mining code may use worker and iframe to run the mining payload in the background to keep the responsiveness of the main thread. Since worker and iframe create new JavaScript contexts, existing hijacked APIs becomes ineffective in the new contexts. Hence, stub.js is required to be executed again and right after the initialization of the new contexts. To achieve it, the worker and iframe object are also hijacked through Web traffic instrumentation.

More specifically, for a worker, we implement a worker agent object to protect the crucial API Worker. When a worker is created, an agent object is returned for replacement. This worker agent has the same interface as the native worker, but it will stitch the stub.js together with the original code to protect the APIs existing in the worker. We also emphasize that any subsequent calls to the worker API within this context will be protected, regardless of how it is called.

In addition to the protection mentioned above, to respond to some existing attack methods [18], such as prototype poisoning, abusing the caller-chain, etc., our work also includes defense against these attacks.

### 3.3 Machine Learning Based Detection

As discussed in Sect. 1, the simple heuristic features used by prior approaches may cause high false positives. This is because applications in the real world may contain instruction patterns similar to mining algorithms, such as video decoding and data encryption/decryption. This scenario makes it difficult to determine the type of programs based on the occurrences of specific instructions without context. To achieve higher accuracy, we mainly improve from two aspects. On the one hand, we add more features through data-driven feature selection; on the other hand, we divide the code into different “modules” by running the clustering algorithm on the function call graph, which helps us reduce data dimensions, improve the performance and enhance resistance to code obfuscation. The overall classification flow is described in Fig. 3.

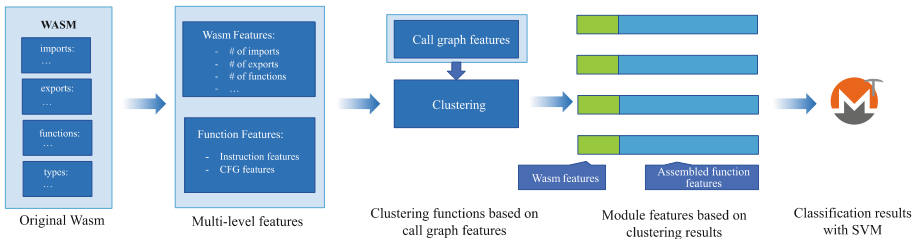


Fig. 3. The classification flow of a WebAssembly module.

**CG and CFG Generation.** It is worth noting that our program analysis is mainly done on WebAssembly code. There are several reasons. First, the asm.js code can be easily converted to WebAssembly bytecodes (e.g., using the “asm2wasm” tool). Second, the

WebAssembly language is well designed. Its bytecodes are simple, clean, and also easier for analysis.

Our analysis is done as follows. First, once the reference (e.g., URL or string) of WebAssembly/asm.js code is obtained, MinerGate constructs the corresponding WebAssembly binary file. Language transformation is also required if the asm.js code is faced. Next, all instructions are carefully analyzed. In a function, adjacent regular instructions (without branch and function invocation instructions) stick together as a basic block. Branch and function invocation instructions link different blocks. Considering the simplicity of WebAssembly bytecodes, this graph construction work can be easily done.

However, there is also a challenge raised in the process. When an indirect function invocation instruction is faced, it is difficult to determine the target function. Our solution is based on the observation: in runtime, when the instruction is executed, the target function's prototype  $F_{target}$  must matches the function prototype  $F_{expected}$  determined by instruction itself. Therefore, MinerGate retrieves  $F_{expected}$ , and scan all functions with proper prototypes to determine the callee function candidates. To avoid false negatives, MinerGate links the function invocation instruction with all function candidates. Our evaluation also shows this simple solution also has relatively low false positives.

**CFG Features.** The critical point in mining code detection is the feature section, because of previous work relied on heuristic methods, we use data-driven feature selection to fill up the missing part of CFG in existing methods by statistics of the graph. Most graph analysis methods rely on graph statistics. Graph statistics can be used to compare graphs, classify graphs, detect anomalies in graphs, and so on. Graph's structure is mapped to a simple numerical space through graph statistic, in which many standard statistical methods can be applied.

In this paper, we introduce graph statistics as an essential part of the analysis of WebAssembly/asm.js. Examples of graph statistics are the number of nodes or the number of edges in a graph, but also more complex measures such as the diameter. Overall, graph statistic can be roughly divided into two categories, global statistics, and nodal statistics. The former describes the global properties of the graph, so only one number is needed for each graph to describe an attribute, and the latter describes the attributes of the nodes in the graph, so each attribute is represented by a vector. In order to analyze the CFG graph as a whole, we use global statistics of the graph as our CFG features, such as graph size, graph volume, graph diameter, etc. When selecting the statistical features of the graph, we mainly consider the work of [3, 14, 33].

**Instruction features.** CryptoNight [27], which is a hash algorithm and heavily used in mining software, explicitly targets mining on general CPUs rather than on ASICs or GPUs. For efficient mining, the algorithm requires about 2 MB of high-speed cache per instance. Cryptography operations, such as XOR, left shift and right shift, are commonly used in CryptoNight algorithm so that we will examine their influences here. In addition to this, we also consider other instructions, not limited to the instructions described earlier, such as various control flow related instructions, memory access instructions, arithmetical operation instructions, and so on.

### 3.4 Data-Driven Feature Selection

We obtained 114 candidate features through the above steps. In our model, we assume that the functions in the mining samples are all related to mining, besides they are mining-related after our manual analysis, and the functions in the benign samples are not related to mining. For the estimation of dependence between features and classes, we use the  $\chi^2$  Test [8], which is commonly used in machine learning algorithms to test dependence between stochastic variables. Following this, we will get scores of features which can be used to select the top N features with the highest values. Part of the top features are shown in the Table 2.

**Table 2.** Top features

Features	Category
Max size of basic blocks	Graph
CFG size	Graph
CFG volume	Graph
Max out degree	Graph
CFG diameter	Graph
Number of loops	Graph
Number of branches	Graph
Number of branches	Instruction
Number of memory instructions	Instruction
Number of arithmetical instructions	Instruction
Number of cryptography instructions	Instruction
Number of instruction <code>get_local</code>	Instruction
Number of instruction <code>set_local</code>	Instruction

We can see from the Table 2 that the graph-related features are more effective than the instruction features, which may be due to the special CFG patterns of the mining code. We can also find that it confirms the previous results [13, 32], cryptography instructions do have influences on the classification results. However, those CFG-related features are more relevant to results. Besides, memory access instructions also showed in the ranking, which is consistent with the fact that the mining code is a memory-intensive application.

Overall, we demonstrate the effects of CFG features and their impact in this section. We will select the top 10 features in the ranking as the basis for subsequent analysis, so each function is represented by a vector of length 10. At this point, we get the features of each function.

**Semantic Signature Matching.** The instruction features or CFG features we discussed earlier can measure the functionality of a piece of code, such as a function or an entire file. The next problem is how to use these features to ensure effectiveness and robustness.



When we examine a payload by analysis of each function, it is difficult to set a proper threshold of malicious functions to discriminate malicious samples. There are many reasons for this dilemma. For example, a library for encryption, it may contain a small number of functions similar to the mining code. On the other hand, malware can also hide in many unrelated code and minimize the number of functions. Similarly, we also face a similar problem when we analyze the payload as a whole.

In this section, we use DBSCAN [28] clustering algorithm to break the solid lines in the analysis. Specifically, we divide the functions into modules according to the call graph (CG), then we generate feature vectors for each module. With clustering functions together, we combine tightly coupled functions into one module, which breaks the boundary between functions, reduces the complexity of data dimension, and enhances the ability against code obfuscation.

DBSCAN is one of the most well-known tools for clustering based on density. The algorithm grows regions with sufficiently high density into clusters and discovers clusters of arbitrary shape in spatial databases with noise. A significant advantage of DBSCAN is that it does not require the number of clusters a priori, unlike k-means, which needs to be specified manually. The number of modules in a payload is uncertain, and the DBSCAN can determine the number of clusters for us. Another advantage is that it does not rely on Euclidean distance, because it is inappropriate to convert the CG to Euclidean distance.

The algorithm requires two parameters:  $\epsilon$ -neighborhood of points and the minimum number of points (*MinPts*) required to form a dense region. In order to apply the algorithm to our domain, we need to redefine the  $\epsilon$ -neighborhood  $N_\epsilon(p)$  of a point (a function in this paper),  $N_\epsilon(p) = \{q \in D \mid \text{if } p \text{ calls } q\}$ , in which  $D$  means the database of the functions.

When *MinPts* = 4, the results of the cluster analysis on the mining payload of CoinHive are shown in Fig. 4. For the sake of brevity, only functions related to `cryptonight_hash_variant_1` are included in the figure. It can be seen that functions related to `cryptonight_hash_variant_1` are divided into two clusters. With manual analysis, it can be seen that the functions in Cluster 1 are mainly related to encryption, and functions in Cluster 2 are mainly related to memory operations. The main reason for this result is that the effect of code is closely related to the functions it calls.

Then we generate the feature vectors for each module with the methods described in Sect. 3.3, which combine with the labels will be used to train the SVM classifier. If the analysis result for a sample contains one or more malicious “modules”, we label the whole sample as malicious.

### 3.5 Security Enforcement

Although the user is executing malicious code while detection is occurring, the main threat of cryptojacking is it occupies a lot of system resources, instead of stealing sensitive information or damaging the system like traditional malware. As long as it can prevent its operation in time, its impact is limited. Our security enforcement is mainly provided in the injected `stub.js` (Sect. 3.2). With detection of the mining code, MinerGate notifies `stub.js` through pre-established WebSocket connection. This connection can be kept alive even when CPU, memory, and network are occupied by mining code.

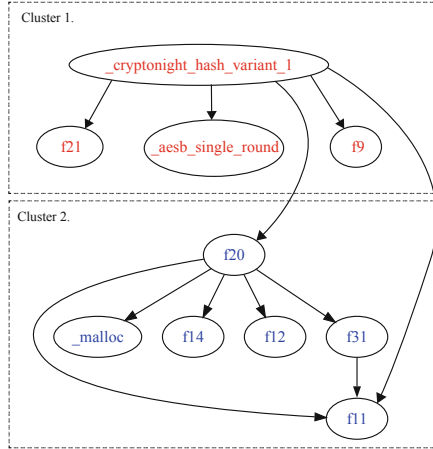


Fig. 4. The clustering result on CoinHive with  $MinPts = 4$ , only includes functions that are related to `cryptonight_hash_variant_1`.

Stub.js can also apply pre-defined security rules. For example, stub.js can directly terminate the execution of the mining code. This is achieved by stopping the worker or removing the iframe with preset callback functions, and mining code running in the main thread of the web page will be closed immediately. Hijacked APIs (e.g., `eval`, `WebAssembly.instantiate`, etc.) in users' browser will refuse to execute code that are marked as untrusted.

The mining code needs to use WebSocket to communicate with the mining pool to obtain the necessary parameters for mining. After discovering the mining payloads, MinerGate can stop the WebSocket connection in the same context by API hooking, so that we can cut off the communications between the miner and the mining pools to forcefully terminate the mining activities.

## 4 Evaluation

In our evaluation, we first verify the correctness of MinerGate by testing MinerGate in a real environment. Then, we check MinerGate's performance, and confirm MinerGate introduces relatively low overhead. Last, we verify the accuracy of MinerGate.

Our test environment consists of PCs with different OS (i.e., Windows 10 version 1809, macOS 10.14.4, and Ubuntu 18.04), and smartphones (Nexus 5 with Android 6).

### 4.1 Dataset

There are currently no reliable labeled mining site datasets or WebAssembly/asm.js datasets. To investigate the deployment of WebAssembly in the real world, we deployed a distributed crawler cluster on Azure using Kubernetes to acquire WebAssembly files. The crawlers in the cluster are built upon Chrome and are driven by the "stub.js" described in Sect. 3.2. The cluster includes 120 crawler instances running on the top of 15 physical

nodes. We crawled the Alexa top 1 M sites and randomly selected 10 different URLs from each top site for the next level of crawling. For each website, we spend up to 30 s to load the page and close the page after 10 s. If WebAssembly is detected on the page, the page will be closed immediately (Table 3).

**Table 3.** Summary of our dataset and key findings

Crawling period	Apr. 25, 2019 - May. 13, 2019
# of crawled websites	10.5 M
# of <i>benign</i> web pages with WebAssembly/asm.js	5,030
# of <i>benign</i> WebAssembly/asm.js from NPM	946
# of <i>malicious</i> mining related web pages	4,659

As a result, we visited a total of 10.5 M pages and found 9,689 web pages containing WebAssembly code, which covers 2,657 registered domains (such as `bbc.co.uk`) and 3,012 FQDNs (such as `forums.bbc.co.uk`), and 1,118 top sites contain the WebAssembly code in their home page. The top 15 categories of websites that have deployed WebAssembly are shown in Table 4.

**Table 4.** Top 15 categories of websites which include WebAssembly.

Categories	#
Adult Content	595
News/Weather/Information	410
Blogs	199
Video & Computer Games	137
Streaming Media	105
Technology & Computing	92
Illegal Content	82
File Sharing	76
Television & Video	61
Sports	38
Weapons	36
Movies	32
Message Boards	31
Shopping	31
Arts & Entertainment	31

To build our training dataset of cryptojacking code, we first match the existing blacklist (uBlock [25], NoCoin [24] and CoinBlockerLists [35]) based on the source URL of WebAssembly/asm.js. If the payloads are from the blacklist URLs, we label the sample as malicious. Some previously unknown mining samples were recognized by reverse engineering analysis with JEB decompiler [21]. Through examination, we found 164 benign WebAssembly samples in 3296 pages (1,735 websites), 55 kinds of malicious WebAssembly, and 6 kinds of malicious asm.js samples in 4659 pages (832 websites) for cryptojacking attacks. We also found that there are 25 undetected malicious WebAssembly samples with the help of VirusTotal [29]. It is worth mentioning that many mining service providers will provide a different bootstrap JavaScript to avoid detection each time they are accessed, but the WebAssembly payloads extracted from them are generally the same. This means that we can analyze the key WebAssembly or asm.js to obtain better analysis results. The top 15 categories of websites which bring Crypto-jacking attacks are shown as Table 5.

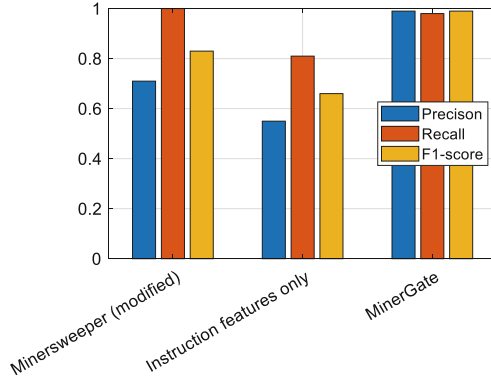
**Table 5.** Top 15 categories of websites which include Cryptojacking.

Categories	#
Adult Content	148
Illegal Content	64
News/Weather/Information	61
File Sharing	42
Technology & Computing	36
Sports	29
Television & Video	27
Streaming Media	26
Video & Computer Games	24
Comic Books/Anime/Manga	22
Arts & Entertainment	16
Movies	14
Web Design/HTML	12
Music & Audio	11
Arts & Entertainment	10

To further build a ground-truth set of non-cryptojacking WebAssembly/asm.js samples, we installed all the packages that be tagged as WebAssembly/asm.js from NPM, which is the largest JavaScript software registry. After the installation is complete, we extract the WebAssembly and asm.js files from the installation folder. The projects we collected include various kinds of libraries and applications, such as video coding, data encryption, data processing, web framework, image processing, physics engine, game framework, and so on. We will publish these samples with labels for future research.

## 4.2 Accuracy

In this section, we examine MinerGate’s classification accuracy on the ground-truth training dataset and compare it with other existing detection techniques. In order to accurately measure the performance of the classifier, we ran 10-fold cross-validation on our dataset. As shown in Fig. 5, the complete MinerGate performs with 99% precision, 98% recall and 99% f1-score. We can also see that the accuracy rate has been greatly improved after adding CFG features and cluster analysis.



**Fig. 5.** Results of Cryptojacking discrimination and comparison to other approaches.

In addition to this, the results of Minesweeper [12] are not satisfactory enough. One reason is that they use Chrome’s undocumented API (`-dump-wasm-module`) to dump WebAssembly. But this API cannot dump WebAssembly loaded by `instantiateStreaming()` or `compileStreaming()`. To this end, we have implemented a modified version of MineSweeper to take advantage of WebAssembly dumped using our system. As shown in Fig. 5, MineSweeper tends to classify samples as malicious, resulting in lower accuracy and high recall.

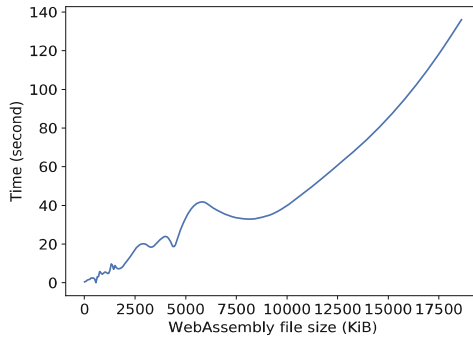
## 4.3 Overhead

First, we test the overhead introduced by MinerGate on benign websites that do not contain WebAssembly/`asm.js` code. We evaluated the overhead of the system by accessing 1,000 benign web pages and measuring the load time of web pages by enabling/disabling the proxy. The overhead is about 6%, and we found that there is only the overhead of a proxy in this case, because our protected code is only triggered if the WebAssembly is loaded.

Then, we test the overhead on infected websites. We still evaluated the overhead by accessing 1,000 malicious web pages and measuring the load time of web pages by enabling/disabling the proxy. We do not prohibit the execution of the mining program during the overhead evaluation, as this behavior itself will speed up the access of the web page. The overhead is less than 9%, the extra overhead here is mainly from the transmission of WebAssembly.

The transparent proxy itself has no complicated operations. It simply inserts our protection code in the response after the browser makes the request. This is different from the instrumentation in the general sense. Therefore, the transparent proxy’s overhead is less than 9% in our evaluation. Since each module of MinerGate is independent, it can be deployed in a distributed manner. To be noticed, both the code injection module and the malicious code analysis module can be independently deployed on multiple machines, so the impact of multiple devices on performance is limited.

Since we have considered performance issues when considering hooks, all code that involves external calls is asynchronous, and only minor performance impacts occur when the program calls a function that is hooked. So overall, our instrumentation will not affect the efficiency of JavaScript. But the time at which the gateway analyzes the code is still important because malicious code can consume a lot of power or block the execution of necessary transactions during the analysis. For background analysis, we plot Fig. 6 which shows the time needed to process different sizes of WebAssembly files.



**Fig. 6.** The time for processing different sizes of WebAssembly in the background.

## 5 Related Work

Until now, there is no practical generic defense solution against Web-based cryptojacking attacks. One of the limitations of existing methods is that the semantic model of the mining payload is not efficient enough to distinguish between malicious mining applications and benign applications. More importantly, there is currently no non-intrusive defense solution, and all existing work requires modifications to the browser and even the operating system. In this paper, we first apply the CFG (control flow graph), CG (call graph) features to the malicious WebAssembly/asm.js classification, which reviews the problem from another perspective. Since the payload analysis is static, the MinerGate provides a lightweight defense and requires no browser modification by deploying the system to the gateway. The results of comparison with other existing related works are shown in the Table 6.

**Table 6.** Comparison with other related works.

Name	Scalable (No browser modification)	JavaScript obfuscation resistance	Security enforcement	Low overhead	Low false positives	Used features
MineSweeper [12]	×	✓	×	✓	×	CPU, WebAssembly Instructions, etc.
SEISMIC [32]	×	✓	×	×	×	WebAssembly Instructions, etc.
BMDetector [17]	×	×	×	✓	×	JavaScript heap and stack info, etc.
Outguard [12]	×	✓	×	✓	×	JavaScript loading, etc.
CMTracker [10]	×	✓	×	✓	×	JavaScript stack info, etc.
MinerGate	✓	✓	✓	✓	✓	CG, CFG, WebAssembly instructions

**Blacklist or Keyword-Based Methods.** Some dedicated extensions [24, 25], browsers [2, 22] provide blacklists and keywords to alleviate cryptojacking by manually running honeypot [23] and collecting URLs on reports to expand the list. However, the updates of blacklists and keywords are hard to keep up with the iterative steps of malicious code, which makes the defense always behind the attack.

**Instruction Features Based Methods.** In the work of Konoth et al. [13], they use static analysis to count the number of cryptographic instructions (`i32.add`, `i32.and`, `i32.shl`, `i32.shr_u`, `i32.xor`) and loops to detect CryptoNight algorithm. The work of Wang et al. [32] is similar, but the number of instructions is calculated by dynamic instrumentation. However, these cryptographic instructions also exist in many benign applications, such as data encryption, image processing, video encoding, game engines and so on, which will make it difficult to classify these samples accurately.

**Stack Dump-Based Methods.** The critical observation of stack dump-based methods is that cryptocurrency miners run mining workloads with repeated patterns. In the work of Hong et al. [10], shows that a regular web page rarely repeats the same calling stack for more than 5.60% of the execution time. However, such performance profile requires modifications to the browser kernel, which makes it impractical. In the work of Liu [17], they extract string features from heap and stack snapshot and use RNN to detect the mining programs. This type of method built on strings or keywords is unreliable and can be easily bypassed by JavaScript code.

## 6 Conclusions and Future Work

With a deeper understanding of the semantics of WebAssembly/asm.js, we designed a novel generic defense solution MinerGate against Web-based cryptojacking attacks. By decentralizing computing tasks to the gateway, we implemented a common protection scheme with the lowest overhead in known scenarios, which does not require modification of the browser. Through data-driven feature selection, we not only further demonstrate the effectiveness of instruction-level features but also indicate the excellent performance of CFG features in malicious code detection.

The main limitations exist in two aspects. First of all, considering that JavaScript is a highly dynamic and continuously evolving language, it is difficult to prove that the APIs we intercept is always complete. On the other hand, since this work uses a machine learning-based method, there is the possibility of constructing adversary samples, and we may need extra work to defend against it.

**Acknowledgments.** This project is supported by National Natural Science Foundation of China (No. 61972224).

## References

1. Ana, A.: Report: Some crypto mining apps remain in Google play store despite recent ban (2018). <https://cointelegraph.com/news/report-some-crypto-mining-apps-remain-in-google-play-store-despite-recent-ban>. Accessed 21 Nov 2019
2. Andrea, M.: Firefox: implement cryptomining URL-classifier (2019). <https://hg.mozilla.org/mozilla-central/rev/d503dc3fd033>. Accessed 01 May 2020
3. Barrat, A., Barthelemy, M., Pastor-Satorras, R., Vespignani, A.: The architecture of complex weighted networks. *Proc. Natl. Acad. Sci.* **101**(11), 3747–3752 (2004)
4. Catalin, C.: New browser attack lets hackers run bad code even after users leave a web page (2019). <https://www.zdnet.com/article/new-browser-attack-lets-hackers-run-bad-code-even-after-users-leave-a-web-page/>. Accessed 01 May 2020
5. Daniel, P.: 8 illicit crypto-mining windows apps removed from microsoft store (2019). <https://www.coindesk.com/8-illicit-crypto-mining-windows-apps-removed-from-microsoft-store>. Accessed 01 May 2020
6. David, H., Luke, W., Alon, Z.: asm.js working draft (2018). <http://asmjs.org/spec/latest/>
7. Eyal, I., Sirer, E.G.: Majority is not enough: Bitcoin mining is vulnerable. *Commun. ACM* **61**(7), 95–102 (2018)
8. Pearson, K.: X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *London Edinburgh Dublin Philos. Mag. J. Sci.* **50**(302), 157–175 (1900). <https://doi.org/10.1080/14786440009463897>
9. Group, W.C.: Webassembly specification (2018). [https://webassembly.github.io/spec/core/\\_download/WebAssembly.pdf](https://webassembly.github.io/spec/core/_download/WebAssembly.pdf). Accessed 01 May 2020
10. Hong, G., et al.: How you get shot in the back: a systematical study about cryptojacking in the real world. In: *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, CCS 2018*, pp. 1701–1713. ACM, New York (2018). <https://doi.org/10.1145/3243734.3243840>. <http://doi.acm.org/10.1145/3243734.3243840>



11. npm Inc.: npm — the heart of the modern development community (2018). <https://www.npmjs.com/>. Accessed 01 May 2020
12. Kharraz, A., et al.: Outguard: detecting in-browser covert cryptocurrency mining in the wild (2019)
13. Konoth, R.K., et al.: Minesweeper: an in-depth look into drive-by cryptocurrency mining and its defense. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, pp. 1714–1730. ACM (2018)
14. Kunegis, J.: KONECT – the Koblenz network collection. In: Proceedings of International Conference on World Wide Web Companion, pp. 1343–1350 (2013). <http://dl.acm.org/citation.cfm?id=2488173>
15. Newman, L.H.: Hack brief: hackers enlisted Tesla’s public cloud to mine cryptocurrencies (2018). <https://www.wired.com/story/cryptojacking-tesla-amazon-cloud/>. Accessed 01 May 2020
16. Lindsey, O.: Cryptojacking attack found on los angeles times website (2018). <https://threatpost.com/cryptojacking-attack-found-on-los-angeles-times-website/130041/>. Accessed 01 May 2020
17. Liu, J., Zhao, Z., Cui, X., Wang, Z., Liu, Q.: A novel approach for detecting browser-based silent miner. In: Proceedings - 2018 IEEE 3rd International Conference on Data Science in Cyberspace, DSC 2018, Guangzhou, China, pp. 490–497. IEEE, June 2018. <https://doi.org/10.1109/DSC.2018.00079>. <https://ieeexplore.ieee.org/document/8411900/>
18. Magazinius, J., Phung, P.H., Sands, D.: Safe wrappers and sane policies for self protecting JavaScript. In: Aura, T., Järvinen, K., Nyberg, K. (eds.) NordSec 2010. LNCS, vol. 7127, pp. 239–255. Springer, Heidelberg (2012). [https://doi.org/10.1007/978-3-642-27937-9\\_17](https://doi.org/10.1007/978-3-642-27937-9_17)
19. Neil, B.: Kaspersky reports 13 million cryptojacking attempts this year, January 2018. <https://www.cryptolinenews.com/2018/12/13-million-cryptojacking-says-kaspersky/>. Accessed 01 May 2020
20. Newman, L.H.: Now cryptojacking threatens critical infrastructure too (2018). <https://www.wired.com/story/cryptojacking-critical-infrastructure/>. Accessed 01 May 2020
21. Nicolas, F., Joan, C., Cedric, L.: Jeb decompiler (2018). <https://www.pnfsoftware.com/jeb/>. Accessed 01 May 2020
22. Opera: Cryptojacking test (2018). <https://cryptojackingtest.com/>. Accessed 01 May 2020
23. Prakash: Drmine (2018). <https://github.com/1lastBr3ath/drmine/>. Accessed 01 May 2020
24. Rafael, K.: NoCoin (2018). <https://github.com/keraf/NoCoin/>. Accessed 01 May 2020
25. Raymond, H.: ublock (2018). <https://github.com/gorhill/uBlock/>. Accessed 01 May 2020
26. Rossberg, A., et al.: Bringing the web up to speed with webassembly. Commun. ACM **61**(12), 107–115 (2018). <https://doi.org/10.1145/3282510>
27. Seigen, Max, J., Tuomo, N., Neocortex, Antonio, M.J.: Cryptonight hash function (2013). <https://cryptonote.org/cns/cns008.txt>. Accessed 01 May 2020
28. Simoudis, E., Han, J., Fayyad, U.M. (eds.): Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD 1996), Portland, Oregon, USA. AAAI Press (1996). <http://www.aaai.org/Library/KDD/kdd96contents.php>
29. VirusTotal: Virustotal (2018). <https://www.virustotal.com/>. Accessed 01 May 2020
30. W3C: Web workers (2015). <https://www.w3.org/TR/workers/>. Accessed 01 May 2020
31. W3C: The websocket api. <https://www.w3.org/TR/websockets/>. Accessed 01 May 2020
32. Wang, W., Ferrell, B., Xu, X., Hamlen, K.W., Hao, S.: SEISMIC: SEcure in-lined script monitors for interrupting cryptojacks. In: Lopez, J., Zhou, J., Soriano, M. (eds.) ESORICS 2018. LNCS, vol. 11099, pp. 122–142. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-98989-1\\_7](https://doi.org/10.1007/978-3-319-98989-1_7)
33. Watts, D.J., Strogatz, S.H.: Collective dynamics of ‘small-world’ networks. Nature **393**(6684), 440 (1998)

34. Wikipedia: Cryptocurrency (2018). <https://en.wikipedia.org/wiki/Cryptocurrency>. Accessed 01 May 2020
35. ZeroDot1: Coinblockerlists (2018). <https://zerodot1.gitlab.io/CoinBlockerListsWeb/index.htm>. Accessed 01 May 2020

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





# Research on Industrial Internet Security Emergency Management Framework Based on Blockchain: Take China as an Example

Haibo Huang<sup>1,2</sup> , Yuxi Gao<sup>2</sup>  , Min Yan<sup>3</sup> , and Xiaofan Zhang<sup>2</sup> 

<sup>1</sup> Beijing University of Posts and Telecommunications, Beijing 100876, China  
poehuang1@163.com

<sup>2</sup> China Industrial Control Systems Cyber Security Response Team, Beijing 100040, China  
gemmagao@126.com

<sup>3</sup> Institute of Software, Chinese Academy of Sciences, Beijing 100093, China

**Abstract.** Building a national unified ISEMS (industrial internet security emergency management system) plays an important role in industrial cybersecurity defense. However, due to technical and management constraints, the current ISEMS has problems such as scattered security organizations, poor sharing channels, and fails to form an overall security guarantee capability for threat reporting, analyzing, warning, and disposing. The blockchain technology has the characters of decentralized trust construction, inter-organizational data sharing, data integrity assurance, data traceability, which just meets the requirements of the emergency management process. This paper analyzes the situation and challenges of ISEMS, describes the system architecture and organizational structure based on the blockchain, and describes the key implementation processes of blockchain-based ISEMS, including threat report, risk analysis, warning release and emergency response.

**Keywords:** Industrial cybersecurity · Emergency management · Consortium blockchain

## 1 Introduction

With the rapid development of global information technology and the deep reform of industrial structure adjustment, China's industrialization and informatization have deepened continuously, and the Industrial Internet has developed rapidly. According to statistics from the MIIT (Ministry of Industry and Information Technology), there are more than 50 Industrial Internet platforms having certain industrial and regional influences by 2019, some of which connected to more than 100,000 industrial equipment. With the rapid development of the industry, security threats are intensified increasingly, and Industrial Internet security events such as supply chain attacks, ransomware attacks, and

---

This work was financially supported by the National Key Research and Development Program of China (2018YFB2100400).

© The Author(s) 2020

W. Lu et al. (Eds.): CNCERT 2020, CCIS 1299, pp. 71–85, 2020.

[https://doi.org/10.1007/978-981-33-4922-3\\_6](https://doi.org/10.1007/978-981-33-4922-3_6)

data leaks are exposed frequently. Meanwhile, China's ISEMS management framework lacks the systematic design. Therefore, it is necessary to construct a comprehensive and secure emergency response mechanism and take closed-loop defense measures to active defense, real-time sensing, and emergency recovery. Building a national unified emergency management system is an important part of the Industrial Internet security defense. It comprehensively analyzes threat information through technology and management methods, builds capabilities such as early warning, notification, emergency handling, and information sharing, also helps emergency department dispatch resources, investigate risk, dispose emergency, to maintain the security of Industrial Internet platforms, networks, controls, equipment, and data.

However, owing to the scattered industrial internet emergency management institutions, the inconsistent sharing channels, and the insufficient risk analysis of the industrial internet infrastructure, it is hard to form a global security capability. The blockchain, which combining data blocks into a "chain" structure in chronological order uses distributed accounting, peer-to-peer communication, cryptographic technology, consensus mechanisms, and the disclosure of intelligent contracts to achieve a decentralized and tamper-proofing data storage [1], and can solve problems such as scattered institutions, unreliable data sources, and inability to achieve multi-party storage in industrial internet emergency management. It also well meets the needs of transparent and credible requirements in multiple parties during the emergency information management process.

## 2 Situation and Challenge

### 2.1 Organizational Structure

Seen from the Fig. 1, China's industrial internet emergency organization is a tree management structure. The root node is the national industrial internet security authority, mainly responsible for the emergency management function including early warning, risk notification, emergency response, information sharing, etc. Secondary nodes are provincial industrial internet security authorities and state-owned enterprises, which responsible for performing its management supervisors. Security vendors and scientific research institutes are also secondary nodes, responsible for reporting risk information and conducting risk research and emergency response. The third-level nodes are mainly city-level emergency management departments, small security vendors and research institutions, of which the function is consistent with the secondary node, and local industrial enterprises which are the main bodies in carrying out the disposal of risk and incident.

In recent years, the MIIT has continuously invested special funds to support the construction of industrial internet threat information sharing and emergency response command platform, through which the country can effectively improve the ability of grasping risk information, carrying out emergency command and response, nevertheless, it has not yet formed a nationwide ISEMS with vertical linkage and horizontal communication.

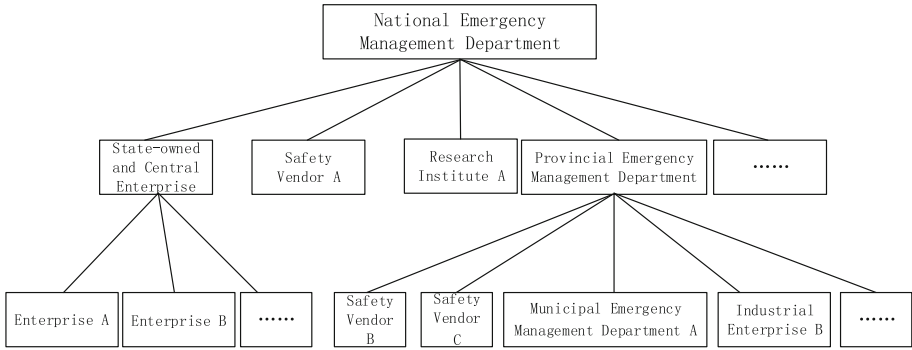


Fig. 1. Management structure of China’s ISEMS

2.2 Technical Status

**Related Work.** In terms of ISEMS research, Zhang Zhen et al. [2] analyzed the content and characteristics of the U.S. cybersecurity emergency management system, and made suggestions on the construction of China’s cybersecurity emergency management system from the legal system, organizational structure, and operating mechanism. The establishment of command systems with incident response, information sharing, and emergency response from the technical level has not been further studied. Liu Feng [3], Zhao Xu et al. [4] proposed technical solutions to the provincial cybersecurity emergency management platform from the aspects of security access design, basic data collection, and command function implementation, but still lacking of consideration on information sharing, multi-party coordination, mutual identity trust, regulatory review, etc. Li Ruoyu et al. [5] established an emergency response system model for governments and enterprises, and pointed out the indispensability of security service vendors in national cybersecurity emergency response work, but did not give specific implementation plans at the operational and technical levels. Since 2003, the U.S. Department of Homeland Security has implemented continuous monitoring, early warning and response to Internet export threats of government agencies through the Einstein Plan. However, due to compatibility and diverse issues, only 68.7% of government agencies have achieved the goals of the Einstein Project by 2019 [6].

**Problems.** The problems in the construction of national ISEMS as follows.

1. Isolated islands of information. The communication among the information systems of institutions and organizations is incomplete. In the early stage of the big data era, the isolated information island problem is common in various industries and fields [7, 8]. Due to historical reasons such as insufficient top-level design and system performance constraints, the governments, security vendors and industrial enterprises have built independent threat information databases, vulnerability databases and emergency disposal systems, leading to an obvious “data chimney” effect, which comprehensively restricts the work efficiency of threat submission, sharing, and emergency disposal, etc.

2. **Poor threat sharing.** The security subject of Industrial Internet has weak willingness to share threat information. On the one hand, due to the high sharing frequency and complex path of industrial internet security data, data leakage may occur in the transmission process or non-legal endpoints; On the other hand, industrial internet security information has its own particular characteristics such as being multi-sourced, heterogeneous and distributed. Data classify measures are deficient to ensure the rationality of the scope of information sharing. In addition, for which the current information sharing rights and responsibilities are not clear and the audit tracking ability is insufficient, both leads to the enterprises unwilling to share information as “private property”, and the competent authorities of industry are afraid of compulsory sharing.
3. **Untrusted data source.** Phishing, extortion, mining, etc. have become an important threat to Industrial Internet Security [9, 10]. In addition to directly attacking industrial enterprises, due to the lack of effective authentication and security measures for information source and transmission, hackers utilize the defects of insufficient end-user’s management ability to spread malicious code embedded in risk information through industrial internet security emergency, causing a more targeted large-scale attack on competent authorities of industry and enterprises.
4. **Inefficient emergency response.** China has not yet established a unified emergency disposal command and communication system. The disposal of major cybersecurity incidents still stays in traditional ways such as SMS and telephone. It is difficult to meet the requirements in timeliness, portability, confidentiality, and other aspects. In addition, due to the lack of recognized evaluation methods, the security technology team cannot get the point in the first time after the security incident and hardly obtain evidence and division of responsibilities. With the combination of above two analysis, the repetitive emergency work has been carried out continuously.
5. **Lack of motivation.** As the main body of information reporting and emergency response, security vendors play an indispensable role in the emergency system, also the key to the effective implementation of the national industrial Internet security emergency. It is difficult to ensure the sustainability simply with the incentive measures of social responsibility. More effective measures must be introduced to improve the positivity of security vendors.
6. **System security factors.** The national ISEMS is intricacy while enormous system, with large cyber-attack surface and high security risk. Once centralized data storage infrastructure being attacked may lead to the collapse of the whole system. Meanwhile, with complex and multi-subject end-user identity, the ineffective management of all users results in the system vulnerability.

### 2.3 Challenge

In view of the issues above, the construction of national ISEMS has the following challenge.

1. **Unified interface standard.** The construction of a unified standard system interface and protocol could realize the interconnection of emergency information, form an emergency coordination and disposal mechanism with timely response and feedback.

which could provide channels for central and local emergency institute and organization to obtain and convey emergency information, realize the interconnection of emergency information systems of superior and subordinate units.

2. Confidentiality, availability and non-repudiation. Traditional information systems are vulnerable to single point of failure due to centralization. Through multi centralized storage deployment Enhance the robustness and usability of the system. In addition, by verifying the identity legitimacy and rationality of the users, the data source can be trusted, managed and traceable. Third, ensure the security of data transmission, storage and sharing, especially the integrity confidentiality and of data.
3. User incentive. In the process of information report and emergency disposal involving security vendors and scientific research institutions, the competitive ranking mechanism can improve the enthusiasm of participation, grasp the technical strength of each unit, so that an appropriate emergency response team could be found timely and accurately in a security incident. Second, for the industry competent departments and industrial enterprises, introduce the reward-punishment and assessment mechanism combined with national laws and industry regulations, implement the responsibility, and ensure the sustainable development of the ISEIMS.
4. Data classification. The system should store all kinds of data information, including system vulnerabilities, early warning notifications, typical cases, knowledge base, etc. In order to ensure the security and controllability of the data as a strategic resource, Data classify and grade according to its ownership, application scope, sensitivity and other dimensions, so as to improve the sharing and circulation of data use while protecting user privacy, realize the effective balance of data privacy and social utility.

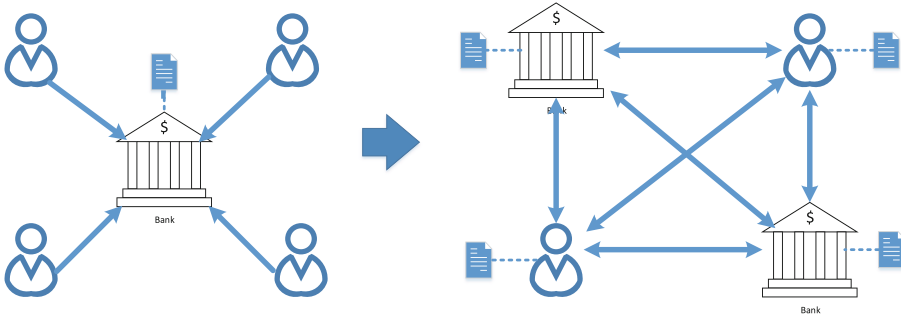
### **3 Overview of Blockchain**

#### **3.1 Principle**

Blockchain technology is a distributed ledger technology that uses the linked data structure to verify, store, generate, update data and ensure its transmission security. It is an integrated application and innovative combination of existing technologies such as distributed data storage, point-to-point transmission, consensus mechanism, encryption algorithm, etc. [11]. Its most significant technical feature is to change the centralization to decentralization, as shown in Fig. 2.

#### **3.2 Research Status**

Blockchain has the technical advantages of decentralization, non-tampering, traceability, high reliability and high availability, began to form distributed collaboration architecture supporting various typical industries [12]. According to the degree of openness, blockchain can be divided into three types: Public Blockchain, Private Blockchain and Consortium Blockchain. Public blockchain is completely decentralized, and also, any user can join the network, access and write data. The typical representatives are bitcoin and Ethereum. Private Blockchain is partial decentralized, and also, only part of users



**Fig. 2.** Centralization and decentralization

can access, read and write data with internal permissions. Consortium Blockchain is multi centralized, and only authorized organizations can join the network. Its organization nodes are fully trusted and strongly scalable. Its scale could rise from institutional enterprises to the national level [13].

With the gradual development of blockchain, the research on its key technologies has shown multiple development directions, Herrera joancommarti [14], Saxena [15] do research on privacy protection issues such as anonymity and hybrid protocol of Bitcoin. Kishigami [16] and others proposed a blockchain-based digital content publishing system to move blockchain technology from theory to practice with intelligent contracts. Paul [17] and others calculated and verified the bitcoin mining energy consumption scheme, and studied the resource loss of blockchain technology. Mougayar [18] and others analyzed the trend of bitcoin vulnerability and countermeasures to study blockchain security technology. In addition, SANA, bjabendu, Jian Chen and others studied the application, management, security and openness of blockchain technology in the Internet of things, big data and other new fields [19–21].

## 4 Consortium-Blockchain-Based ISEMS

### 4.1 Base Model

The earliest form of blockchain is public blockchain, but the public blockchain is completely decentralized and difficult to supervise, which is different from China's governance structure. Consortium Blockchain is a form of "supervision friendly" blockchain, which is easy to pass the access system and use contract automation supervision to meet regulatory needs. Generally, the industry is oriented to institutions and enterprises, which need to solve the trust problems among them, and require the organizations that set up the blockchain to conduct identity authentication. The number of members in the blockchain can be controlled, and the characteristics of the Consortium Blockchain fully meet these needs. The Consortium Blockchain adopts the federal access mechanism with certain trust premise, which has a large space in the selection of efficient consensus algorithm and is easy to find a balance between security and performance. In recent years, various industries are actively exploring the "blockchain +" industry application mode. Based on the blockchain as a service (BaaS), the rapid construction



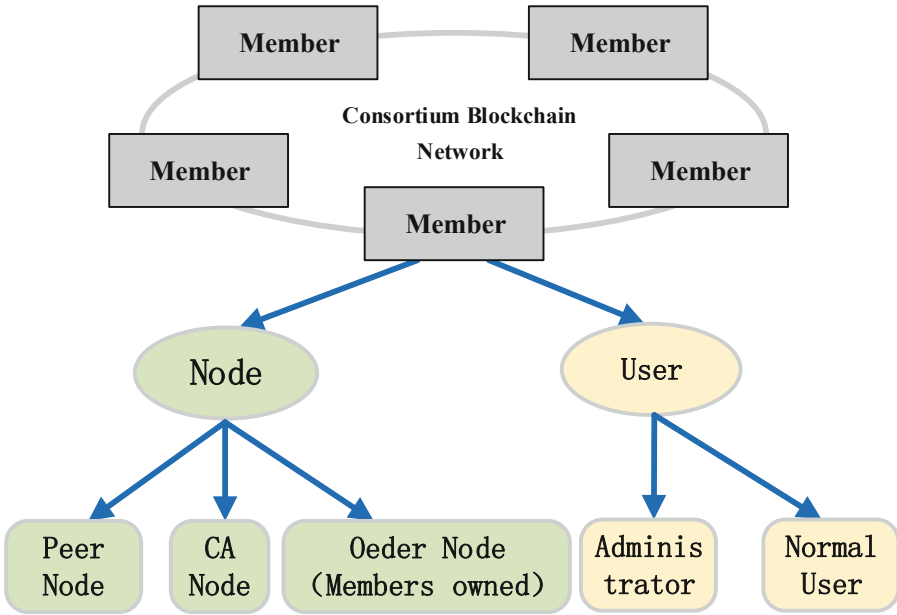
of blockchain network and the implementation of industry application are gradually deepened. By deeply combining blockchain technology with cloud computing, BaaS platform integrates the underlying computing resources, blockchain service functions and upper business functions through the centralized management, realizes the available and dynamic expansion of the underlying blockchain framework with virtualization container, support the ability of multi-user, multi-chain, shared storage, log monitoring, etc., and greatly reduces blockchain barriers.

In this scheme, the Hyperledger Fabric Consortium Blockchain is proposed as the technology selection to design the ISEMS architecture. Fabric is the most widely used project in the hyper ledger blockchain open source project, aiming to promote the cross-industry application of blockchain, and its architecture model is shown in Fig. 3. Fabric Consortium Blockchain network is composed of members, which refers to the organization, also composed of several organizations with cooperative relationship. The users in the Consortium Blockchain belong to the members of the blockchain, which can be divided into two types, administrator and ordinary user. Administrator is the manager of blockchain, who can choose to join, exit the chain and install the intelligent contract. The user is the initiator of the transaction, and can interact with the blockchain network through the client or SDK. The nodes in the Consortium Blockchain refer to the physical resources actually running in the Consortium Blockchain. Each member of the blockchain has one or more peer nodes and CA nodes. Peer node is the node that each member can realize ledger storage, which includes endorsement node, bookkeeping node and master node. Endorsement refers to the process that a specific peer node executes a series of transactions and verifies their validity, and returns a successful or failed endorsement response to the members who generate the transaction proposal. The function of CA node is to provide members in Fabric network with identity information based on digital certificate. The order node is jointly owned by the members of the blockchain. It is mainly responsible for collecting and sorting the received transactions of protection endorsement signature, generating blocks in sequence and broadcasting the transactions, in order to ensure that the nodes in the same chain receive the same messages and have the same logical order.

In the process of ISEIM, the organization is scattered and diverse. Based on the Consortium Blockchain, it can solve the problems, such as the organization is not mutual trust, the data source is not credible, not achieving multi-party storage, etc.

## 4.2 Architecture

The technical architecture of the ISEMS is shown in Fig. 4, which includes the underlying infrastructure, the intermediate service layer, and the calling API provided by the upper application system. In order to quickly start the consortium blockchain, the basic underlying blockchain framework uses the Swarm or K8s group management technology and container management technology to build the Fabric blockchain network framework, and automatically starts and manages the CA and peer nodes for blockchain members. The blockchain service layer management includes five modules of basic services, contract management, security management, operation monitoring, and query engines. Among them, the basic service module implements storage management, pluggable consensus mechanism, and network communication services. The contract management

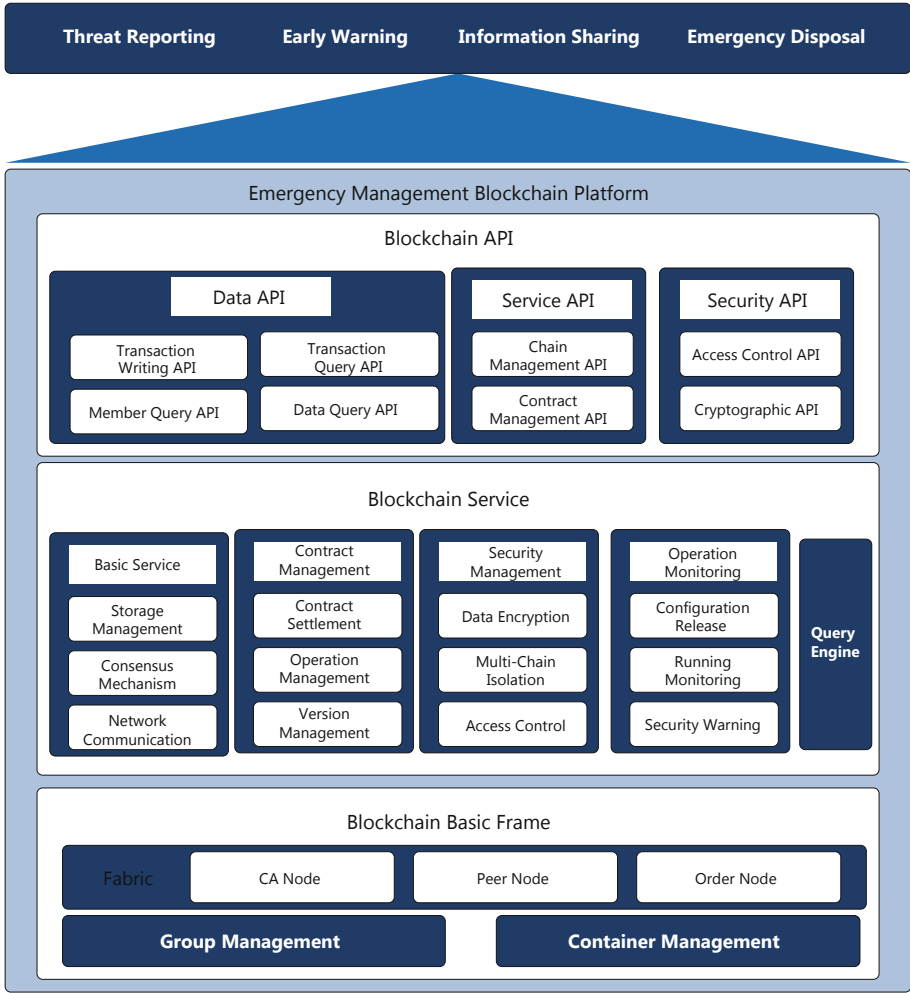


**Fig. 3.** Consortium blockchain network architecture model

module implements intelligent installation, operation, and version management. The security management module implements security mechanisms such as data encryption, access control, and multi-chain isolation. Core modules, such as operation monitoring and query engines, provide basic services for upper data API and blockchain service API interfaces. The blockchain API layer provides blockchain transaction read-write API, chain and contract management API, access control API and encryption authentication API, etc. It provides call interfaces for application requirements such as upper-level risk reporting, early warning release, information sharing, and emergency disposal.

### 4.3 Organization Structure

Organizational Structure of ISEMS based on consortium Blockchain is shown in Fig. 5. In the business scenario of ISEIM, the organization nodes involved are not completely parallel in function positioning. For example, local emergency management departments are responsible for reviewing the risk and vulnerability information reported by the regional security vendors, industrial enterprises and research institutions, and reporting to the central authorities only after passing the review. Therefore, for ISEIM business is multi-level and needs timely supervision, this scheme combines multi-chain and cross-chain technology to build multi-level consortium blockchain. Details can be seen in Fig. 5. The first-level members are composed of national and provincial industrial internet security emergency management department, security enterprises, state-owned enterprises and central enterprises. The second-level members are composed of provincial and municipal industrial internet security emergency management department, security



**Fig. 4.** Technical framework of ISEMS based on consortium blockchain

vendors, research institutions, etc. The provincial and municipal departments exist in both the first level and the second level. Each member of the primary and secondary consortium blockchain has its own administrator and ordinary user group. The administrator is responsible for consortium blockchain management, contract management and other functions. The subordinate local organization can set multiple users to report risks, receive early warnings and dispose emergencies. The administrator of national industrial internet security emergency management department is also responsible for blockchain management and contract management. Different users can respectively call intelligent contracts to implement information sharing, emergency strategy release, early warning, etc. For the local industrial Internet security emergency management departments,

it is also necessary to create multiple users for risk reporting, information receiving, reviewing and releasing.

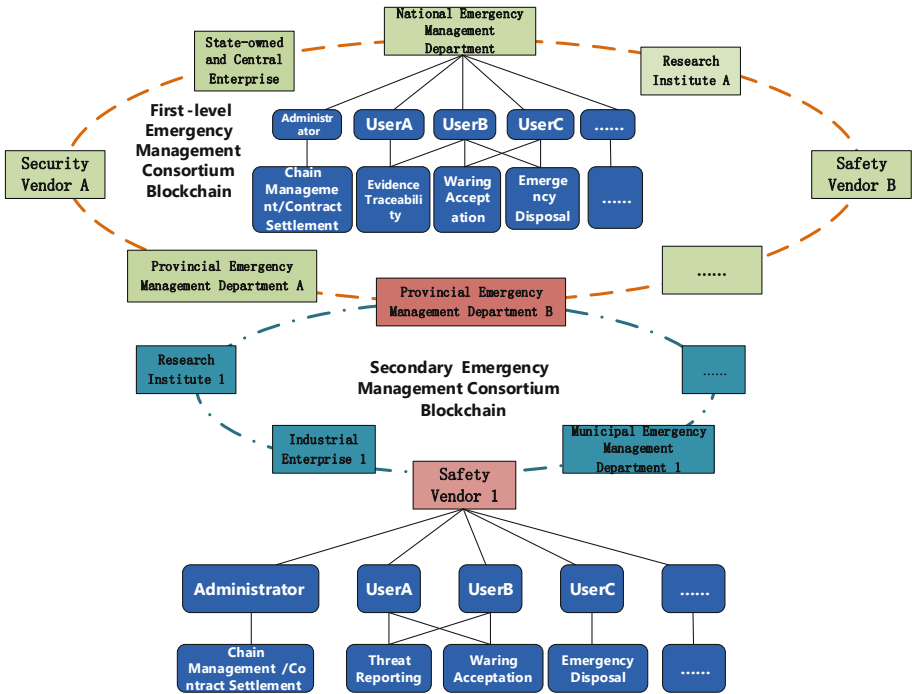


Fig. 5. Organizational structure of ISEMS based on consortium blockchain

#### 4.4 Critical Process

**Threat Reporting.** The threat reporting process is generally handled by members of the secondary consortium blockchain such as local security vendors, research institutions and industrial enterprises. As shown in Fig. 6, after the consortium members find the vulnerability, Threat reporting subsystem call the blockchain smart contract through the risk reporting API to write the risk data to blockchain ledger. At the same time, according to the agreement in the endorsement strategy of the intelligent contract, they first submit it to the default endorsement node, i.e. the local emergency management department for review. After the review is passed, the risk information will be written into the ledger and synchronized to the members of the secondary consortium blockchain. In addition, the local emergency management department will submit the risk information to the first-level consortium blockchain and synchronously submit to the central emergency management department and other local emergency management departments for information sharing, so as to complete the reporting of emergency information under abnormal conditions. Compared with the traditional risk threat reporting process, the

blockchain-based reporting, using the tamper-proof capability of the blockchain, can spread and synchronize to the peer timely.

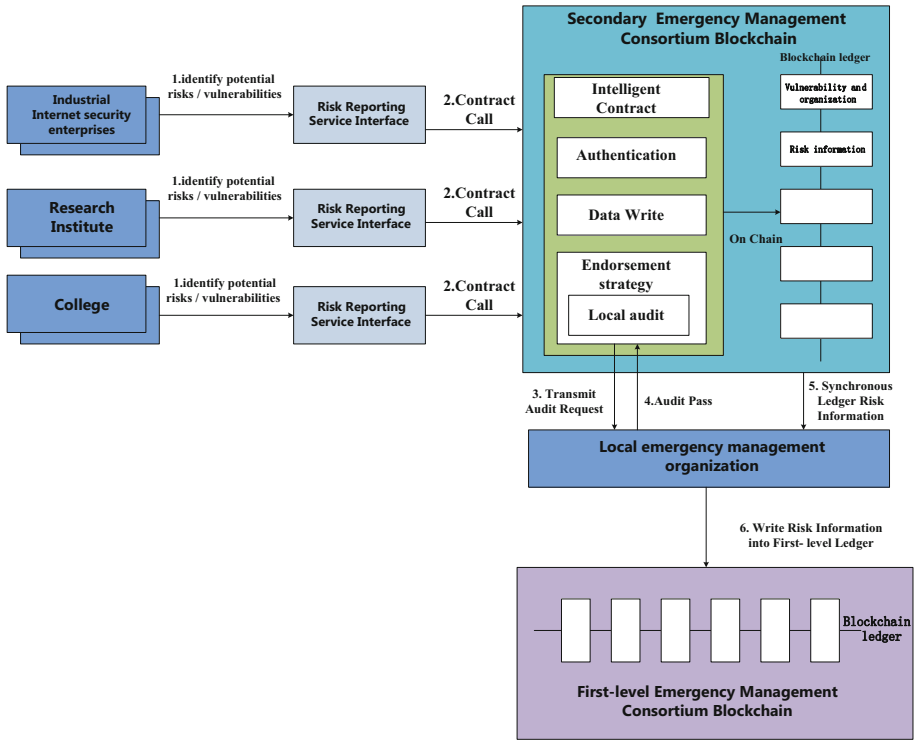


Fig. 6. Threat reporting process based on blockchain

**Risk Analysis.** The traditional risk analysis requires the central emergency management department to collect related risk data and organize relevant experts to carry out risk analysis and prediction. However, the risk analysis source data is too scattered to mobilize these resources to carry out analysis in time. The distributed blockchain-based risk analysis can realize risk vulnerability analysis and the training of distributed shared risk model locally. The online incremental learning of monitoring data is realized by capturing the data characteristics of each participant. Finally, each node can synchronize the updated risk model parameters or gradients to improve the risk prediction accuracy, as shown in Fig. 7.

**Warning Release.** When industrial internet security emergency event occurs, the early warning release and sharing system can quickly release vulnerabilities, notifications and policies to local competent departments or enterprises at all levels. In order to share emergency event knowledge base to specific members, and ensure members' identity trusted, the consortium blockchain firstly implements the identity authentication of members. Based on the CA digital certificate mechanism, it realizes the identification and authority

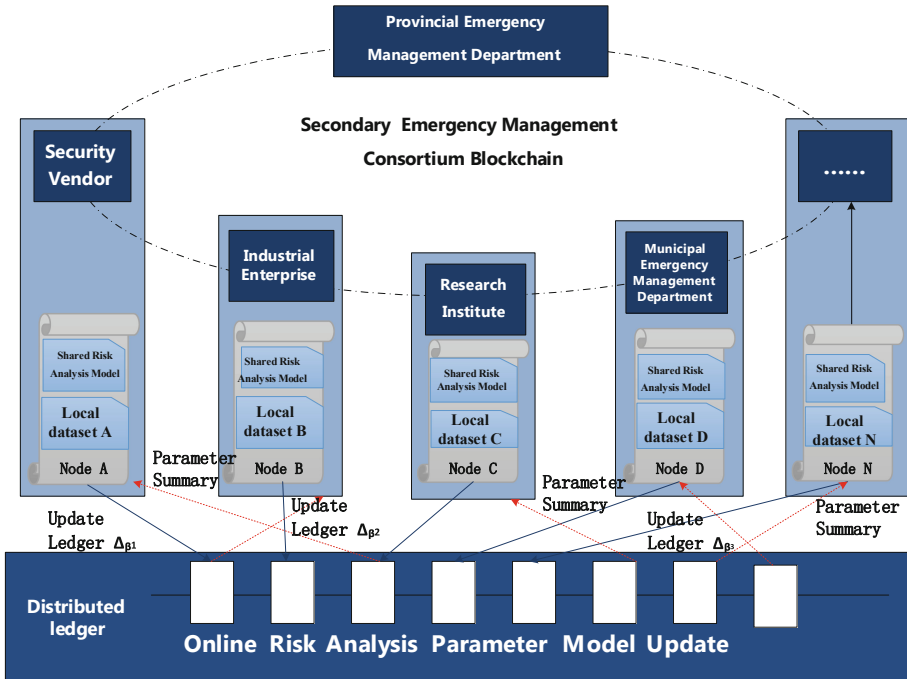


Fig. 7. Risk analysis process based on blockchain

control for members, so that early warning can be managed and controlled. Secondly, based on the consortium blockchain, the parallel uplink authority management supporting the complex business among multiple organizations is implemented, as shown in Fig. 8. By building different sub chains in the consortium blockchain and assigning different roles to its members, the authority control of the sub chain management and the ledger data reading and writing is carried out, so that the early warning information can be updated in time and synchronized to the relevant organizations, the scope of the early warning release is controlled, and the access of the non-authorized organizations to the relevant information is prevented.

**Emergency Response.** Emergency response mainly includes evidence collection, tracing and coordination. traditional methods take a lot of valuable time to find the responsible person and technical support. Meanwhile, it is unable to quickly locate whether the support has good technical reserves in this risk field. In order to mobilize the enthusiasm of various organizations in ISEM and maintain the normal operation of the emergency management blockchain platform, a competition incentive mechanism is introduced to reward enterprises that can timely report vulnerabilities and analysis results. Through the scoring mechanism in blockchain, Management department can independently select security vendors or research institutions with higher score to support offline emergency response, and timely assist industrial enterprises in upgrading the system and vulnerability database. Detailed incentive model is referred in Fig. 9.

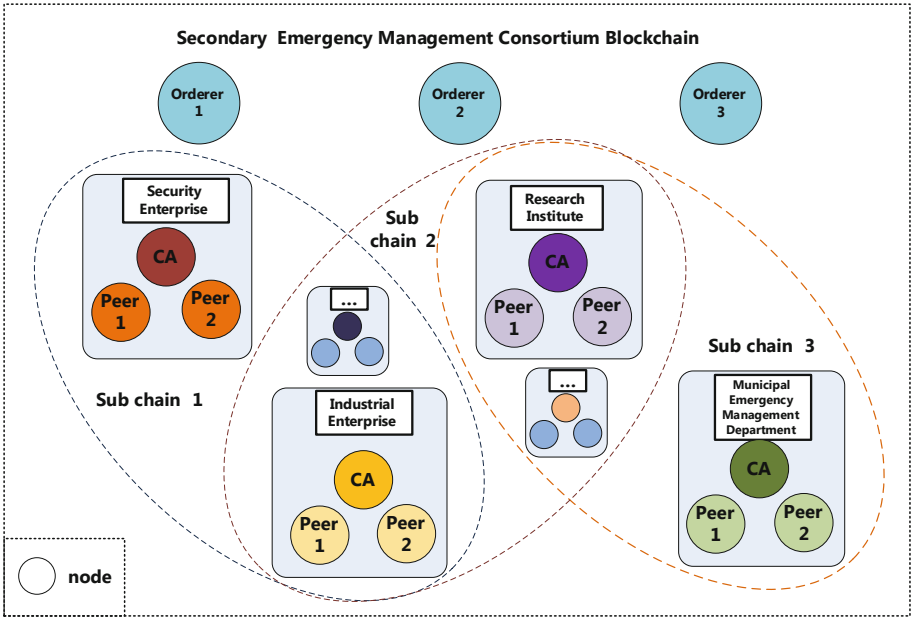


Fig. 8. Multi-organization sub chain division structure

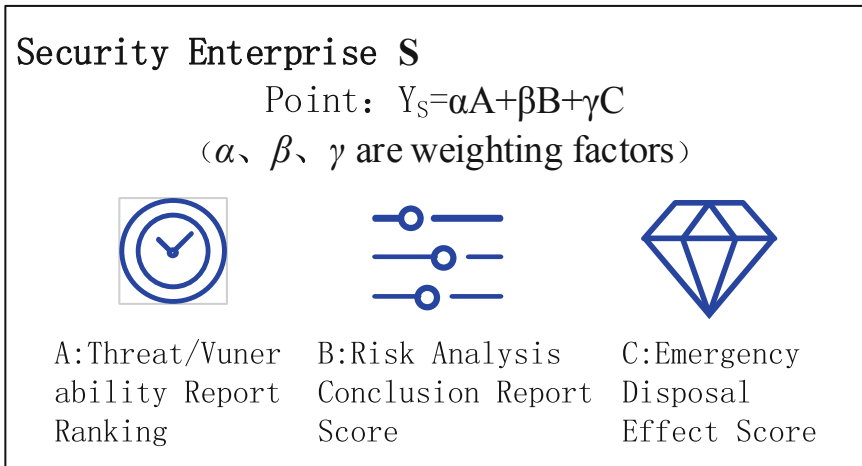


Fig. 9. User bonus point model

First of all, the report time of risk and vulnerability can be recorded on the blockchain ledger. Local emergency management organizations utilize the untampered and untraceable characteristics of blockchain to accurately and fast trace and rate the enterprise who report risk. Secondly, each organization can adjust the accuracy of the risk prediction model according to the local risk source data, improve the analysis model with

adjusted accuracy. The institutions report the trained risk prediction model in time will be rewarded with points. Third, when the emergency management organization assigns the emergency assistance tasks, enterprises with high points and outstanding technical advantages will have the priority. Industrial enterprises can also score on the chain for the effect of support organizations' disposal, and the score results will be distributed to support enterprises in the form of points. The points will ultimately affect the industry influence of enterprises, provide basis for national projects, awards and honor declaration, and form a benign incentive for enterprises and research institutions to actively report, analyze and deal with safety risks.

## 5 Summary and Prospect

This paper analyzes the situation and challenges of ISEMS, including organizational structure and technical status. Meanwhile, the principle and situation of blockchain and the challenge of building consortium-blockchain-based ISEMS are briefly introduced. Besides, this paper describes the system architecture and base model of the ISEIM based on the blockchain, and describes the key blockchain-based implementation processes, including threat report, risk analysis, warning release and emergency response. In the future, we can further study the balance between the realization of enterprise data privacy and the enhancement of data's social utility based on blockchain technology, so that it could expand the upper application and play effectiveness in the fields of data classification, classification and sharing.

## References

1. China blockchain technology and Industry Development Forum. White paper on China blockchain technology and application development. Department of information technology and software services, Ministry of industry and information technology (2016)
2. Zhang, Z., Sun, B., Li, B.: The US cybersecurity emergency management system and its enlightenment. *Intell. J.* (3), 94–98 (2018)
3. Liu, F.: Cybersecurity situation awareness and emergency response platform solutions. *Inf. Technol. Stand.* **405**(09), 18–20 (2018)
4. Zhao, X., Wen, J.: Research on provincial cybersecurity emergency management platform based on security access design. *Laboratory Research and Exploration*, vol. 37, no. 268 (06), pp. 300–303 (2018)
5. Li, R., Jia, R.: Research on cybersecurity emergency response system. *Network Security Technology and Application*, p. 2 (2019)
6. Zhang, X., Xiao, Y.: Overview of the construction of cyberspace situational awareness, early warning and protection system in the United States and Its Enlightenment to China. *Confidential Science and Technology*, no. 67(04), pp. 22–28 (2016)
7. Wu, H.: Research on the information sharing mechanism of Chinese government in the era of big data (2017)
8. Kang, K.: Analysis of isolated information island in the field of e-government (2016)
9. Abhishek, G., Alagan, A., Glauccio, C.: Prevailing and emerging cyber threats and security practices in IoT-Enabled smart grids: a survey. *J. Netw. Comput. Appl.* **132**, 118–148 (2019)
10. Shi, Y.: Research on security defense technology of IT/OT integration in industrial internet environment. *Inf. Technol. Netw. Secur.* **38**(7), 1–5 (2019)



11. Zhou, P., Tang, X., Li, B.: Research Report on China's blockchain technology and application development. China blockchain technology and Industry Development Forum (2018)
12. Yang, L., Zhang, C., Wang, F.: Overview of blockchain technology research and application. *Contemp. Econ.* **4**, 126–128 (2018)
13. Zhang, S., Yang, Y.: Block chain technology foundation and application. *Inf. Secur. Res.* **4** 33(06), 89–94 (2018)
14. Herrera-Joancomartí, J.: Research and challenges on bitcoin anonymity. In: Garcia-Alfaro, J., et al. (eds.) *DPM/QASA/SETOP -2014. LNCS*, vol. 8872, pp. 3–16. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-17016-9\\_1](https://doi.org/10.1007/978-3-319-17016-9_1)
15. Saxena, A., Misra, J., Dhar, A.: Increasing Anonymity in Bitcoin (2014)
16. Kishigami, J., Fujimura, S., Watanabe, H., et al.: The blockchain-based digital content distribution system. In: 2015 IEEE Fifth International Conference on Big Data and Cloud Computing (BDCloud). IEEE (2015)
17. Paul, G., Sarkar, P., Mukherjee, S.: Towards a more democratic mining in bitcoins. In: International Conference on Information Systems Security (2014)
18. Mougayar, W.: Why Fragmentation Threatens the Promise of Blockchain Identity (2016). <https://www.coindesk.com/fragment-blockchain-identity-market>
19. Sana, M., Ahmad, K., Zanaab, S.: Securing IoTs in distributed blockchain: analysis, requirements and open issues. *Future Gener. Comput. Syst.* **100**, 325–343 (2019)
20. Bhabendu, K.M., Debasish, J., Soumyashree, S.P.: Blockchain technology: a survey on applications and security privacy challenges. *Internet Things* **8**, 100107 (2019)
21. Chen, J., Lv, Z., Song, H.: Design of personnel big data management system based on blockchain. *Future Gener. Comput. Syst.* **101**, 1122–1129 (2019)

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





# Research Status and Prospect of Blockchain Technology in Agriculture Field

Dawei Xu<sup>1,2(✉)</sup>, Weiqi Wang<sup>2</sup>, Liehuang Zhu<sup>1</sup>, and Ruiguang Li<sup>2,3</sup>

<sup>1</sup> School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China  
3220195131@bit.edu.cn

<sup>2</sup> College of Cybersecurity, Changchun University, Changchun, China

<sup>3</sup> National Computer Network Emergency Response Technical Team/Coordination Center of China, Beijing, China

**Abstract.** Agriculture 4.0 is the era of integrating intelligent technologies in agriculture. Problems such as low informatization, food safety, high management cost and imbalance between supply and demand in agriculture have greatly hindered the development of agriculture. The various properties of blockchain technology can make up for the lack of agricultural mechanism, and the fusion of the two is a hot issue in the application of blockchain. Blockchain technology has already had some application cases in the field of agriculture. Based on the research status of Chinese and foreign scholars in this field, this paper firstly introduces the basic overview of blockchain. Then, with agricultural supply chain and agricultural product traceability as the core, it describes the application of blockchain technology in the agricultural field, and further explores solutions to different application problems. Finally, combined with the practical application of “agriculture + blockchain”, relevant Suggestions are proposed to provide reference for cross-disciplinary research.

**Keywords:** Blockchain · Agriculture · Supply chain · Tracing

## 1 Introduction

Blockchain technology is considered as the key technology leading intelligent communication and information sharing, is also a hot research field in the current academic circle with the topics mainly focusing on technical basis, security analysis and scenario application.

Agriculture is one of the most important fields in the world. However, the development of agriculture is restricted by its weak foundation, high cost, low efficiency and difficult management. In recent years, to solve the problems existing in the field of agriculture, the research of applying blockchain technology to it has been increasing gradually. In blockchain articles, there are many reviews on the nature of technology, and a few reviews the typical application [1]. This paper first introduces the basic overview of blockchain; Then, with agricultural supply chain and agricultural product traceability as the core, describes the research status and development of blockchain technology in

the agricultural field, and further explores solutions to different application problems; Finally, combined with the practical application of “agriculture + block chain” to put forward relevant suggestions.

## 2 Overview of Blockchain

### 2.1 Core of Blockchain Technology in Agriculture

The data layer uses hash pointer and follows a certain time sequence to arrange each block consisting of head and body into a chain structure. Each new block needs to pass the consensus verification of 51% nodes on the chain and load the current data status into the state library. Encryption technology is required to provide privacy protection in the blockchain, and the most representative ones are public key cryptography and hashing algorithm.

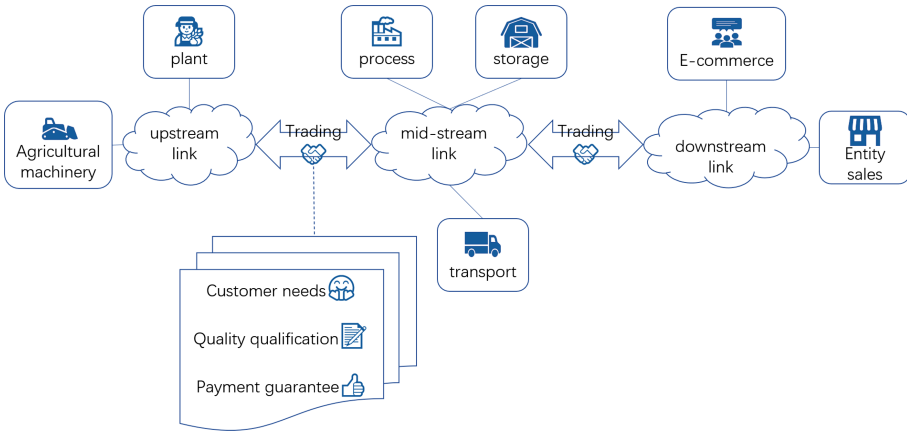
Consensus mechanism is the key to determining which miners have the right to generate new blocks. Since it does not rely on the centralized institutions, consensus algorithm is needed as the basis for judgment. This algorithm includes: POW, POS, DPOS, POA, and PBFT. Smart contract of the control layer is the key to the blockchain. It executes the corresponding code program through the computer. It's a script that can self-guarantee and execute a contract without the participation of a third party and geographical restrictions.

### 2.2 Application of Blockchain in Agriculture

The blockchain technology applying in the agriculture field, which point-to-point implementation of transparent transactions are conducive to the collection of agricultural data; Distributed ledger of blockchain enables all participating nodes to share and store unclassified data synchronously, which solves the problem that information cannot be received in real-time among multiple processes; Under the framework of blockchain, the behaviors of participants in the agricultural chain are encouraged and restricted to increase the authenticity of data. In this paper, the research status and development of blockchain in the field of agricultural supply chain and the traceability of agricultural products are mainly described, solutions to different application problems are further explored and relevant suggestions are proposed.

## 3 Current Situation of Agricultural Supply Chain

Current production, supply and marketing in agricultural model is a linear from the producers to the retailers, shown in Fig. 1. Among them, the upstream contains crops and the use of agricultural machinery supervision, middle includes packaging agricultural products processing, cargo transportation, warehousing and logistics, and the downstream contains e-commerce sales and entity. Through trade connected all links, during the transaction should be familiar with the actual needs of customers, passed by quality qualification, provided appropriate payment guarantee. Still many problems in the agricultural supply chain in reality, which fail to meet the conditions that each link should have. Can used to integrate the upper, middle, lower and trading links of the agricultural supply chain into the blockchain, so as to improve the trouble.



**Fig. 1.** Agricultural supply chain model

### 3.1 Upstream Link

The upstream link, also called the production link, which includes all the agricultural activities carried out in the farm.

#### 3.1.1 Agricultural Greenhouse Technology

Agricultural greenhouse technology is a technology to improve crop yield by controlling environmental factors and is also the key to agricultural production. Patil et al. [2] provide a security framework combining private chain with IoT devices, to provide a secure monitoring communication platform for intelligent greenhouse cultivation to ensure safe communication of devices. Authentication and access control at the physical layer using the immutability of the maintenance ledger; At the communication layer, distributed structures and consensus mechanisms are used to ensure transmission security; There are timestamps and signatures at the data layer to maintain the data authenticity; The interface layer takes the anonymity of the blockchain to solve the changing of attacker’s ID.

#### 3.1.2 Breeding Information Management Technology

After data collection, store massive data efficiently and safely is more difficult, which puts forward higher requirements for breeding information management technology, and it is necessary to ensure its safety.

Zhang et al. [3] using improved lightweight block chain technology update GSBCP platform software, it covers the whole breeding process from breeding materials to field layout and data collection and analysis. SACBDIBT’s storage structure was established, data were divided and stored in multiple databases according to breeding process and location, and summary information was stored in blockchain. When accessing breeding data, system reasonably allocates computing resources and storage space, provides the most idles server as the main server, then encrypts and saves the information in the block chain to improve the security of data.

### 3.1.3 Risk Control Technology

Technology cannot withstand the damage of various natural disasters to crops. Due to the weak risk resistance of agriculture itself, the existing risk control technology is not perfect, once the damage is accompanied by huge losses. When the database shows that rainfall in farmer's area is below the insured threshold, Kim et al. [4] keeps costs low by using smart contracts to automatically process claims.

## 3.2 Mid-Stream Link

The middle link is related to many users and transactions. The required information includes correct information of agricultural products, processing history of each node to control the production process, etc., which puts forward higher requirements for logistics management technology and quality detection technology.

### 3.2.1 Logistics Management Technology

Logistics help enterprises to realize the whole business effectiveness of the supply chain, is important to the middle link. As the existing logistics management cannot meet the flexibility and efficiency required by the enterprise supply chain, Private Chain or consortium chain can be used to protect personal data, and the validity of data can be maintained by consensus mechanism and intelligent contract. Li Xiaoping et al. [5] took consortium chain as the underlying technology to build LSSC platform, provided consistent interface program for unified format definition of operational data, to realize the assumption of intelligent monitoring and real-time information sharing of agricultural products.

### 3.2.2 Quality Inspection Technology

Quality inspection technology is an important measurement technology in the process of agricultural goods transportation. The transparency of existing technologies fails to satisfy people's needs. The authentication function and non-tampering of blockchain are used to ensure the authenticity of information provided.

Lucena et al. [6] using Hyperledger tracking the source of the delivery batch and establish the special communication channel, the contract has a separate Node. Js process as open application program interface of the business, by Passport. Js configuration of open source authentication middleware access security protection. The results demonstrate that the blockchain technology meets the potential demand for certification, and is expected to increase the value of goods, as real-time sharing reduces disputes and information asymmetry between supply chains.

## 3.3 Downstream Link

The downstream link, also called the sales link, is the process of currency value exchange between agricultural products and users. The e-commerce involved in this link is faced with the problem of information security and sharing, and the efficient flow of resources is difficult. Therefore, blockchain technology is utilized to adjust.

Huang Wei et al. [7] electricity in rural areas and the value chain of logistics enterprise to carry on the conformity and reconstruction, through the blockchain, dynamic laser anti-counterfeit labels and dynamic image recognition technology to build information tracking and anti-fake model, with the smart contracts to improve the level of automation and prove the authenticity of the product itself and its flow, build public distributed mutualism mode, prompt information safe, efficient, reliable delivery, receipt and payment to ensure electrical business, logistics and customer benefit maximization.

### 3.4 Trading Link

As a bridge between production and sales, the trading link needs to investigate market information, predict customer demand and changes in advance, and communicate with producers in real time, to achieve the balance between supply and demand.

HuoHong et al. [8] with integrated supply chain perspective, with the public, private, consortium chain, a three-state cycle of product, information and economic benefits is formed in the system. Consensus and through the permissions assigned to protect privacy, simultaneously clear regulatory subject realized the agricultural product quality traceability, without intermediary participation, consumer can be directly to produce feedback, taste, quality requirements. Optimize the interest demands of different subjects, ensure the quality, and coordinate the supervision cost and benefit distribution.

## 4 Current Status of Agricultural Products Traceability

The traceability system of agricultural products is mainly responsible for tracking the quality and safety of agricultural products from production to consumption. Its framework is shown in Fig. 2. Agricultural products go through multiple transfers before consumption. To accurately identify the quality of agricultural products, effective detection and prevention of product safety problems and accountability, establishing a reliable traceability system on blockchain is essential.

### 4.1 Solves the Problem of Accountability for Agricultural Products

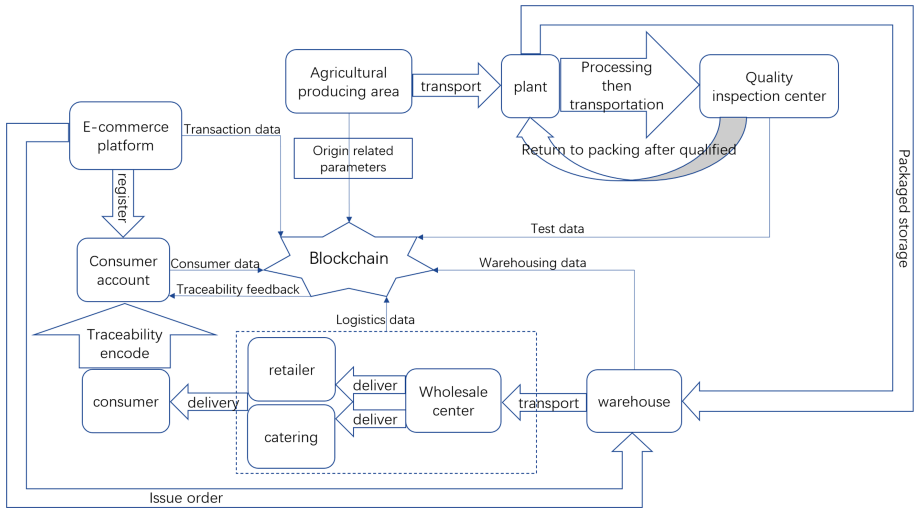
Due to multiple transfers, people cannot determine the accuracy of information, and the retail industry and enterprises become the most responsible persons. To solve such problems, scholars have been exploring with the help of blockchain and IoT.

Liu Zongmei [9] used consortium chain to build a food traceability platform of “blockchain +RFID”, formulated a list of violations to remove malicious nodes in time, and the multi-port could efficiently query the source and destination of products with legal benefits and accurately locate fault points.

### 4.2 Solves the Problem of Traceability Trust for Agricultural Products

#### 4.2.1 Reputation Evaluation Mechanism

Reputation system is an effective method to solve trust problems. Malik et al. [10] proposed the trust management framework of supply chain application based on blockchain, assigned credit scores to participants and products, conducted qualitative security analysis on threats in the reputation system, and improved the credibility of products.



**Fig. 2.** Agricultural product traceability framework

### 4.2.2 Reverse Technical Constraints

When ethics cannot constrain the operational norms in agriculture, laws and technologies can be used for reverse restriction. The smart contract of blockchain forces users to fulfill the contract.

Pearson et al. [11] applied block chain to provide the encryption security of transactions and related metadata of the entire supply chain, including origin, contract, process steps, environmental changes, microbial records, etc., and the records are unchangeable, so as to meet the DLT requirements of data standardization in the field of agricultural products, and the whole product supply chain can be securely linked, to reverse the constraints of each link behavior.

## 5 Summary and Prospect

By discussing the current research papers, journals and projects related to blockchain in the agricultural field, it is found that the current research mainly focuses on supply chain, agricultural product traceability. In the research, most of them discuss the technical restrictions and loopholes related to blockchain, and a few of them restrict it morally and legally in combination with the regulatory system.

In the future, blockchain will have more applications in the agricultural field. From a technical perspective, the throughput of the system can be improved by optimizing the consensus algorithm, thus accelerating the upload process. Periodically clean expired information on the chain to reduce the accumulation of data volume; Strengthen data source management to improve the lack of credibility in the chain link; Update the encryption mechanism in time to increase the security of information; Collect data in the same way, unify data standards and improve data quality; Improve the privacy protection mechanism and increase the security of the member information on the chain.

From the perspective of management, it can constantly improve the management mechanism in the field of agriculture and provide security guarantee. Increase the extension of rural finance and rural insurance to provide farmers with operational funds and security; improve the knowledge level of farmers, so that farmers can correctly understand the blockchain; reduce traceability costs, promote more product labeling information, etc.

Blockchain and agricultural applications need to be further integrated, and technology and management combined to improve the system. For example, smart contracts are combined with laws to limit the scope of contract execution, and contracts are used to confirm whether personnel follow the system, thus forming a two-way and mutually beneficial situation. The combination of blockchain and agriculture in the future remains to be explored by researchers.

## References

1. Ren, M., Tang, H.B., You, W.: Survey of applications based on blockchain in government department. *Comput. Sci.* **45**(02), 1–7 (2018). (in Chinese)
2. Patil, A.S., Tama, B.A., Park, Y., Rhee, K.-H.: A framework for blockchain based secure smart green house farming. In: Park, J.J., Loia, V., Yi, G., Sung, Y. (eds.) *CUTE/CSA-2017. LNEE*, vol. 474, pp. 1162–1167. Springer, Singapore (2018). [https://doi.org/10.1007/978-981-10-7605-3\\_185](https://doi.org/10.1007/978-981-10-7605-3_185)
3. Zhang, Q., Han, Y.Y., Su, Z.B., Fang, J.L., Liu, Z.Q., Wang, K.Y.: A storage architecture for high-throughput crop breeding data based on improved blockchain technology. *Comput. Electron. Agric.* **173**(6) (2020). <https://doi.org/10.1016/j.compag.2020.105395>
4. Kim, H.M., Laskowski, M.: *Agriculture on the blockchain: sustainable solutions for food, farmers, and financing*. Social Science Electronic Publishing (2017)
5. Li, X.P., Wang, Y.Y.: Construction of logistics service supply chain information platform based on blockchain technology. *Logistics Technol.* **38**(05), 101–106 (2019). (in Chinese)
6. Lucena, P., Binotto, A.P.D., Momo, F.S., Kim, H.M.: A case study for grain quality assurance tracking based on a blockchain business network. In: *Symposium on Foundations and Applications of Blockchain (FAB 2018)* (2018)
7. Huang, W., Chang, R.R., Chang, R.: The symbiotic development of rural e-commerce and logistics from the perspective of block chain tech. *J. Commercial Econ.* **6**, 118–121 (2019). (in Chinese)
8. Huo, H., Zhan, S.: Construction of a whole-process supervision system for the quality and safety of agrifood from the perspective of integrated supply chain. *Forum Sci. Technol. China* **8**, 105–113 (2019). (in chinese)
9. Liu, Z.M.: Research on “blockchain + RFID” enabling food traceability platform. *Food Mach.* **6**, 1–8 (2020). (in Chinese)
10. Malik, S., Dedeoglu, V., Kanhere, S.S., Jurdak, R.: TrustChain: trust management in blockchain and IoT supported supply chains. In: *IEEE International Conference on Blockchain. Semantic Scholar, Atlanta* (2019). <https://doi.org/10.1109/blockchain.2019.00032>
11. Pearson, S., May, D., Leontidis, G., Swainson, M., Brewer, S., Bidaut, L., et al.: Are distributed ledger technologies the panacea for food traceability? *Glob. Food Secur.* **20**, 145–149 (2019). <https://doi.org/10.1016/j.gfs.2019.02.002>



**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# **Denial-of-Service Attacks**



# Practical DDoS Attack Group Discovery and Tracking with Complex Graph-Based Network

Yu Rao<sup>1</sup>, Weixin Liu<sup>2</sup>(✉), Tian Zhu<sup>1</sup>, Hanbin Yan<sup>1</sup>, Hao Zhou<sup>1</sup>, and Jinghua Bai<sup>2</sup>

<sup>1</sup> National Computer Network Emergency Response Technical Team/Coordination Center of China (CNCERT/CC), Beijing 100029, China

<sup>2</sup> NSFOCUS Tianshu Lab of NSFOCUS Information Tech Co., Ltd., Beijing 100089, China  
liuweixin@nsfocus.com

**Abstract.** In recent years, a large number of users continuously suffer from DDoS attacks. DDoS attack volume is on the rise and the scale of botnets is also getting larger. Many security organizations began to use data-driven approaches to investigate gangs and groups beneath DDoS attack behaviors, trying to unveil the facts and intentions of DDoS gangs. In this paper, DDoSAGD - a DDoS Attack Group Discovery framework is proposed to help gang recognition and situation awareness. A heterogeneous graph is constructed from botnet control message and relative threat intelligence data, and a meta path-based similarity measurement is set up to calculate relevance between C2 servers. Then two graph mining measures are combined to build up our hierarchical attack group discovery workflow, which can output attack groups with both behavior-based similarity and evidence-based relevance. Finally, the experimental results demonstrate that the designed models are promising in terms of recognition of attack groups, and evolution process of different attack groups is also illustrated.

**Keywords:** Botnet · Graph mining · DDoS · Attack group discovery · Community detection

## 1 Introduction

Among many network attack methods, DDoS (Distributed Denial of Service) has always been regarded as the effective weapon of hacker attacks due to its low attack threshold and high damage. Compared with other attack methods, the technical requirements and cost in launching an attack of DDoS are very low. In the past three years, the situation of DDoS attacks is still grim. In late February 2018, the world-renowned open source project hosting site GitHub suffered a DDoS attack with a peak value of 1.35 Tbps, which has reached a record high, marking the official entry of the DDoS attacks into Tb level. Super-large DDoS attacks have been increasing steadily year by year after a sharp increase in 2018. The ultra-large-scale attacks above 300 Gbps in 2019 increased by more than 200 times as compared with 2018 [1]. Botnets and Internet of Things are hot words for DDoS attacks in recent years. The active botnet family is further concentrated

on the IoT platform, which mainly includes Gafgyt and Mirai. DDoS attacks have also become one of the important methods for attackers to use IoT devices.

At the same time, with the rise of big data technology and threat intelligence, many security agencies began to use data-driven methods to mine the gang behaviors behind DDoS attacks. NSFOCUS has found 60 DDoS gangs in 2019, and up to 15 gangs have attack resources of greater than 1000, and the largest attack gang contains 88,000. The highest proportion of IoT devices in a single gang of DDoS gangs reaches 31% [1]. An in-depth analysis on gang behaviors in network security data is also made in *2018 Active DDoS Attack Gang Analysis Report* [2] and *2018 Website Attack Situation and "Attack Gang" Mining Analysis Report* [3] released in 2018 by Cncert. The gang analysis behind DDoS can help regulators and security researchers understand the attack trends and the overall situation.

In this article, DDoS gangs are analyzed based on control instruction propagation logs and threat intelligence data of a botnet. Articles with similar goal of this article include Zhu Tian's group analysis of DDoS based on network attack accompanying behavior [5], and Application of community algorithm based on malicious code propagation log by Wang Qinqin [6], and IP Chain-Gang analysis by NSFOCUS based on DDoS logs [4, 7]. Existing DDOS gang analysis mostly focuses on the behaviors of attacking resources, searching for communities in big data. Gang analysis based on the behaviors of attacking resources has two disadvantages. The first is the detection accuracy of the attack behavior data. DDoS is always accompanied by normal user behaviors with high traffic, while some of them are very hard to be distinguished. The second is the problem of unity of data. The gang analysis based on the behavior of attacking resources usually originates from large-scale behavior similarities and community structure of attacking resources, lacking the correlation of small scale but strong evidences. Therefore, for the purpose of uncovering attack gangs, it is necessary to not only perform clustering at the behavior level, but also combine the control messages of the attack resources and related samples/domain names.

This article presents a DDoS attack group discovery framework based on complex graphs. Entities and relations are extracted from botnet control messages and threat intelligence data of a botnet, and the constructed underlying heterogeneous graph contains a DDoS behavior relationship and an intelligence association relationship. Then the control end is taken as the key entity, to set a series of meta paths, establish the similarity relationship between the control ends, and form a homomorphic graph with the control end as the node and the similarity as the relationship. Finally, the DDoS gang is calculated through the hierarchical community algorithm.

The main contributions of this article are as follows:

- This paper proposed a heterogeneous graph construction method based on control instruction logs and threat intelligence data of a botnet, fused behavioral relations and intelligence association relations, and constructed the underlying graph.
- This article proposed a meta path-based similarity graph construction method with the control end as the core. At the same time, the hierarchical similarity interval can ensure that the subsequent group discovery can distinguish the scale similarity from the evidence/intelligence similarity.

- This article proposed a hierarchical DDoS attack gang discovery method, and in combination with the advantage of Louvain algorithm for mining community structure and the advantage of Connected Component for mining strong evidence relationship, this article obtained a more complete gang structure, and retained the results of hierarchical community analysis to assist in security operations.

The structure of this article is as follows: Sect. 1 is the introduction, introducing related work and main research ideas; Sect. 2 is a technical route and data background, introducing workflow and data overview; Sect. 3 is a detailed elaboration of the DDoS attack group discovery framework; and Sect. 4 is the experimental results, introducing research results and cases.

## 2 Methodology and Background

The dataset of this paper is the botnet control message logs from January 2018 to December 2019. The botnet control message logs contain a C2 (Command & Control) server, a C2 family, a bot list, attack target information and attack time. Botnet refers to the use of one or more propagation methods to infect a large number of hosts with bot virus, thus forming a one-to-many control network between the controller and infected hosts. Botnets rely on large-scale DDoS attacks or bitcoin mining for profit. This paper only focuses on DDoS attacks in botnets.

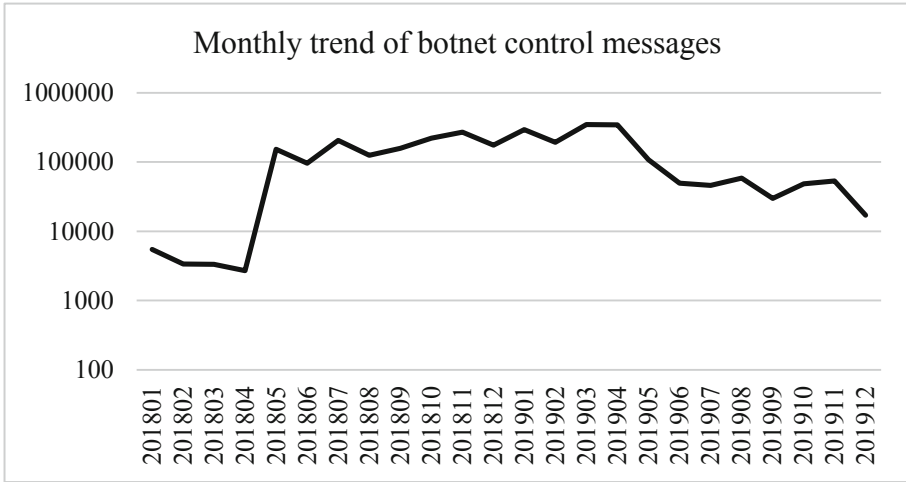
The preparation of the dataset consists of three modules, including data import, threat intelligence correlation and data storage/calculation.

**Data Import:** The dataset used in this paper originates from the evaluation dataset provided by the National Computer Network Emergency Technology Processing and Coordination Center (CNCERT/CC).

**Threat Intelligence Correlation:** In this step, network entities are extracted from botnet control messages, and related intelligence information, including related domain names and related samples, is queried from various external intelligence sources and data sources.

**Data Storage/Calculation:** Hadoop is used to store the large-scale sample data, and Spark is used for calculation.

**Data Overview:** The test dataset contains 3005888 botnet controlling messages of 5225 C2 servers. The monthly trend of botnet controlling messages is shown in Fig. 1 and the distribution of active months among C2 servers is shown in Table 1. C2 servers cover 34 botnet families. The top three botnet families are DDoS.Linux.Gafgyt, DDoS.Linux.BillGates and DDoS.Linux.Xorddos. The active period of C2 servers can reach a maximum of 20 months and a minimum of less than 1 month, with an average active time of 1.4 months.



**Fig. 1.** Monthly trend of botnet control messages

**Table 1.** Active period (in month) distribution of C2 servers

Active months	C2 count
1	4093
2	740
3	103
4	49
6	24

### 3 DDoSAGD – DDoS Attack Group Discovery Framework

In this paper, the DDoS Attack Group Discovery (DDoSAGD) Framework is developed to unveil DDoS attack groups with behavior-based similarity and evidence-based relevance from DDoS attack logs and threat intelligence data. The DDoSAGD framework provides principles and practices for attack group discovery, including three phases: heterogeneous graph data modeling, relevance measurement, and community detection.

#### 3.1 Basic Concepts and Fundamental Graph Initialization

In this section, the construction details of the graph model are introduced and the relevant definitions are clarified.

**Definition 1 DDoS Attack Group:** The core of a DDoS attack group is C2 servers. Bots and other attack resources are related to C2 servers. The C2 server set is the most critical part of a DDoS attack group.

Graphs are used to represent the interactions among different entities in the real world. In this paper, we regard the network entities, such as C2 servers, victim IP addresses, bots in DDoS attacks as nodes in the graph. Those nodes are extracted from the behavior logs or related intelligence. We assign each node/entity with a globally unique identifier (ID) and attach attributes to them. Moreover, we divide the entities into the following two categories.

**Definition 2 Critical Entity:** Critical entities are core members in an attack scenario. Specifically, the critical entities in the DDoS attack scenario are C2 servers.

**Definition 3 Associated Entity:** Associated entities are related to critical entities. In the DDoS scenario, C2 servers are critical entities while bots, victim targets and related domains are all associated entities.

Table 2 lists entities involved in the DDoS scenario. To be specific, the ‘EVENT’ entities are extracted according to attack targets and time characteristics. Within an empirical attack cycle, which is usually no longer than 24 h, an ‘event’ refers to a DDoS attack launched by a bunch of Internet resources aiming at a certain victim. It is noted that, if that victim is attacked by the same cluster of resources after more than 24 h from the last attack, it will be regarded as another event.

**Table 2.** Entities in DDoS attack scenario

DDoS entity	Entity type
C2	Critical Entity
BOT	Associated Entity
TARGET	Associated Entity
EVENT	Associated Entity
DOMAIN	Associated Entity
SAMPLE	Associated Entity
PDB	Associated Entity

We extract three different types of relations among these entities, namely, behavioral relations, associated relations and correlated relations.

**Definition 4 Behavioral Relation:** Behavioral relations are extracted from behavior logs or alerts and can represent the attacks or communications between entities.

**Definition 5 Associated Relation:** Associated relations are extracted from external intelligence or knowledge base and can represent the affiliation or usage relations between entities. Such relations are often related to knowledge, rather than behaviors.

The two relations above construct a heterogeneous graph in DDoS attack scenario. For further analysis on similarity, correlated relations are defined to calculate the similarity among entities of the same type.

**Definition 6 Correlated Relation:** Correlated relations depict the relevance of a pair of entities with the same type. Relevance measurement comes from comparative analysis on behavioral relations, association relations and attributes between a pair of entities with the same type.

### 3.2 Meta Path-Based Similarity Framework

The main task of DDoS attack group discovery is to cluster the critical entities based on correlated relations. Specifically, the correlated relation between two entities is calculated through meta path-based similarity in the heterogeneous graph. Table 3 lists the heterogeneous relations, including behavioral relations and associated relations in the heterogeneous graph model constructed for the DDoS attack scenario.

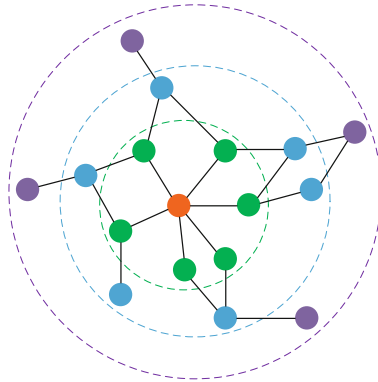
**Table 3.** Relations in DDoS attack scenario

d	Relation	Relation type
d1	C2-TARGET	Behavior/Behavioral Relation
d2	C2-BOT	Behavior/Behavioral Relation
d3	C2-EVENT	Behavior/Behavioral Relation
d4	C2-DOMAIN	Association/Associated Relation
d5	C2-SHA56	Association/Associated Relation
d6	SHA256-PDB	Association/Associated Relation
d7	SHA256-SHA256	Association/Associated Relation
d8	C2-MD5	Association/Associated Relation
d9	MD5-PDB	Association/Associated Relation

#### **Correlated Relations Based on Meta Path (C2-C2@SIM[Associated Entity]).**

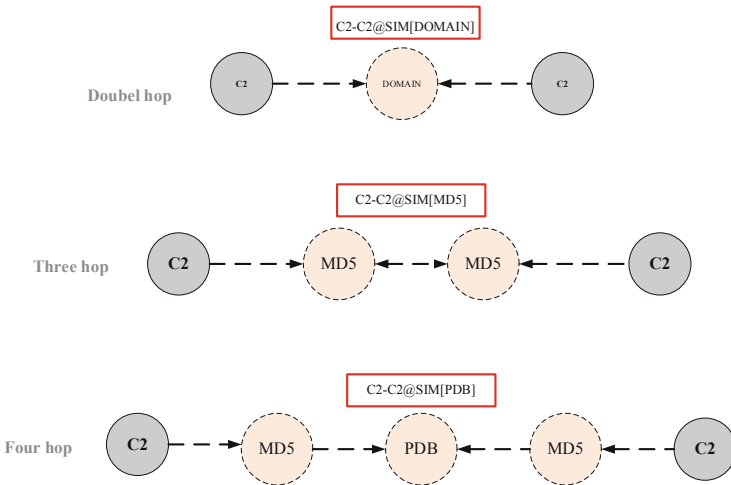
Considering the multi-hop similarity theory, a certain entity can have 1-hop neighbors, 2-hop neighbors and even n-hop neighbors in a graph. Figure 2 shows an example of k-hop neighbors. In this figure, the orange circle represents an entity, the green ones represents the 1-hop neighbors, the blue ones represent the 2-hop neighbors and the purple ones represent the 3-hop neighbors. This theory can be applied to the DDoS attack scenario to extract multidimensional correlated relations between C2 servers.





**Fig. 2.** Multi-hop in graph (Color figure online)

The correlated relations are illuminated in Fig. 3 and Table 4. For example, C2-C2@SIM[DOMAIN] represents the correlated relation between a C2 server and its 2-hop neighbors based on the associated domains while C2-C2@SIM[MD5] represents the correlated relation between a C2 server and its 3-hop neighbors, which is based on the similarity of the associated MD5 samples. Finally, C2-C2@SIM[PDB] represents the correlated relation between a C2 server and its 4-hop neighbors. This relation is based on two types of associated relations, that is, malware samples associated to C2 servers and the PDB paths associated to MD5 samples.



**Fig. 3.** Multi-hop in DDoS attack

**Table 4.** Relations of C2

Relation (C2-C2@SIM[Associated_Entity])
C2-C2@SIM[PDB]
C2-C2@SIM[SHA256]
C2-C2@SIM[SHA256&BDFF]
C2-C2@SIM[MD5]
C2-C2@SIM[TARGET]
C2-C2@SIM[EVENT]
C2-C2@SIM[BOT]
C2-C2@SIM[DOMAIN]

We determine whether two critical entities belong to the same attack group according to these nine correlated relations. Specifically, given two C2 servers C2\_1 and C2\_2 with a kind of associated entity A, A\_set1 and A\_set2 are subsets of A, which contain all the associated Class A entities of C2\_1 and C2\_2 respectively. As shown in Eq. 1, suppose that a correlated relation exists between C2\_1 and C2\_2 based on Class A entity if the number of Class A entities related to both C2 servers is greater than  $n$ , or the Jaccard similarity is greater than  $t$ . It is noted that C2-C2@SIM[A] is a Boolean variable, where the true value represents that the two C2 servers are relevant while the false value represents that they are irrelevant.

$$C2-C2@SIM[A] = bool(A_{set1} \cap A_{set2} > n) || bool(Jaccard(A_{set1}, A_{set2})) \quad (1)$$

$$Jaccard(A_{set1}, A_{set2}) = \frac{|A_{set1} \cap A_{set2}|}{|A_{set1} \cup A_{set2}|} \quad (2)$$

**Similarity of C2s(C2-C2@SIM).** We utilize the attention mechanism to aggregate the multi-dimensional correlated relations, since different types of relations are not equally important when the similarity between C2 servers is calculated. According to Eq. 3,  $\omega_e$  is weight of C2-C2@SIM[A<sub>e</sub>], and a homogeneous graph can be constructed for the following community detection and DDoS attack group discovery. Figure 4 shows the process of similarity construction.

$$C2-C2@SIM = \omega_0 C2-C2@SIM[A_0] + \omega_1 C2-C2@SIM[A_1] + \dots + \omega_q C2-C2@SIM[A_q] = \sum_{e=0}^q \omega_e C2-C2@SIM[A_e] \quad (3)$$

Considering that more relations may exist beyond those listed in Table 3 and Table 4, our following group discovery framework is designed to be extensible, so that users can add or remove relations to customize the system.

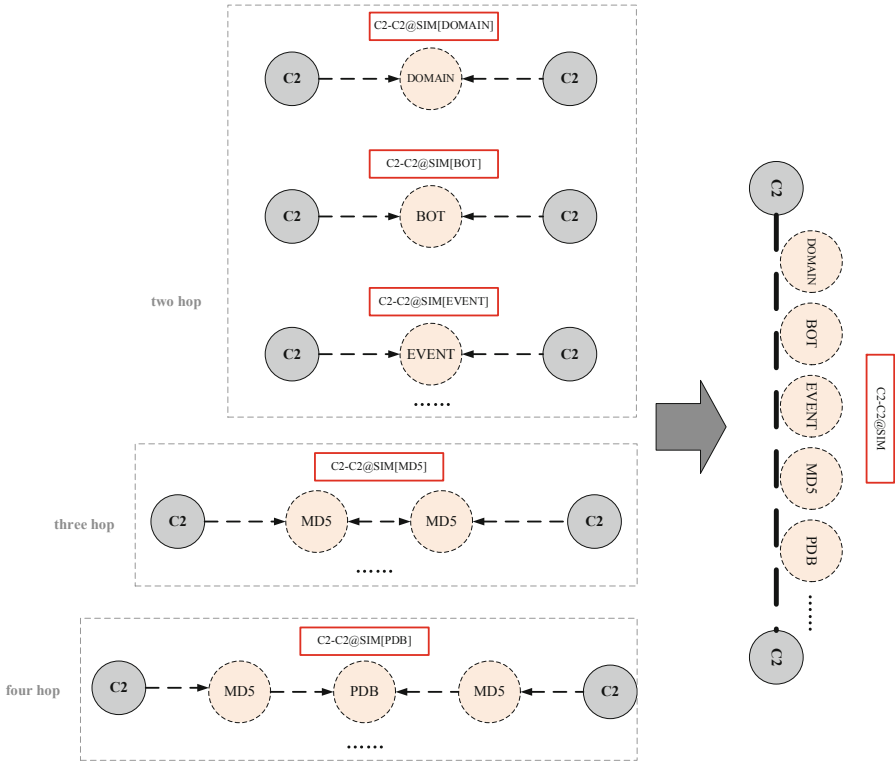
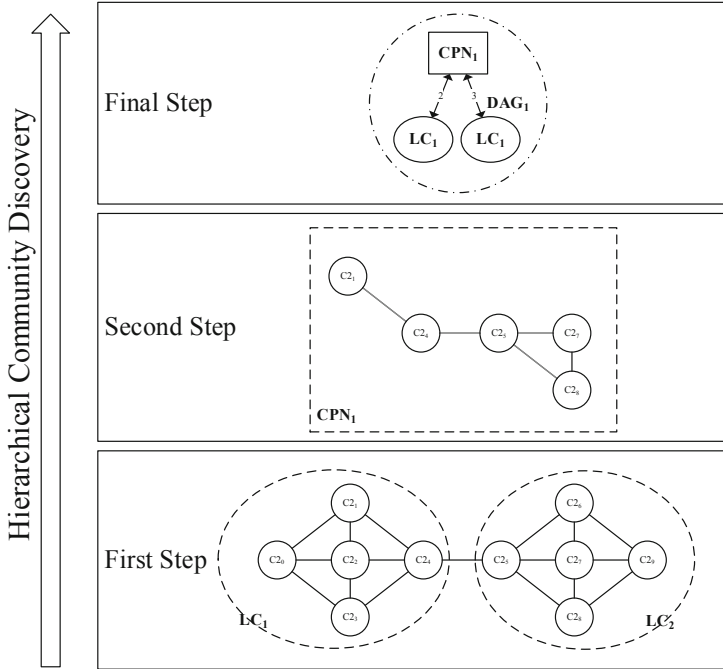


Fig. 4. Similarity calculation of C2s

### 3.3 Hierarchical Community Discovery

Considering the goal of DDoS attack graph discovery, we aim to find groups with several characteristics:

- **Behavior-based similarity:** Attackers in a specific group should have similarity in their large-scale attack behaviors, for instance, in a certain time period, bots should be controlled by the same set of C2 servers, bots or C2 servers should participate in the same set of attack events. Behavioral similarity is adopted to measure whether the entities in a specific group may have the same temporary goal beneath their attacks. Only entities with behavioral similarities above threshold will be considered into the same group.
- **Evidence-based relevance:** Unlike behavior-based similarity from large-scale attacks or connections, evidence-based relevance is built to extract relevance from small-scale relations with high confidence. For example, in a certain time period, two C2 servers are both resolved by the same domain names, or both have network connections from the same malware samples or malware samples with a high similarity. Evidence-based relevance may appear in small scales, but they should not be neglected in our grouping strategy, due to the fact that they are strong evidence of same attack resources and attacking methods.



**Fig. 5.** Hierarchical community discovery workflow.

Hence, we need to establish a community detection workflow to capture closeness from behavior and evidence. Meanwhile, similarity connections between  $C_2$  servers build up a large weighted correlated graph, and time-efficiency should be taken into good consideration. Various unsupervised learning techniques are available for community detection, but none of them can capture large-scale behavior closeness and small-scale strong relevance at the same time. A 3-step workflow is set up to accomplish our attack group discovery, as shown in Fig. 5.

First, we choose to use the Louvain method to discover groups in  $C_2$  servers from their behavior similarity, considering Louvain’s efficient handling of large networks. In this step, Louvain [9] will output community results with the best modularity.

Second, we run Connected Component Algorithm [10] on the super graph of  $C_2$  vertices connected by strong evidence. As a result, the super graph will be spitted in to several components, in which any two vertices are connected to each other by paths, and which is connected to no additional vertices in the super graph.

Last, we merge overlapping Louvain’s communities and Connected Component’s component result. A community and a component will be merged into a DAG group if they both have the same  $C_2$  vertices. Each final DAG group consists of a set of  $C_2$  vertices.

This workflow will be illustrated in detail below:

**First Step: Louvain.** The Louvain algorithm was proposed in 2008, which is one of the fastest modularity-based algorithms and works well with large graphs. Modularity is defined as follows:

$$Q = \frac{1}{2m} \sum_{i,j} \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \quad (4)$$

Modularity  $Q$  [9] has a value between  $-1$  and  $1$ , which measures the quality of relation density within communities.  $A_{ij}$  represents the weight of the edge between  $i$  and  $j$ ,  $k_i = \sum_j A_{ij}$  is the sum of the weights of the edges attached to vertex  $i$ ,  $c_i$  is the community to which vertex  $i$  is assigned, the  $\delta$ -function  $\delta(u, v)$  is  $1$  if  $u = v$  and  $0$  otherwise and  $m = \frac{1}{2} \sum_{ij} A_{ij}$ .

The method consists of repeated applications of two steps. At the beginning, each node of the graph is considered as a community. The first phase is a repeated and sequential assignment of nodes to their neighbor communities, favoring local optimizations of modularity score, until no further improvement can be achieved. The second phase is the definition of a new coarse-grained network based on the communities found in the first phase. These two phases are repeated until no further modularity-increasing reassignments of communities are possible.

At the end of the Louvain process, we can derive communities of C2 vertices ( $LC_i, i = [1, l]$ ,  $l$  is the number of communities) with the best global modularity.

**Second Step: Connected Component.** Connected Component is a simple algorithm with time efficiency of  $O(n \log n)$ ,  $n$  is the number of nodes in the graph. Nodes in a component are connected by paths while different components have no overlapping nodes. It works well in large-scale networks. Hence, we can extract subgraphs of C2 vertices connected by strong evidence from fundamental graph into a super graph EG (evidence graph). Running Connected Component on EG will help us find out the components ( $CPN_i, i = [1, p]$ ,  $p$  is the number of components) within which all possible evidence paths are considered.

**Final Step: Merging Communities and Components.** In this Step, components from the second step and communities from the first step are taken as nodes, links will be established if any two components and communities have the common C2 vertices. We simply run Connected Components algorithm on this graph, which results in several subgraphs. After correlating subgraphs with former community-related C2 and component-related C2, we can obtain our final DDoS attack groups DAG. Each attack group consists of a set of C2 vertices.

## 4 Evaluation

The evaluation process is illustrated as follows. Firstly, we extract entities and events from input data sources, then we are able to grasp the trend of active entities/events and construct our fundamental graph and similarity graph. Secondly, we run hierarchical community discovery on graphs and evaluate the effectiveness of DDoS attack groups. Finally, we conduct an in-depth analysis on several typical DDoS attack groups.

### 4.1 Statistics of Input Data and Graphs

According to our extraction strategy and attack event definition, we are able to know the scale of entities participating in attack events and resources attackers used, as well as the trend from different perspectives.

**Monthly Trend.** After we extract entities and attack events, we can derive the statistics of active entities and activities in each month.

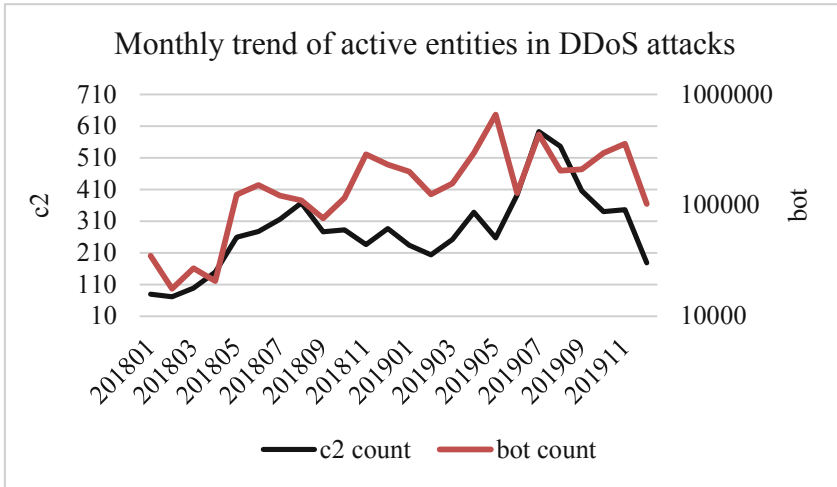
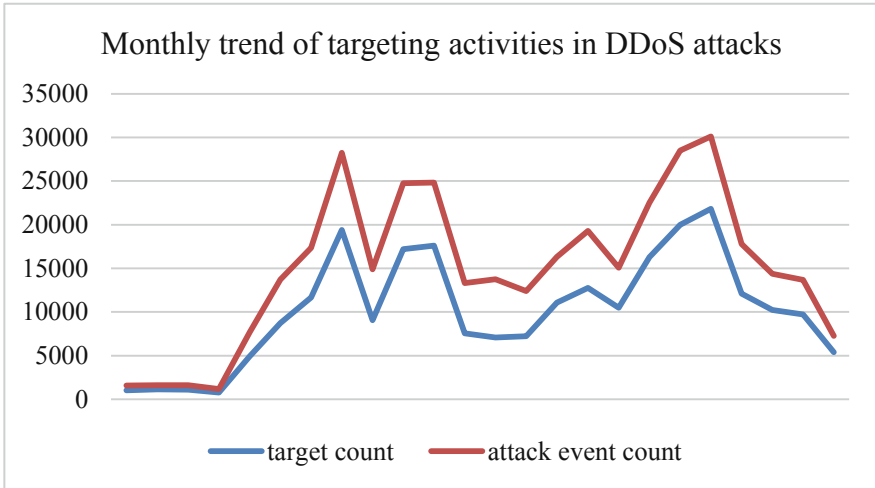


Fig. 6. Monthly trend of active entities in DDoS attacks

Figure 6 shows us the number of active C2 servers and bots in each month. Though the trend of each month is moving up and down, the overall trend is upwards in the scale of C2 servers and bots. Meanwhile, the control ability of botnets is enhanced a lot in 2019 than in 2018. In May 2019, 392 C2 servers control botnets of over 0.65 million bots.

Figure 7 illustrates the trend of targeting activities in DDoS attacks. An interesting fact is that targeting activities reach a peak of each year in August in both 2018 and 2019. In August 2019, active botnets conduct over 30K attack events targeting 21K destination. On average, each target suffers from DDoS attacks for approximately a day and a half.

**Graph Construction.** Table 5 and Table 6 show the scale of a fundamental graph built from entities and relations from behavior data and threat intelligence data. Vertices of types C2, BOT, TARGET origin from botnet communication logs, while types DOMAIN, SHA256, PDB origin from passive DNS data and threat intelligence data.



**Fig. 7.** Monthly trend of targeting activities in DDoS attacks

**Table 5.** Vertex types in fundamental graph

Vertex type	Count
BOT	3542413
DOMAIN	502354
TARGET	212273
SHA256	29254
C2	5225
PDB	59

**Table 6.** Edge types in fundamental graph

Edge type	Count
SHA256-PDB	316
C2-TARGET	539176
C2-SHA256	29885
SHA256-SHA256	9675
C2-DOMAIN	523410
C2-BOT	5862597

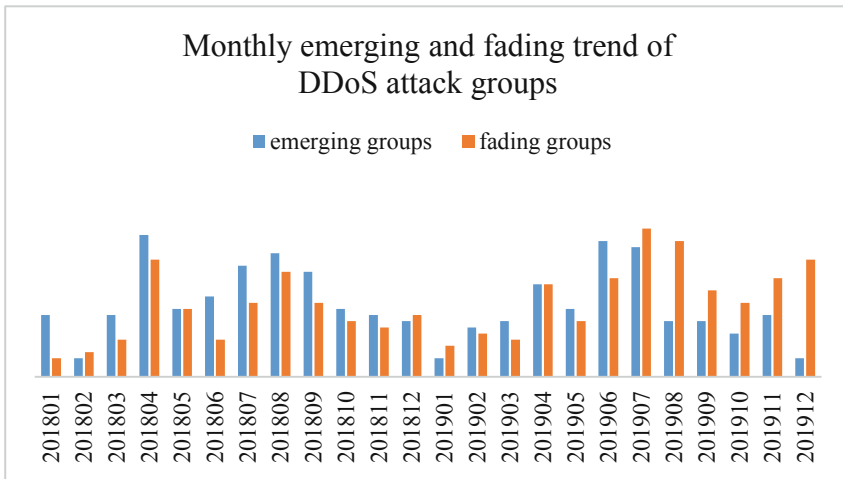
Table 7 shows scale of the graph we construct from meta path-based similarity between C2 servers. The following DDoS attack group discovery is based on this graph.

**Table 7.** Scale of similarity graph

Name	Type	Count
CC	vertex	5225
CC-CC@SIM	edge	13161

### 4.2 Situation Awareness of DDoS Attack Groups

After DDoSAGD framework’s process, we get the result of 282 DDoS attack groups. Each DDoS attack group contains more than one C2 server. DDoS attack groups’ characteristics vary a lot in lifecycle length and scale of attack resources.

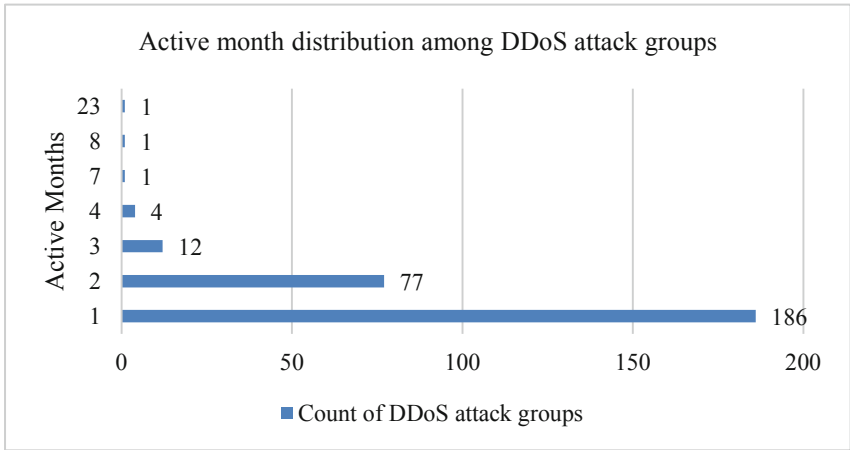


**Fig. 8.** Monthly emerging and fading trend of DDoS attack groups

**Lifecycle of DDoS Attack Groups.** Analysis results reveal the fact that most DDoS attack groups stay active for a relatively short time period, only 19 groups remain active after three months. Meanwhile, attackers can utilize only no more than 10 C2 servers to gain possession of over 27K vulnerable machines or devices to be their botnet army in a very short time.

Figure 8 tells the fact that DDoS attack groups keep emerging and fading in every month. Figure 9 shows active month distribution among all DDoS attack groups. Most attack groups disappear in less than 3 months and the largest group remains active for 23 months.





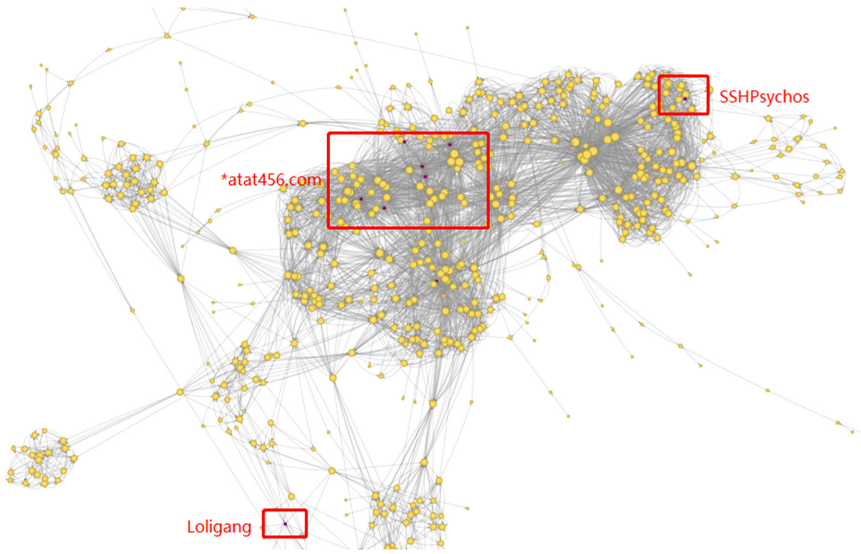
**Fig. 9.** Active month distribution of DDoS attack groups

**Statistics of Top 10 DDoS Attack Groups:** In Table 8, we display the top 10 DDoS attack groups by ranking bots in possession by each one of them.

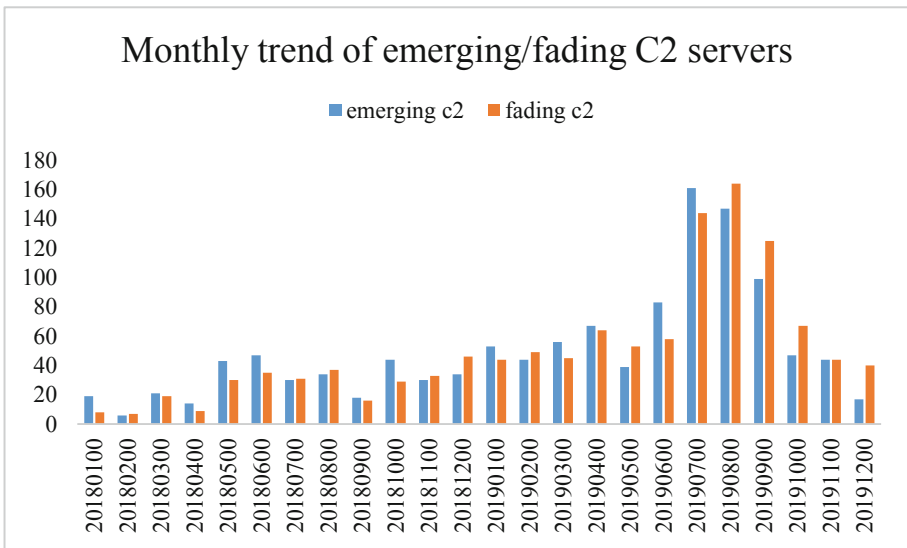
**Table 8.** Statistics of Top 10 DDoS attack groups

Group ID	bot	c2	target	days	domain	sha256
G769	2,768,880	1,197	101,750	647	344,867	6,687
G36341	52,621	23	3,779	48	59	116
G31508	35,698	21	1,809	29	40	93
G1226	27,177	3	1,442	24	5	12
G1291	24,104	3	1,338	34	3	26
G904	23,341	19	7,754	65	15	145
G704	22,605	5	2,781	39	63	37
G1376	21,953	7	2,118	101	2	180
G1466	18,333	2	466	29	0	17
G18837	15,331	28	5,414	66	30	426

**Typical DDoS Attack Group Analysis.** The largest DDoS attack group G769 we discover is found to be related to multiple DDoS attack groups unveiled by different security organizations, such as SSHPsychos or Group 93 [11] from Cisco Talos Group, Loligang [12] from Tencent Threat Intelligence Center and malicious IPs (related to \*atat456.com domains) referred to by many security researchers [13].



**Fig. 10.** Subgraphs related to DDoS attack group uncovered by external security researchers



**Fig. 11.** Monthly trend of emerging/fading C2 servers of DDoS Attack group G769

Figure 10 depicts the relations between C2 servers, in which the red frames are subgraphs related to attack groups recognized by external security researchers in different times. Our approach can construct behavior similarity and threat intelligence relevance for C2 servers, hence be able to correlate them in the same DDoS attack group, confronting the fact that real world attackers keep switching C2 servers to evade detection.

Figure 11 supports this point of view by showing the monthly trend of emerging and fading C2 servers of G769.

## 5 Conclusion

In this paper, a practical attack group framework DDoSAGD is proposed to unveil the facts beneath DDoS attack behaviors. DDoSAGD takes the advantage of a graph theory, and adopts dual community detection methods to discover groups in DDoS attacks. DDoSAGD overcomes the difficulty in discovering attack groups in a long period. Through an in-depth analysis on and comparison with external uncovered attack groups, results verify that our approach is both applicable and efficient in the real world.

**Acknowledgements.** This work was supported in part by National Key R&D Program of China under Grant No. 2017YFB0803005.

## References

1. NSFOCUS. DDoS Attack Landscape, pp. 3–6. NSFOCUS, Beijing (2019). <https://nsfocusglobal.com/2019-ddos-attack-landscape-report>
2. CNCERT/CC. Analysis report of active DDoS attack gang in 2018, p. 3. CNCERT/CC, Guangzhou (2019). <https://www.cert.org.cn/publish/main/upload/File/20190131.pdf>
3. CNCERT/CC. Analysis report on website attack situation and “attack Gang” mining in 2018, pp. 21–38. CNCERT/CC, Guangzhou (2019). <https://www.cert.org.cn/publish/main/upload/File/2018threats.pdf>
4. Yang, H., Sun, X., Zhao, R.: Behavior Analysis of IP Chain-Gangs, pp. 7–22. NSFOCUS, Beijing (2018). [https://nti.nsfocusglobal.com/pdf/Behavior\\_Analysis\\_of\\_IP\\_Chain\\_Gangs.pdf](https://nti.nsfocusglobal.com/pdf/Behavior_Analysis_of_IP_Chain_Gangs.pdf)
5. Zhu, T., Yan, H., Zhu, L.: DDoS attack gang analysis method based on network attack accompanying behavior: China, cn108173884a (2018)
6. Wang, Q., Zhou, H., Yan, H., Mei, R., Han, Z.: Network security situation analysis based on malicious code propagation log. *J. Inf. Secur.* **4**(05), 14–24 (2019)
7. Zhao, T., Qiu, X.: Detection of IP Gangs: Strategically Organized Bots. Springer, New York (2018)
8. Santanna, J.J., De Schmidt, R.O., Tuncer, D., et al.: Booter blacklist: unveiling DDoS-for-hire websites. In: 2016 12th International Conference on Network and Service Management (CNSM), Montreal, QC, pp. 144–152 (2016)
9. Blondel, V.D., et al.: Fast unfolding of communities in large networks. *J. Stat. Mech.: Theory Exp.* **10**(2008), P10008 (2008)
10. Shapiro, L.G.: Connected component labeling and adjacency graph construction. *Mach. Intell. Pattern Recogn.* **19**(19), 1–30 (1996)
11. <https://blogs.cisco.com/security/talos/sshpsychos>
12. [https://mp.weixin.qq.com/s/jPA0lCbSi\\_JLkEn3WoMH7Q](https://mp.weixin.qq.com/s/jPA0lCbSi_JLkEn3WoMH7Q)
13. <https://blog.malwaremustdie.org/2015/07/mmd-0037-2015-bad-shellshock.html>

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# **Hardware Security Implementation**



# Research on the Remote Deployment Design of OTN Electrical Racks

Tianpu Yang<sup>(✉)</sup>, Junshi Gao, Haitao Wang, Guangchong Dai, and Rui Zhai

China Mobile Group Design Institute Co., Ltd., Beijing 10080, China  
yangtianpu@cmdi.chinamobile.com

**Abstract.** The rapid development of 4G and multimedia services drives the exponential increase of the demand for transmission bandwidth. The OTN technology therefore emerges. In recent years, the number of OTN devices in backbone and core equipment rooms has increased sharply. However, due to factors such as equipment room planning, air conditioner, and power supply, new electrical racks cannot be installed in the same equipment room as original optical racks during OTN expansion of 80-wavelength systems. The remote deployment of OTN electrical racks has certain impact on OTN system indicators, OM/OD, and OTU optical-layer parameters. This document analyzes the factors that are affected by the remote deployment of OTN electrical racks, creates simulation models based on scenarios, and provides suggestions on the remote deployment design of OTN electrical racks.

**Keywords:** OTN · Capacity expansion · Remote deployment

## 1 Background

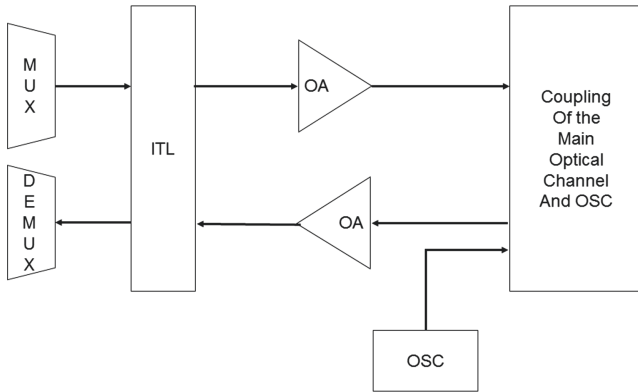
### 1.1 Current Situation

The rapid development of 4G and multimedia services drives the exponential increase of the demand for transmission bandwidth. The OTN technology therefore emerges. Especially in recent years, the number of OTN devices in China Mobile's backbone and core equipment rooms has increased sharply. For example, the inter-province backbone transport network has been constructed from phase 8.2 to phase 12, covering over 2,000 electrical racks on the entire network. However, due to factors such as equipment room planning, air conditioners, and power supplies, new 100G × 80-wavelength OTN systems cannot be installed in the same equipment room as the original optical racks during capacity expansion. As a result, the OTN electrical racks need to be remotely deployed.

### 1.2 Inevitability of Remote Electrical Rack Deployment

As networks develop rapidly, OTN devices are no longer integrated but separated. As the number of OTN devices increases sharply, the power consumption in equipment rooms increases rapidly, and the equipment room footprint is prominently insufficient. It is inevitable that OTN electrical racks are remotely deployed during capacity expansion.

**Optical and Electrical Racks Have Been Separated.** As shown in Fig. 1, the industry's current AWG demultiplexing/multiplexing mode at the optical layer consists of optical-layer boards such as AWG multiplexers/demultiplexers, comb filters, and optical amplifiers (OAs). These boards are connected using optical fibers rather than backplane buses. The backplanes of the optical-layer boards only need to be responsible for power supply and communication control. Therefore, the required structure is simple.



**Fig. 1.** Main optical-layer architecture of AWG multiplexers/demultiplexers

The OTN architecture provides an over 64 Tbit/s switching capability, over 1 Tbit/s in each slot. The backplanes are powerful in grooming. If optical-electrical integration is still used, slot waste will cause loss of the electrical-layer grooming capability. To address this issue, optical-electrical separation can be used so that optical and electrical racks can play their respective advantages to achieve the optimal combination of performance and costs.

**Rapid Development of Transmission Requirements Brings Sharp Increase of OTN Quantity.** The rapid development of 4G and broadband has driven the rapid increase of the traffic of transmission networks, which will be further boosted by 5G and 4 K applications. Currently, a provincial backbone network has three to four planes. Optical racks deployed in one cabinet at core sites support two to four optical directions. Considering that 50% wavelengths need to be added and dropped, four to eight OTN electrical racks need to be installed. The more services, the more optical directions and systems, and the more OTNs.

**Power Consumption Increase of the Entire OTN Device Requires More Cabinets for Installing OTN Devices.** Service development drives technology improvement. The improvement of cross-connect capacity and integration leads to the continuous increase of the power consumption of the entire device. After the single-wavelength bandwidth of OTN reaches 100G, the single-bit power consumption is continuously reduced, but the device capacity increases from less than 5 Tbit/s to over 20 Tbit/s, causing the rapid power consumption increase of the entire device. However, the heat

dissipation and maintenance of the current operators' equipment rooms are outdated. The heat dissipation conditions of transmission equipment rooms cannot support higher power consumption. Therefore, boards need to be split into multiple electrical cabinets. As a result, a large number of OTN devices are installed in more OTN electrical racks. For example, a fully configured Huawei OSN 9800 U64 subrack cannot be installed in an equipment room supporting a maximum power consumption of 5000 W. Four U32 subracks need to be installed instead.

**Poor Equipment Room Planning, Making Capacity Expansion Restricted by Cabinet Space and Power Supplies.** Some equipment rooms are not well planned. For example, one equipment room houses multiple types of devices, such as transmission devices and IP devices, or houses devices of the national backbone network, provincial backbone network, and local metro network. In addition, the development among some sites is unbalanced. For example, a provincial core site accesses multi-layer services at the same time, such as those from the national backbone, provincial backbone, and local metro networks. There are five to six rings and multiple optical directions, requiring more than 50 OTN electrical cabinets. Therefore, as services develop, cabinet space or power supply becomes insufficient. As a result, new equipment rooms must be constructed. After capacity expansion, the electrical racks can only be remotely deployed.

## 2 Implementation Mode and Impact

The remote deployment of OTN devices prolongs the distance between OTU and OM/OD boards, increasing the attenuation and affecting the receive optical power of OTU boards. For example, if a pair of multiplexer/demultiplexer is added for optical-layer regeneration between the existing optical layer and a remote OTN electrical rack, system performance indicators will be affected. In addition, the remote electrical racks cause inconvenience to fiber patch cord maintenance.

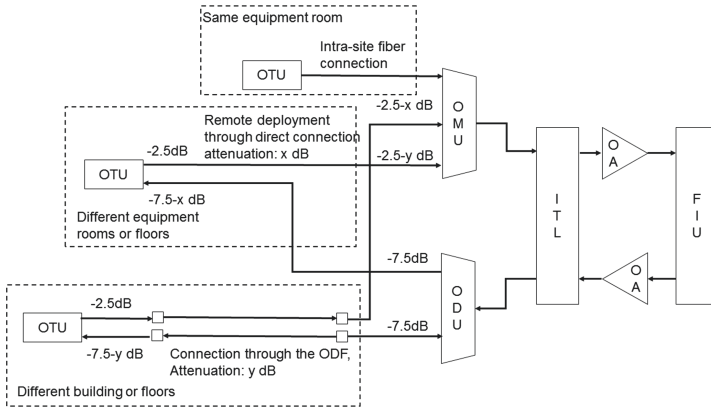
### 2.1 Implementation Modes

Based on factors such as environments and distances, optical and electrical racks can be remotely deployed in the ways described in the following sections.

**Direct Connection Using Optical Fibers/Cables.** Direct connection using optical fibers or cables is the most common mode, which is used in the following scenarios:

- An electrical rack and an optical rack in the same equipment room are directly connected using an optical fiber routed along the fiber trough. In this case, fiber attenuation can be ignored.
- An electrical rack and an optical rack in different equipment rooms on the same floor are directly connected through an optical fiber. In this case, fiber attenuation is large.
- An electrical rack and an optical rack on different floors or in different buildings are connected using ODFs. In this case, both fiber attenuation and connector attenuation must be considered (Fig. 2).



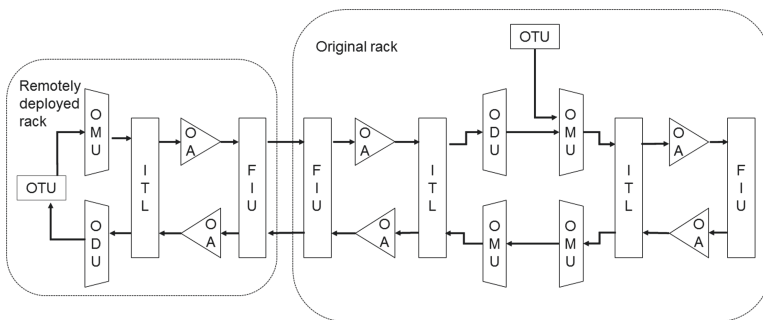


**Fig. 2.** Direct connection between OTU boards and multiplexers/demultiplexers using optical fibers

Devices will generate alarms and cannot function properly in the following direct connection scenarios:

- The remote electrical rack and the original optical rack are deployed in different buildings or at different sites. The fiber attenuation is greater than 2.5 dB.
- The optical fibers between the remote rack and the original rack cannot be routed properly, or the number of optical fibers is insufficient.

**Adding OMSs.** Optical multiplex sections (OMSs) are added to connect the original optical rack and the remote electrical rack. To be specific, a pair of optical racks that share a pair of fiber cores is added, and the new optical racks and original racks are connected using LC-LC fiber patch cords, as shown in Fig. 3.



**Fig. 3.** Adding OMSs

**Adding WSS Boards for Demultiplexing/Multiplexing.** A pair of WSS board is added between the original optical rack and the remote electrical rack for demultiplexing/multiplexing, as shown in Fig. 4.

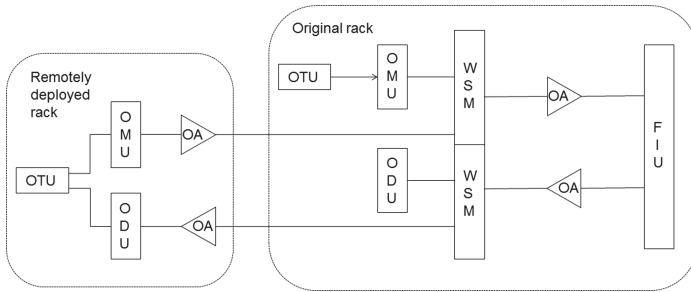


Fig. 4. Adding WSS boards for demultiplexing/multiplexing

## 2.2 Impact on System OSNR

The remote deployment of electrical racks has certain impact on system OSNR:

- In direct connection mode, the attenuation caused by the remote deployment of electrical racks occurs before the OTU board. As a result, the received signals and noise of the OTU board are amplified or reduced at the same level, without affecting the OSNR.
- Adding an OMS to separate optical and electrical racks is equivalent to adding an OMS with 10 dB loss between them, which decreases the OSNR by 0.3 to 1 dB.
- Adding WSS boards to separate optical and electrical racks introduces fixed noise, which has impact on the OSNR.

## 2.3 Impact on the Receive Optical Power of the OM/OD or OTU Boards

The optical and electrical racks of OTN devices are separated, and the electrical racks are remotely deployed. This affects the receive optical power of the OM/OD and OTU boards.

**Direct Connection Using Optical Fibers/Cables.** In the direct fiber connection mode, the input optical power of the OM port and the receive optical power of the OTU board decrease. The attenuation less than 2.5 dB has no impact on the system. When the attenuation is greater than 2.5 dB, the input optical power of the OTU board becomes excessively low. As a result, an alarm may be reported.

- Permitted attenuation between the wavelength-dropping port of the demultiplexer and the input port of the OTU board: The output single-wavelength optical power of the OA is 1 dBm, and the insertion loss of the demultiplexer is less than 6.5 dB. Considering the flatness, the output optical power of the wavelength-dropping port of the demultiplexer is  $-7.5$  dBm. It is recommended that the receive optical power of the optical port on the OTU board be greater than or equal to  $-10$  dBm. Otherwise, an alarm indicating abnormal optical power may be generated. Therefore, the permitted attenuation between the demultiplexer and the OTU board must be less than 2.5 dB.

- Permitted attenuation between the output port of the OTU board and the multiplexer: The transmit optical power of the OTU board is  $-2.5$  dBm, the insertion loss of the multiplexer is less than  $8$  dB, and the gain of the OA on the transmit side is about  $20$  dB. The single-wavelength output optical power of the OA is  $1$  dBm. Assume that the minimum single-wavelength input optical power is  $-19$  dBm. Considering the impact of  $3$  dB flatness, the permitted attenuation is calculated as follows:  $-2.5 - 8 - 3 - (-19) = 5.5$  dB.

**Adding OMSs.** The input optical power at the OM is decreased to  $-7.5$  dB. The impact on the system optical power is like that optical signals pass through one more OTM site. Since OAs are added, there is no impact on the receive optical power of OTU boards.

**Adding WSS Boards for Demultiplexing/Multiplexing.** In this mode, OAs are added. The gain of the OAs and the VOA at the input ports of the OAs are adjusted to achieve the optimal optical power. The impact on the optical power at the receive end of the OM/OD and OTU boards does not need to be considered. Only the impact on the OSNR needs to be considered.

## 2.4 Other Impacts

The remote deployment of electrical racks also has the following impacts:

- For fiber routing across floors or buildings, each OTU port requires two fibers. For the remote deployment of electrical racks in an 80-wavelength system, 160 fibers are required.
- Fault locating: Once a remote optical fiber is faulty, it is difficult to locate and rectify the fault. In this case, the faulty optical fiber needs to be re-routed and replaced.
- Attenuation caused by fiber aging: When the remote deployment distance is long, the impact of fiber attenuation caused by aging and increased connector attenuation must be considered.

## 3 Applicability Analysis

### 3.1 Advantages and Disadvantages

**Direct Connection Using Optical Cables.** The mode of connecting remotely deployed OTN electrical racks using optical cables has the following advantages:

- Optical cables have a high protection level and strong tension and compression resistance capabilities.
- Optical cables have low requirements on terrain. They can be buried underground or routed through pipes.
- After the deployment is complete, the attenuation of the optical cables changes slightly and is not affected by future construction.
- Optical cables can be repaired once they are broken.

- However, the mode of connecting remotely deployed OTN electrical racks using optical cables has the following disadvantages:
- Optical cables have high costs and the construction period is long.
- ODFs are required to connect racks, which increases the connector attenuation.

In summary, this mode requires that optical cables have high reliability and deployment flexibility. The attenuation caused by the optical cables and ODFs is within the permitted range, which has no impact on the OSNR. Fiber cores need to be reserved to prevent abnormal attenuation of some fiber cores.

**Direct Connection Using Optical Fibers.** The mode of connecting remotely deployed OTN electrical racks using optical fibers has the following advantages:

- Optical fibers can be used to directly connect racks without ODFs, and extra connector loss does not need to be considered.
- The cost is lower than that of optical cables.

However, the mode of connecting remotely deployed OTN electrical racks using optical fibers has the following disadvantages:

- Optical fibers have a poor protection capability and require dedicated cable troughs.
- Optical fibers may need to be customized based on required lengths.
- If an optical fiber is faulty, a new optical fiber is required to replace the faulty one.

In this mode, bundle optical fibers or armored optical fibers with an enhanced protection capability are required for direct connections, avoiding connector insertion loss. If the attenuation is within the permitted range, the OSNR will not be affected. If the attenuation is beyond the permitted range, the direct fiber connection mode cannot be used.

**Adding OMSs.** After the OMSs are added on the transmission network, the system OSNR will decrease, but the intra-site optical power does not need to be considered.

### 3.2 Scenario Analysis

Solutions vary depending on the scenarios where the OTN electrical racks are remotely installed.

Scenario 1: The OSNR margin of the system is large, and the equipment rooms housing the optical and electrical racks are far from each other.

In this case, the attenuation may easily exceed the threshold. Since the OSNR margin of the system is large, adding OMSs is the optimal mode. In the engineering design, the impact on the OSNR must be considered. The intra-site fiber attenuation and construction impact do not need to be considered.

Scenario 2: The OSNR margin of the system is small, and the equipment rooms housing the optical and electrical racks are far from each other.

In this case, the attenuation of the fibers for direct connections usually exceeds the threshold. So the direct fiber connection mode cannot be used. If OMSs are added, the OSNR margin of the system will be affected and the system performance will deteriorate. Therefore, this mode is not recommended. In this scenario, the space and layout of the equipment rooms need to be adjusted, or the planning and design need to be modified so that the optical and electrical racks are placed in the same equipment room.

Scenario 3: The OSNR margin of the system is large, and the equipment rooms housing the optical and electrical racks are near to each other.

The attenuation in this scenario can be controlled within the required range. The OSNR margin of the system is large. After OMSs are added, the system has sufficient margin. Both modes are applicable. An optimal mode can be selected based on other relevant factors. If the optical and electrical racks are deployed on different floors or two adjacent buildings and fiber troughs are available, the direct fiber connection mode is preferred. Optical and electrical racks are deployed in different buildings and areas (for example, across a street), there is no fiber trough, but there is space for optical racks in the equipment room. In this scenario, it is recommended that OMSs be added for remote deployment.

Scenario 4: The OSNR margin of the system is small, and the equipment rooms housing the optical and electrical racks are near to each other.

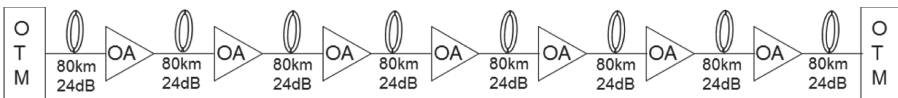
The OSNR margin is small, and OMSs cannot be added for remote deployment. In this scenario, the direct fiber connection mode is recommended so that the attenuation is controlled within the required range.

## 4 Test Models and Simulation

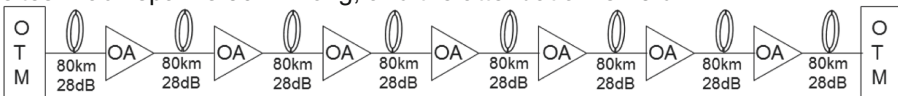
### 4.1 Scenario-Based Modeling

Based on the preceding analysis, this section builds simulation models for scenarios with different OSNR margins in the OTN system. The simulation results are obtained based on the impact of different remote deployment modes on the OSNR.

Model with a large OSNR margin: There are eight OA spans between OTM sites. Each span is 80 km long, and the attenuation is 24 dB.



Model with a small OSNR margin: There are eight OA spans between OTM sites. Each span is 80 km long, and the attenuation is 28 dB.



Other simulation conditions:

- The equipment rooms housing optical and electrical racks are over 1 km away from each other. The equipment rooms are connected using optical cables. Multiple fiber patch cords may exist in the sites. Each site has two ODF connectors. The connector loss is 2 dB ( $0.5 \times 4$ ), the fiber loss is 0.4 dB/km, and the extra loss of pigtails is 0.5 dB. The total loss is 3 dB.
- The equipment rooms housing optical and electrical racks are less than 1 km away from each other, and are directly connected using optical fibers. There is no connector loss, fiber attenuation, or fiber layout loss. The total loss is 1 dB. If an ODF is used to connect the equipment rooms, the connector loss is 1 dB ( $0.5 \times 2$ ), and the fiber loss is 0.5 dB. Considering the fiber layout loss, the total loss is 2 dB.

#### 4.2 Solution-Based Simulation for Different Scenarios

**OSNR simulation for non-remote electrical racks (Table 1).**

**Table 1.** OSNR simulation for non-remote electrical racks

Simulation model	Incident optical power	Receive-end OSNR
$8 \times 24$ dB	1 dBm	20.6 dB
$8 \times 28$ dB	1 dBm	16.7 dB

**OSNR simulation for direct fiber connection of an electrical rack (Table 2).**

**Table 2.** OSNR simulation for direct fiber connection of remote deployment (loss: < 2.5 dB)

Simulation model	Incident optical power	Receive-End OSNR
$8 \times 24$ dB	1 dBm	20.6 dB
$8 \times 28$ dB	1 dBm	16.7 dB

**OSNR simulation for direct fiber connection of an electrical rack (Table 3).**

**Table 3.** OSNR simulation for direct fiber connection of remote deployment (loss: 5 dB)

Simulation model	Incident optical power	Receive-end OSNR
$8 \times 24$ dB	1 dBm	20.4 dB
$8 \times 28$ dB	1 dBm	16.6 dB

### OSNR simulation for remote deployment through added OMSs (Table 4).

**Table 4.** OSNR simulation for remote deployment through added OMSs (added span: 10 dB)

Simulation model	Incident optical power	Receive-end OSNR
$8 \times 24$ dB	1 dBm	20.2 dB
$8 \times 28$ dB	1 dBm	16.5 dB

### 4.3 Conclusion

The simulation results are summarized as follows:

- When the racks are directly connected using optical fibers, the attenuation is within 2.5 dB, and the receive-end OSNR remains unchanged.
- When the racks are directly connected using optical fibers, the attenuation is greater than 2.5 dB, which has slight impact on the OSNR but great impact on the receive optical power of the OTU boards. When the system optical power fluctuates, alarms are easily generated.
- Adding OMSs to connect racks affects the OSNR. When the OSNR margin is large, the OSNR is decreased by about 0.5 dB. When the OSNR margin is small, the OSNR is decreased by 0.2 dB, which still has great impact on the system.
- If the racks are near to each other and the attenuation is less than 2.5 dB, the direct fiber connection mode is recommended, which has no impact on the OSNR.
- If the racks are far away from each other, the attenuation is greater than 2.5 dB, and the OSNR margin is large, it is recommended that OMSs be added to connect the racks.
- If the racks are far away from each other, the attenuation is greater than 2.5 dB, and the OSNR margin is small, remote deployment is not recommended. Instead, the equipment room and planning should be adjusted to house the racks together.

## 5 Application on Live Networks

The OTN electrical rack remote deployment solution described in this document has been applied in phase II of China Mobile's ITMC private network. The following uses the remote electrical rack deployment in the Xili equipment room in Shenzhen as an example.

### 5.1 Live Network Description

As shown in Fig. 5, the Xili equipment room has five directions. MS 5030 is connected to the Guanlan equipment room; MS 5031 and MS 5033 are connected to equipment rooms in Hong Kong to form a four-point WDM ring network; MS 5019 is connected to an equipment room in Qinghedong of Guangzhou; MS 5009 is connected to an equipment room in Ganzhou of Jiangxi province.

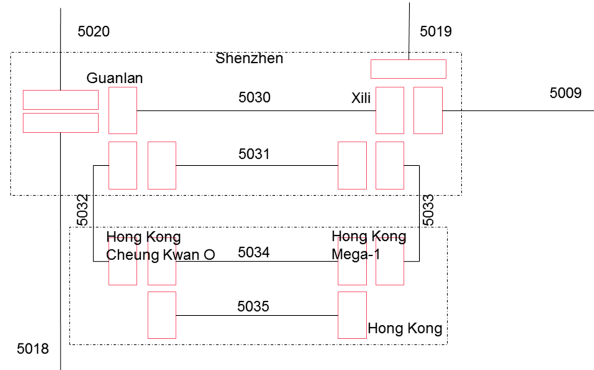


Fig. 5. Network architecture of the Xili equipment room in Shenzhen

### 5.2 Solution Selection

Table 5 lists the performance parameters, current-period performance parameters, OSNR thresholds, and OSNR margin of each MS during the network construction in phase 9.2.

Table 5. Live-network performance parameters

MS No.	MS name	Designed OSNR	Live-network OSNR	OSNR threshold	Current OSNR margin
5009	Xili – Ganzhou	18.6	19	18.4	0.6
5019	Xili – Qinghedong	21.7	21.2	18.4	2.8
5030	Xili – Guanlan	31.5	30.3	18.4	11.9
5031	Xili – Guanlan	26.1	23.3	18.4	4.9
5033	Xili – Hong Kong Mega I	23.1	26	18.4	7.6

According to the survey on the live network, if the direct fiber connection mode is used, the attenuation of the connection between the new electrical rack (OTU board) and the old optical rack (M40V/D40) at the local site is less than or equal to 1.5 dB. Based on the specifications of second-generation 100G, Table 6 lists the performance changes of each MS.

MS 5009 determines the distance between the new electrical rack and the old optical rack in Xili, because among the five MSs, the OSNR margin of MS 5009 is the lowest.

Solution 1: Use optical fibers/cables for direct connection. When the second-generation 100G boards are used, the OSNR margin of MS 5009 is 3.4 dB. If this solution is used, the incident optical power of the OTU boards in MS 5009 can remain unchanged, that is, the OSNR on the live network should be 19 dB. The OSNR margin of 100G channels on the live network is 0.6 dB. If the second-generation 100G boards



**Table 6.** Performance parameters after direct fiber connection

MS No.	MS name	OSNR threshold	OSNR at 1.5 dB connection loss	OSNR margin at 1.5 dB connection loss	Connection insertion loss at OSNR margin greater than 1 dB	Max. insertion loss supporting service provisioning
5009	Xili – Ganzhou	15.6	19	3.4	2.5	2.5
5019	Xili – Qinghedong	15.6	21	5.4	2.5	2.5
5030	Xili – Guanlan	15.6	30.3	14.7	2.5	2.5
5031	Xili – Guanlan	15.6	22.1	6.5	2.5	2.5
5033	Xili – Hong Kong Mega I	15.6	24.5	8.9	2.5	2.5

are used for new services, the OSNR threshold of the second-generation 100G HDFEC boards is expected to be 15.6 dB, and the OSNR margin will be 3.4 dB.

**Solution 2: Add OMSs.** The new electrical rack and old electrical rack are connected through an optical rack, which is similar to an inter-office transfer ring. Simulate the scenario where the line attenuation between the new electrical rack and the old electrical rack is 10 dB, the end-to-end OSNR from the new electrical rack in Xili to Ganzhou is 17.3 dB, the OSNR threshold is 15.6 dB, and the OSNR margin is 1.7 dB.

**Solution 3: Adjust the position of equipment rooms.** If the new equipment room is located together with the old equipment room and the fiber distance and attenuation remain unchanged, the OSNR of each MS remains unchanged.

Based on the preceding analysis, when the equipment rooms cannot be adjusted, the actual indicators are consistent with the analysis results if solution 1 is used.

## 6 Conclusions and Suggestions

When the equipment room space and power consumption on the live network are increasingly limited, separate deployment of optical and electrical racks needs to be considered during live network planning and design, and relevant rules need to be specified in advance. This prevents implementation failures and O&M difficulties in the future and prevents service performance from being affected by remote deployment.

Based on the comparison and analysis of remote deployment modes and scenarios, performance simulation, and live network implementation solutions, it is recommended that the remote deployment modes be selected as follows:

- If optical and electrical racks are close to each other and the OSNR margin is large, use optical fibers to directly connect the racks or add OMSs to connect the racks.
- If optical and electrical racks are close to each other and the OSNR margin is small, use optical fibers to directly connect the racks.
- If optical and electrical racks are far away from each other and the OSNR margin is large, add OMSs to connect the racks.
- If optical and electrical racks are far away from each other and the OSNR margin is small, adjust the equipment room plan or service plan, because remote deployment is not applicable.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# **Intrusion/Anomaly Detection and Malware Mitigation**



# An Effective Intrusion Detection Model Based on Pls-Logistic Regression with Feature Augmentation

Jie Gu<sup>1,2</sup>(✉)

<sup>1</sup> Postdoctoral Research Station, Agricultural Bank of China, Beijing 100005, China  
gujie@pku.edu.cn

<sup>2</sup> School of Electronics Engineering and Computer Science, Peking University,  
Beijing 100871, China

**Abstract.** Computer network is playing a significantly important role in our society, including commerce, communication, consumption and entertainment. Therefore, network security has become increasingly important. Intrusion detection systems have received considerable attention, which not only can detect known attacks or intrusions, but also can detect unknown attacks. Among the various methods applied to intrusion detection, logistic regression is the most widely used, which can achieve good performances and have good interpretability at the same time. However, intrusion detection systems usually confront with data of large scale and high dimension. How to reduce the dimension and improve the data quality is significant to improve the detection performances. Therefore, in this paper, we propose an effective intrusion detection model based on pls-logistic regression with feature augmentation. More specifically, the feature augmentation technique is implemented on the original features with goal of obtaining high-qualified training data; and then, pls-logistic regression is applied on the newly transformed data to perform dimension reduction and detection model building. The NSL-KDD dataset is used to evaluate the proposed method, and the empirical results show that our proposed method can achieve good performances in terms of accuracy, detection rate and false alarm rate.

**Keywords:** Feature augmentation · Intrusion detection · Logistic regression · Partial least square · Network security

## 1 Introduction

With the rapid development of internet, networks are becoming more and more important in our daily life. Organizations rely heavily on networks to do on-line transactions, and also, individuals are dependent on networks to work, study and entertain. In a word, networks are an essentially indispensable part in modern society. However, this over-dependence on networks might have potential risk, because considerable information that relates to organization operation and individual activities is accumulated and stored. It would cause huge losses, when the networks are been invaded or attacked.

© The Author(s) 2020

W. Lu et al. (Eds.): CNCERT 2020, CCIS 1299, pp. 133–140, 2020.

[https://doi.org/10.1007/978-981-33-4922-3\\_10](https://doi.org/10.1007/978-981-33-4922-3_10)

Intrusion detection systems are the most widely used tool to protect information from being compromised. Intrusion detection has been long considered as a classification problem [1, 2]. Various statistic-based and machine-learning-based methods have been applied to improve the performances of intrusion detection systems [3, 4]. However, machine learning-based methods for intrusion detection suffer criticisms [5]. Though many machine-learning-based detection methods, such as support vector (SVM) machine and artificial neural network (ANN), could achieve better detection performances, the detailed procedures of the detection process remain unknown. It is called the black-box which is not favorable for practical applications. Moreover, machine-learning-based detection methods are common time-consuming. For example, the training complexity of SVM cannot be tolerable when confront with large-scale and high dimension dataset. However, the statistic-based detection methods could cover these shortages to a large extent in terms of the model interpretation and training speed. Therefore, it can be inferred that when compared to machine-learning-based intrusion detection approaches, statistic-based intrusion detection method have some advantages, that is, good interpretability and fast training speed.

Among these statistic-based detection methods, logistic regression is the most widely used classification approach, which could achieve good detection performances [6–8]. It is worthy to noting that logistic regression could model the correlations among feature and take into account of the joint effects between features to produce a decision boundary to separate different classes effectively. Therefore, logistic regression can be considered as an effective detection method. However, we should also realize that to achieve further improvement in detection performance, it may not be sufficient to use logistic regression alone. Review of related work in intrusion detection indicates that data quality data quality has been considered as a critical determinant [9].

Therefore, in our study, we propose an effective intrusion detection framework based on pls-logistic regression with feature augmentation. Specifically, the feature augmentation technique is used to improve the data quality, and pls-logistic regression is chosen to reduce the dimension and build the intrusion detection model using the transformed data. The reminder of this paper is organized as follows. In Sect. 2, we give a brief overview of feature augmentation and pls-logistic regression. Section 3 describes the details of the proposed intrusion detection model. Section 4 presents the experiment settings, results and discussions. Finally, Sect. 5 comes to conclusion.

## 2 Methodology

To better illustrate the proposed detection model, firstly, we briefly review the main principles of the feature augmentation [10] in Sect. 2.1, as well as the pls-logistic regression classification model [11] in Sect. 2.2.

### 2.1 Feature Augmentation

Following Fan et al. (2016), suppose we have a pair of random variables  $(\mathbf{X}, Y)$  with  $n$  observations, where  $\mathbf{X} \in \mathbb{R}^p$  denotes the original features and  $Y \in \{0, 1\}$  denotes the corresponding binary response. The logarithm marginal density ratio transformation is used

as the feature augmentation technique to transform the original features. Specifically, for  $X_j, j = 1, 2, \dots, p$  in  $\mathbf{X}$ , denote by  $f_j, g_j$  the class conditional densities, respectively, for class 1 and class 0, that is,  $(X_j|Y = 1) \sim f_j$  and  $(X_j|Y = 0) \sim g_j$ . Denote by  ${}^1X_j = \{X_{ij}|Y_i = 1, i = 1, 2, \dots, n\}$  and  ${}^0X_j = \{X_{ij}|Y_i = 0, i = 1, 2, \dots, n\}$ . Then,  $f_j, g_j$  is obtained by kernel density estimation on  ${}^1X_j$  and  ${}^0X_j$ , and denote the estimates by  $\hat{f}_j$  and  $\hat{g}_j$ , respectively. Thus, the feature augmentation for  $X_j$  using logarithm marginal density ratio transformation is shown as follows:

$$X'_j = \log \hat{f}_j(X_j) - \log \hat{g}_j(X_j), \tag{1}$$

where  $X'_j$  denotes the transformed feature for the  $j$  th feature  $X_j$ .

### 2.2 Pls-Logistic Regression Classification Model

Suppose we have a pair of random variables  $(X, Y)$ , where  $X \in \mathbb{R}^p$  denotes the original features and  $Y \in \{0, 1\}$  denotes the corresponding binary response. The procedures of pls-logistic regression is depicted as follows:

- Step 1.** Perform univariate logistic regression on each feature to obtain  $p$  coefficients denoted by  $\omega^1 = (\omega_1, \omega_2, \dots, \omega_p)$ . Denote the normalized  $\omega^1$  by  $\bar{\omega}^1$ .
- Step 2.** Extract the first pls component  $t_1$  by  $t_1 = \mathbf{X} \cdot \bar{\omega}^1$ .
- Step 3.** Perform OLS regression of  $X$  against  $t_1$ . Denote the residual of  $X$  by  $\mathbf{X}^*$ .
- Step 4.** Perform logistic regression on each feature of  $\mathbf{X}^*$  against  $t_1$  to obtain the  $p$  coefficients of features in  $\mathbf{X}^*$ , denoted by  $\omega^2$ , and then normalize  $\omega^2$  to  $\bar{\omega}^2$ .
- Step 5.** Extract the second pls component  $t_2$  by  $t_2 = \mathbf{X}^* \cdot \bar{\omega}^2$ .
- Step 6.** Repeat Step 3, Step 4 and Step 5 until the stopping criteria are satisfied.
- Step 7.** Denote by  $t_1, t_2, \dots, t_h$  the final extracted pls components. Perform the logistic regression on these pls components to build the classification model.

## 3 Proposed Intrusion Detection Model: Fa-Plslogistic

In this section, we present the main procedures of our proposed intrusion detection model based on pls-logistic with feature augmentation. By embedding the data quality improvement technique into pls-logistic, we can obtain an effective intrusion detection with good performances and less complexity. First, we perform feature transformations on the original features to obtain high-quality training data that can significantly improve the detection performances. Then, the pls-logistic regression is perform on the newly transformed data to conduct dimension reduction and build the intrusion detection model. For clarity, the detailed procedures are summarized as follows:

- **Step 1. Data transformation**  
Perform feature transformations on the original data to obtain high-qualified training data.
- **Step 2. Detection model building**  
Use the newly obtained data from Step 1 to train pls-logistic-based classifier and build the intrusion detection model.

- **Step 3. Intrusion detection**

For a new testing sample, it is first transformed by the logarithm marginal density ratio transformation illustrated in Sect. 2.1; then, the transformed data is fed into the built intrusion detection model to classify it as either an intrusion or a normal.

## 4 Experimental Setting

### 4.1 Dataset Description

In our study, the NSL-KDD dataset is used to evaluate the performance of the proposed intrusion detection model. The NSL-KDD dataset is a modified version of KDD 99 dataset which is considered as the benchmark dataset in intrusion detection domain. However, the KDD 99 dataset suffers from some drawbacks [12, 13]. For example, there are redundant and duplicate records which cause the classifier would be biased towards these more frequent records. The NSL-KDD dataset was proposed by [14] by removing all the redundant samples and reconstituting the dataset, making it more reasonable not only in data size, but also in data structure. The NSL-KDD dataset contains TCP connections that consist of 41 features and one labeling feature.

### 4.2 Experimental Results and Discussion

In order to prevent the dominance of features with large ranges, we normalize the data into a range of [0, 1] before conducting the experiments. To evaluate our proposed detection model, the 10-fold cross validation has been adopted and the performance is evaluated by the following measurements according to the confusion matrix presented in Table 1.

**Table 1.** Confusion matrix

		Predicted	
		Attack	Normal
Actual	Attack	TP	FN
	Normal	FP	TN

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}, \text{Detection rate (DR)} = \frac{TP}{TP+FN}, \text{False alarm rate (FAR)} = \frac{FP}{TN+FP}$$

To verify the effectiveness of our proposed intrusion detection model, we first compare the detection performance of Fa-plslogistic with that of the naïve-plslogistic detection model (pls-logistic regression on original data without feature transformation). The 10-fold cross validation results of these two detection models on NSL-KDD dataset with regard to accuracy, DR, FAR and training time are summarized in Table 2.

As the results shown in Table 2, our proposed intrusion detection model takes clear advantages over the naïve-plslogistic detection model, indicating that the data quality improvement technique can greatly boost the detection performance. More specifically,

**Table 2.** Performances of proposed methods

Metric	PLS-logistic (with feature augmentation)	PLS-logistic (without feature augmentation)
Accuracy (%)	97.39(0.33)	91.29(6.03)
DR (%)	96.95(0.32)	88.59(12.84)
FAR (%)	2.23(0.45)	6.35(3.05)
Training time (in sec)	98.66	137.51

the accuracy and detection rate of our proposed model both exceed 96%, while naïve-plslogistic only achieves 91.29% and 88.59%, respectively. Besides, in terms of false alarm rate, our proposed method is below 2.3%, while naïve-plslogistic is over 6%. Moreover, the performances of our proposed is also more robust than that of naïve-plslogistic.

To further demonstrate the advantages of our proposed method, the training time required by Fa-plslogistic and naïve-plslogistic is also compared in Table 2. As shown, the training time of our proposed method is superior to that of naïve-plslogistic. Specifically, naïve-plslogistic demands about 1.39 times as much training time as Fa-plslogistic does. Thus, it can be inferred that our proposed method is much more concise than naïve-plslogistic, which can reduce the training time.

Therefore, according to the comparison results, it can be concluded that our proposed intrusion detection model is more effective than naïve-plslogistic and can achieve better detection performances.

Standard errors are in the parentheses in percentage form.

In addition, we examine which features are influential on the intrusion detection. Here, for simplicity, the feature whose coefficient is greater than 1 after standardization is considered to be important. Thus, the influential features recognized during the 10-fold cross-validation are shown in Table 3.

According to the results in Table 3, the important features for intrusion detection are listed in descending order by frequency: land, su\_attempted, num\_failed\_logins, src\_bytes, urgent, hot, num\_root, num\_compromised, root\_shell, is\_guest\_login and dst\_bytes. These features are helpful in practice to efficiently detect network intrusion and attacks.

Furthermore, in order to better interpret the effectiveness of our proposed method in intrusion detection, performance comparisons between our proposed model and other existing methods in intrusion detection using NSL-KDD dataset are conducted. The comparison results are summarized in Table 4.

From the comparison results shown in Table 4, our proposed method outperforms other intrusion detection methods with regard to detection accuracy. However, it should be noted that Table 4 just provides a snapshot of performance comparison between our proposed method and other detection methods. Thus, it can be claimed that our proposed method always performs better when compared to any other methods. Nevertheless, from the results above, we can make a conclusion that our proposed method



**Table 3.** Influential features for intrusion detection

K-fold	Influential feature
1	src_bytes, land, hot, su_attempted, num_root
2	land, num_failed_logins
3	land, urgent, num_failed_logins
4	src_bytes, land, hot, root_shell, su_attempted, is_guest_login
5	dst_bytes, num_compromised
6	src_bytes, land, num_compromised, root_shell, num_root
7	land, urgent, num_failed_logins, num_compromised, su_attempted, num_root
8	src_bytes, urgent, hot, num_failed_logins, su_attempted
9	su_attempted
10	src_bytes, land, urgent, hot, num_failed_logins, su_attempted, is_guest_login

**Table 4.** Performance comparisons of proposed method and other detection methods

Method	Accuracy (%)
GHSOM [15]	96.02
A-GHSOM [16]	96.63
Naïve Bayes + N2B [17]	96.50
AdaBoost [17]	90.31
Proposed method	97.39

still possesses advantages in intrusion detection and can provide inspirations for the following researches.

## 5 Conclusion

Intrusion detection system is critical to network security. In this paper, we proposed an effective intrusion detection model based on pls-logistic with feature augmentation. Though the pls-logistic classifier might achieve a good performance, the detection capacity is much more dependent on the quality of the training data. Therefore, in order to increase the detection capacity, we use the logarithm marginal density ratio transformation on the original data to obtain high-quality training data for pls-logistic before building the intrusion detection model. Empirical results on NSL-KDD dataset show that our proposed intrusion detection model is effective and can achieve good and robust detection performances.

**Acknowledgments.** This research was financially supported by National Natural Science Foundation of China (Grant No. 72001222, 61832001, 61702016).

## References

1. Kumar, G., Thakur, K., Ayyagari, M.R.: MLEsIDSs: machine learning-based ensembles for intrusion detection systems—a review. *J. Supercomput.* **76**(11), 8938–8971 (2020). <https://doi.org/10.1007/s11227-020-03196-z>
2. Bamakan, S.M.H., Wang, H., Yingjie, T., Shi, Y.: An effective intrusion detection framework based on MCLP/SVM optimized by time-varying chaos particle swarm optimization. *Neurocomputing* **199**, 90–102 (2016)
3. Moustafa, N., Hu, J., Slay, J.: A holistic review of network anomaly detection systems: a comprehensive survey. *J. Netw. Comput. Appl.* **128**, 33–55 (2019)
4. Tsai, C.F., Hsu, Y.F., Lin, C.Y., Lin, W.Y.: Intrusion detection by machine learning: a review. *Expert Syst. Appl.* **36**(10), 11994–12000 (2009)
5. Sommer, R., Paxson, V.: Outside the closed world: on using machine learning for network intrusion detection. In: 2010 IEEE Symposium on Security and Privacy, pp. 305–316 (2010)
6. Wang, Y.: A multinomial logistic regression modeling approach for anomaly intrusion detection. *Comput. Secur.* **24**(8), 662–674 (2005)
7. Mok, M.S., Sohn, S.Y., Ju, Y.H.: Random effects logistic regression model for anomaly detection. *Expert Syst. Appl.* **37**(10), 7162–7166 (2005)
8. Ji, S.Y., Choi, S., Jeong, D.H.: Designing an internet traffic predictive model by applying a signal processing method. *J. Netw. Syst. Manag.* **23**(4), 998–1015 (2015)
9. Aburomman, A.A., Reaz, M.B.I.: A survey of intrusion detection systems based on ensemble and hybrid classifiers. *Comput. Secur.* **65**, 135–152 (2017)
10. Fan, J., Feng, Y., Jiang, J., Tong, X.: Feature augmentation via nonparametrics and selection (FANS) in high-dimensional classification. *J. Am. Stat. Assoc.* **111**(513), 275–287 (2016)
11. Bastien, P., Vinzi, V.E., Tenenhaus, M.: Pls generalised linear regression. *Comput. Stat. Data Anal.* **48**(1), 17–46 (2005)
12. Mahoney, M.V., Chan, P.K.: An analysis of the 1999 DARPA/lincoln laboratory evaluation data for network anomaly detection. In: Vigna, G., Kruegel, C., Jonsson, E. (eds.) RAID 2003. LNCS, vol. 2820, pp. 220–237. Springer, Heidelberg (2003). [https://doi.org/10.1007/978-3-540-45248-5\\_13](https://doi.org/10.1007/978-3-540-45248-5_13)
13. Bamakan, S.M.H., Wang, H., Yingjie, T., Shi, Y.: An effective intrusion detection framework based on MCLP/SVM optimized by time-varying chaos particle swarm optimization. *Neurocomputing* **199**, 90–102 (2016)
14. Tavallaee, M., Bagheri, E., Lu, W., Ghorbani, A.A.: A detailed analysis of the KDD CUP 99 data set. In: Proceedings of the 2009 IEEE Symposium on Computational Intelligence in Security and Defense Applications, pp. 1–6. IEEE (2009)
15. Yu, Z., Tsai, J.J., Weigert, T.: An adaptive automatically tuning intrusion detection system. *ACM Trans. Auton. Adapt. Syst.* **3**(3), 10 (2008)
16. Ippoliti, D., Zhou, X.: A-GHSOM: an adaptive growing hierarchical self-organizing map for network anomaly detection. *J. Parallel Distrib. Comput.* **72**(12), 1576–1590 (2012)
17. Panda, M., Abraham, A., Patra, M.R.: Discriminative multinomial naive bayes for network intrusion detection. In: Proceedings of 2010 Sixth International Conference on Information Assurance and Security, pp. 5–10. IEEE (2010)

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





# DeepHTTP: Anomalous HTTP Traffic Detection and Malicious Pattern Mining Based on Deep Learning

Yuqi Yu<sup>1</sup>(✉), Hanbing Yan<sup>1</sup>(✉), Yuan Ma<sup>3</sup>, Hao Zhou<sup>1</sup>, and Hongchao Guan<sup>2</sup>

<sup>1</sup> National Computer Network Emergency Response Technical Team/Coordination Center of China, Chaoyang District, Beijing, China

{yyq, yhb}@cert.org.cn

<sup>2</sup> Beijing University of Posts and Telecommunications, Haidian District, Beijing, China

<sup>3</sup> Chongqing Municipal Public Security Bureau, Yuzhong District Branch, Chongqing, China

**Abstract.** Hypertext Transfer Protocol (HTTP) accounts for a large portion of Internet application-layer traffic. Since the payload of HTTP traffic can record website status and user request information, many studies use HTTP protocol traffic for web application attack detection. In this work, we propose DeepHTTP, an HTTP traffic detection framework based on deep learning. Unlike previous studies, this framework not only performs malicious traffic detection but also uses the deep learning model to mine malicious fields of the traffic payload. The detection model is called AT-Bi-LSTM, which is based on Bidirectional Long Short-Term Memory (Bi-LSTM) with attention mechanism. The attention mechanism can improve the discriminative ability and make the result interpretable. To enhance the generalization ability of the model, this paper proposes a novel feature extraction method. Experiments show that DeepHTTP has an excellent performance in malicious traffic discrimination and pattern mining.

**Keywords:** Bi-LSTM · Attention mechanism · Anomalous HTTP traffic detection · Malicious pattern mining

## 1 Introduction

According to the 2018 Internet Security Report released by China National Computer Network Emergency Response Technical Team/Coordination Center (CNCERT/CC) [1], website attacks and exploits occur frequently. How to improve the ability of web attack detection is one of the urgent problems in the field of network security.

Among various network protocols, Hypertext Transfer Protocol (HTTP) occupies a considerable proportion of the application layer traffic of the Internet. Since HTTP traffic can record website access states and request content, it provides an excellent source of information for web application attack detection [2–4]. We focus on HTTP traffic mainly for three reasons. 1) Although protocol HTTPS is used by 57.4% of all the websites [5], HTTP traffic still accounts for a large proportion of network traffic. Research [6] shows that for smaller B2B websites, the uptake of HTTPS is low. Because they lack awareness

of the streaming importance of SSL. Also, the perceived complexity of switching to HTTPS is high. 2) A large majority of malware uses HTTP to communicate with their C&C server or to steal data. Many web application attacks use HTTP, such as Cross-site scripting attack (XSS), SQL injection, and so on. 3) The HTTP protocol is transmitted in clear text, which makes it easier to analyze network behaviors.

In this paper, we design DeepHTTP, a complete framework for detecting malicious HTTP traffic based on deep learning. The main contributions are as follows.

Firstly, unlike researches that only detect malicious URLs (Uniform Resource Locators) [7, 8], we extract both URL and POST body (if the HTTP method is POST) to detect web application attacks. This is of great help to portray network behavior more comprehensively.

Secondly, we perform an in-depth analysis of the types and encoding forms of HTTP traffic requests, then propose an effective method to extract content and structure features from HTTP payload (in this paper, “payload” refers to URL and POST body). Content and structure features are used for classification.

Thirdly, the detection model AT-Bi-LSTM is Bidirectional Long Short-Term Memory (Bi-LSTM) [9] with attention mechanism [10]. Since each HTTP request follows the protocol specification and grammar standards, we treat elements in traffic payload as vocabulary in natural language processing and use Bi-LSTM to learn the contextual relationship. The attention mechanism can automatically dig out critical parts, which can enhance the detection capabilities of the model. Due to the introduction of attention mechanism, the model is more interpretable than other deep learning models.

Finally, we design a module for malicious pattern mining. The “malicious pattern” is essentially a collection of strings representing web attacks. Specifically, we cluster malicious traffic entries and perform pattern mining for each cluster. Then we can generate new rules based on the mined malicious patterns. New rules will be configured into detection systems to capture specific types of web attacks.

In a word, DeepHTTP is a complete framework that can automatically distinguish malicious traffic and perform pattern mining. We set up a process that can verify and update data efficiently. The model is updated periodically so that it can adapt to new malicious traffic that appears over time.

The rest of this paper is organized as follows. Section 2 gives a summary of the relevant research. Section 3 briefly introduces the system framework and data preprocessing methods. The proposed model is introduced in detail in Sect. 4, including the malicious traffic detection model and pattern mining method. We launched a comprehensive experiment to demonstrate the effectiveness of the model. The experimental results are shown in Sect. 5. Section 6 gives the conclusions and future works.

## 2 Related Work

### 2.1 Malicious Traffic Detection

In recent years, quite a few researches are aiming for detecting anomaly traffic and web application attacks. Communication traffic contains lots of information that can be used to mine anomaly behaviors. Lakhina et al. [58] perform a method that fuses

information from flow measurements taken throughout a network. Wang et al. [59] propose Anagram, a content anomaly detector that models a mixture of high-order n-grams designed to detect anomalous and “suspicious” network packet payloads. To select the important features from huge feature spaces, Zseby et al. [60] propose a multi-stage feature selection method using filters and stepwise regression wrappers to deal with feature selection problem for anomaly detection. The methods mentioned above care less about the structural features of communication payloads which are important for distinguishing anomaly attacking behaviors and mining anomaly patterns. In this paper, we put forward a structure extraction approach, which can help enhance the ability to detect anomaly traffic. The structure feature also makes an important role in pattern mining.

Existing approaches for anomalous HTTP traffic detection can be roughly divided into two categories according to data type: feature distribution-based methods [11, 12] and content-based methods [13]. Content-based methods can get rid of the dependency of artificial feature extraction and is suitable for different application scenarios. Nelms T et al. [14] use HTTP headers to generate control protocol templates including URL path, user-agent, parameter names, etc. Because Uniform Resource Locator (URL) is rich in information and often used by attackers to pass abnormal information, identifying malicious URLs is a hot studied problem in the security detection [8, 15, 16]. In this paper, we use both URL and POST body (if the HTTP method is POST) to detect web attacks. We do not use other parameters in the HTTP header because these fields (like Date, Host, and User-agent, etc.) have different value types and less valid information.

Various methods have been used for detection. Juvonen and Sipola [18] propose a framework to find abnormal behaviors from HTTP server logs based on dimensionality reduction. Researchers compare random projection, principal component analysis, and diffusion map for anomaly detection. Ringberg et al. [19] propose a nonparametric hidden Markov model with explicit state duration, which is applied to cluster and scout the HTTP-session processes. This approach analyses the HTTP traffic by session scale, not the specific traffic entries. Additionally, there are also many kinds of research based on traditional methods such as IDS (intrusion detection system and other rule-based systems) [3, 20–22]. Since malicious traffic detection is essentially an imbalanced classification problem, many studies propose anomaly-based detection approaches that generate models merely from the benign network data [17]. However, in practical applications, the anomaly-based detection model usually has a high false-positive rate. This problem undoubtedly increases the workload of manual verification.

With the rapid development of artificial intelligence, deep learning has been widely used in various fields and has a remarkable effect on natural language processing. Recently, deep learning has been applied to anomaly detection [8, 23–25]. Erfani et al. [25] present a hybrid model where an unsupervised DBN is trained to extract generic underlying features, and a one-class SVM is trained from the features learned by the DBN. LSTM model is used for anomaly detection and diagnosis from System Logs [24]. In this article, we use deep learning methods to build detection models to enhance detection capabilities.

## 2.2 Pattern Mining Method

In addition to detecting malicious traffic and attack behaviors, some researches focus on pattern mining of cluster traffic. Most existing methods for traffic pattern recognition and mining are based on clustering algorithms [26, 27]. Le et al. [27] propose a framework for collective anomaly detection using a partition clustering technique to detect anomalies based on an empirical analysis of an attack's characteristics. Since the information theoretic co-clustering algorithm is advantageous over regular clustering for creating a more fine-grained representation of the data, Mohiuddin Ahmed et al. [28] extend the co-clustering algorithm by incorporating the ability to handle categorical attributes which augments the detection accuracy of DoS attacks. In addition to the clustering algorithm, JT Ren [29] conducts research on network-level traffic pattern recognition and uses PCA and SVM for feature extraction and classification. I. Paredes-Oliva et al. [30] build a system based on an elegant combination of frequent item-set mining with decision tree learning to detect anomalies.

The signature generation has been researched for years and has been applied to protocol identification and malware detection. FIRMA [31] is a tool that can cluster network traffic clusters obtained by executing unlabeled malware binaries and generate a signature for each cluster. Terry Nelms et al. [32] propose ExecScent, a system that can discover new C&C domains by building adaptive templates. It generates a control protocol template (CPT) for each cluster and calculates the matching score to find similar malware. These tools have proven to automatically generate valid signatures, but the process still needs to define the composition of the initial signature or template in advance. As far as we know, signature generation is rarely used in web attack detection. The study of pattern mining for malicious traffic is not yet mature.

In recent years, the attention-based neural network model has become a research hotspot in deep learning, which is widely used in image processing [33], speech recognition [34], and healthcare [35]. Attention mechanism has also proved to be extremely effective. Luong et al. [36] first design two novel types of attention-based models for machine translation. Since the attention mechanism can automatically extract important features from raw data, it has been applied to relation Classification [37] and abstract extraction [38]. To the best of our knowledge, as for HTTP traffic detection and pattern mining, proposed models rarely combine sequence models with attention mechanism. Hence, in this paper, we build a model based on attention mechanism, which can get rid of the dependency of artificial extraction features and do well in pattern mining.

## 3 Preliminaries

### 3.1 DeepHTTP Architecture and Overview

The "Rule Engine" mentioned in this paper is an engine that consists of many rules. Each rule is essentially a regular expression used to match malicious HTTP traffic that matches a certain pattern. Generally, the expansion of the rule base relies on expert knowledge. It requires high labor costs. And the malicious traffic that the "Rule Engine" can detect is limited. Therefore, we additionally introduce a deep learning model based on the attention mechanism, which can identify malicious traffic entries that are not detected

by the “Rule Engine”. Also, the pattern mining module can automatically extract the string patterns in the traffic payload, which can greatly reduce the workload of rule extraction.

In this paper, rules can be roughly divided into seven categories according to the type of web application attack: File Inclusion (Local File Inclusion and Remote File Inclusion), framework vulnerability (Struts2, CMS, etc.), SQL Injection (Union Select SQL Injection, Error-based SQL Injection, Blind SQL Injection, etc.), Cross-Site Scripting (DOM-based XSS, Reflected XSS, and Stored XSS), WebShell (Big Trojan, Small Trojan and One Word Trojan [39]), Command Execution (CMD) and Information Disclosure (system file and configuration file).

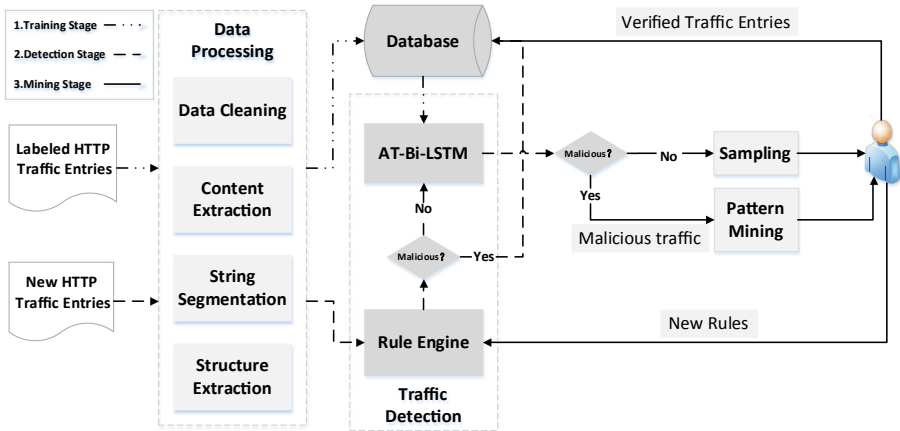


Fig. 1. DeepHTTP architecture.

DeepHTTP is a complete framework that can detect web application attacks quickly and efficiently. In this section, we introduce three stages of DeepHTTP (see Fig. 1), which are training stage, detection stage, and mining stage.

- **Training stage.** The core task of this phase is to train the model (AT-Bi-LSTM). It includes data processing and model training. First, we put the labeled dataset into the data processing module to obtain content and structure features of traffic payload. After that, we divide the processed formatted data into training, test, and verification sets and store in the database. To enhance the robustness of the model, we build data sets containing positive and negative samples in different proportions and use cross-validation to train the model.
- **Detection stage.** The pre-trained model and the “Rule Engine” are used for anomaly traffic detection. After data processing, new HTTP entries are first entered into the “Rule Engine” for detection. For the entries which are detected by the engine, we labeled the data and update them directly into the database. Other traffic entries will be entered into the pre-trained model for detection. Anomaly traffic entries detected by AT-Bi-LSTM will be used in the mining stage.



- **Mining stage.** The main works of this phase are verifying the anomalous traffic labeled by the model and then mining malicious patterns. Generally speaking, there are a large number of traffic entries that the model identifies as malicious. To improve efficiency, we first cluster and sample the data. Specifically, malicious traffic will be divided into different clusters by clustering. In each cluster, we mine malicious patterns based on attention mechanism and then generate new rules. Simultaneously, we sample a small number of entries from each cluster and perform manual verification. Verified malicious data will be updated regularly to the database and new rules will be updated regularly to “Rule Engine”.

DeepHTTP is a complete closed-loop workflow. The detection model and “Rule Engine” complement each other. The timing update and feedback mechanism can continuously improve the detection ability of the system, which is the main reason for the practicability of the framework. Data processing, traffic detection model, and pattern mining method are critical parts in DeepHTTP, which will describe in the later sections.

### 3.2 Data Preprocessing

**Data Collection.** The study spends nearly half a year to collect actual traffic. Nearly 1.5 million malicious HTTP traffic samples are accumulated through vulnerability scanning, rule filtering, and manual verification. After sampling and deduplication, we eventually collect 10, 645, 12 malicious samples.

- **Rule-based collection method.** Specifically, we collect network traffic from the university network monitoring system and filter out HTTP traffic. To protect the privacy of teachers and students, we remove sensitive content from the data. Then, we use the “Rule Engine” mentioned in Sect. 3.1 to identify malicious traffic.
- **Tools-based collection method.** In order to enrich the type of malicious traffic, we use kali [40], Paros [41], W3AF [42] to perform simulation attack and vulnerability scanning. We collect relevant traffic as malicious traffic samples.
- **Model-based collection method.** As described in Sect. 3.1, after manual verification, malicious traffic entries detected by AT-Bi-LSTM are periodically updated to the data set.

**Data Cleaning.** We parse HTTP traffic packets and extract Uniform Resource Locator (URL) and POST body (if the request method is POST). Then, we mainly perform the following data cleaning operations:

- URL decoding: Since URL data often been encoded, we perform URL decoding.
- Payload decoding: Many strings in traffic payload are encoded by different encoding methods, like MD5, SHA, and Base64, etc. For these strings, we identify the encoding type and replace them with the predefined flag (see Table 1).
- We replace garbled characters and invisible characters with null characters.
- Since the binary stream data in the Post request body does not contain semantic information, we replace this kind of data with the predefined flag (see Table 1).

**String Segmentation.** Text vectorization is the key to text mining. Numerous studies use n-grams [43] to extract the feature of payloads [44–46]. This method can effectively capture the byte frequency distribution and sequence information, but it is easy to cause dimension disaster. To prevent dimensional disaster, we split the string with special characters. The special characters refer to characters other than English letters and numbers, such as “@”, “!”, “#”, “%”, “^”, “&”, “\*”, “?”, etc. Here is an instance. Suppose the decoded data is: “/tienda1/publico/vaciar.jsp <EOS> B2 = Vaciar carrito; DROP TABLE usuarios; SELECT \* FROM datos WHERE nombre LIKE”. “<EOS>” is the connection symbol. After string splitting, the data is denoted as: “/tienda1 /public /vaciar. jsp <EOS> B2 = Vaciar carrito; DROP TABLE usuarios; SELECT \* FROM datos WHERE nombre LIKE”. Strings are connected by spaces. Another benefit of this approach is that it makes the results of malicious pattern mining more understandable. In this example, the malicious pattern we want to obtain from the data is {“SELECT”, “FROM”, “WHERE”}. However, if we use n-grams ( $n = 3$ ) or character-based method [39], the result may be denoted as {“SEL”, “ELE”, ..., “ERE”} or {“S”, “L”, ..., “R”}, which is not intuitive.

**Structure Feature Extraction.** To better measure the similarity of URLs, Terry Nirm, etc. [32] use a set of heuristics to detect strings that represent data of a certain type and replaces them accordingly using a placeholder tag containing the data type and string length. Inspired by this, the paper uses a similar way to extract structure features from HTTP payload. The “structure feature” mentioned in this paper refers to string type other than the meaning of the string itself. We replace string with predefined flags according to their data type. The types of data we currently recognize include hash (MD5, SHA, and Base64), hexadecimal, binary, Arabic numerals and English alphabet (upper, lower and mixed case) .etc. The main replacement rules are shown in Table 1.

**Table 1.** Characters replacement rules.

Encoding type	Replacement string
MD5 hash	‘MD5_HASH’
SHA hash	‘SHA_HASH’
Base64	‘BASE64_ENCODE’
Hexadecimal	‘HEXADECIMAL’
Encryption	‘ENCRYPTION’
Binary	‘BINARY’

```

/mobile/notify?verifytype=4&verifycontent=68247&tenantid=3c5fee3560000218bf9c5d7b5d3524e
/wwwwww/wwwwww?wwwwwwwwwwww=D&wwwwwwwwwwwwwww=DDDDD&wwwwwwwww=MD5_HASH
-----
/mobile/notify?templettype=8&articlecontent=486975&password=8efe04d797dad53d5c43d21a0d320eab
/wwwwww/wwwwww?wwwwwwwwwwww=D&wwwwwwwwwwwwwww=DDDDD&wwwwwwwww=MD5_HASH
    
```

**Fig. 2.** An example of structure extraction.

Here is an example of a structure feature extraction (see Fig. 2). Since the encoding type of the string “3c5fee3560000218bf9c5d7b5d3524e” is MD5 (We use “hashID” [47] to identify the different types of hashes.), we replace it with “MD5\_HASH”. For those string not belong to any special type, we replace each character in the string with the specified character. “D” for Arabic numeral and “W” for the English alphabet (not case sensitive). Since the string “68247” consists of five Arabic numerals, we replace it with five “D”. Obviously, by extracting structural features, we can easily find requests with different content but almost the same in data type.

## 4 Our Approach

### 4.1 Anomaly HTTP Traffic Detection

The goal of the proposed algorithm is to identify anomaly HTTP traffic based on semantics and structure of traffic entries. Figure 3 shows the high-level overview of the proposed model. The model (AT-Bi-LSTM) contains five components: input layer, word embedding layer, Bi-LSTM layer, attention layer and output layer.

**Problem Definition.** Let  $\mathbf{R} = \{R_1, R_2, \dots, R_i, \dots, R_N\}$  be the set of HTTP traffic entries after data processing. For each traffic entry  $R_i (i = 1, 2, \dots, N)$ , there are two sequences  $S_i^1 = \{c_{11}, c_{12}, c_{13}, \dots, c_{1n}\}$  and  $S_i^2 = \{c_{21}, c_{22}, c_{23}, \dots, c_{2n}\}$ , which respectively represent content sequence and structure sequence. Because structure sequence is derived from content sequence, the length of both sequence is equal to  $n$ .

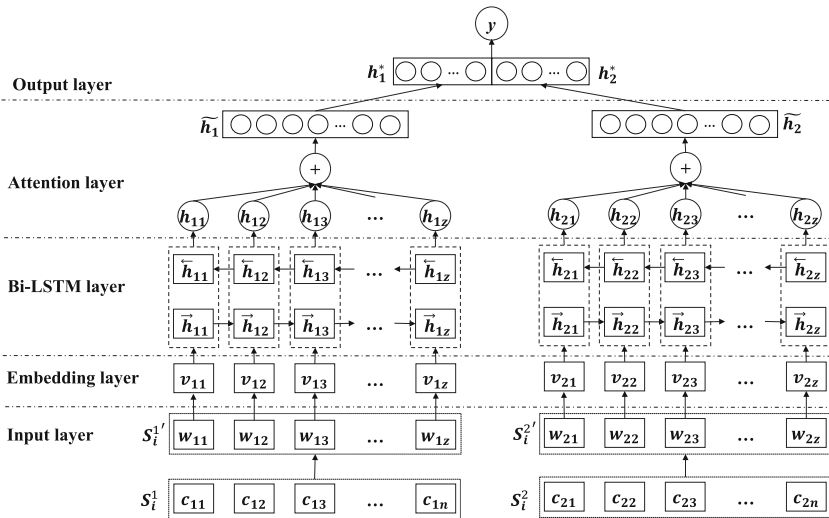


Fig. 3. Model architecture.

**Input Layer.** In this paper, we use the content and structure sequence after word segmentation as a corpus, and select words that are common in the corpus to build a vocabulary according to term frequency inverse document frequency (TF-IDF) [48]. Then, the unique index is generated for each word in the vocabulary. We convert the word sequences ( $S_i^1$  and  $S_i^2$ ) to final input vectors ( $S_i^{1'}$  and  $S_i^{2'}$ ), which are composed of indexes. The length of input vector is denoted as  $z$ , which is a hyper-parameter (the fixed length in this paper is set to 300 because the proportion of sequence length within 300 is 0.8484). The excess part of input sequence is truncated, and the insufficient part is filled with zero. Formally, the sequence of content can be converted to  $S_i^{1'} = \{w_{11}, w_{12}, w_{13}, \dots, w_{1z}\}$  and the sequence of structure can be expressed as  $S_i^{2'} = \{w_{21}, w_{22}, w_{23}, \dots, w_{2z}\}$ . Here is an example. Given a sequence of content: {'/', 'admin', '/', 'caches', '/', 'error\_ches', '.', 'php'}. The input vector with fix length can be denoted as [23, 3, 23, 56, 23, 66, 0, 0, ..., 0]. Since the index of 'admin' in vocabulary is 3, the second digit in the vector is 3. And since the length of this sequence is less than fixed length, the rest of the vector is filled with zeros.

**Embedding Layer.** Take a content sequence of  $i$ -th traffic entry as an example. Given  $S_i^{1'} = \{w_{11}, w_{12}, \dots, w_{1k}, \dots, w_{1z}\}$ , we can obtain vector representation  $v_{1k} \in R^m$  of each word  $w_{1k} \in R^1 (k = 1, 2, \dots, z)$  as follows:

$$v_{1k} = \text{ReLU}(W_e w_{1k} + b_e) \quad (1)$$

where  $m$  is the size of embedding dimension,  $W_e \in R^{m \times 1}$  is the weight matrix, and  $b_e \in R^m$  is the bias vector. Rectified Linear Unit (ReLU) is the rectified linear unit defined as  $\text{ReLU}(v) = \max(v, 0)$ , where  $\max()$  applies element-wise to vector.

**Bidirectional Long Short-Term Memory.** We employ Bidirectional Long Short-Term Memory (Bi-LSTM), which can exploit information both from the past and the future to improve the prediction performance and learn the complex patterns in HTTP requests better. A Bi-LSTM consists of a forward and backward LSTM. Given embedding vector  $\{v_{11}, v_{12}, \dots, v_{1k}, \dots, v_{1z}\}$  of content sequence of  $i$ -th traffic entry  $R_i$ , the forward LSTM  $f$  reads the input sequence from  $v_{11}$  to  $v_{1z}$ , and calculates a sequence of forward hidden states  $(\vec{h}_{11}, \vec{h}_{12}, \dots, \vec{h}_{1k}, \dots, \vec{h}_{1z})$  ( $\vec{h}_{1k} \in R^p$ ) and  $p$  is the dimensionality of hidden states). The backward LSTM  $\bar{f}$  reads the input sequence in the reverse order and product a sequence of backward hidden states  $(\bar{h}_{11}, \bar{h}_{12}, \dots, \bar{h}_{1k}, \dots, \bar{h}_{1z})$  ( $\bar{h}_{1k} \in R^p$ ). The final latent vector representation  $h_{1k} = \begin{bmatrix} \vec{h}_{1k}; \bar{h}_{1k} \end{bmatrix}^T$  ( $h_{1k} \in R^{2p}$ ) can be obtained by concatenating the forward hidden state  $\vec{h}_{1k}$  and the backward one  $\bar{h}_{1k}$ . We deal with the embedding vector of structure sequence in the same way.

**Attention Layer.** In this layer, we apply attention mechanism to capture significant information, which is critical for prediction. General attention is used to capture the relationship between  $h_t$  and  $h_i (1 \leq i < t)$ :

$$\alpha_{ti} = h_t^T W_a h_i \quad (2)$$

$$\alpha_t = \text{softmax}([\alpha_{t1}, \alpha_{t2}, \dots, \alpha_{t(t-1)}]) \quad (3)$$

where  $W_\alpha \in R^{2p \times 2p}$  is the matrix learned by model,  $\alpha_t$  is the attention weight vector calculated by softmax function. Then, the context vector  $c_t \in R^{2p}$  can be calculated based on the weights obtained from Eq. (3). The hidden states from  $h_1$  to  $h_{t-1}$  can be calculated by the following formulas:

$$c_t = \sum_i^{t-1} \alpha_{ti} h_i \quad (4)$$

We combine current hidden state  $h_t$  and context vector  $c_t$  to generate the attentional hidden state as follows:

$$\tilde{h}_t = \tanh(W_c[c_t; h_t]) \quad (5)$$

where  $W_c \in R^{r \times 4p}$  is the weight matrix in attention layer, and  $r$  is the dimensionality of attention state.  $\tilde{h}_1$  and  $\tilde{h}_2$  can be obtained using Eq. (2) to Eq. (5), which denote the attention vector of content and structure sequence learned by the model.

**Output Layer.** Before feeding the attention vector into softmax function, the paper apply dropout regularization randomly disables some portion of attention state to avoid overfitting. It is worth noting that we concatenate vector of content and structure to generate output vector for prediction. The classification probability is calculated as follows:

$$p = \text{softmax}(w_s[h_1^*; h_2^*] + b_s) \quad (6)$$

where  $h_1^*$  is the output of  $\tilde{h}_1$  after dropout strategy,  $h_2^*$  is the output of  $\tilde{h}_2$ .  $w_s \in R^{q \times r}$  and  $b_s \in R^q$  are the parameters to be learned.

$$\hat{y} = \text{argmax}(p) \quad (7)$$

where  $\hat{y}$  is the label predicted by the attention model.

**Objective Function.** The paper calculate the loss for all HTTP traffic entries using the cross-entropy between the ground truth  $y_i \in (0, 1)$  and the predicted  $p_i (i = 1, 2, \dots, N)$ :

$$L = -\frac{1}{N} \sum_{i=1}^N y_i \log(p_{i1}) + (1 - y_i) \log(1 - p_{i1}) \quad (8)$$

where  $N$  is the number of traffic entries,  $p_{i1}$  denotes the probability that the  $i$ -th sample is predicted to be malicious.

We train the model to minimize the objective function so that the model automatically learns the appropriate parameters. The model can automatically learn the feature expression of input data without manual feature extraction. In addition to outputting the judgment results, the model will also output attention weights which will be used as important inputs for the pattern mining part. The introduction of attention mechanism makes this model more explanatory than other deep learning models.

## 4.2 Mining Stage

The function of this module is to interpret the results of the model and extract the string pattern. For malicious traffic that is not detected by the rules engine but is discriminated by the model, we perform pattern mining and verification. Figure 4 shows the architecture of the mining stage.

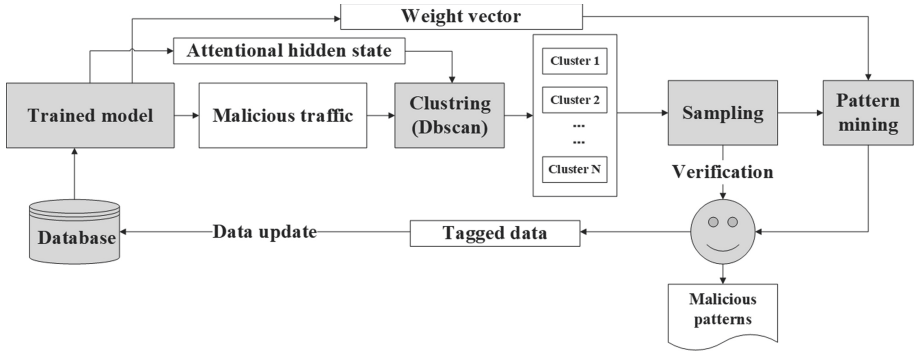


Fig. 4. The architecture of mining stage.

**Clustering.** We cluster traffic entries that were flagged as malicious by AT-Bi-LSTM. Specifically, we feed the attentional hidden state (obtained by Eq. (5) in Sect. 4.1) into the clustering model. The clustering method we apply is DBSCAN [49], a density-based clustering algorithm, which does not require prior declaring the number of clusters. After clustering, we obtain several clusters. Traffic entries in each cluster are similar in content or structure.

**Tag Verification.** In practical applications, there are massive suspicious HTTP requests every day. There is no doubt that manual verification requires a lot of time and effort. In this paper, we use clustering and sampling to reduce the workload. After clustering, we sample some entries from each cluster for verification. If the predicted labels of these samples are consistent with the ground-truth, then all the prediction results in this cluster are considered correct.

**Pattern Mining.** This module can mine the string pattern of the payload of malicious traffic. Experts generate new rules based on the results of pattern mining, which can reduce the workload of manual extraction. As mentioned in Sect. 3.1, the attention weight vector obtained in the attention layer can reflect the crucial parts of the payload. Therefore, for each malicious traffic entry, we dig out the key parts according to the corresponding attention weight vector. The greater the weight is, the more important the word is.

Specifically, given a cluster with  $N$  traffic entries  $T = \{t_1, t_2, \dots, t_N\}$ , we perform pattern mining according to the following steps:

- **Get a keyword set according to attention weight.** AT-Bi-LSTM can output the attention weight vector (obtained by Eq. (3)). For each traffic entry  $t_i (i = 1, 2, \dots, N)$ , we get  $n$  keywords  $K_i = \{k_1, k_2, \dots, k_n\}$  according to its weight vector. The greater the weight, the more important the word is. At last, we can obtain a set of keywords  $K = \{K_1, K_2, \dots, K_N\}$  identified by the model.
- **Extracting frequent patterns.** The goal of this step is to unearth words that not only frequently occur in this cluster but also recognized by the model as key parts. We calculate the co-occurrence matrix of keywords in set  $K$ . If we discovery several words in keywords set  $K$  to appear together frequently, then the combination of these words can represent a malicious pattern. The malicious pattern can be used as an effective basis for security personnel to extract new filtering rules.

## 5 Evaluation

### 5.1 Dataset

We use the method mentioned in Sect. 3.2 to build the HTTP traffic dataset. For the collected data, we perform manual verification and tagging. Finally, the total number of labeled data is 2,095,222, half of them are malicious traffic entries. The types and quantities of tagged malicious samples are shown in Table 2. Moreover, we prepare five million unmarked HTTP traffic for model testing.

**Table 2.** Distribution of malicious traffic entries.

Data type	Number
Deserialization	6014
CMS	5836
File inclusion	46438
SQL injection	463776
Webshell	288050
XSS	127750
Sensitive data exposure	16656
Middleware vulnerability	47614
Struts2 vulnerability	42477
Botnet	19901
Total	1064512

### 5.2 Validation of Structural Feature Extraction Method

To verify the effectiveness of the structural feature extraction method, we compare the convergence speed and detection ability of the model trained by different features.

We record the loss and accuracy of each iteration of the model and draw the loss curve and the accuracy curve (Fig. 5). To balance the memory usage and model training efficiency, the best batch size is set to 200. As we observe from the figure, the model trained based on content and structural features converge faster. In other words, after fusing structural features, the learning rate has been enhanced, and it can reach the convergence state faster.

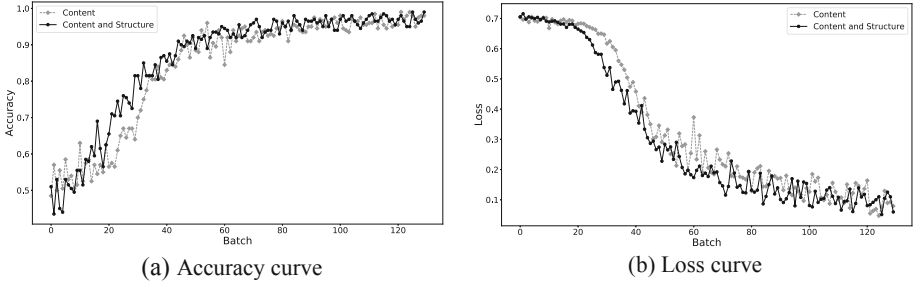


Fig. 5. Accuracy curve and loss curve.

Moreover, in unbalanced dataset, we compare the effects of models trained by different features. As shown in Table 3, the model trained based on content and structure features performs better. The reason is that structural features increase the generalization ability of the model.

Table 3. Performance of models trained by different features

Different features	Precision	Recall	F1-score	AUC
Content feature	0.9856	0.8765	0.9278	0.9382
Structure feature	0.9633	0.6643	0.7863	0.8320
Content and structure features	0.9560	0.9608	<b>0.9584</b>	<b>0.9795</b>

### 5.3 Model Comparison

We use 3-gram [43], TF-IDF [48], Doc2vec [50] and Character\_level feature extraction method [8, 39] to obtain the feature vector of the payload. Then, we compared the effects of models between classic machine learning methods and machine learning models, including Support Vector Machine(SVM) [51], Random Forest(RF) [52], eXtreme Gradient Boosting(XGBoost) [53], Convolutional neural networks (CNNs) [54], Recurrent neural networks (RNNs) [55, 56], Long short term memory (LSTM) [57] and the proposed model AT-Bi-LSTM.



**Detection in Labeled Dataset.** We sample 1.1 million traffic entries from labeled dataset (as described in Sect. 3.2) to build a balanced dataset (550,000 for normal samples and 550,000 for malicious samples). To approximate the actual situation, we also sample 1.1 million traffic entries from labeled dataset to build an unbalanced dataset (1,000,000 for normal samples and 100,000 for malicious samples). Then the data set is divided into training set, test set and verification set according to the ratio of 6:2:2. The evaluation metrics consist of precision, recall, F1-score.

**Table 4.** Model performance in labeled dataset.

Dataset	Classifier	Precision	Recall	F-score
Balanced dataset	3-gram_TF-IDF_SVM	0.9607	0.9564	0.9585
	3-gram_TF-IDF_RF	0.9518	0.9269	0.9378
	3-gram_TF-IDF_XGBoost	0.9755	0.9683	0.9717
	Doc2vec_SVM	0.9365	0.9201	0.9274
	Doc2vec_RF	0.9646	0.9444	0.9534
	Doc2vec_XGBoost	<b>0.9810</b>	<b>0.9753</b>	<b>0.9781</b>
	Doc2vec_CNN	0.9611	0.9467	0.9538
	Doc2vec_LSTM	0.9765	0.9538	0.9650
	Doc2vec_Bi-LSTM	0.9852	0.9791	0.9821
	Character_Level_CNN	0.9556	0.9461	0.9508
	Character_Level_LSTM	0.9895	0.9847	0.9870
	Character_Level_Bi-LSTM	<b>0.9954</b>	<b>0.9921</b>	<b>0.9937</b>
	AT-Bi-LSTM	<b>0.9979</b>	<b>0.9963</b>	<b>0.9970</b>
Unbalanced dataset	3-gram_TF-IDF_SVM	0.6573	0.5987	0.6266
	3-gram_TF-IDF_RF	0.7036	0.6835	0.6934
	3-gram_TF-IDF_XGBoost	0.7499	0.6937	0.7207
	Doc2vec_SVM	0.7531	0.6111	0.6061
	Doc2vec_RF	0.8212	0.7484	0.7675
	Doc2vec_XGBoost	0.8844	0.8570	0.8683
	Doc2vec_CNN	0.8823	0.7851	0.8308
	Doc2vec_LSTM	0.8921	0.8235	0.8564
	Doc2vec_Bi-LSTM	0.9011	0.8221	0.8597
	Character_Level_CNN	0.9365	0.9342	0.9353
	Character_Level_LSTM	0.9485	0.9456	0.9470
	Character_Level_Bi-LSTM	0.9545	0.9574	0.9559
	AT-Bi-LSTM	<b>0.9661</b>	<b>0.9609</b>	<b>0.9635</b>

We can conclude the following conclusions according to Table 4. First, in the balanced dataset, Doc2vec\_XGBoost, Character\_Level\_Bi-LSTM, and AT-Bi-LSTM perform well. However, in the imbalanced dataset, the detection capabilities of Doc2vec\_XGBoost is not as good as deep learning models. Second, although the character-level deep learning models are comparable to AT-Bi-LSTM, the model proposed in this article is superior in interpretability. Finally, AT-Bi-LSTM is superior to all baseline models in almost all metrics. In unbalanced data sets, the superiority of the proposed model is even more pronounced.

At the same time, we record the training time of each model (see Fig. 6). Doc2vec-based deep learning models take more time because using Doc2vec to obtain sentence vectors requires additional training time. Because CNN has faster training speed, the training time of Character\_Level\_CNN is the least. The training time of AT-Bi-LSTM is at the middle level. It is acceptable in practical application.

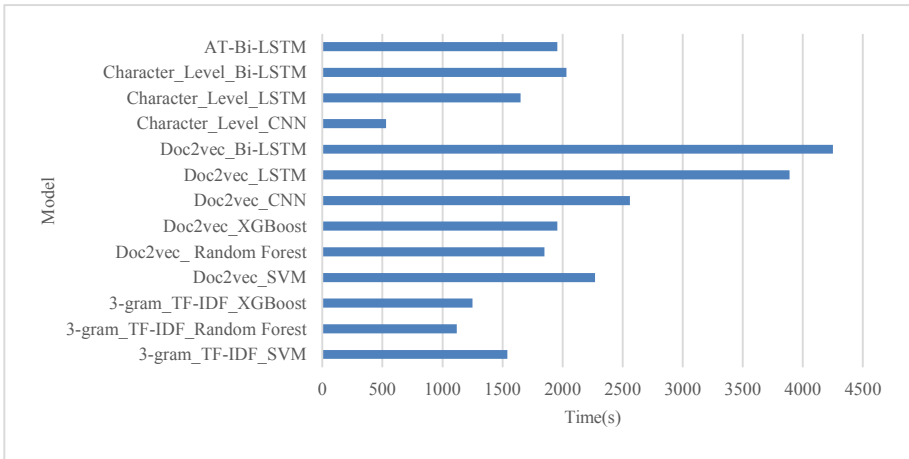


Fig. 6. Training time of models.

**Detection in Unlabeled Dataset.** We conduct comparative experiments using five million unlabeled traffic entries. Rules in “Rule Engine” are derived from expert knowledge so that we use the rules engine to verify the validity of the detection model. The explanation of the assessment indicators is as follows:

$N_M$ . The number of malicious entries detected by the model.

$N_{RE}$ . The number of malicious entries detected by the “Rule Engine”.

$N_M \cap RE$ . The number of malicious entries detected by both the model and the “Rule Engine”.

$M-RE$ . A collection of malicious entries detected by the model but not detected by the “Rule Engine”.

$N_{TP}$ . The number of true positive samples in the  $M-RE$ .

$N_{FP}$ . The number of false positive samples in the  $M-RE$ .

$N\_TP$  and  $N\_FP$  are depend on manual verification.

$Rule\_Coverage\_Rate (RCR) = N\_M \cap RE / N\_RE$ . It represents the coverage of the model towards the “Rule Engine”.

$False\_Rate (FR) = N\_FP / N\_M$ . It means the false rate of the model.

$New\_Rate (NR) = N\_TP / N\_M$ . It represents the ability of the model to identify malicious traffic outside the scope of the “Rule Engine”.

We adopt the “Rule Engine” to extract malicious entries across the overall unlabeled traffic set. The amount of malicious traffic entries detected by “Rule Engine” ( $NMT\_RE$ ) equals to 217100. The result of model evaluation in the unlabeled dataset is shown in Table 5. According to the value of RCR, Doc2vec\_Bi-LSTM, Character\_level\_CNN and AT-Bi-LSTM can basically cover the detection results of the “Rule Engine”. However, Doc2vec\_Bi-LSTM and Character\_level\_CNN have a higher false rate. Overall, AT-Bi-LSTM is superior to other models.

**Table 5.** Model results in the unlabeled dataset.

Model	$N\_M$	$N\_RE$	$N\_M \cap RE$	$N\_TP$	$N\_FP$	RCR	FR	NR
3-gram_TF-IDF_SVM	231246	217100	98965	38957	93324	0.4558	0.4036	0.1685
3-gram_TF-IDF_RF	234796	217100	102578	39875	92343	0.4725	0.3933	0.1698
3-gram_TF-IDF_XGBoost	265478	217100	119867	48057	97554	0.5521	0.3675	0.1810
Doc2vec_SVM	250164	217100	117687	47895	84582	0.5421	0.3381	0.1915
Doc2vec_RF	302546	217100	116598	48965	136983	0.5371	0.4528	0.1618
Doc2vec_XGBoost	348951	217100	124263	53248	171440	0.5724	0.4913	0.1526
Doc2vec_CNN	458964	217100	169542	91458	197964	0.7809	0.4313	0.1993
Doc2vec_LSTM	486525	217100	189981	90259	206285	0.8751	0.4240	0.1855
Doc2vec_Bi-LSTM	589647	217100	200143	99653	289851	<b>0.9219</b>	0.4916	0.1690
Character_level_CNN	653287	217100	198756	180145	274386	<b>0.9155</b>	0.4200	<b>0.2758</b>
Character_level_LSTM	325648	217100	165478	55641	104529	0.7622	0.3210	0.1709
Character_level_Bi-LSTM	295876	217100	187569	31542	76765	0.8640	<b>0.2594</b>	0.1066
AT-Bi-LSTM	428270	217100	206809	110974	110487	<b>0.9526</b>	<b>0.2580</b>	<b>0.2591</b>

#### 5.4 Malicious Pattern Mining

As mentioned before, one of the highlights of AT-Bi-LSTM is that it can automatically identify the malicious part of each traffic request according to attention weight vector. This is also the difference between this model and the traditional fingerprint extraction methods [14, 31]. As described in Sect. 4.2, we first cluster the malicious entries detected by AT-Bi-LSTM but not detected by the “Rule Engine”, then we perform pattern mining for each cluster.

Given a cluster that consists of several traffic of cross-site scripting attack (see Fig. 7). We can get keywords for each entry according to its attention weight vector.

```

/cgi-bin/wa.exe?SHOWTPL=<script>alert(/openvas-xss-test/)</script>
/webContent/fck/wlkt.htm?xss_test"></textarea></script><script>prompt(42873);</script>
adminDirHand="></script><script>alert(1);</script>
itemid=1527"></textarea></script><scr<script>ipt>alert(2014)</scr<script>ipt>
/survey/list.jsp?s_id=f65b</textarea></script><a href=//eye.hihop.cn/>webscan</a>
    
```

Fig. 7. Traffic samples of cross-site scripting attack.

For instance, the first traffic entry in Fig. 7 is “/cgi-bin/wa.exe?<EOS>SHOWTPL=<script> alert(/openvas-xss-test/)</script>”. The visualization of its attention vector is shown in Fig. 8. The color depth corresponds to the attention weight  $\alpha_t$  (Eq. 3). The darker the color, the greater the weight value. Obviously, top 10 keywords for this entry are {‘:’, ‘exe’, ‘<’, ‘script’, ‘>’, ‘alert’, ‘)’, ‘/’, ‘openvas’, ‘xss’}. Based on this string pattern, we can generate a rule that identifies such malicious traffic.

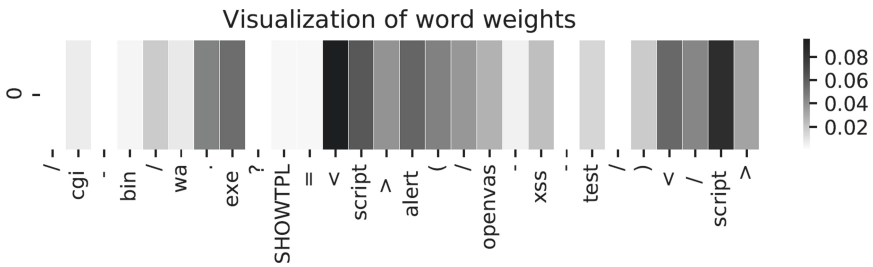


Fig. 8. Visualization of attention.

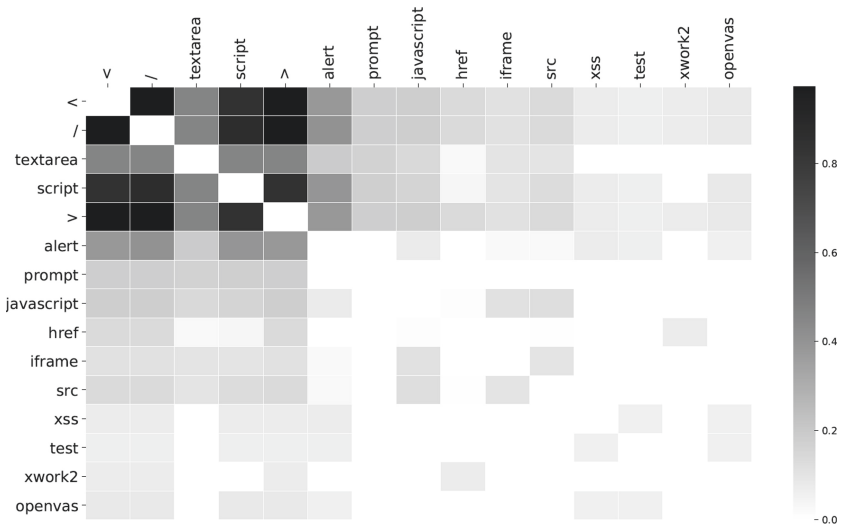


Fig. 9. Visualization of pattern mining.

To further illustrate the performance of the proposed model in malicious pattern mining, we visualize the pattern mining results of this cluster (see Fig. 9). The darker the color of the square is, the more times the words appear together. Hence, the pattern of these traffic can be denoted as {“<”, “/”, “script”, “>”, “textarea”, “prompt”, “javascript”, “alert”, “iframe”, “src”, “href”}.

## 6 Conclusion

This paper presents DeepHTTP, a general-purpose framework for HTTP traffic anomaly detection and pattern mining based on deep neural networks. We build AT-Bi-LSTM, a deep neural networks model utilizing Bidirectional Long Short-Term Memory (Bi-LSTM), which can enable effective anomaly diagnosis. Besides, we design a novel method that can extract the structural characteristics of HTTP traffic. DeepHTTP learns content feature and structure feature of traffic automatically and unearths critical section of input data. It performs detection at the single traffic level and then performs pattern mining at the cluster level. The intermediate output including attention hidden state and the attentional weight vector can be applied to clustering and pattern mining, respectively. Meanwhile, by incorporating user feedback, DeepHTTP supports database updates and model iteration. Experiments on a large number of HTTP traffic entries have clearly demonstrated the superior effectiveness of DeepHTTP compared with previous methods.

Future works include but are not limited to incorporating other types of deep neural networks into DeepHTTP to test their efficiency. Besides, improving the ability of the model to detect unknown malicious traffic is something we need to further study in the future. With the increasing popularity of encrypted traffic, the detection of encrypted traffic attacks is also our future research direction.

**Acknowledgement.** This work was supported by the National Key R&D Program China under Grant 2018YFB0804701.

## References

1. China Internet Network Security Report. <https://www.cert.org.cn/publish/main/upload/File/2019-year.pdf>
2. Estévez-Tapiador, J.M., García-Teodoro, P., et al.: Measuring normality in HTTP traffic for anomaly-based intrusion detection. *Comput. Netw.* **45**(2), 175–193 (2004)
3. Jamdagni, A., Tan, Z., Nanda, P., He, X., Liu, R.P.: Intrusion detection using GSAD model for HTTP traffic on web services. In: *Proceedings of the 6th International Wireless Communications and Mobile Computing Conference*, pp. 1193–1197. ACM (2010)
4. Tombini, E., Debar, H., Mé, L., Ducassé, M.: A serial combination of anomaly and misuse IDSes applied to HTTP traffic. In: *Proceedings of the 20th Annual Computer Security Applications Conference (ACSAC)*, pp. 428–437. IEEE Computer Society, Washington, DC, USA (2004)
5. w3techs Homepage. <https://w3techs.com/technologies/details/ce-httpsdefault>. Accessed 21 Jan 2020

6. Tony Messer's blog published in Pickaweb. <https://www.pickaweb.co.uk/blog/local-business-seo-stats-chart-and-infographic/>
7. Le, H., et al.: URLNet: learning a URL representation with deep learning for malicious URL detection. arXiv preprint [arXiv:1802.03162](https://arxiv.org/abs/1802.03162) (2018)
8. Saxe, J., Berlin, K.: eXpose: a character-level convolutional neural network with embeddings for detecting malicious URLs, file paths and registry keys. arXiv preprint [arXiv:1702.08568](https://arxiv.org/abs/1702.08568) (2017)
9. Sundermeyer, M., Schlüter, R., Ney, H.: LSTM neural networks for language modeling. In: Proceedings of INTERSPEECH, pp. 601–608 (2012)
10. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. arXiv preprint [arXiv:1409.0473](https://arxiv.org/abs/1409.0473) (2014)
11. Lakhina, A., Crovella, M., Diot, C.: Mining anomalies using traffic feature distributions. SIGCOMM Comput. Commun. Rev. **35**(4), 217–228 (2005)
12. Samant, A., Adeli, H.: Feature extraction for traffic incident detection using wavelet transform and linear discriminant analysis. Comput.-Aided Civil Infrastruct. Eng. **15**(4), 241–250 (2010)
13. Swarnkar, M., Hubballi, N.: OCPAD: one class Naive Bayes classifier for payload based anomaly detection. Expert Syst. Appl. **64**, 330–339 (2016)
14. Nelms, T., Perdisci, R., Ahamad, M.: Execscent: mining for new C&C domains in live networks with adaptive control protocol templates. Presented as part of the 22nd {USENIX} Security Symposium ({USENIX} Security 2013), pp. 589–604 (2013)
15. Ma, J., Saul, L.K., Savage, S., et al.: Beyond blacklists: learning to detect malicious web sites from suspicious URLs. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1245–1254. ACM (2009)
16. Ma, J., Saul, L.K., Savage, S., et al.: Identifying suspicious URLs: an application of large-scale online learning. In: Proceedings of the 26th Annual International Conference on Machine Learning, pp. 681–688. ACM (2009)
17. Bortolameotti, R., van Ede, T., et al.: DECANter: DETECTION of anomalous outbounD HTTP TRaffic by passive application fingerprinting. In: Proceedings of the 33rd Annual Computer Security Applications Conference (ACSAC 2017), pp. 373–386. ACM, New York (2017)
18. Juvonen, A., Sipola, T., Hämäläinen, T.: Online anomaly detection using dimensionality reduction techniques for HTTP log analysis. Comput. Netw. **91**, 46–56 (2015)
19. Ringberg, H., Soule, A., et al.: Sensitivity of PCA for traffic anomaly detection. SIGMETRICS Perform. Eval. Rev. **35**(1), 109–120 (2007)
20. El-Alfy, E.S.M., Al-Obeidat, F.N.: A multicriterion fuzzy classification method with greedy attribute selection for anomaly-based intrusion detection. Procedia Comput. Sci. **34**, 55–62 (2014)
21. Estévez-Tapiador, J.M., et al.: Measuring normality in HTTP traffic for anomaly-based intrusion detection. Comput. Netw. **45**(2), 175–193 (2004)
22. Mahoney, M.V., Chan, P.K.: Learning rules for anomaly detection of hostile network traffic. In: Third IEEE International Conference on Data Mining, pp. 601–604 (2003)
23. Radford, B.J., Apolonio, L.M., et al.: Network Traffic Anomaly Detection Using Recurrent Neural Networks. CoRR abs/1803.10769 (2018)
24. Du, M., Li, F., et al.: DeepLog: anomaly detection and diagnosis from system logs through deep learning. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS 2017), pp. 1285–1298. ACM, New York (2017)
25. Erfani, S.M., et al.: High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. Pattern Recogn. **58**, 121–134 (2016)
26. Chiou, T.-W., Tsai, S.-C., Lin, Y.-B.: Network security management with traffic pattern clustering. Soft Comput. **18**(9), 1757–1770 (2014)

27. Le, T.T., Millaire, A., Asseman, P., De, G.P., ThAcAry, C., Ducloux, G.: Novel approach for network traffic pattern analysis using clustering based collective anomaly detection. *Ann. Data Sci.* **2**(1), 111–130 (2015)
28. Ahmed, M., Mahmood, A.N.: Network traffic pattern analysis using improved information theoretic co-clustering based collective anomaly detection. In: Tian, J., Jing, J., Srivatsa, M. (eds.) *SecureComm 2014. LNICST*, vol. 153, pp. 204–219. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-23802-9\\_17](https://doi.org/10.1007/978-3-319-23802-9_17)
29. Ren, J.T., Ou, X.L., Zhang, Y., Hu, D.C.: Research on network-level traffic pattern recognition. In: *Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems*, pp. 500–504 (2002)
30. Paredes-Oliva, I., Castell-Uroz, I., Barlet-Ros, P., Dimitropoulos, X., SolA-Pareta, J.: Practical anomaly detection based on classifying frequent traffic patterns. In: *2012 Proceedings IEEE INFOCOM Workshops*, pp. 49–54 (2012)
31. Rafique, M.Z., Caballero, J.: FIRMA: malware clustering and network signature generation with mixed network behaviors. In: Stolfo, S.J., Stavrou, A., Wright, C.V. (eds.) *RAID 2013. LNCS*, vol. 8145, pp. 144–163. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-41284-4\\_8](https://doi.org/10.1007/978-3-642-41284-4_8)
32. Nelms, T., Perdisci, R., et al.: ExecScent: mining for new C&C domains in live networks with adaptive control protocol templates. In: *Presented as part of the 22nd USENIX Security Symposium (2013)*, pp. 589–604. USENIX, Washington, D.C. (2013)
33. Xu, K., Ba, J., Kiros, R., et al.: Show, attend and tell: neural image caption generation with visual attention. In: *International Conference on Machine Learning*, pp. 2048–2057 (2015)
34. Chorowski, J.K., Bahdanau, D., et al.: Attention-based models for speech recognition. In: *Advances in Neural Information Processing Systems*, pp. 577–585 (2015)
35. Ma, F., Chitta, R., Zhou, J., et al.: Dipole: diagnosis prediction in healthcare via attention-based bidirectional recurrent neural networks. In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1903–1911. ACM (2017)
36. Luong, M.T., Pham, H., Manning, C.D.: Effective approaches to attention-based neural machine translation. *arXiv preprint arXiv:1508.04025* (2015)
37. Zhou, P., Shi, W., Tian, J., et al.: Attention-based bidirectional long short-term memory networks for relation classification. In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (vol. 2: Short Papers)*, pp. 207–212 (2016)
38. Ren, P., Chen, Z., Ren, Z., et al.: Leveraging contextual sentence relations for extractive summarization using a neural attention model. In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 95–104 (2017)
39. Zhang, H., et al.: Webshell traffic detection with character-level features based on deep learning. *IEEE Access* **6**, 75268–75277 (2018)
40. Kali Official Website. <https://www.kali.org/>. Accessed 21 Jan 2020
41. Chinotec Technologies Company: Paros - for web application security assessment. <http://www.parosproxy.org/index.shtml>. Accessed 21 Jan 2020
42. Riancho, A.: Web Application Attack and Audit Framework. <http://w3af.sourceforge.net>. Accessed 21 Jan 2020
43. Damashek, M.: Gauging similarity with n-grams: language-independent categorization of text. *Science* **267**(5199), 843–848 (1995)
44. Kloft, M., Brefeld, U., Düessel, P., Gehl, C., Laskov, P.: Automatic feature selection for anomaly detection. In: *Proceedings of the 1st ACM Workshop on Workshop on AISec (AISec 2008)*, pp. 71–76. ACM, New York (2008)
45. Wang, K., Stolfo, S.J.: Anomalous payload-based network intrusion detection. In: Jonsson, E., Valdes, A., Almgren, M. (eds.) *RAID 2004. LNCS*, vol. 3224, pp. 203–222. Springer, Heidelberg (2004). [https://doi.org/10.1007/978-3-540-30143-1\\_11](https://doi.org/10.1007/978-3-540-30143-1_11)

46. Zolotukhin, M., Hämmäläinen, T., Kokkonen, T., Siltanen, J.: Analysis of HTTP requests for anomaly detection of web attacks. In: 2014 IEEE 12th International Conference on Dependable, Autonomic and Secure Computing, pp. 406–411 (2014)
47. hashID, a tool written in Python 3. <https://github.com/psypana/hashID>
48. Joachims, T.: A probabilistic analysis of the Rocchio algorithm with TFIDF for text categorization. In: Proceedings of International Conference on Machine Learning, pp. 143–151 (1996)
49. Ester, M., Kriegel, H.P., Sander, J., et al.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: KDD, vol. 96, no. 34, pp. 226–231 (1996)
50. Le, Q., Mikolov, T.: Distributed representations of sentences and documents. In: International Conference on Machine Learning (2014)
51. Cortes, C., Vapnik, V.: Support-vector networks. *Mach. Learn.* **20**(3), 273–297 (1995)
52. Ho, T.K.: Random decision forests. In: Proceedings of 3rd International Conference on Document Analysis and Recognition, vol. 1. IEEE (1995)
53. Chen, T., Guestrin, C.: Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM (2016)
54. Lawrence, S., Giles, C.L., Tsoi, A.C., Back, A.D.: Face recognition: a convolutional neural-network approach. *IEEE Trans. Neural Netw.* **8**(1), 98–113 (1997)
55. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. *Nature* **323**(6088), 533–536 (1986)
56. Schuster, M., Paliwal, K.K.: Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* **45**(11), 2673–2681 (1997)
57. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* **9**(8), 1735–1780 (1997)
58. Lakhina, A., Crovella, M., Diot, C.: Characterization of network-wide anomalies in traffic flows. In: Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement (IMC 2004), pp. 201–206. ACM, New York (2004)
59. Wang, K., Parekh, J.J., Stolfo, S.J.: Anagram: a content anomaly detector resistant to mimicry attack. In: Zamboni, D., Kruegel, C. (eds.) RAID 2006. LNCS, vol. 4219, pp. 226–248. Springer, Heidelberg (2006). [https://doi.org/10.1007/11856214\\_12](https://doi.org/10.1007/11856214_12)
60. Iglesias, F., Zseby, T.: Analysis of network traffic features for anomaly detection. *Mach. Learn.* **101**(1–3), 59–84 (2015)

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





# **Social Network Security and Privacy**



# A Label Propagation Based User Locations Prediction Algorithm in Social Network

Huan Ma<sup>1</sup> and Wei Wang<sup>2</sup>(✉)

<sup>1</sup> National Computer Network Emergency Response Technical Team/Coordination Center of China, Beijing, China

<sup>2</sup> Harbin Engineering University, No. 145, Nantong Street, Harbin, China  
w\_wei@hrbeu.edu.cn

**Abstract.** Network community detection is an important service provided by social networks, and social network user location can greatly improve the quality of community detection. Label propagation is one of the main methods to realize the user location prediction. The traditional label propagation algorithm has the problems including “location label countercurrent” and the update randomness of node location label, which seriously affects the accuracy of user location prediction. In this paper, a new location prediction algorithm for social networks based on improved label propagation algorithm is proposed. By computing the K-hop public neighbor of any two point in the social network graph, the nodes with the maximal similarity and their K-hopping neighbors are merged to constitute the initial label propagation set. The degree of nodes not in the initial set are calculated. The node location labels are updated asynchronously is adopted during the iterative process, and the node with the largest degree is selected to update the location label. The improvement proposed solves the “location label countercurrent” and reduces location label updating randomness. The experimental results show that the proposed algorithm improves the accuracy of position prediction and reduces the time cost compared with the traditional algorithms.

**Keywords:** Social network · Location prediction · Label propagation · Social relationships · User location probability

## 1 Introduction

As social networks with location-based information are increasingly popular, the users' location in social network attracts more attention than before. Location information can help to shorten the gap between the virtual and the real world, such as monitoring residents' public health problems through online network [1], recommending local activities or attractions to tourists [2, 3], determining the emergency situation and even the location of the disaster and so on [4–6]. In addition, users' offline activity area and trajectory can also be analyzed through their locations in social networks. Due to the increasing awareness of privacy protection, people will cautiously submit their personal location information or set the visibility of the location of the message in social networks,

which make it difficult to acquire their real location information. Therefore, how to accurately predict the actual location information of social network users is an important and meaningful research question.

This paper proposes location prediction algorithm for social network users based on label propagation, which solves the following two key problems:

- (1) The accuracy of the traditional label propagation algorithm is not high in the user location prediction, and “countercurrent” phenomenon will appear in the iterative process, which will lead to the increase of the time overhead.
- (2) Improve the accuracy of social network users’ location prediction by using their offline activity location.

## 2 Related Work

There are three scenarios for user location prediction in social networks, such as user’s frequent location prediction, prediction of the location of messages posted on the user’s social network, and forecasts of the locations mentioned in messages. The main methods of location prediction include location prediction based on the content of message published by users, user friend relationships, and so on.

Laere et al. chose two types of local vocabulary and extracted valid words to predict the location of users [7]. Ren [8] and Han et al. [9] were inspired by the frequency of reverse documents, using the reverse position frequency (ILF) and the reverse city frequency (ICF) to select the position of the vocabulary, they assumed that the location vocabulary should be distributed in fewer locations, but with large ILF and ICF values. Mahmud et al. [10] applied some column heuristics to select local vocabulary. Cheng [1] makes the position word distribution conform to the spatial change model proposed by the Backstrom [11], secondly they make local or non-local mark on 19,178 dictionary words, and use the Labeled Vocabulary Training classification model to discriminate all words in the tweet dataset.

Backstrom [12] established probability models through physical distances between users to express the possibility of relationships between users, which has no effect on the position prediction of friends considering different degrees of tightness. Kongl [13] on the basis of Backstrom work by adding the weight of the edge to predict the user’s position, where the weight of the edge is determined by a social tight coefficient. Li [14] considered the location of user neighbors, and captures the information of users’ neighbors that intuitively consider the location of users. The user location is allocated randomly, then the user’s location is iteratively updated from the user’s neighbors and the location name mentioned, and then the parameters in the update are improved by measuring the prediction error of the known location of the user. Davis Jr et al. [15] thought that the most frequent user’ locations that appear in the user’s social network as a basis for predicting their location. Jurgens et al. [16] extend the concept of location prediction into location label propagation, which is made by the location of the label space to explain the location of label propagation, they think that the position of the user through the iterative process that many times.

Li et al. [17] thought that the literature assume the user has only one home location is a defect, they think that users should have the relationship with a number of positions, so they have defined the location information of a user and user set as the set of locations, and these users about the system is not only a geographical location the range is not a point, is not a temporary and user related position, but a long-term position, so they set up a MLP in the paper (Multiple Location Profiling Model) to establish a model containing a plurality of position information of the position of archives to the user, and this model is to the location file according to the target user relationships and their tweets content released.

The label propagation algorithm can effectively deal with large data sets, so in this paper, we are in the position of the user prediction based on label propagation algorithm, but with the label propagation algorithm in-depth study, we found that the label propagation algorithm will position the label “countercurrent” label update and node location is random, this algorithm cannot guarantee the accuracy of prediction of the position of the user, in order to improve the accuracy of location prediction algorithm and reduce the time overhead, this paper pro-poses a label propagation based on user location prediction algorithm (Label Propagation Algorithm-Location Prediction, LPA-LP).

### 3 Related Concept and Problem Definition

**Definition 1** Social Network. A social network can be represent by a graph  $G = (V, E, A)$ , where  $V$  represents the collection of the users who are in the social network, and  $n = |V|$ .  $E$  represents the collection of the relationship between users and  $m = |E|$ , and  $A$  represents the collection of the activities and  $a = |A|$ . Beyond that,  $L$  represents the set of locations, including users’ locations and activities’ locations, and  $n_l = |L|$ ,  $U_0$  is the set of the users whose locations are known, on the contrary,  $U_n$  is the set of users whose locations are unknown.

**Definition 2** Shortest Path Length. It refers to the shortest path between the two nodes  $i$  and  $j$  in the social network graph. It means the minimum number of paths through the node  $i$  to the node  $j$ . It can be used  $d(i, j)$  to represent the shortest path length between two nodes.

**Definition 3**  $K$ -Hopping Neighbors. It means that the user to its neighbor needs a  $k$  hopping to achieve, that is to say, the shortest path length of the two node is  $k$ .

**Definition 4**  $K$ -Hopping Public Neighbors.  $G = (V, E, A)$  is a social network diagram, where  $V$  represents the user set in the graph,  $E = (v_i, v_j, w_{ij})$  represents the set of relations between the user nodes with weights,  $w_{ij}$  represents the weight of the edges between nodes. The  $k$ -hopping public neighbors set of the nodes is defined as follows:

$$\Gamma_k(v_i, v_j) = \{v | d(v_i, v) = d(v_j, v) = k\} = \Gamma(v_i, k) \cap \Gamma(v_j, k), k \geq 1 \quad (1)$$

In the formula (1),  $\Gamma(v_i, k)$  represents the set of  $k$ -public neighbor of node  $v_i$ , and represents the set of node  $v_j$ , represents the set of  $k$ -public neighbor between  $v_i$  and  $v_j$ .

**Definition 5.** Similarity of  $k$ -hopping public neighbors. The value of  $k$  is determined by the network itself, it can be defined on formula (2).

$$\bar{k} = \frac{\sum_{i \neq j} k_{\max} |\Gamma(i) \cap \Gamma(j)|}{|V|} \quad (2)$$

In the formula (2),  $k_{\max} |\Gamma(i) \cap \Gamma(j)|$  represents the max public neighbor hops between two nodes. The  $k$  value in the network refers to the average of any two nodes in the network. The similarity of the two node  $k$ -hopping public neighbors is defined by formula (3).

$$S(v_i, v_j) = \frac{|\Gamma(v_i, k) \cap \Gamma(v_j, k)|}{|\Gamma(v_i, k) \cup \Gamma(v_j, k)|}, k \geq 1 \quad (3)$$

**Definition 6** Similarity of Nodes. It means denominator size of the similarity between the  $k$ -hopping public neighbors between nodes subtracts the two nodes. It can be defined by formula (4).

$$\gamma = \frac{|\Gamma(v_i, k) \cap \Gamma(v_j, k)|}{|\Gamma(v_i, k) \cup \Gamma(v_j, k)| - 2}, k \geq 1 \quad (4)$$

**Definition 7.** The max degree between nodes and users set. If the user is divided into different sets  $L_1, L_2, \dots, L_e$  according to their locations, nodes are set up by users who are not labeled as location labels. The max degree of users divided into different sets according to their location is the degree and the maximum of some nodes in the nodes. It can be defined by formula (5).

$$d(v_i, L_i) = \max\{d(v_i, L_1), d(v_i, L_2), \dots, d(v_i, L_e)\} \quad (5)$$

**Definition 8**  $K$ -Hopping Weight. We believe that the most important impact on user location is its 1 hop neighbors. Moreover, the offline location of users also has a great impact on user location, and its weight can also be set to 1. For  $k > 1$ , when setting the weight of the edge, it will be attenuated according to the speed of  $1/5$ , that is, the weight of the edge of the 1 hop neighbor is 1, the weight of the 2 neighbors is 0.8, and so on.

Now given the location prediction problem definition: In the social network  $G$ , the unknown location information of the user  $u$ , according to the location information and the users of their  $k$ -hopping neighbors, to predict the unknown location information of the user  $u$  in the prediction of the probability of the position of  $L$ .

## 4 Label Propagation Based User Location Prediction Algorithm

In this section, a correlation algorithm for location prediction for users of unknown location information in social networks is proposed. This paper proposed a location prediction algorithm based on label propagation (Label Propagation Algorithm-Location Prediction, LPA-LP), the algorithm is mainly divided into two parts, one part is to run before the label propagation algorithm of data preprocessing algorithm, the other part is the use of label propagation of location prediction algorithm.

---

### Algorithm 1 Data Preprocessing

---

Input: G-dataset, L-location label set

Output: pre-processed user sets C

Initialization: set C is empty

while  $S(v_i, v_j) \geq \gamma$  do

  for each  $v \in U_n$  do

    if  $\Gamma(v) \neq \emptyset$  do

      set  $k = 1$

$\Gamma(v)_i \leftarrow BFS(v)$

$k = i$

    end for

      get  $\Gamma_k(v_i, v_j)$  for every nodes based on formula(1)

      select all nodes with  $k_{max}$

      calculate the similarity of k-neighbors based on formula(3)

      choose the k-neighbors as the begin set

      update the nodes' label in the begin set

      calculate node similarity based on formula(4)

    end while

---

Algorithm 1 is pretreated before running the label propagation algorithm to initialize the data set, according to the Definition 5, the node with its maximum similarity and the  $k$  hop neighbor as the set of starting processing for the user location prediction, and according to the known label to the data in the collection of the label, which is in order to be able to quickly and accurately using the label propagation algorithm for unknown location information in a social network user node location prediction. After preprocessing the data set, location prediction algorithm based on label propagation can be used to predict the location of users who have not tagged location labels in the processed data set. Algorithm 2 gives a description of the location prediction algorithm (LPA-LP) based on the label propagation.

---

**Algorithm2LPA-LPalgorithm description**


---

Input:pre-processed user sets  $C_i$   
 Output:estimated user location, E  
 Initialization:set E is empty  
   choose the begin set  $C_i$   
   while the nodes label change do  
     for each node do  
       sort the nodes as the k to every set  $C_i$  to made F  
       end for  
       for  $v \in F$  do  
          $F_v(t) = f(F_{v_{i1}}(t), \dots, F_{v_{im}}(t), F_{v_{m+1}}(t-1), \dots, F_{v_m}(t-1))$   
         update the nodes' label in the begin set build E'  
       end for  
       E=E'  
     end while

---

In Algorithm 2 location prediction algorithm based on label propagation in the iterative process of user location labels are updated, and the location information of the user location information of neighbors and user participation in the offline activities are taken into account, which significantly improves the prediction accuracy of the locations of users, and in the operation of label propagation algorithm for data sets are preliminary the treatment improve the performance of the label propagation algorithm of user location prediction algorithm, the following will be proved by experiments.

## 5 Experiment Result and Analysis

In this section, we will analyze the experimental results, the experimental results are divided into two parts, one part is the results of algorithm time overhead and the other is the accuracy of user locations prediction algorithm.

### 5.1 Data Set Description

In this paper, we use the dataset is NLPiR microblogging corpus. We extracted several datasets from the dataset. In order to compare the accuracy of the improved algorithm for user location prediction and improve the execution efficiency of the algorithm, we extract different scale datasets from the data set to compare the experimental results. The detail of our data sets are described in Table 1.

### 5.2 Experimental Results Analysis

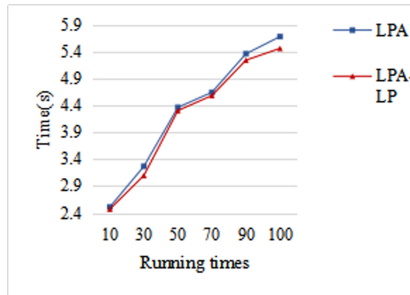
The location prediction algorithm based on the label propagation (LPA-LP) is an improvement on the preprocessing of the data set and the selection strategy of the location label in the iterative process. It can avoid the “countercurrent” phenomenon of the position label and reduce the randomness to update the location tag, and improve the efficiency and the accuracy of the prediction. The whole experiment is divided into two

**Table 1.** Data sets description

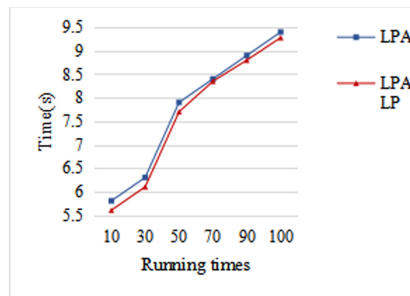
Dataset	Users number	Relations number	Activities number
A	2748	12121	452
B	4025	61881	983
C	5194	77567	2392
D	9940	107979	4938

parts. The first part is using label propagation algorithm to predict user location on these four datasets of different sizes. The second part is using LPA-LP algorithm to predict location on four different scale datasets.

In the process of user location prediction, probabilistic LPA algorithm and LPA-LP algorithm with random or update the node label to a certain extent, the running times of the two algorithms may produce different results, so the choice between the four data sets of different size on the running times of experimental results for the 10, 30, 50, 70, 100 and mean value. The time required for the experiment to run on different scale data sets is shown in Fig. 1, 2, 3 and 4.



**Fig. 1.** Time overhead comparison with dataset A



**Fig. 2.** Time overhead comparison with dataset B



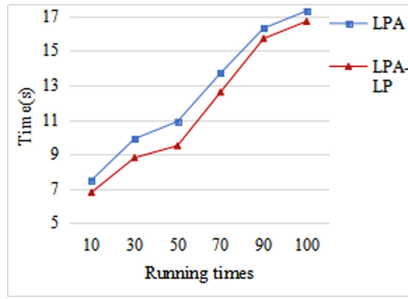


Fig. 3. Time overhead comparison with dataset C

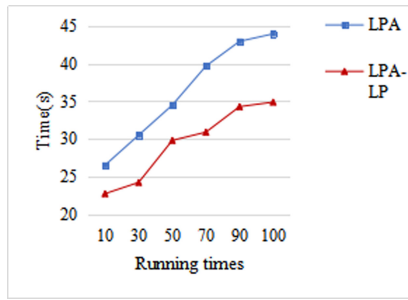


Fig. 4. Time overhead comparison with dataset D

From these four figures, we can know that the running time of different dataset is similar between the improved algorithm LPA-LP and the algorithm LPA when the dataset have less than 5000 nodes, when the nodes are more than 9000 in dataset, we can see that the running time of the improved algorithm LPA-LP is obviously less than the algorithm LPA. It shows that the LPA-LP algorithm can be effectively applied to large-scale data sets.

In addition to comparing the running time of the algorithm, it is necessary to compare the accuracy of the algorithm. The results of the experiment are shown in Table 2.

Table 2. Algorithm accuracy comparison

Dataset	Accuracy	
	LPA	LPA-LP
A	59.3%	64.4%
B	62.2%	67.3%
C	66.5%	69.5%
D	70.7%	78.4%

## 6 Conclusion

This paper proposes a location prediction algorithm for social network users based on label propagation. The algorithm first obtains  $k$ -hop public neighbors at any two points in the social network graph, and uses the node with the largest similarity and its  $k$ -hop neighbors as the initial set of label propagation, and calculates the degree of the node to these sets. In each iteration, the node adopts the strategy of asynchronous update, and selects the node with the highest degree to update the position label, so as to avoid the “countercurrent” phenomenon of the position label and reduce the possibility of randomly updating the position label. Relevant experiments show that the algorithm proposed in this paper improves the accuracy of user location prediction and reduces the time cost of the algorithm.

## References

1. Cheng, Z., Caverlee, J., Lee, K.: You are where you tweet: a content-based approach to geo-locating Twitter users. In: *The 19th ACM Conference on Information and Knowledge Management*, pp. 759–768. ACM, Toronto (2010)
2. Yuan, Q., Cong, G., Ma, Z., et al.: Who, where, when and what: discover spatio-temporal topics for Twitter users. In: *The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 605–613. ACM, Chicago (2013)
3. Noulas, A., Scellato, S., Lathia, N., et al.: Mining user mobility features for next place prediction in location-based services. In: *13th Industrial Conference on Data Mining*, pp. 1038–1043, IEEE, New York (2013)
4. Rakesh, V., Reddy, C.K., Singh, D., et al.: Location-specific tweet detection and topic summarization in Twitter. In: *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 1441–1444. ACM, Niagara (2013)
5. Ao, J., Zhang, P., Cao, Y.: Estimating the locations of emergency events from Twitter streams. *Procedia Comput. Sci.* **31**, 731–739 (2014)
6. Lingad, J., Karimi, S., Yin, J.: Location extraction from disaster-related microblogs. In: *Proceedings of the 22nd International Conference on World Wide Web*, pp. 1017–1020. ACM, Rio de Janeiro (2013)
7. Van Laere, O., Quinn, J., Schockaert, S., et al.: Spatially aware term selection for geotagging. *IEEE Trans. Knowl. Data Eng.* **26**(1), 221–234 (2014)
8. Ren, K., Zhang, S., Lin, H.: Where are you settling down: geo-locating Twitter users based on tweets and social networks. In: Hou, Y., Nie, J.-Y., Sun, L., Wang, B., Zhang, P. (eds.) *AIRS 2012. LNCS*, vol. 7675, pp. 150–161. Springer, Heidelberg (2012). [https://doi.org/10.1007/978-3-642-35341-3\\_13](https://doi.org/10.1007/978-3-642-35341-3_13)
9. Han, B., Cook, P., Baldwin, T.: Geolocation prediction in social media data by finding location indicative words. In: *24th International Conference on Computational Linguistics*, pp. 1045–1062. ACM, Mumbai (2012)
10. Mahmud, J., Nichols, J., Drews, C.: Where is this tweet from? Inferring home locations of Twitter users. In: *Sixth International AAAI Conference on Weblogs and Social Media*, pp. 73–77. AAAI, Dublin (2012)
11. Backstrom, L., Kleinberg, J., Kumar, R., et al.: Spatial variation in search engine queries. In: *Proceedings of the 17th International Conference on World Wide Web*, pp. 357–366. ACM, Beijing (2008)

12. Backstrom, L., Sun, E., Marlow, C.: Find me if you can: improving geographical prediction with social and spatial proximity. In: Proceedings of the 19th International Conference on World Wide Web, pp. 61–70. ACM, North Carolina (2010)
13. Kong, L., Liu, Z., Huang, Y.: SPOT: locating social media users based on social network context. *Proc. VLDB Endow.* **7**(13), 1681–1684 (2014)
14. Li, R., Wang, S., Deng, H., Wang, R., Chang, K.C.: Towards social user profiling: unified and discriminative influence model for inferring home locations. In: The 18th International ACM SIGKDD Conference, pp. 1023–1031. ACM, Beijing (2012)
15. Davis, Jr C., Pappa, G., de Oliveira, D., de L Arcanjo, F.: Inferring the location of twitter messages based on user relationships. *Trans. GIS* **15**(6), 735–751 (2011)
16. Jurgens, D.: That’s what friends are for: inferring location in online social media platforms based on social relationships. In: Seventh International AAAI Conference on Weblogs and Social Media, pp. 237–240. AAAI, Massachusetts (2013)
17. Li, R., Wang, S., Chang, C.: Multiple location profiling for users and relationships from social network and content. *Proc. VLDB Endow.* **5**(11), 1603–1614 (2012)

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





# Personalized Differentially Private Location Collection Method with Adaptive GPS Discretization

Huichuan Liu, Yong Zeng<sup>(✉)</sup>, Jiale Liu, Zhihong Liu, Jianfeng Ma, and Xiaoyan Zhu

Xidian University, Xi'an 710126, Shaanxi, China

hc\_liu@stu.xidian.edu.cn, {yzeng, liuzhihong, jfma, xyzhu}@mail.xidian.edu.cn, liujialehenu@163.com

**Abstract.** In recent years, with the development of mobile terminals, geographic location has attracted the attention of many researchers because of its convenience in collection and its ability to reflect user profile. To protect user privacy, researchers have adopted local differential privacy in data collection process. However, most existing methods assume that location has already been discretized, which we found, if not done carefully, may introduces huge noise, lowering collected result utility. Thus in this paper, we design a differentially private location division module that could automatically discretize locations according to access density of each region. However, as the size of discretized regions may be large, if directly applying existing local differential privacy based attribute method, the overall utility of collected results may be completely destroyed. Thus, we further improve the optimized binary local hash method, based on personalized differential privacy, to collect user visit frequency of each discretized region. This solution improve the accuracy of the collected results while satisfying the privacy of the user's geographic location. Through experiments on synthetic and real data sets, this paper proves that the proposed method achieves higher accuracy than the best known method under the same privacy budget.

**Keywords:** Local differential privacy · Geographical location · Privacy security

## 1 Introduction

With the development of mobile Internet technology, various mobile platforms such as mobile phones, tablets, smart watches and other devices have brought many conveniences and joys to people's lives. Sensors such as Accelerometer, GPS, Gyroscope and Magnetometer could capture information about the user's surroundings and provide a richer and more interesting interface for human-computer interaction. Among them, geographic location sensing has been widely equipped on smart devices. As a form of information that could reflect the user's trajectory and lifestyle, it is widely used by major application service providers in the recommendation system to provide users with personalized advertisement.

However, due to the sensitivity of the geographic location itself, and the fact that background applications may collect user data at any time, the uploaded user trajectory data may reflect the user's sensitive information, such as the user's income, beliefs, daily habits, illness and other information [1]. Users may dislike their private data that could expose their activity being analyzed. Besides that, improper data management may result in the disclosure of user privacy data, thereby causing legal problems.

In order to ensure privacy of user uploaded data in analysis process, many researches have been conducted and most differential privacy based methods for solving privately analysis can mainly be divided into two categories. The first category [2–6] is to disturb the collected data before data sharing and publishing. This type mainly uses differential privacy settings. The other category [7–9] mainly focuses on the data collection process and disturbs the data before users upload their private data. Among them, the former category couldn't provide protection against database intrusions or application service providers' threats to user privacy. In reality, the database interface provided by the server is very likely to have problems. For example, in March 2018, a security breach on Facebook enables third-party application software to download unpublished private photos of users without permission, affecting up to 6.8 million users. It is conceivable that with the expansion of business and the growth of code volume, security vulnerabilities are inevitable. The privacy protection of the second category, which is based local differential privacy model, can also essentially prevent third-party analysts from threatening privacy, and it can also prevent the inappropriate use of user privacy data by the enterprise itself, so it has a stronger privacy protection. In this paper, we follow the second category research line and adopt a variant of local differential privacy as our privacy model.

Most existing attribute collection methods [10–12] assume that the user attributes to be collected are discrete, which means, for GPS data, the continuous GPS signal must be quantified before being applied to an existing collection method. But in fact, due to the non-uniformity of the geographical location itself, completely uniform density quantization without any knowledge of the whole user density distribution, will cause very low signal-to-noise ratio. In addition, in order to provide more fine-grained geographic location collection, the number of quantized geographic location areas is large, so local differential privacy based location collection methods would cause overwhelming noise, completely destroying the utility of the data collection results.

This paper proposes a new geographic location collection method. The method is divided into two modules, each of which takes exclusive user sets as input. The first module is a location division module, which is responsible for sending location-related query requests to users in the corresponding user set. On the premise of localized differential privacy, the location area is divided, in the form of quadtree, to establish a quantitative level of location. The second module is the location collection module. It collected the its users' disturbed location set on the division results of the first module, and estimate the true user location distribution as the final result. The main innovations of our method are as follows:

**Adaptive location discretization.** Unlike the previous work, the method in this paper does not need to assume that the input geographical location are discrete. We propose a local differential privacy based method that can interactively make queries to users and could adaptively discretize the GPS data according to the user access density of each

region. This module divides the area as finely as possible while ensuring the signal-to-noise ratio of the collected result, which balances the graininess of region and signal-to-noise ratio.

Adoption of personalized differential privacy. In our experiments, we found that the geographic location collection scheme that conforms to local differential privacy introduces a lot of noise and makes the overall utility of the final collection results low. Therefore, we adopt the personalized local differential privacy model and modified existing attribute collection algorithms, achieving collection result with higher utility.

## 2 Related Work

Since local differential privacy needs to disturb user data before the user uploads the data, a mechanism that conforms to local differential privacy generally runs on the user side. Local differential privacy will disturb each user's data, and the variance of the noise of the aggregate result is proportional to the number of samples. In order to avoid noise overwhelming the real signal results, the data collection method that conforms to local differential privacy will only count the frequency of frequent item sets. In order to reduce the impact of noise on the data collection process, and to optimize the communication overhead and computational efficiency, researchers have conducted a lot of researches on the implementation of data collection mechanisms that conform to local differential privacy. Here we briefly introduce the design of methods that have inspired our work.

In 2014, a statistical method RAPPOR that conforms to local differential privacy is proposed. This method encodes the user's attribute set through the bloom filter and randomly disturbs all bits of the bloom filter. On the basis of local differential privacy, the disturbed bloom filter is uploaded to the data collector. On the collector side, the collector sums the set times of all bits of the bloom filter uploaded by all users, and use the least square method to estimate the frequency of occurrence of each attribute. In 2016, RAPPOR [8] was further improved, no longer need to assume that user attributes belong to a known limited set, so that RAPPOR can count the frequency of frequent occurrences of arbitrary unknown attributes. Their improved method is comprised of two modules. The first module is the same as the original RAPPOR method, using bloom filter results to estimate the frequency of attributes. The second module is used to calculate attribute sets that belong to frequent items. It cuts the string encoding of all attribute names into multiple fixed-length character segments, and uses the expected maximum algorithm to estimate the probability of occurrence of all character segment pairs. The connection of the character combination is stored in a graph. Each character segment corresponds to a node in the graph. When the occurrence probability of the character segment pair exceeds a certain threshold, the two nodes are connected. Since all character segments of each frequent element must also be frequent, fully connected subgraphs of a specific length in the graph then correspond to frequent item sets. Finally, the first module could estimate the frequency of items in the frequent attribute set.

In 2015, a local differential privacy based method—binary local hashing method [9] is proposed, which is completely different from RAPPOR and based on the principle of compressed sensing theory. This method randomly generates a  $\pm 1$  vector with a fixed length of  $m$  for each attribute of the user attribute set, and uses this vector as the binary

representation of the attribute. Since the expectation of two different vector dot product is 0, and the dot product of the same vector is  $m$ , the method randomizes the input vector while keeping the expectation of each value in the vector unchanged, and then sums all the uploaded user vector. And by dot multiplying the sum vector with any representation vector of an attribute, we can get an unbiased estimate of the frequency of the attribute.

In 2017, researchers [10] summarized methods such as random response, RAPPOR, and binary local hash method, and proposed an error analysis framework for automatically optimizing random response probability parameters. But these two methods can only estimate the attribute frequency of a known and limited set, and cannot deal with the unknown or unlimited number of attribute sets.

In 2018, a frequent item set discovery framework, called PrivTrie [11], based on prefix trees was proposed. They believed that the reason RAPPOR improved method [8] has excessive computational overhead and sensitivity to noise interference, is that graph is not suitable for storing the relationship between character segments. Therefore, they propose to use the prefix tree structure to describe the coupling relationship between character segments. In addition, their paper proposes a method that can make the same query to users of different branches of the prefix tree at the same time and still ensure differential privacy security. It can make more query requests to a limited set of users, thereby improving the accuracy of estimated attribute frequency.

In addition, in 2016, researchers [12] first applied the concept of local differential privacy to the field of geographic location collection research, and its scheme adopted a binary local hash method for location data collection. As the direct use of localized differential privacy would result in low signal-to-noise ratio, researchers proposed the concept of personalized local differential privacy, which is different from local differential privacy in that the new concept only requires that the probability distribution on the user-specified attributes are approximate rather than the whole attribute set. In addition, the scheme assumes that all geographic locations have been quantified as discrete areas. This scheme is a geographic location collection scheme based on the concept of local differential privacy derivation, which is known to have high data utility. Therefore, we use this work as a comparison to verify the utility of the data collection results of our work, and in paper, we refer to it as PSDA.

### 3 System Overview

In order to guarantee the user's data privacy during data collection, our method adopts the local differential privacy [13] as the privacy protection model. The principle of localized differential privacy is to randomly disturb the user's data before uploading it. After the collector collects a certain number of users' disturbed data, the collector then estimates the distribution of real users. There are mainly two problems in the scheme design:

- (1) Suppose the size of the user set to be collected is  $N$ , the noise magnitude added by local differential privacy is orders of, and the noise added by centralized differential privacy is generally a constant. Therefore, compared to centralized differential privacy based method, data collection methods that conform to local differential privacy need to be designed to ensure that the attribute whose frequency is to be estimated must be frequent. As a result, before estimating the frequency of geographic

location access, our method first needs to calculate the frequent item sets, and the process of calculating frequent item sets also needs to satisfy the local differential privacy.

- (2) There are huge differences in user attitudes towards privacy. On the one hand, capturing this difference meets the personalized privacy requirement; on the other hand, it adaptively reduces the magnitude of added noise. Therefore, in our method, it is necessary to adopt a privacy concept that can reflect the privacy protection needs of different users according to the characteristics of geographic location data, so as to improve the availability of data.

In response to the problem in (1), our method first divides the user set into two disjoint set, the first set is used to calculate frequent itemsets of geographic location. As original GPS data is continuous, and there is a certain unevenness in the distribution, so first of all, it is necessary to quantify the continuous geographic location into discrete areas, and adjust the quantization granularity of different areas according to each area’s user access frequency. More fine-grained quantification need to be performed on the area with higher user access frequency; the second user set is used to collect the disturbed frequency of user visits in each geographical area, and estimate the true geographic distribution of users.

In response to the problem in (2), our method adopts the concept of personalized local differential privacy, using the tree structure to organize the calculated frequent area sets, and allows users to personalize their privacy requirement, which can greatly improve the accuracy of the estimation result.

In terms of system architecture, this chapter is divided into a geographic location division module and a geographic location collection module. The relationship between these two modules is shown in Fig. 1.

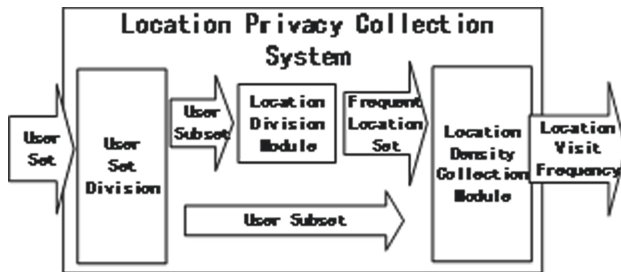


Fig. 1. Architecture of our location privacy collection method

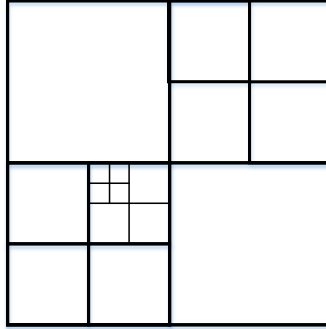
## 4 Module Design

### 4.1 Location Division Module

This section introduces the design of the geographical location division module. The map division method used in our method uses the quadtree division method adopted by



previous researchers [4, 14, 15], and the division method is shown in Fig. 2. The largest square represents the entire map. By recursively dividing the map with a quadtree, multiple quantization areas are obtained. In the location density collection module, the results of the map division will be used to quantify the continuous geographic location of the user, and then the data collection method for discrete attributes can be adopted.



**Fig. 2.** Schematic diagram of geographical location division method based on quadtree

Because the local differential privacy based method can only collect the frequency of frequent itemsets, it is necessary to ensure that the frequency of user access in each sub-region finally obtained is higher than a certain threshold to reduce the impact of noise. Therefore, the problems solved in this section are summarized as follows: Under the limitation of local differential privacy, the map is reasonably segmented using a limited set of user data, so that the number of users in each sub-region is higher than a certain threshold and as close as possible to the threshold.

Before introducing the scheme, first we introduce the data structure used in the algorithm. The `TreeNode` structure is used to store tree node information, where `Cell` represents the area corresponding to the tree node, `children` represents the child nodes of the tree node, `number` represents the number of users who select the node as a geographic protection area. As our Location Density Module exploits personalized differential privacy, `user_index` is used to store the set of users who designate this `TreeNode` as their privacy protection area. `Count` is used to save the weights of the four child nodes of the node, and `parent` represents the parent node of the node.

```

struct TreeNode {
    Cell c
    TreeNode* [] children
    int number
    int[] users_index
    CellCount count
    TreeNode* parent
}

```

The algorithm for segmenting the map is shown in Algorithm 1. It draws on the design of Privtree [11], which was designed for calculating frequent discrete attribute,

and we modified the algorithm process to make it adaptively calculating discretization level of continuous GPS data.

---

**Algorithm 1** :  $\text{DivideTreeNode}(\text{rt}, \mathbf{D}, \varepsilon, \text{batch\_size})$

---

**Algorithm 1** :  $\text{DivideTreeNode}(\text{rt}, \mathbf{D}, \varepsilon, \text{batch\_size})$

---

**Input:** the tree root node  $\text{rt}$ , user subset  $\mathbf{D}$ , local differential privacy budget  $\varepsilon$ ,  $\text{batch\_size}$   
**Output:** map division tree rooted with  $\text{rt}$

- 1:  $F = \emptyset$
- 2:  $\text{CS} = \text{set of four sub-areas of rt}$
- 3:  $\text{count} = 0$
- 4: **while**  $\mathbf{D} \neq \emptyset$ :
- 5:     **choose**  $\text{batch\_size}$  users from  $\mathbf{D}$ , represented as  $\mathbf{G}$
- 6:     **delete**  $\mathbf{G}$  elements from  $\mathbf{D}$
- 7:      $F = F \cup \mathbf{G}$
- 8:     **for every** user  $u$  in  $\mathbf{G}$  **do**
- 9:          $\text{count} += \text{IsInCell}(\text{r.Cell}, u.\text{Location}, \varepsilon)$
- 10:         **if**  $\text{evaluate}(\text{count}, F.\text{size}) > \text{threshold}$  **then**
- 11:             **for every** cell  $\text{cnode}$  in  $\text{CS}$  **do**
- 12:                  $\text{root.Children.append}(\text{DivideTreeNode}(\text{cnode}, \mathbf{D}, \varepsilon))$
- 13:             **break**
- 14:     **return**  $\text{root}$

---

Lines 1–3 are the initialization of parameters. Lines 5–7 indicate that  $\text{batch\_size}$  users are randomly sampled from the set of users assigned to the current node. In the 9–10 line,  $\text{IsInCell}$  is used to simulate the process of making a query request to the sampled user, and the implementation of the  $\text{IsInCell}$  function is given in Algorithm 2. Line 10 simulates the process that the data collector uses the  $\text{evaluate}$  function to remove noise and determine whether the frequency of user access to the node is a certain threshold. We choose  $\max(0.001|\mathbf{D}|)$  as threshold, among which, means the variance of evaluate result. Since the evaluate result follows normal distribution, its variance could be calculated easily. If evaluate result is greater than the threshold, then in line 12, corresponding areas to the child nodes are further recursively divided; if it is less, return to line 5, adds more users, and repeat the process of lines 7–13 until  $\mathbf{D}$  is the empty set.

---

**Algorithm 2 :** IsInCell(Cell  $c$ , Location  $l$ , double  $\varepsilon$ )

---

**Input:** Location Area  $c$ , user location  $l$ , local differential privacy budget  $\varepsilon$   
**Output:** 0 or 1

1: sample from the distribution, and get the result  $\mathbf{b}$

$$\Pr[\text{output} = 1] = \begin{cases} p = \frac{e^{\frac{\varepsilon}{2}}}{e^{\frac{\varepsilon}{2}} + 1}, & \text{if } l \in c \\ q = \frac{1}{e^{\frac{\varepsilon}{2}} + 1}, & \text{if } l \notin c \end{cases}$$

2: **return**  $\mathbf{b}$

---

The information collection process given in Algorithm 2 exploits the randomized response mechanism, which has been proved to satisfy local differential privacy [7]. We simply show the proof of local differential privacy here.

There are four situations here, which are:

$$\frac{\Pr[\text{output}(l) = 1]}{\Pr[\text{output}(l') = 1]} = \begin{cases} 1, & \text{if } l \in c \text{ and } l' \in c' \\ e^{\frac{\varepsilon}{2}}, & \text{if } l \in c \text{ and } l' \notin c' \\ e^{-\frac{\varepsilon}{2}}, & \text{if } l \notin c \text{ and } l' \in c' \\ 1, & \text{if } l \notin c \text{ and } l' \notin c' \end{cases}$$

Thus we can easily see that each IsInCell algorithm satisfies  $0.5\varepsilon$ -local differential privacy. Furthermore, in algorithm 1, every user sent bit vector contains at most one 1-bit, and all others 0-bit, so algorithm 1 satisfies  $\varepsilon$ -local differential privacy. On the server side, The implementation of evaluate function is

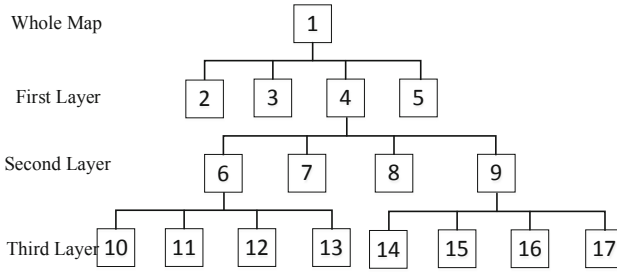
$$\text{evaluate}(\text{count}, n) = \frac{\text{count} - n \cdot q}{p - q}$$

Finally, the algorithm given in Algorithm 1 can get the quadtree corresponding to the map area division, and the leaf nodes in the tree have a one-to-one correspondence with each quantized area.

## 4.2 Personalized Location Privacy

Since the map has been recursively divided into multiple areas, and the areas are in a tree-like, hierarchical relationship, our method allows users to specify their privacy protection areas. Note that user-specified privacy protection areas are considered not to be private data and it could be obtained directly by the server. Assume that the geographical division module divides the map as shown in Fig. 3.

In order to reduce the error caused by quantization, the user's location data will only be quantized to any element in the set of leaf nodes, in our example, {2, 3, 5, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17} numbered nodes corresponding areas. Assume that a user is quantified to area 11, he is allowed to choose his privacy protection level in 4 levels of differential privacy protection. The numbers of the privacy protection areas



**Fig. 3.** Example of map division result

corresponding to the four levels are 11, 6, 4, 1, respectively, that is, a user can choose any ancestor node of his location node as his privacy protection area.

For example, when the user selects area 4 as its privacy protection area, according to the definition of personalized local differential privacy, we need to ensure that on all leaf nodes under area 4, including {7, 8, 10, 11, 12, 13, 14, 15, 16, 17}, local differential privacy needs to be satisfied. The advantage of personalized differential privacy is that the user’s data collection process only needs to ensure the differential property in the privacy protection area specified by the user, which doesn’t need to consider the probability distribution on the locations outside the privacy protection area, in this example, {2, 3, 5} area.

### 4.3 Location Density Collection Module

Since the privacy protection areas designated by users are different, firstly, users are divided according to their designated differential privacy protection areas, and a data collection method is called individually for each user subset. This section introduces the design of location collection module.

This module uses the improved method of the binary local hash method proposed by researchers [9, 10, 12] and in order to improve the utility of collection results, this module exploit personalized differential privacy model. Assuming that each user to be collected has designated his privacy protection area, suppose the geographic location of a user  $u$  is  $u.l$  and privacy protection area is  $u.L$ . The collection process is shown in Algorithm 3.

---

**Algorithm 3 :** Location Collection Process

---

**Input:** Collector designated parameter  $\mathbf{m}$  and  $\mathbf{g}$ , quantization location set  $\mathbf{D}$ , user set  $\mathbf{U}$

**Output:** All quantization locations' estimated frequency

- 1:  $d=|\mathbf{D}|$
- 2: collector generate a  $m \times d$  sized matrix  $M$ , each item in matrix is randomly chosen from  $\{1,2,3,\dots,g\}$ , and each column corresponds to a location
- 3: collector initializes a zero matrix  $z$ , sized  $m \times g$
- 4: collector initializes a  $d$  sized zero vector  $f$ , to save all locations' estimated frequency
- 5: **for** every user  $u$  in  $\mathbf{U}$  **do**
- 6: collector randomly generates a number  $j$  from  $\{1,2,3,\dots,m\}$
- 7: collector sends  $j$ -th row of  $M$  to user  $u$
- 8: user  $u$  computes  $r=\text{LocalRandomize}(u.l, u.L, M_{j,\cdot})$ , and sends  $r$  to collector
- 9: collector computes  $z[j][r]=z[j][r]+1$
- 10: **for** every location  $l$  in  $\mathbf{D}$  **do**
- 11:  $f_l = \text{EstimzteFrequency}(M_{\cdot,l}, z)$
- 12: **return**  $f$

---

In the first step, the collector generates a random matrix. It should be noted that this matrix does not need to be kept secret. It can be obtained by sharing the key between the data collector and the user and generated from a random stream, which reduces communication overhead of sending the  $j$ -th row of matrix  $M$  in the row 7. The matrix  $z$  in the second step is used to save the user's aggregate statistical results. Steps 6 to 9 are basically the same as the binary local hash mechanism [9, 12]. The difference is that the return value  $r$  of `LocalRandomize` in our method is no longer, but a value in  $\{1, 2, 3, \dots, g\}$ . Corresponding to that, in step 7, our method takes  $r$  as an index, add 1 to the  $r$ -th column of the  $j$ -th row of the aggregate result  $z$ .

The implementation of `LocalRandomize` and `EstimateFrequency` are shown in Algorithm 4 and Algorithm 5 respectively.

---

**Algorithm 4 :** LocalRandomize

---

**Input:** user location  $l$ , user designated privacy protection area  $L$ ,  $j$ -th row  $R$  of matrix  $M$ , location quantization set  $\mathbf{D}$

**Output:** disturbed user location index from  $\{1,2,3,\dots,g\}$

- 1:  $e=R[l]$
- 2: user randomizes  $z$  following the distribution, and get the result  $v$

$$\Pr[v = z] = \begin{cases} \frac{e^e}{e^e + g - 1}, & z = e \\ \frac{1}{e^e + g - 1}, & z \neq e \end{cases}$$

- 3: **return**  $v$

---

Since for every user, randomized response mechanism is invoked, and the proof is the same as Algorithm 2.

---

**Algorithm 5 :** EstimateFrequency

---

**Input:** the location encoding  $c$ , aggregate matrix  $z$ , user number  $N$  that designate the location as their privacy protection area

**Output:** the location’s estimated visit frequency

- 1:  $p = \frac{e^\epsilon}{e^\epsilon + g - 1}, q = \frac{1}{g}$
- 2: count=0
- 3: **for**  $i=0; i < c.size; i++$ :
- 4:     count +=  $z[i][c[i]]$
- 5: **return**  $\frac{\text{count} - N \cdot q}{p - q}$

---

The basic idea of the frequency estimation process in Algorithm 5 is the same as the randomize response mechanism. The difference is that the user aggregation result here is a matrix instead of a vector. Since each column of the random matrix generated by the collector can be regarded as a encoding of a location area, each element is randomly chosen from  $\{1, 2, \dots, g\}$ . So when estimating the frequency, only the same indexed aggregation value as the target encoding needs to be count. So in line 4, we first take the value of the column  $c[i]$ , and use  $c[i]$  as index to take the corresponding aggregation frequency value in  $z$ . After eliminating the bias in line 5, we can get the estimated frequency of the target attribute.

It should be noted that in our method, Location Collection Process needs to be invoked for every set of users that designate the same privacy protection area. But this wouldn’t be a efficiency bottleneck, because every user still only needs to participates in one collection. After all users location data has been collected, add the estimated results in each collection and then the total corresponds to the location’s real visit frequency.

## 5 Experimental Validation

### 5.1 Experiment Setup

In our experiment, we use Brinkhoff [16] and the Portugal taxi trajectory dataset as the users’ location data set.

Brinkhoff is trajectory generator that has been widely adopted as benchmark [17, 18]. It takes the map in the real world as the input, and establishes a trajectory generator, which can generate trajectory data sets of any size according to the characteristics specified by the user. In the experiment, the German Oldenberg is used as the map, and a total of 1,000,000 trajectory data are generated as the trajectory data set of the experiment.

Protugal taxi trajectory dataset was drawn from the ECML/PKDD 2015, and we randomly chose 1,000,000 trajectory data from original 1,673,686 trajectories.

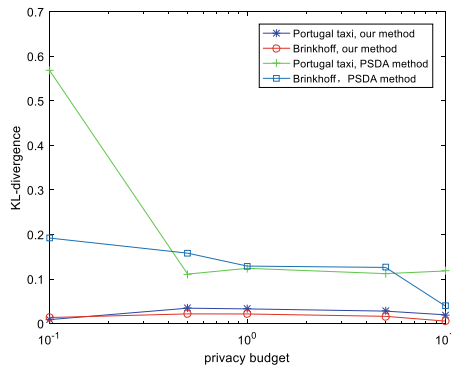
Since the goal of our method is to collect the users’ geographic location data as accurately as possible, we compare the collected user location distribution with real user

data distribution to evaluate the geographic location collection method proposed in this paper. The evaluation indicators adopted in this article are the same as PSDA work and are as follows:

- (1) **KL distance.** We calculate the distribution of the original data set on the geographical location division results, and then calculate the distribution of the collected location access probability distribution. In order to measure the distance between the two distributions, KL divergence is used as the evaluation metric.
- (2) **The accuracy of top-K areas with the highest density.** We calculate the K locations with the highest frequency of density in the original data set, then calculate the K locations with the highest frequency of access in the estimation result, and calculate the accuracy of the estimation result.

## 5.2 Experiment Results

The performance of this scheme and PSDA scheme on the KL distance evaluation index on different data sets is shown in Fig. 4.



**Fig. 4.** KL divergence between original data set and collected results.

It can also be verified that under the same local differential privacy budget, our method could achieve lower KL divergence and higher top-K accuracy than PSDA method. In addition, it should be noted that in Fig. 5, when differential privacy budget, the geographical location is divided and the size of the division location set is less than  $K = 100$ , so the accuracy rate of the K regions with the highest access density is 100%. It can be seen that the accuracy of the experimental results in Fig. 6 does not increase with the increase in differential privacy budget. According to the analysis, there are two reasons for this phenomenon:

- (1) The top-K indicator only cares about the frequency of the area with a larger frequency, and the collection result of the area with a higher frequency itself has higher signal-to-noise and is less affected by noise.

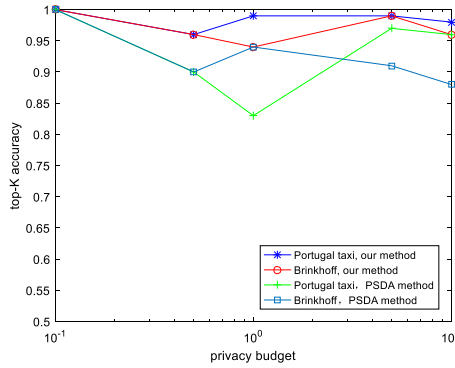


Fig. 5. Top-K accuracy of collected location results. K = 100.

(2) The location collection module also uses the result of the geographic location division module. As the differential privacy overhead increases, the variance of the noise also decreases, so the threshold of the leaf nodes in the division process also decreases. As a result, the leaf nodes are further divided, making the location set larger. In the experiments of the Portuguese taxi data set, the change of the size of the divided location set with the differential privacy budget is shown in Fig. 6.

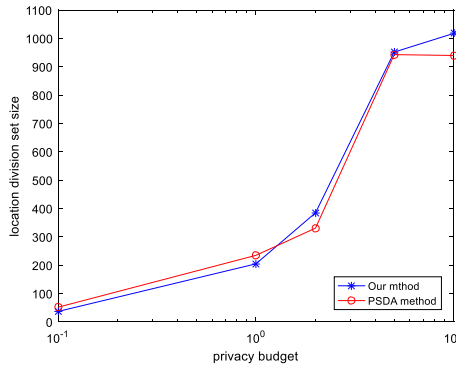


Fig. 6. Change of location division result size with differential privacy budget.

It can be seen from Fig. 6 that the size change of the location set obtained by this scheme and PSDA scheme is basically the same. When the differential privacy budget is low, the number of geographically divided areas is also low, which can compensate for the increase in noise, even if signal-to-noise ratio of each collected location density reduces. It should be noted that in the experiments corresponding to Fig. 4 and Fig. 5, PSDA scheme also has this effect, but because the noise amplitude of their method grows too fast, the change in the size of the location set is not fast enough to compensate for the increase of noise. Therefore, its accuracy shows a significant downward trend, which



also proves that the method proposed in our paper could achieve better collected results utility.

In order to further illustrate the influence of the original location set size and location division results size on the accuracy of the final collection results, experiments are carried out on different sizes of original datasets. The experimental results are shown in Table 1. Note that original data size's unit is million.

**Table 1.** Change of evaluation with dataset size (batch\_size = 1000)

Original dataset size/million	0.2	0.4	0.6	0.8	1
KL divergence	0.0208	0.0639	0.0868	0.136	0.222
Top-K	0.93	0.96	0.97	0.98	0.97
Location division set size	133	538	1546	2653	5239

As can be seen from the results in Table 1, as the scale of the data set increases, the number of regions obtained by dividing the map by the location division module has increased significantly, and the relative proportion of the growth rate is far faster than the growth rate of the scale of the data set, resulting in that the signal-to-noise ratio averaged in each area is reduced. With the increase in the size of the data set, the KL divergence indicator showed a significant increase, but the top-k accuracy rate remained almost unchanged. The reason for this result is that the KL divergence represents the accuracy of the collection results of all regions, and the top-K accuracy represents the accuracy of the collection results of high-frequency sub-regions, so the latter itself is less affected by noise. In summary, it can be concluded that if the goal of collecting data only considers high-frequency attributes, the system can achieve high-precision collection results without special settings; if the data to be collected needs to consider the frequency of all attributes, we need to adjust the size of batch\_size according to the size of the user set to be collected, so that the number of regions divided by the geographic location division module increases in proportion to the size of the data set, so as to ensure the relative stability of the signal-to-noise ratio.

## 6 Conclusion

In this paper, we explain the necessity of privately collecting user locations from the perspective of users and service providers, and then divides the private collection method into a location division module and a location density collection module, and explains functions and principles of the two modules. Finally, the utility and accuracy of the method are tested using the Brinkhoff trajectory generator and the Portugal taxi trajectory data set. The results shows that out method could achieve better utility than the best method known so far.

## References

1. Fawaz, K., Feng, H., Shin, K.G.: Anatomization and protection of mobile apps' location privacy threats. In: 24th USENIX Security Symposium, pp. 753–768. USENIX (2015)
2. Chen, R., Fung, B., Desai, B.C.: Differentially private trajectory data publication. arXiv preprint [arXiv:1112.2020](https://arxiv.org/abs/1112.2020) (2011)
3. Chen, R., Acs, G., Castelluccia, C.: Differentially private sequential data publication via variable-length n-grams. In: Proceedings of the 2012 ACM Conference on Computer and Communications Security, pp. 638–649. ACM (2012)
4. Zhang, J., Xiao, X., Xie, X.: PrivTree: a differentially private algorithm for hierarchical decompositions. In: Proceedings of the 2016 International Conference on Management of Data, pp. 155–170. ACM (2016)
5. He, X., Cormode, G., Machanavajjhala, A., Procopiuc, C.M., Srivastava, D.: DPT: differentially private trajectory synthesis using hierarchical reference systems. In: Proceedings of the VLDB Endowment, pp. 1154–1165. Springer (2015)
6. Gursoy, M.E., Liu, L., Truex, S., Yu, L., Wei, W.: Utility-aware synthesis of differentially private and attack-resilient location traces. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, pp. 196–211. ACM (2018)
7. Erlingsson, Ú., Pihur, V., Korolova, A.: RAPPOR: randomized aggregatable privacy-preserving ordinal response. In: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, pp. 1054–1067. ACM (2014)
8. Fanti, G., Pihur, V., Erlingsson, Ú.: Building a rappor with the unknown: privacy-preserving learning of associations and data dictionaries. *Proc. Priv. Enhanc. Technol.* **2016**(3), 41–61 (2016)
9. Bassily, R., Smith, A.: Local, private, efficient protocols for succinct histograms. In: Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing, pp. 127–135. ACM (2015)
10. Wang, T., Blocki, J., Li, N., Jha, S.: Locally differentially private protocols for frequency estimation. In: 26th USENIX Security Symposium, pp. 729–745. USENIX (2017)
11. Wang, N., et al.: PrivTrie: effective frequent term discovery under local differential privacy. In: 2018 IEEE 34th International Conference on Data Engineering, pp. 821–832. IEEE (2018)
12. Chen, R., Li, H., Qin, A.K., Kasiviswanathan, S.P., Jin, H.: Private spatial data aggregation in the local setting. In: 2016 IEEE 32nd International Conference on Data Engineering, pp. 289–300. IEEE (2016)
13. Warner, S.L.: Randomized response: a survey technique for eliminating evasive answer bias. *J. Am. Stat. Assoc.* **60**(309), 63–69 (1965)
14. Samet, H.: The quadtree and related hierarchical data structures. *ACM Comput. Surv.* **16**(2), 187–260 (1984)
15. Ho, S.S., Ruan, S.: Differential privacy for location pattern mining. In: Proceedings of the 4th ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS, pp. 17–24. ACM (2011)
16. Brinkhoff, T.: Generating network-based moving objects. In: Proceedings of the 12th International Conference on Scientific and Statistical Database Management, pp. 253–255. IEEE (2000)
17. Agarwal, P.K., Fox, K., Munagala, K., Nath, A., Pan, J., Taylor, E.: Subtrajectory clustering: models and algorithms. In: Proceedings of the 37th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, pp. 75–87, May 2018
18. Orakzai, F., Calders, T., Pedersen, T.B.: k/2-hop: fast mining of convoy patterns with effective pruning. *Proc. VLDB Endow.* **12**(9), 948–960 (2019)

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# **Systems Security**



# Analysis on the Security of Satellite Internet

Huan Cao<sup>1</sup>, Lili Wu<sup>2</sup>, Yue Chen<sup>2</sup>, Yongtao Su<sup>1</sup>, Zhengchao Lei<sup>2</sup>(✉),  
and Chunping Zhao<sup>3</sup>

<sup>1</sup> Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China

caohuan@ict.ac.cn

<sup>2</sup> National Computer Network Emergency Response Technical Team/Coordination Center of China, Beijing 100029, China

leizhengchao@cert.org.cn

<sup>3</sup> Beijing Sylincom Technology Co., Ltd., Beijing, China

**Abstract.** Satellite Internet (SI) is a new way to provide internet access all over the world. It will bring great convenience to international communication. Compared with the traditional communication networks, SI has a significant change in network architecture and communication model, which will have an important impact on national information network security. For example, the global interconnected SI consists of a large number of small satellites and each satellite has multi-beams to cover a vast area, which leads to the disorderly flow of information across the border, and greatly increases the difficulty of network protection. Therefore, it is necessary to closely track the development of SI and analyze security problems brought by SI. In this paper, we analyze the security risks of SI from the perspective of national security, network security and equipment security, and thirteen security issues have been summarized to provide reference for the healthy development of SI industry.

**Keywords:** Satellite internet · Network security

## 1 Introduction

In recent years, the world's space powers have proposed low-earth-orbit (LEO) satellite constellation plans, which has triggered a boom in satellite internet (SI) development. Concerning the development of SI, the white paper published by China Center for Information Industry Development (CCID) points out that the world is on the eve of the dense launch of man-made satellites [1]. It is estimated that the low Earth orbit (LEO) satellites will deploy a total of about 57000 by 2029. A space resource race of satellite orbits is quietly beginning, countries all over the world have joined in the space race of SI, and the earth surface will be covered by a large number of LEO satellites intensively. Therefore, security problems brought by this will become a new challenge [2–4]. With the construction of SI becoming a national strategy all over the world, the industry has entered a period of rapid market growth [5, 6], and it specifically reflected in the following aspects:

© The Author(s) 2020

W. Lu et al. (Eds.): CNCERT 2020, CCIS 1299, pp. 193–205, 2020.

[https://doi.org/10.1007/978-981-33-4922-3\\_14](https://doi.org/10.1007/978-981-33-4922-3_14)

- Fighting for frequency and orbit resources: The competition for frequency and orbit resources among countries has become increasingly white-hot. According to the data submitted to international telecommunications union (ITU), satellite companies in France, United States and United Kingdom have the largest number of resources such as key frequency bands and orbital heights. For example, OneWeb has submitted at least seven materials of network resources to ITU, including THEO, STRIPE, 102, etc., covering 8425 km / 8575 km, 1200 km and other medium and low orbital altitude, as well as Ku / Ka / V and other frequency bands; SpaceX submitted twelve materials to ITU, including usasat-ngso-3a-r / 3b-r / 3C, 3D / 3E / 3F / 3G / 3H / 3I, usasat-ngso-3j / 3K / 3l, covering 345.6–1325 km orbital altitude and Ku / Ka / V frequency bands.
- Large-scale network deployment: SI constellation construction has entered the stage of large-scale network deployment. SpaceX plans to launch 42000 satellites, and 482 broadband Starlink satellites have been launched by June 5, 2020. In addition, OneWeb has launched 74 satellites in the past two years [7].
- International operation: The service providers of SI have been striving for landing rights in countries around the world. For example, OneWeb initially obtained market access authorization in about 19 countries in 2019.

SI can be mainly used for emergency rescue, rural and remote area coverage, maritime market (including cruise ships, merchant ships, fishing boats, yachts, etc.), aviation market, military and government applications [8]. Compared with the terrestrial mobile communication system (TMCS), the SI will face the following new security challenges:

- Due to the limited computing and storage capacity, the satellites in SI constellation don't support high-complexity encryption protocols and algorithms, resulting in the weak protection of traffic data.
- The topological structure of the LEO satellite networks are constantly changing, the openness of the satellite's orbit makes it very difficult to be supervised.
- Communication satellite is a highly integrated product, its components are supplied by many manufacturers. There may be security holes and design defects in all aspects of integration. Especially, the technology of on-orbit satellite reprogramming is not mature, which makes it very difficult to make up for the security holes of on-orbit satellites.
- Satellite communication has the characteristics of wide coverage [9], which can broadcast data to a large number of user terminals in a large range. When the SI network is attacked, the impact is greater than that of the TMCS, so it is easier to become the target of hackers.

In summary, the security problems faced by the SI are more severe than those of the TMCS. If the SI is attacked, it will have a wider range of influence and cause greater damage. Therefore, it is necessary to carry out the research on the security problems faced by the SI.

## 2 Related Work

The research on the security problems of SI is still in its infancy. 3GPP puts forward the network architecture of non-terrestrial networks (NTN) [10], but there is no systematic analysis on the security problems of NTN. Sat5G analyzes the security threats of the integration of satellite network and 5G networks, mainly including the following three main aspects [11]:

1. Security threats of satellite connections as transport network for backhaul  
One of the main security threats perceived by the terrestrial network is the tampering or eavesdropping of the data transmitted (the control plane signaling or the user plane data) over the backhaul connection. In addition, another threat perceived by terrestrial networks in case of sharing of the satellite network is the tampering and eavesdropping of traffic via the shared network.
2. Security threats of satellite connections as transport network among 5G core networks  
In this case, the two terrestrial networks usually are not in the same trust domain, and the intermediate satellite network is not considered to be part of the trust domain of either of the two terrestrial networks. At the same time, it is very common for satellite networks to be shared among multiple terrestrial networks. The security threats perceived by the terrestrial network are tampering, eavesdropping and unauthorized traffic redirection (i.e. traffic ‘hijacking’) [12, 13].
3. Security threats to content delivery via satellite

Security threats related to content delivery networks (CDN) are DDOS at-tacks; Content leakages, such as unauthorized access to content, which is aggravated by local caching and the use of MEC servers; Deep linking, in this case, all media slices can be accessed by accessing a manifest file due to use MPEG DASH.

However, Sat5G has made a preliminary analysis of the security issues of SI, but it is not comprehensive enough. This paper summarizes and analyzes the security issues faced by SI in the future from the aspects of national security, network security and equipment security based on the existing research.

## 3 Analysis of SI Security

### 3.1 Overview of Security Issues of SI

The system architecture of SI can be divided into user segment, space segment and ground segment. The user segment includes various satellite terminals; the space segment includes satellite constellation [14], which can be divided into constellation with inter satellite link (ISL) and constellation without ISL [15]; the ground segment includes gateway station (GS), operation management and control system (OMCS), measurement and control system (MCS), network management system (NMS), etc. According to the characteristics of SI, the possible security problems in SI are summarized in Table 1.

**Table 1.** Security issues of SI.

Classification	ID	Security problem	Description
National security	(1)	National and military security threats	<ul style="list-style-type: none"> <li>• Illegal organizations can steal strategic information of target countries by deploying earth observation payload on LEO satellites</li> <li>• LEO satellite provides communication platform for future information warfare weapons</li> </ul>
	(2)	Frequency and orbit resource preemption	To occupy limited orbit resources by planning LEO satellite constellation
	(3)	Interference in astronomical exploration	The launch of a large number of LEO satellites can cause serious interference to astronomical observation
Network security	(4)	Identity impersonation	<ul style="list-style-type: none"> <li>• Disguised as a satellite terminal (ST) to access the SI and destroy the network</li> <li>• Disguised as a satellite to trick legitimate STs into accessing a false network to obtain the ST's location or identification information</li> </ul>
	(5)	Data eavesdropping	Illegal organizations illegally receive and analyze transmitted traffic data or signaling data through wireless links (feedback link, user link, ISL)
	(6)	Data integrity issues	Modify, insert, replay, delete user or signaling data to destroy data integrity
	(7)	Information interception	Illegal interception of user location or identification information transmitted by ST through wireless links
	(8)	Signal interference	Attackers interfere with satellite wireless links by emitting high-power electromagnetic waves
	(9)	Denial of service	Interfere with satellite or gateway, and interfere with data or signaling physically or by protocol, which makes SI unable to provide normal services for legitimate ST
	(10)	Anonymous attack	Attackers attack the satellite node in space, but the satellite cannot determine the attackers

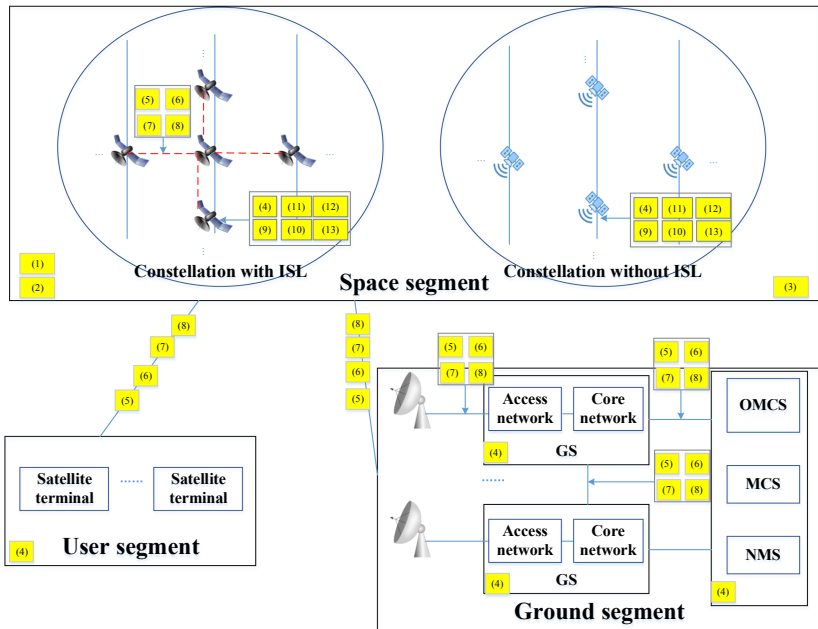
*(continued)*



**Table 1.** (continued)

Classification	ID	Security problem	Description
	(11)	Malicious occupation of satellite bandwidth resources	Sending illegal signals to the satellite through wireless link, because the satellite will not check the legitimacy of the signals, so the illegal signals will occupy the bandwidth resources of the satellite
Equipment security	(12)	Malicious satellite control	By issuing malicious instructions or injecting viruses to satellite nodes from ground facilities or space to achieve the goal of controlling satellites
	(13)	Malicious consumption of satellite resources	Malicious consumption of satellite propellant resources to achieve the goal of reducing satellite life

The distribution of the above thirteen security issues in the SI is shown in Fig. 1.



**Fig. 1.** The distribution of security issues in SI system

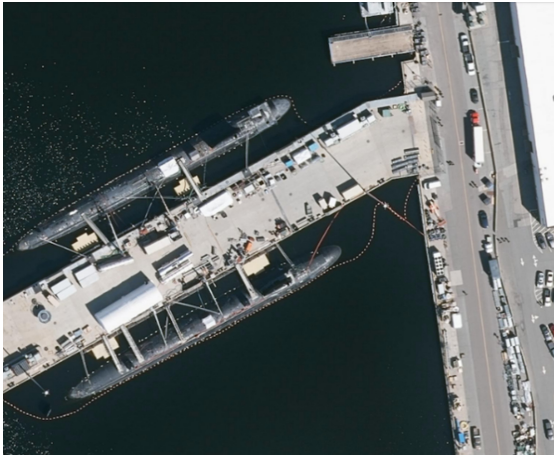
## 3.2 National Security Issues

### National and Military Security

The security threats include national strategic information security and military security threats.

#### *National Strategic Information Security*

SI involves a large number of satellites, and the orbit altitude is concentrated between 300 km and 2000 km. If the corresponding satellites equipped with high-resolution scanning observation payloads, such a large number of satellites will expose the important military infrastructure of countries all over the world and threaten national security. Recently, earth observation industry (EOI) company is promoting the development of a new very low earth orbit (VLEO) satellite constellation. Its propulsion system and innovative design will enable the satellite to run in a very low orbit. In order to support continuous monitoring service, the initial constellation consists of 30 satellites with an average revisit time of two hours. The company plans to launch its first satellite by the end of 2022. EOI company's mission is to enable defense and intelligence agencies and commercial customers to easily access ultra-high resolution images at affordable prices to support a range of applications such as resource management, environment and disaster assessment, asset monitoring, logistics planning, infrastructure mapping, public security, homeland security, insurance and real estate. For example, Fig. 2 shows the image of naval ship captured by EOI company's VLEO satellite.



**Fig. 2.** Image of naval ship captured by EOI's VLEO satellite

*Military Security Threat*

1. The low cost and high launch success rate (LSR) of SI will pose new challenges to the technical field of anti-satellite weapons.

The existing kinetic energy anti-satellite weapons (KEAW) rely on the momentum of high-speed moving objects to destroy the target, which has a strong lethality for satellites with high cost and low LSR. However, for the large-scale and low-cost SI, traditional KEAW are facing many challenges. Taking Starlink constellation of SpaceX as an example:

- a. The traditional KEAW are all disposable. It means that a large number of KEAW need to be manufactured and maintained to deal with the threat of the Starlink constellation of 42000 satellites, and the cost will be astronomical.
- b. The traditional KEAW adopts the hard-kill method. The method will generate a large number of space debris, which may hit more satellites, causing uncontrollable and irreversible chain reaction, making the whole earth surrounded by satellite debris.
- c. If we give up the hard-kill method and study more advanced weapons such as soft-kill method, it will cost a lot of money to tackle key technical problems, and the development cycle will very long.

2. The cooperative operation of SI and drone swarm will pose great challenges to the national defense system.

With the accelerated evolution of the war form, a large number of intelligent equipment appear in the war. As an important part of intelligent warfare, Unmanned Aerial Vehicle (UAV) cluster warfare poses great challenges to the traditional defense system. However, the UAV cluster warfare relies on the communication link among each UAV to achieve real-time information interaction, and also relies on the control system to achieve collaborative command, so the overall viability and combat ability of the UAV cluster depends on the security and controllability of the communication link and control system. If the communication link or the control system is jammed by enemy, the UAV cluster will likely to be completely annihilated. Starlink constellation can make up for this defect, it can provide large bandwidth, low delay and wide coverage communication services through a large number of LEO satellites without being affected by any terrain and climate. It can help the UAV cluster get rid of the dependence on land-based communication system, and significantly improve the overall combat effectiveness of the cluster through flight control, situation awareness, information sharing, target allocation and intelligent decision-making, which makes it more difficult for the national defense system to deal with the threat of UAV cluster warfare.

### **Frequency and Orbit Resource Preemption**

According to international regulations, all countries have the right to explore outer space peacefully. Radio frequencies and satellite orbits are limited natural resources and must be used equally, reasonably, economically and effectively. Effective interference control mechanisms should be adopted to make full use of frequency and orbit resources. Effective interference control mechanisms should be adopted to make full use of the limited resources. According to the ITU rules [16], orbit resources are mainly allocated in the principle of first come, first served, and the later declarers cannot cause adverse interference to the satellites of the first declarers. The LEO constellation system should not only launch the satellite in accordance with ITU regulations, but also provide relevant services to the public in accordance with the specified time and proportion, so as to legalize the frequency usage. In other words, the development and utilization of LEO constellation can not only occupy the limited space resources of LEO satellite, but also help to seize the priority use right of spectrum, which has an important impact on the channel use range of battlefield communication.

Generally, at the end of a geostationary earth orbit (GEO) satellite's life, it will increase more than 200 km by using its final energy and enter the grave orbit to complete its self-destruction, so as to release the original working orbit. But the LEO satellite communication system is different, it needs many small satellites to maintain a complete network to provide communication services. When some small satellites cannot work normally or reach the end of their life, it is necessary to launch new satellites to supplement the network, so they will always occupy the frequency and orbit resources. However, the near earth orbit resources can only hold about 60000 satellites, the United States, the United Kingdom, France, Canada, Norway, the Netherlands and other countries have taken the lead in the deployment of SI. SpaceX alone plans to launch 42000 LEO satellites, which will further compress the available orbit resources of other countries.

In addition, all countries except the United States have gaps in the supervision technology of SI, and corresponding laws are not perfect. Once the design and construction of SIs are completed, it will lead to difficulties for countries to effectively supervise the communication services provided by SI, leaving huge security loopholes.

### **Interference in Astronomical Exploration**

Due to the huge scale of SI constellations, astronomical observation will become more difficult as small satellites in constellations are launched one after another. Starlink will launch 42000 satellites, with an average of about 400 satellites observed at any time and at any place. Although they are invisible to the naked eye in orbit, they have a great influence on the astronomical research of optical, infrared and radio telescopes, and are easy to leave traces in the astronomical images. The large-scale integrated Sky Survey Telescope (such as China sky eye) will be greatly affected, which will reduce our ability to observe and warn near Earth Asteroids.

In addition, LEO constellations have the most interference for astronomers who detect dark matter and dark energy, because the signals detected by related instruments are very weak. A large number of LEO satellites will interfere with the space observation of various countries to a certain extent when passing over them, and affecting the corresponding research.

### 3.3 Network Security Issues

#### Identity Impersonation

Due to the lack of identity authentication mechanism in SI's user link, feedback link and ISL, there are three problems of identity impersonation in the following aspects:

1. If the transmission mechanism adopted by the communication system is public, the attacker can calculate the uplink signal according to the downlink signal of the satellite, and then use the satellite communication equipment to disguise as a legitimate ST to access the network and illegally obtain network services.
2. The attacker disguised himself as a satellite network and induced legal STs to access the satellite network to obtain relevant user identification information and location information.
3. The attacker disguised himself as adjacent satellites in the same orbit or different orbit to induce the target satellite to establish an ISL with it, so as to obtain the relevant data transmitted by the ISL.

#### Data Eavesdropping and Data Integrity Attack

Due to the openness of wireless communication of user link, ISL and feed link of SI, the data transmitted through satellite network can be easily eavesdropped. In addition, data encryption will increase the cost of satellite terminal equipment and reduce the utilization rate of satellite link resources. Many satellite communication networks do not encrypt the transmitted data, so it is very easy to cause data leakage.

The most possible attack methods are as follows:

1. The attacker uses a kind of satellite data receiving card to steal data, which is similar to the computer network card with low cost.
2. The attacker makes use of retired equipment abandoned by manufacturers to perform network attacks.
3. The attacker can use the VLEO or LEO satellite in the overseas satellite constellation to eavesdrop the service data on the user link and feeder link of the domestic satellite system.
4. If the satellite constellation built in a country has an ISL, which uses microwave communication, the attacker can control the foreign satellite to approach the target satellite as close as possible to implement data eavesdropping.

Data eavesdropping is often combined with data integrity attack. The attacker often implements data eavesdropping, then inserts, modifies, falsifies the stolen data, and finally send it to the data receiver to achieve the purpose of destroying data integrity.

#### Information Interception

If the orbit of a foreign satellite is lower than that of a domestic satellite (for example, the lowest orbit of SpaceX is about 300-500km), the attacker can use the attack means similar to the terrestrial pseudo base station to carry out network attack. For example, the

torpedo attack method in 4G system can be used in SI, the attacker can use legitimate ST to launch multiple paging to the attacked ST, and this will expose the user identification information, which can be intercepted by the LEO satellite and terrestrial equipment owned by the attacker, so as to track the user's location and bring great security threat.

### **Signal Interference**

This kind of attack is the most common but effective, and it is often used in wars. Interference can be divided into blocking interference and noise interference. Strong interference signals will cause the satellite to be unable to receive the signal normally and provide the service for the legitimate ST.

The possible attack methods are as follows:

1. If the orbit of the overseas satellite is lower than that of the domestic satellite, the attacker can deliberately transmit signals on the working frequency band of the feeder link, user link or ISL of the domestic satellite system to cause interference or interruption of the domestic satellite service.
2. Satellite transponders can be divided into on-board processing transponders and bent-pipe transparent transponders. On-board processing transponders can rely on channel coding, advanced modulation technology, spread spectrum technology, coherent intermodulation product cancellation, etc. to resist interference attacks. But the bent-pipe transponder has a simple structure and does not process any communication signals, so it is easy to encounter signal interference attack. Attackers can interfere with satellites by transmitting signals from high-power transmitters.

### **Denial of Service**

The attack mode against terrestrial network is also applicable to satellite network, such as DDoS attack. Attackers make use of software to simulate massive satellite terminals to send legitimate and bogus requests, which leads to the failure of satellites to provide effective services to legitimate STs. This kind of attack is difficult to defend due to the diversity of satellite communication links. Each ST's client has a receiving and transmitting system. If the transceiver fails to get effective processing when it has problems, it will lead to unstable connection and generate a large number of connection requests. In addition, access requests will also increase greatly when satellite links suffer from signal fading caused by severe weather. However, satellites cannot blindly defend these requests, and the system design of satellite system will not defend these requests as well as the network firewall. Because satellites cannot distinguish whether these requests come from legitimate STs or malicious attackers, which leads to denial of service problems.

### **Anonymous Attack**

Space belongs to the global commons and has no national boundaries. Therefore, it is possible for attackers to launch anonymous attacks against the target satellite in space. Moreover, it is difficult for the attacked satellite to determine and trace the attacks due to the long distance and limited information. On the one hand, there are many factors that lead to satellite failure, such as changes in space environment, design defects, device

problems and even debris effects. Attacks are not the only reason for the failure of satellites in orbit. On the other hand, it is difficult for the ground station to accurately judge what is happening in space limited by distance, meteorology, technical capability and other conditions. The combined effect of these factors enables the attacker to find reasonable excuses to deny the attack.

### **Malicious Occupation of Satellite Bandwidth Resources**

Satellite is a typically resource limited system, on-board computing resources, wireless resources are very scarce, so they are not suitable for complex communication payloads. Most of the on-orbit satellites adopt the bent-pipe transponder without signal unpacking, so it is not possible to determine whether the received data is from a legitimate user. When the attacker sends his own illegal signal, the satellite will still forward the signal to the GS. At this time, if the attacker builds a receiving system to demodulate, decode the data and extract useful data, the purpose of privately communicating with the aid of the satellite is achieved, and a complete method of stealing satellite resources is formed. Moreover, attackers will use their own encryption algorithm to effectively encrypt communication data.

## **3.4 Equipment Security Issues**

### **Malicious Satellite Control**

Due to the lack of network security standards for commercial satellites, coupled with the complex supply chain of satellites, satellite manufacturing uses ready-made technologies to maintain low cost. The wide availability of these components means that hackers can analyze their security vulnerabilities. In addition, many components use open source technology, and hackers may insert backdoors and other vulnerabilities in satellite software, making satellites vulnerable to security risks that are maliciously controlled by attackers.

The means to control the satellite maliciously are as follows:

1. The attacker can capture the target satellite in space and drag the captured satellite out of the working orbit, causing the whole satellite constellation unable to provide continuous services. Moreover, the attacker can inject virus into the captured target satellite after it has been dragged off the working orbit, and then it will be pushed back to the working orbit, causing the virus to spread throughout the whole SI. The technology of capturing on-orbit satellites is already available and has been used in the service of extending the life of on-orbit satellites in orbit. Once this technology is used by hackers, the target satellites can be captured arbitrarily.
2. Satellites are usually controlled by the GSs. These stations are vulnerable to the same network attacks as computers. Although the satellite control attack is not as simple as stealing other people's email, but it can be realized. If there are security loopholes that can be exploited by hackers in the GS, the hackers may invade these stations, and then they can send malicious instructions to control the satellite, or they can use special tools to trick the satellite, and finally achieve the purpose of attacking the SI. For example, the attacker can carry out further attacks after controlling the target

satellite: the attacker can use the broadcast channel of the target satellite to send a large amount of garbage data or spread viruses to the whole SI; shutting down the target satellite to make it unable to provide normal services; if the hackers control the target satellite and it has a propeller device, they can change the orbit of the satellite and hit it on the ground, other satellites or even the international space station.

### Malicious Consumption of Satellite Resources

Attackers can also directly affect the life of satellites by consuming propellants, depleting the write life of charged erasable programmable read-only memory (EEPROM) and other attacks.

## 4 Conclusion

The rapid development of SI has brought some security risks. On the one hand, we should actively develop SI industry, giving full play to the unique advantages of SI, which is not affected by geographical obstacles and disasters; on the other hand, in view of the different levels of security threats faced by SI, it is necessary to carry out forward-looking research on the satellite network security, so as to fill in the regulatory gaps.

**Acknowledgement.** This research was supported by GF science and technology foundation enhancement program, National computer network and information security management center (XDA-Y07-02).

## References

1. CCID. research on the development of China's satellite Internet industry. In: CCID. <https://zhuanlan.zhihu.com/p/144513640> (2020)
2. De Azúa, J.A.R., Calveras, A., Camps, A.: Internet of Satellites. IoSat), Analysis of Network Models and Routing Protocol Requirements. (2018)
3. Su Y., Liu Y., Zhou Y., Yuan J., Cao H., Shi J.: Broadband LEO Satellite Communications: Architectures and Key Technologies (2019).
4. Cao H., Su Y., Zhou Y., Hu J.: QoS Guaranteed Load Balancing in Broadband Multi-Beam Satellite Networks (2019).
5. Anpilogov V.R., Gritsenko A.A., Chekushkin Y.N., Zimin I.V.: A Conflict in the Radio Frequency Spectrum of LEO-HTS and HEO-HTS Systems (2018).
6. Tong, X., et al.: Normalized Projection Models for Geostationary Remote Sensing Satellite: A Comprehensive Comparative Analysis (January 2019). IEEE Trans. Geosci. Remote Sens. **57**(57), 9643 (2019)
7. Foust J.: SpaceX's space-Internet woes: Despite technical glitches, the company plans to launch the first of nearly 12,000 satellites in 2019 (2019).
8. 3GPP: Study on using Satellite Access in 5G. TR 22.822 (2018). <https://3gpp.org/DynaReport/38-series.htm>.



9. Sacchi, C., Rossi, T., Murrioni, M., Ruggieri, M.: Extremely High Frequency (EHF) Bands for Future Broadcast Satellite Services: Opportunities and Challenges. *IEEE Trans. Broadcast.* **65**(65), 609 (2019)
10. 3GPP.: Solutions for NR to support non-terrestrial networks (NTN). TR 38.821 R16 (2018). <https://3gpp.org/DynaReport/38-series.ht>.
11. Sat5G. Extending 5G Security to Satellites. D4.5 (2019). <https://www.sat5g-project.eu/public-deliverables>.
12. 3GPP.: Security architecture and procedures for 5G system. TS 33.501 (2019). <https://3gpp.org/DynaReport/33-series.htm>.
13. 3GPP.: System Architecture for the 5G System (5GS); Stage 2. TS 23.501 (2019). <https://3gpp.org/DynaReport/23-series.htm>.
14. Jiang J., Yan S., Peng M.: Regional LEO Satellite Constellation Design Based on User Requirements (2018).
15. Xia S., Jiang Q., Zou C., Li G.: Beam Coverage Comparison of LEO Satellite Systems Based on User Diversification (2019).
16. ITU.: Radio Regulations 2016 Edition (2016).

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





# A Survey on Cyberspace Search Engines

Ruiguang Li<sup>1,2</sup>(✉), Meng Shen<sup>1</sup>, Hao Yu<sup>1</sup>, Chao Li<sup>2</sup>, Pengyu Duan<sup>1</sup>,  
and Lihuang Zhu<sup>1</sup>

<sup>1</sup> School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China  
lrg@cert.org.cn

<sup>2</sup> National Computer Network Emergency Response Technical Team/Coordination,  
Center of China, Beijing, China

**Abstract.** This paper introduces the concept of cyberspace search engine, and makes a deep survey on 5 well-known search engines, say Shodan, Censys, BinaryEdge, ZoomEye and Fofa, by querying official websites, analyzing APIs, and making academic research. We discuss the following items in details: Supporting internet protocols, Total amounts of detected devices, Device information, Scanning frequency, System architecture, The third party databases, Probes distribution, etc. We give a comprehensive comparison of the detecting abilities and working principles of the cyberspace search engines.

**Keyword:** Cyberspace search engines

Cyberspace search engines, such as Shodan, Censys, BinaryEdge, ZoomEye and Fofa, are new Internet applications in recent years. They search various types of online devices in cyberspace, such as webcams, routers, intelligent refrigerators, industrial control devices, etc. They are becoming powerful tools to detect network resources. At present, mastering the network resources is valuable for cyberspace governance and network security protection. Therefore, global security companies and scientific research institutions pay great attention on the development and utilization of cyberspace search engines. This paper will carry out a comprehensive investigation and analysis on the detection capabilities and working principles of 5 well-known search engines.

## 1 Introduction

Network resources exploration is to send probe packets to the remote network devices, and to receive and analyze the response data, so as to get the information of remote devices, such as opening ports and services, operating systems, vulnerability distribution, device types, organizations, the geographical position, and so on. The detecting protocols are mainly on the transport layer and the application layer in the TCP/IP stacks. The detection methods of transport layer include SYN scan, TCP connection scan, UDP scan, FIN scan, ICMP scan, etc. Application layer detection mainly uses the special fields of internet protocols, special files, hash values, certificates, and so on.

The working principles of cyberspace search engines are very different from the Web search engines such as Google, Baidu. Web search engines collect, store and analyze

Web page for information querying, while the cyberspace search engines adopt the network resource detecting technology. By sending the detection packet to the remote devices, it can obtain the important information of the target, and conduct comprehensive analysis and display. Global security companies and research institutions have developed a number of search engines, in which the following are most well-known: Shodan ([www.shodan.io](http://www.shodan.io)) Censys (Censys.io) from the US, BinaryEdge ([www.binaryedge.io](http://www.binaryedge.io)) from Europe, and ZoomEye ([www.zoomeye.org](http://www.zoomeye.org)) Fofa ([www.fofa.so](http://www.fofa.so)) from China. Some of these engines are commercially available, while others offer none-profit services.

We are very interested in the detection abilities and the working principles of these search engines, so we made a comprehensive investigation on Shodan, Censys, BinaryEdge, ZoomEye, Fofa, by querying official websites, analyzing APIs, and making academic research. The main contents include: Supporting internet protocols, Total amounts of detected devices, Device information, Scanning frequency, System architecture, The third party databases, Probes distribution, etc.

## 2 Supporting Internet Protocols

Mastering various types of Internet protocol formats is the basis for the exploration of cyberspace search engines. Different devices in the internet have different protocols. In order to facilitate the comparative study, we first carry out a classification of various network devices.

We got all types of devices from the search engine’s official websites, and classify all devices into 11 categories: Network Equipments, Terminal, Server, Office Equipment, Industrial Control Equipment, Smart Home, Power Supply Equipment, Web Camera, Remote Management Equipment, Blockchain, Database, shown as Fig. 1.

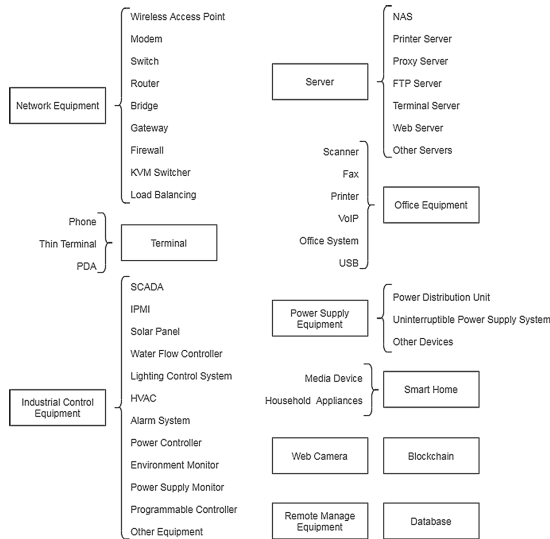


Fig. 1. Device categories

On this basis, we obtained the lists of all engines' supporting protocols from the official websites, user manuals, the APIs, and some technical forums. We classify them into 11 categories according to Fig. 1, shown as Table 1, where "-" means there is no such agreement.

**Table 1.** Supporting internet protocols

	Shodan	Censys	ZoomEye	Fofa	BinaryEdge
Network equipment	10	1	54	7	8
Terminal	19	1	227	6	13
Server	67	10	154	20	63
Office Equipment	12	5	31	6	11
Industrial Control Equipment	26	5	16	23	17
Smart Home	9	-	3	7	9
Power Supply Equipment	4	1	3	2	4
Web Camera	3	-	8	-	3
Remote Management Equipment	13	5	31	8	11
Blockchain	5	-	4	21	4
Database	17	6	19	16	15
Total	185	34	550	116	158

Shodan's API interface contains supporting protocols that can be directly queried [1]. Censys's protocols information comes from the official forum [2]. ZoomEye's protocols information comes from the NMAP-Services file in the user's manual [3]. Fofa's protocols information comes from the technical forum [4]. BinaryEdge's protocols information comes from the API documentation [5]. As you can see in the table, Shodan and ZoomEye have mastered more types of network protocols, covered all protocol categories, and presumably have better device detecting capabilities. Due to the different statistical caliber of network protocols, there may be some deviation in the comparison results.

### 3 Total Amounts of Detected Devices

Based on the analysis in Sect. 2, we investigate the total numbers of detected devices of different search engines. Typically, the official websites will claim the total numbers of detected devices, but sometimes we need to do more auxiliary analyzing.

The total amount of Shodan comes from the official website query tool CLi.shodan.io [6]. All the data records after January 1, 2009 can be inquired by the command line tool, so we can calculate the total number of detected devices.

The official website of Censys provides data statistics function [7]. We divide the IPv4 address space into 256 parts, and retrieve each address block with Censys, and calculate the manufacturer's brands of specific types in the returned results, and then obtain the total number as a summary. The total amount of ZoomEye, Fofa and BinaryEdge are from the official website [5, 8, 9].

**Table 2.** Comparison of the total amount of detectable devices

	Shodan	Censys	ZoomEye	Fofa	BinaryEdge
Total amounts	436489751	111368143	1190860679	270363	89871839

The total numbers of detected devices for each engine are shown in Table 2. As you can see from the table, ZoomEye (nearly 1.2 billion) and Shodan (over 0.4 billion) have the strongest detecting capabilities.

It should be noted that, because of the lack of industry standards in the field of network devices classification, there are statistical caliber of the comparison results.

## 4 Device Information

Cyberspace search engines need to present the detected device information in a comprehensive way for users to use. One device stands for a file or a record. By analyzing the files or the records, we can get the device information architecture. Typically, the device information architecture includes such important information as domain names, opening ports, services, geographic locations, countries, device types, affiliation, and so on.

We collect, analyze and draw the device information architecture of the above search engines, and make a comparison. We can classify all the device information into: Equipment information, location information, port information, loopholes, probe point information, tag information, network equipment information, WEB information, file transfer, email protocol information, remote access to information, database information, industrial control protocol information, message queues, clustering information. This will be of great value to developers and users of the cyberspace search engines.

Taking Censys as an example, by analyzing the official documents of Censys [10], we get the tree diagram of Censys' device information architecture, as shown in Fig. 2. All these information will be reflected on Censys' web pages. In the below figure, the vulnerability information and probe point information are represented as dotted lines because Censys does not provide such information.

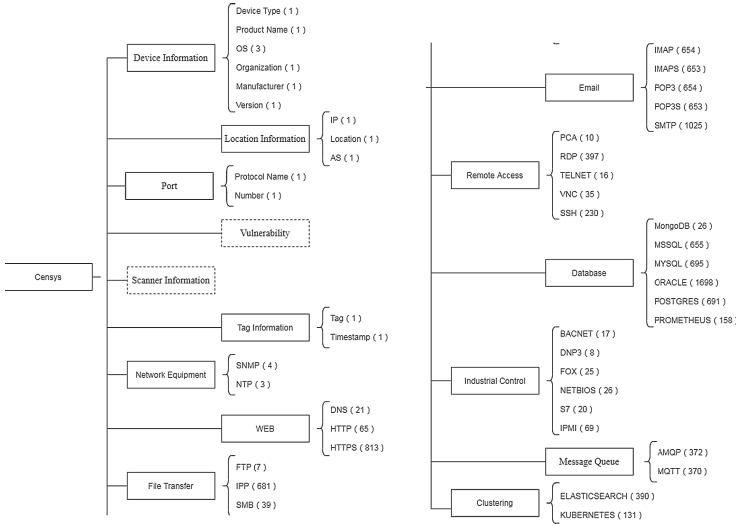


Fig. 2. Device information architecture for censys

## 5 Scanning Frequency

The cyberspace search engines constantly scan and probe the whole network, discover the new connected devices, and periodically update the detected devices. As a complete scan of the whole network consumes lots of computing and storage resources, so search engines usually set a scanning frequency. Scanning frequency is an important index for the detecting ability. The higher the frequency, the stronger the search engines’ performance.

We measured the scanning frequencies of Shodan, Censys, ZoomEye and Fofa. More than 130 IP addresses (opening HTTP, HTTPS, TELNET, FTP and SSH services) were randomly selected. By checking the update status of these IP addresses every day, we can get the scanning intervals of each engines, as shown in Table 3 below.

Table 3. Comparison of scanning frequencies

Protocol (port)	Shodan	Censys	ZoomEye	Fofa
HTTP (80/TCP)	10 days	2 days	389 days	39 days
TELNET (23/TCP)	24 days	2 days	-	-
HTTPS (443/TCP)	9 days	1 day	26 days	102 days
FTP (21/TCP)	13 days	2 days	173 days	74 days
SSH (22/TCP)	10 days	3 days	24 days	60 days

In the above table, “-” means it hasn’t been scanned for a long time. As can be seen from the table, that the scanning frequencies of Shodan and Censys are significantly

higher than that of ZoomEye and Fofa. We can include that Shodan and Censys have more powerful performance.

### 6 System Architecture

We are very interested in the system architectures of the cyberspace search engines, so we conducted an extensive academic research. Typically, the architecture of search engine can be divided into three modules: information acquisition module, data storage module and information retrieval module. The information acquisition module is responsible for collecting the information of various devices in the cyberspace. The data storage module is responsible for storing the massive device information collected, and the information retrieval module is responsible for providing statistical and querying services.

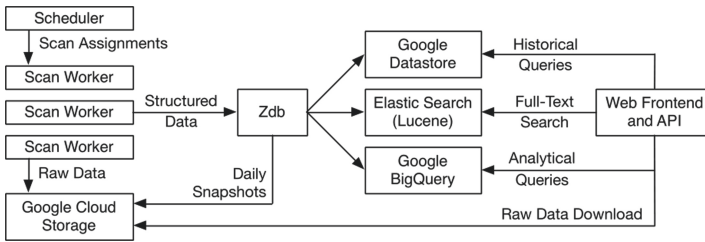


Fig. 3. Censys system architecture1

Figure 3 shows the system architecture of Censys [11]. In the above figure, the Scan Worker is responsible for information acquisition. The Scheduler allocates scanning tasks to multiple scanning modules. The scanning module will save the detection results to Zdb database, and all the information will be stored in Google Cloud. In the information retrieval module, Censys provides elastic Search for full-text retrieval. Google Datastore offers history retrieval and Google BigQuery offers statistics retrieval.

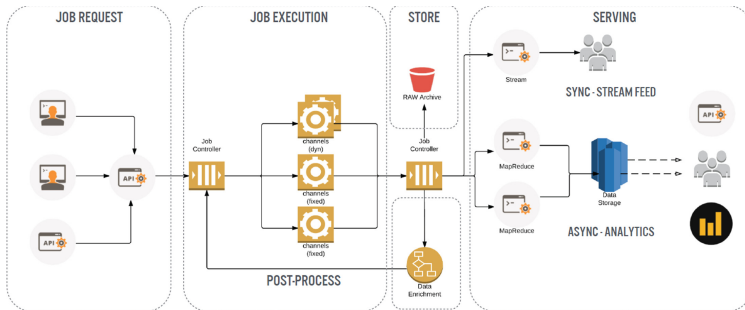


Fig. 4. BinaryEdge system architecture2

BinaryEdge system architecture is shown in Fig. 4 [12], which is divided into four parts: task submission, task execution, storage and service. Task submission uses HTTP,

command line, third party and other forms of API for data acquisition. In the task execution stage, the task is sent to multiple channels, including port scanning, Screen shotter, OCR and other technologies. In the storage stage, the collected information will be divided into original data and processed data, and stored in the database. During the service stage, the processed data will be sent to users through a real-time information flow, or deeply analyzed by MapReduce, Kibana, or InfluxDB.

## 7 Third Party Databases

Many cyberspace search engines work with third-party databases, such as IP databases, domain name databases, and geographic location databases. We investigated the third-party databases associated with commercial search engines, as shown in Table 4 below:

**Table 4.** Search engines associate third-party databases

	Shodan	Censys	ZoomEye	Fofa	BinaryEdge
IP database	Randomly generated	Randomly generated	-	-	-
Domain database	-	Alexa	-	-	Passive DNS
Address database	-	GeoIP	IPIP	GeoIP	GeoIP

In the table, the IP addresses of Shodan and Censys are randomly generated and do not rely on the third-party IP database. We haven't found the information of ZoomEye, Fofa and BinaryEdge. As for the domain name database, Censys used the domain data provided by Alexa Top 1 Million Websites, while BinaryEdge used the passive DNS resolution service. We haven't found the information of Shodan, ZoomEye and Fofa. As for geographic location databases, Censys, Fofa and BinaryEdge all use the database of GeoIP, while ZoomEye uses the database of IPIP.net.

## 8 Probes Distribution

Cyberspace search engines often need to deploy many probes because there are many security devices (such as firewalls) in cyberspace, making it difficult to detect the network edges. Only by deploying widely distributed probes, can we minimize the impact of security devices and find more edge nodes as possible.

We conducted an extensive research, focusing on the open-source tools and third-party organizations. GreyNoise and BinaryEdge have done well.

GreyNoise is a tool for collecting and analyzing scanning traffics [13]. It found the probes of 96 search engines, including Shodan, Censys, BinaryEdge and ZoomEye, as shown in Table 5 below.



**Table 5.** Probes distribution marked by GreyNoise

	Shodan	Censys	BinaryEdge	ZoomEye
United States	31	398	368	-
Canada	-	-	37	-
Britain	1	-	236	-
Netherlands	10	-	86	-
Iceland	2	-	-	-
Romania	1	-	-	-
Greece	-	-	1	-
Germany	-	-	239	-
India	-	-	29	-
Singapore	-	-	27	-
Japan	-	-	-	16

BinaryEdge recorded the contents of received packets(including IP, ports and payloads) which it received by deploying honeypots all around the world. Because the honeypots do not actively interact with other devices, the data received in the honeypots are most likely sent by the probes. Table 6 shows the global probe distribution of Shodan, Censys and BinaryEdge recorded by BinaryEdge during a period of 2000 days.

**Table 6.** Probes distribution marked by BinaryEdge

	Shodan	Censys	BinaryEdge
The United States	17	321	146
Canada	-	-	24
The British	1	-	90
In the Netherlands,	11	-	36
Iceland	2	-	-
Romania	1	-	-
Germany	-	-	115
India	-	-	8
Singapore	-	-	9

## 9 Conclusion

We made a comprehensive research and analysis on the well-known cyberspace search engines such as Shodan, Censys, BinaryEdge, ZoomEye and Fofa. We deeply analyze the items of Supporting internet protocols, Total amounts of detected devices, Device information, Scanning frequency, System architecture, The third party databases, Probes distribution. This paper give an objective evaluation of the detecting abilities and the working principles of the cyberspace search engines by querying official websites, analyzing APIs, and making academic research. We believe this paper will greatly help those who are developing and using cyberspace search engines.

## References

1. <https://api.shodan.io/shodan/protocols>
2. <https://support.censys.io/hc/en-us/articles/360038762031-What-does-Censys-scan->
3. [https://www.zoomeye.org/doc? The channel = user# d - service](https://www.zoomeye.org/doc?The+channel+=+user#d+service)
4. <https://www.freebuf.com/articles/ics-articles/196647.html>
5. <https://docs.binaryedge.io/modules/>
6. <https://cli.shodan.io>
7. [https://censys.io/ipv4/report? Q = &](https://censys.io/ipv4/report?Q=&)
8. <https://www.zoomeye.org/component>
9. <https://fofa.so/library>
10. [https://censys.io/ipv4/help/definitions? Q = &](https://censys.io/ipv4/help/definitions?Q=&)
11. Durumeric, Zakir, et al. "A search engine backed by Internet-wide scanning." Proceedings of the 22ND ACM SIGSAC Conference on Computer and Communications Security.
12. <https://www.slideshare.net/balغان/binaryedge-presentationbsides>
13. <https://greynoise.io/>

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





# Brief Introduction of Network Security Asset Management for Banks

Yumo Wang<sup>(✉)</sup> and Qinghua Zhang

China Everbright Bank CO., Ltd, Beijing 100034, China  
wangyumo@cebbank.com

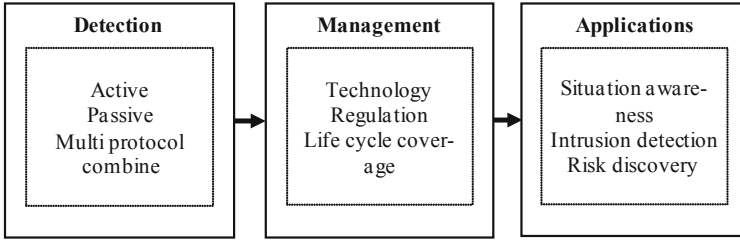
**Abstract.** During the digital development process, enterprises have accumulated a lot of network asset including hardware, software and websites. Effective management of network asset can reduce the internet risk. Network asset is the primary object of information security. Therefore, the essential content of enterprise information security operation is ensuring the security of network assets sufficiently. This paper has investigated researches about detection, management and applications of network assets. The difficulty and current solutions have been summarized by the review. Moreover, this paper puts forward a solution of network asset management according to the bank situation.

**Keywords:** Network asset · Host detection · Security management

## 1 Introduction

With the development of technologies in information security, the demand of management of network assets is increasing in banks. During the digital construction of banks, many network assets have been accumulated including domain name, IP, system, application and so on. The network assets are the main object of information security management in banks. The purpose of managing security assets is to support the information security operation of banks, so it is essential to collect and analyze the security information of network assets. This paper conducts the research on network asset management from three aspects: detection, management and applications. A construction method of security network assets management for bank is proposed (Fig. 1).

According to the controllability of assets, the network security assets of banks are usually divided into two parts: internet and intranet. From the perspective of safety management, both of internet and intranet assets are protection objects that need to be focused on. In the management of network security assets, there are generally three aspects: detection, management and application. Detection means to discovery the security assets in cyberspace. Effective management of assets can only be achieved by timely and accurately detection. At the same time, the method of detection and monitoring are similar. Periodic updating of asset information is also an important part of safe asset management. Management means to clearly counting the proven safety assets, accumulating the detection results, so as to form an asset library that can support information security operation and provide data support for the further development of security works. The



**Fig. 1.** Network asset management process

most important part of asset management is conducting two aspects of constructions: information and regulation. Application means using the managed network security asset data in multiple dimensions in order to embody value of it. The most typical application scenario is active risk discovery. The ability of active risk discovery for security assets can make security operation more accurate and targeted.

In view of the previous three aspects of network security asset management design, this paper conducts a literature review.

## 2 Detection of Network Security Assets

Detection is the starting point of network security asset management. At present, there are three common asset detection methods: active, passive and information hunting based on search engine. With the help of network scanning tools, the active way can obtain information by scanning the host, which has strong pertinence, but it will occupy part of the resources of the target host. The passive way means to the aggregation of transaction information through the carding of network traffic, which an important method in the asset discovery of intranet. Information hunting based on search engine is a non-invasive asset detection method, which can expand the collection field. However, it also depends on the data collection ability of the searching platform [1]. The detection work needs to consider different levels of assets. For the IaaS level, it mainly relies on scanners, network detection, NAT mapping table and other methods to detect network security assets. For PaaS and SaaS level, methods like traffic carding, DNS domain name aggregation are used to gather asset application information [2]. Through the acquisition of network fingerprints, the details of assets can be collected in order to identify website components, application services, communication protocols, which is able to assist the identification of vulnerabilities [3]. The design of active scanning scanner for IP requires different port scanning of TCP and UDP protocols to obtain more comprehensive host information [4, 5]. There are four scanning methods for asset discovery: ICMP, TCP connect, TCP SYN and TCP FIN. In practice, these methods are usually combined to obtain more accurate asset opening information [6]. The complex and changeable asset information needs to be monitored dynamically, and the comprehensive information including multiple dimensions as host, system, internal information should be gathered [7]. In the complex network environment, big data technology can support the asset discovery process and provide technical means for the excavation of massive information [8]. At the same time, vulnerability is also a key information in asset scanning. Periodically vulnerability

mining is very important for banks [9]. In order to discovery vulnerability efficiently, automatic tools like Nmap, Xscan, Goby, Nessus are needed to discover the assets and vulnerabilities [10, 11].

### 3 Management of Network Security Assets

Management is the core content of network security asset management. The method of network security asset management can be divided into two aspects: technology construction and regulation construction. Empirically regulation is more important than technology in network asset management. At present, there are many problems in network security asset management, including insufficient handover, lack of sharing mechanism between different systems, untimely updating, and lack of evaluation process [12]. Though, it is very important for banks to overcome many obstacles in asset management, the management work should be appropriate considering the current situation of banks [13]. Technology construction is an indispensable method for the current network security asset management, which makes the assets fine management and strengthens the achievements of regulation construction [14]. Cloud platform is able to make the deployment of network asset management system more efficient and enable the dynamic monitoring update of asset information [15]. The integrated asset management platform usually includes account management, IP address information, resource check, electronic reminder, baseline verification, vulnerability scanning and other functions to achieve comprehensive technical function support [16, 17]. The management of network security assets needs to cover the whole life cycle of assets. The detection and monitoring needs to contain several processes like asset addition, allocation, change and invalidation. Network security assets need dynamic management, especially focusing on the changes of assets in its whole life cycle. In particular, it is necessary to check and recover the assets in time when it is out of use [18, 19]. The asset information management system based on block-chain technology makes the asset information more complete and consistent. The unchangeable characteristic of block-chain makes the asset data management process more reliable and controllable [20]. The management of network security asset data also requires multi-source fusion technology to integrate data from different sources in order to gather comprehensive information of the asset. Based on the segmentation and vectorization of address information, the cosine similarity between feature vectors is applied to assist the automatic matching and fusion of asset information [21, 22].

### 4 Applications of Network Security Assets

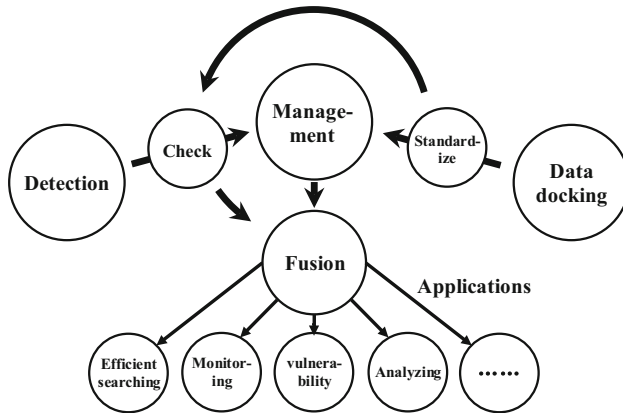
Application for security operations reflects the true value of network security asset management. The purpose of network security assets management is to find risks actively. Situational awareness system is a very practical tool in the current information security operation whose construction progress is highly associated with asset management. To enable active risk detection, many functional parts rely on the network asset management including attack detection, software and hardware information collection, external threat information and so on [23, 24]. This kind of active risk discovery has a good effect on the security of dumb terminals. For example, asset monitoring for dumb terminals such

as video monitoring equipment can assist in detecting network intrusion [25]. Artificial intelligence is a potential technology in situation awareness in which asset data plays an important role and can provide data materials for situation awareness work [26]. Big data technology can also assist the network asset management in security operation. Big data technology provides sufficient storage and rapid searching for massive asset information data and enables multiple applications [27]. Big data technology provides an over-all support for comprehensive asset information management and risk discovery [28]. Vulnerability management also needs network asset management system. The whole processes of vulnerability management starts from discovering assets and includes classification, scanning, repair, tracking, compliance and periodically repetition. In the case of the asset management of FIFTH THIRD BANK in the United States, both management of network security assets security and level of compliance continuity should be paid attention in order to provide a more comprehensive guarantee for the business [29]. Asset lifecycle management can also make each data clear and controllable by assisting the work of data privacy protection which should cover the process generation, use and extinction [30]. Based on the analysis of the network flow, asset baseline is established in order to focus on the dynamic changes in data to guarantee the security of assets [31].

## 5 Design of Network Security Assets Management System

Based on the analysis of the relevant literature on network security asset management, current technologies and theories of network security asset management are isolated, which may be caused by the complexity of asset. Discrete management can be flexibly applied in small-scale and relatively monotonous information management but it is difficult to support complex scenarios such as information security operation with many factors. Therefore, the key of effective management of network security assets is the fusion of multi-source data. Large number of fragmented asset data need to be gathered and mixed together in order to obtain the whole picture of assets. Common asset information includes hardware, software, network, application system, organization management and so on, which involves many aspects of information about network assets (Fig. 2).

Key marking of security assets need to be focused on and be supplemented when necessary. The lack of key attribute marks will hinder the of asset management. For instance, the lack of information of the person in charge of a system will make the responsibility identification unclear. Information attributes can be roughly divided into five aspects: network, software, application, management and vulnerability. In practice, due to the partial accumulation of asset information, the management of security assets does not need start with nothing. Asset information with different attributes is generally stored in different departments of a bank. Therefore, the core problem of banks in asset management is to integrate the fragmented information comprehensively and integrate it to support the security operation. For the supplement of asset information, both detection and docking should be considered. Detecting and supplementing asset information is as important as integrating asset information from multiple channels. Moreover, asset detection is also a method of asset monitoring, which is the most important step in the whole life cycle management to protect asset information timely and accurately.



**Fig. 2.** Design of network asset management

The purpose of safety asset management is to find risks actively. In the multi-dimensional application of network assets, it can include: asset governance, asset full perspective, vulnerability warning, compliance inspection and so on. Asset governance means to discover unregistered assets, which is the most practical application in safe asset management. The asset full perspective means the association and display of asset data from different sources in order to provide multi-directional information for security operation. Vulnerability warning means to match the system, middleware, database, framework and other asset data in vulnerability notification. Auto POC verification tool can make the vulnerability matching more effectively. Compliance inspection means using the recorded asset information to automatically check whether assets meet the baseline regulation. With the support of comprehensive, timely and accurate asset information, security operation can be carried out more effectively.

## 6 Conclusions

Based on the literature review of bank safety asset management, this paper summarizes the detection, management and multi-dimensional application of asset information. A network asset management method suitable for banks is put forward. The conclusions are as listed as follows:

- 1) The detection of network security assets is the starting point. Comprehensive, timely and multi-dimensional detection methods can make the asset management work more effective.
- 2) Management of network security assets is the core. With the support of technology construction and regulation construction, network security assets can make the information security operation easier.
- 3) The aim of asset management is to discover risks actively and multi-dimensional application reflects the true value of management achievement. The network risk facing banks can be minimized.

- 4) At present, banks need to take the problem of fragmental management of data into consideration in network security asset management. It is a practical solution to fully and timely docking and fusing multi-source information from different systems.

## References

1. Wang, C., Guo, Y., Zhen, S., Yang, W.: Research on network asset detection technology. *Comput. Sci.*, 24–31 (2018)
2. Zhang, H., Wang, S., Jin, H., Deng, X.: Detection of operator network asset security management and control technology and solutions. *Guangdong Commun. Technol.* 5–9 (2019)
3. Yao, M., Lu, N., Bai, Z., Liu, Y., Shi, W.: Building method of device fingerprint search engine for network asset vulnerability assessment. *J. Electron.* 2354–2358 (2019)
4. Pei, Z., Li, B., Wang, X.: Logic processing design of IP and port scanning system. *Network Secur. Technol. Appl.* 26–27 (2017)
5. Ding, Y., Gao, Q., He, L.: Design and realization of assets sacn system based on complement protocol. *J. Shanghai Univ. Technol.* 196–200(2010)
6. Yu, X.: Design and realization of TCP/IP network scan strategy. *J. Wuhan Vocational Techn. College*, 54–56 (2009)
7. Li, J., Liu, P., Cai, G.: Dynamic network asset monitoring based on traffic perception. *Inf. Secur. Res.* 523–529 (2020)
8. Deng, X., Jin, H., Wang, S., Zhang, H.: Research on active discovery of IP assets in enterprise open network environment. *Guangdong Commun. Technol.* 2–4 (2019)
9. Lin, P.: Research on web risk scanning of internet information assets of postal enterprises. *Postal Res.* 15–17 (2008)
10. Chen, Z.: Case analysis and practice of web penetration. *Network Secur. Technol. Appl.* 22–24 (2020)
11. Wang, K., Li, Z., Wang, R., Gao, W., Wang, W., Wang, J.: Vulnerability scanning based on Nmap&Nessus. *Commun. Power Technol.* 135–136 (2020)
12. Zou, H.: Exploring the strategy of strengthening network asset management in the communication industry. *Modern State-owned Enterprise Res.* 49–50 (2015)
13. Li, Y.: On the role and importance of IP address planning and management in large and medium-sized enterprises. *Commun. World*, 20–21 (2019)
14. Wang, W.: Study on computer network security management and maintenance in hospital informatization construction. *Technology*, 115–116 (2020)
15. Zhang, X., Yuan, S., Ma, Z., Zhang, M., Gao, F.: Cloud-oriented asset security management scheme. *Post Telecommun. Des. Technol.* 12–15 (2019)
16. Xiao, Y., He, M., Wang, L.: Application research and practice of telecom operators' network asset security management technology. *Guangdong Commun. Technol.* (2018)
17. Song, J., Tang, G.: Research and application of network security situational awareness technology. *Commun. Technol.* 1419–1424 (2018)
18. Yang, X.: Thoughts on implementing dynamic management of network assets in the communication industry. *Chinese Foreign Entrepreneurs*, pp. 68–69 (2014)
19. Xie, R.: Lean management of optical network assets. *Commun. Enterprise Manage.* 24–27 (2017)
20. Zhang, S.: Network security technology based on blockchain. *Inf. Technol. Inform.* 129–131 (2019)



21. Chen, J.: Pre-matching scheme of network asset resources based on weighted cosine similarity. *Telecommun. Technol.* 46–49 (2018)
22. Lei, B.: About operation and maintenance management of IP addresses in enterprise networks. *Network Secur. Technol. Appl.* 106–107 (2019)
23. Yue, J.: Building an e-government network health evaluation platform based on situational awareness technology. *Inform. China*, 44–48 (2018)
24. Xia, Z., Li, L.: Research and design of network security situational awareness system. *Inf. Commun.* 147–148 (2017)
25. Li, H., Huang, X.: Illegal access detection of wireless routing equipment based on asset identification technology. *China Secur.* 101–105 (2019)
26. Xiao, X., et al.: A review of research on security situation prediction technology based on artificial intelligence. *Inf. Security Res.* 506–513 (2020)
27. Zhao, C., Sun, H., Wang, G., Lu, X.: Network security analysis of power information system based on big data. *Electron. Des. Eng.* 148–152 (2019)
28. Ma, Y.: Research on information security and protection of computer networks in the era of big data. *Wind Sci. Technol.* 82 (2020)
29. Hua, R.: Vulnerability management five: ten best practices. *Instrument and Instrument*, 60–62 (2016)
30. Liu, Z.: Theory and practice of Internet finance users' privacy data security. *Secur. Cyberspace*, 11–15 (2020)
31. Cai, G., Liu, P., Li, C.: Analysis of traffic security baseline of government websites. *Inf. Secur. Res.* 537–542 (2020)

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





# Embedded Security-Critical Device Resource Isolation

Xuguo Wang<sup>(✉)</sup>, Shengzhe Kan, and Yeli Xu<sup>(✉)</sup>

AsiaInfo Technologies (Chengdu), Inc., Chengdu 610000, China  
{wangxg6, xuy15}@asiainfo-sec.com

**Abstract.** At present, embedded devices have been widely used in people's daily life, which makes more convenience for the public. However, embedded devices still have security problems, such as automatic driving accidents that may cause casualties. In the field of embedded device security, there are many studies, for instance, OPENTEE for ARM handheld devices, providing a secure execution environment for payment devices, and SGX for Intel desk top devices, for security-critical applications, such as bank teller systems, build a safe operating environment. But it is a lack of correlation between these new and existing technologies. In our opinion, through the combination of mature technology accumulation and embedded devices, the antivirus industry can create a more secure user environment. In this paper, we propose a source isolation method to make the resources used by key processes exclusive. This method uses antivirus software and builds a more comprehensive embedded security system in critical security applications. The experimental results show that the proposed method is effective and safe.

**Keywords:** Embedded system · Security-Related · Resource isolation · Virus

## 1 Introduction

Embedded devices are becoming more and more popular in people's daily life [11], from aerospace, submarine missiles, to TV phones, watches and earphones, etc. They are everywhere, and their roles are becoming more and more important. For example, to reduce vehicles Automatic braking system for security risks.

An embedded device is usually composed of one or more hardware main bodies, and there is usually an embedded system specially developed for hardware running in the embedded device, and the system runs specifically for the device. These programs are divided into critical tasks and non-critical tasks according to different functions. For example, in the on-board embedded system, the automatic braking system is a critical task [14], and the on-board imaging system is a non-critical task. The execution of the critical task must be guaranteed, and it must be completed within the specified time from response to execution and completion. Otherwise serious consequences will occur, and non-critical tasks will be delayed for a few cycles without serious consequences. This article mainly focuses on the security issues of embedded systems with very high penetration rate-on-board systems [13].

The original in-vehicle systems were not connected to the Internet. They usually included imaging systems, air conditioning and ventilation systems, etc. In the disconnected era, people did not have an intuitive understanding of the consequences of such devices being controlled by hackers. However, with the development of in-vehicle systems, for example, Tesla and other electric vehicle manufacturers have taken the in-vehicle systems as their main selling point. They are more humane and smarter, and even allow the driver to let go of his hands and let the in-vehicle system replace people to realize autonomous driving. Such systems are usually networked, and they are connected to non-specially constructed general-purpose networks, so the vulnerabilities in the system will be easily enlarged. According to the description in [2], such systems are more resistant to network security. Poor, and the particularity of the system makes it impossible to use common system security defense measures [3], what will happen after being controlled by hackers? At the Def Con hacking conference in August 2015, hackers claimed that they had 6 ways to control Tesla's vehicle and make it stop. Just imagine if the vehicle happened to be driving on a highway and suddenly stopped, it is likely that there will be more cars connected. Collision, causing immeasurable losses.

Do we could simply install anti-virus software in the vehicle system to solve these problems [9]? the answer is negative. Because embedded devices have relatively large limitations in storage capacity and computing performance [1], and modern on-board systems are a very complex hybrid system, they have as many as 80 control units on average [10]. Distributed in the management of multiple key tasks in the system, the operation of anti-virus software will inevitably consume storage, and may also lock up the resources occupied by key tasks. If the execution of key tasks is affected, such as an automatic braking system, it will be slightly delayed by 1 s. Zhong, stop the car, this may have caused the car crash.

In many hardware platforms, key tasks can be placed in a specific execution environment. For example, on the ARM platform, the key tasks are executed in the security domain of OPENTEE, and on the X86 platform, the key tasks are placed in the security domain of SGX. To execute, in this article, we pass Put the antivirus software in an isolated client similar to OpenTEE or SGX to run, and allocate independent resources, such as memory, cache, etc., for this isolated client. The isolated client is built through hardware virtualization features. Hardware virtualization has a very broad foundation. It is implemented on both ARM and X86 platforms, so the cost of implementation and promotion and use are of practical significance.

## 2 Related Work

Virtualization technology is the foundation of resource isolation technology. This technology guarantees from the bottom layer that the upper layer virtual machines can run their own processes and perform various tasks without interfering with each other. Resource isolation mainly isolates resources are CPU, memory, network and other resources. There are two types of resource isolation according to the implementation scheme: software isolation and hardware isolation.

### 2.1 Resource Software Isolation

Software resource isolation technology, also known as the first type of virtualization technology, mainly isolates the resources of multiple virtual machines through software simulation. This technology has the following problems and the current research status of solving these problems:

- 1) The issue of the authority of privileged instructions. Before the concept of hardware virtualization, the execution of all instructions was simulated by software, including privileged instructions. The way of simulation execution greatly reduced the performance of the system.
- 2) Compression of the process address space. Although the process can be run effectively through software simulation, the run is based on the premise that the address space is compressed, because the process-specific address is allocated to the kernel for use. If you want to fully control the process execution, you must compress the address. space. The consequence of compressing the address space is that the virtual machine’s own management program needs to reserve part of the address space to store important data structures, such as IDT, GDT, etc., through these structures to manage the running image better and easier. However, if a large number of programs are compiled before that, these compressed addresses must be accessed according to the execution mode of the operation. In order to be able to process the request, performance will inevitably decrease, and if not processed, the system will crash.
- 3) Interrupt virtualization. The current research on this aspect has not proposed a good method to improve the way of software simulation to improve the efficiency of the mechanism.

The first type of virtualization technology mentioned above has insurmountable performance defects due to its own architecture. As shown in Fig. 1, VMM running in the user mode is the key to this problem and cannot be improved. This architecture cannot meet the real-time characteristics of the vehicle system.

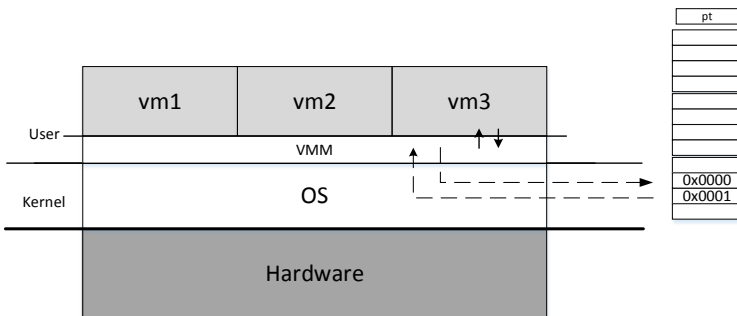


Fig. 1. Basic architecture of resource software isolation

### 2.2 Resource Hardware Isolation

At present, there are not many studies on resource hardware isolation technology for embedded devices, and most of them are concentrated in some research universities.

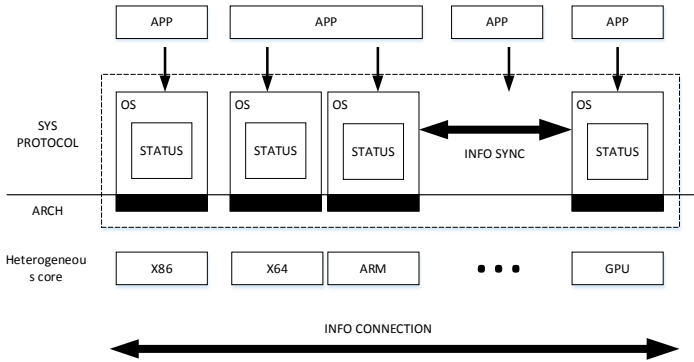


Fig. 2. Multi-core multi-operating system model

Based on the above-mentioned embedded multi-core operating system resource isolation model, this paper proposes a client solution based on a single multi-core, supporting ARM virtualization extension features, running a complete operating system per core, and running a specific strategy for each operating system.

### 2.3 Current Status of Research on Resource Isolation of Vehicle System

At present, most vehicle systems use SoC systems [13], in which the processor architecture uses ARM architecture [12], which provides conditions for us to implement resource isolation and sharing. By isolating the ARM processors of the SoC system according to different usage requirements For example, the driving control system is assigned to run in an independent large core, so that non-critical tasks will not interfere with him, and the control system can respond to user control in a very short time as a key task. Isolating the multimedia system into a dedicated multimedia core can achieve a better experience without affecting key tasks such as the control system.

At present, there are very few studies on resource isolation of vehicle systems that use ARM hardware virtualization extension technology. In the paper of [16], ARM’s TrustZone is used to track and protect the system, but TrustZone cannot provide a complete system simulation. Need to make a lot of changes to the original system, it is more difficult to reform the system. In [17] the paper also uses TrustZone to provide a security zone for the system. The TrustZone method is more suitable for use in the context of processes, such as payment systems, fingerprint recognition systems, etc.

### 2.4 Current Status of Anti-virus Software Running on Vehicle Systems

Due to the particularity of the on-board system itself, it requires extremely high security. Few manufacturers consider security. By hardware isolation of resources, we can provide

a completely isolated environment for the control system and install antivirus the software provides an operating environment.

The client computer installed with anti-virus software is a low-priority client. It runs only when the vehicle is stopped or is being maintained or repaired. Through such a symbiosis method, it can not only provide strong security for the vehicle, but also requests for exclusive resources that would destroy critical mission.

### 3 Modeling of Vehicle Resource Isolation System

#### 3.1 Isolation Model Design

Based on the hardware characteristics described in the ARM platform architecture analysis above, the following hardware isolation model of vehicle system resources is designed.

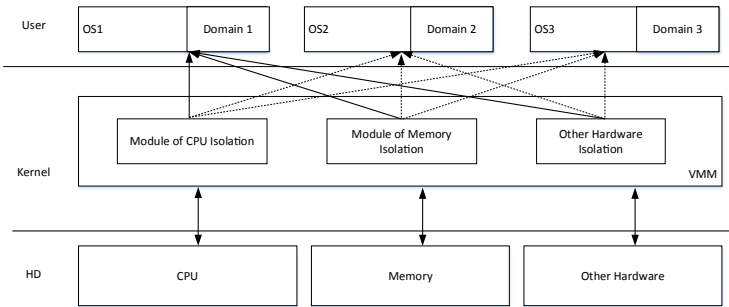


Fig. 3. Resource hardware isolation model

The entire general model is divided into three layers, namely: 1. Hardware isolation layer; 2. VMM management layer; 3. Application program operation layer.

All operating systems on this vehicle-mounted system platform run on the user layer, and the CPU isolation problem of the operating system is completed through the ARM virtual extension mechanism, so that all user-mode operating systems have their own CPU, memory and hardware devices, which are isolated from each other. Do not interfere.

#### 3.2 Implementation and Verification of on-Board System

In the system design, we divide the system into three layers: 1. Hardware abstraction layer. 2. The virtual machine management layer is the VMM layer. 3. The application layer is the system where the control system, multimedia system and anti-virus software are installed. This chapter mainly designs the virtual machine management layer, namely the VMM layer, which provides a complete virtual environment for various applications through the VMM layer, and assigns different permissions to different applications.

### 3.2.1 CPU Isolation Design

The isolation of the physical processor is the most critical design [15], because if the processor resource isolation design is not perfect, then there may be the possibility of non-critical tasks affecting critical tasks, which will greatly affect the stability and reliability of the system. A big threat, especially when used in vehicle systems, has unimaginable consequences (Fig. 4).

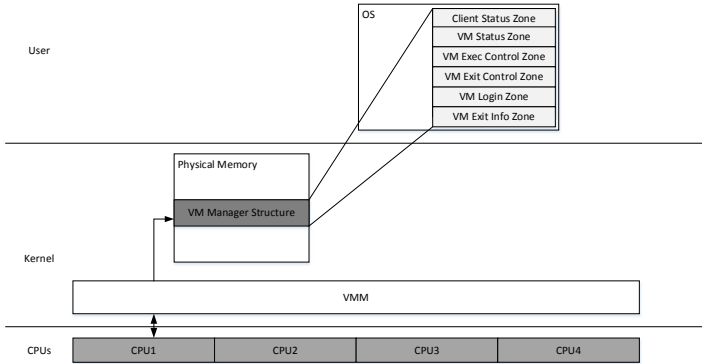


Fig. 4. Processor isolation model

Most modern ARM processors are multi-core architectures, such as ARM-v7 or ARM-v8 processors that can provide virtualization extensions, most of which are multi-core processors, and each processor can run in parallel and used in isolation, So that the applications in each client are isolated from each other without interference.

The processor isolation architecture is as shown in the figure above. The VMM corresponding to each physical processor is described by a virtual machine management structure, which includes the following parameters: 1. Client status area. 2. Host status area. 3. The VM executes the control domain. 4. The VM exits the control domain. 5. The VM enters the control domain. 6. The VM exits the information domain.

Each running client is described by this abstract structure, stored in memory, and pointed to by a special register.

### 3.2.2 Device Isolation Design

The vehicle system has about 80 devices with an average value that need to be controlled, so the design of device isolation is also particularly important, especially the allocation of different device controls for different clients, and the static setting of priorities.

The isolation design of the entire in-vehicle equipment is less difficult, but the workload is huge, because the consistency of the peripheral equipment is poor, and various bus standards are used, such as AHB (high-performance bus), ASB (system bus) and APB (peripheral) Bus), etc., so our current isolation scheme only isolates the AHB standard, and other standards will be considered and supported later.

1. AHB Device principle (Fig. 5)

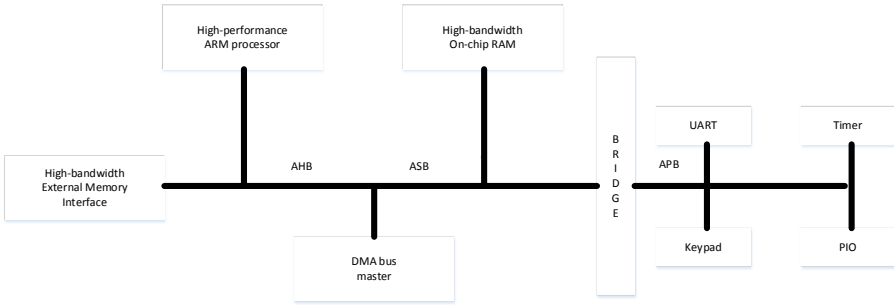


Fig. 5. ARM's typical AMBA bus architecture

From the bus architecture in Fig. 6, AHB is in the high-speed device link. It is used by most SoC systems and is mainly used in high-performance, high-clock frequency system structures, such as high-speed RAM, NAND FLASH, etc. In the connection between DMA and Bridge, this bus is also a device bus used for critical tasks, so this isolation prioritizes the completion of device isolation on the AHB bus to ensure the reliability and availability of key tasks for vehicle equipment.

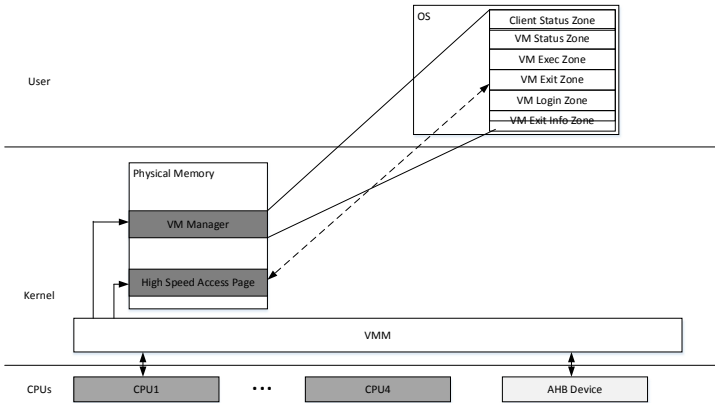


Fig. 6. AHB Device Isolation

A research statement on software reliability mentioned that errors in device drivers account for 70% or more of the entire system. The isolation of the above device drivers can also effectively reduce the system error rate and improve system reliability.



## 2. High-speed memory isolation

ARM has added the feature of Second-stage of translation (secondary translation) to the hardware, and supports the conversion of the physical addresses of all clients into actual physical addresses, instead of using shadow page tables (Fig. 7).

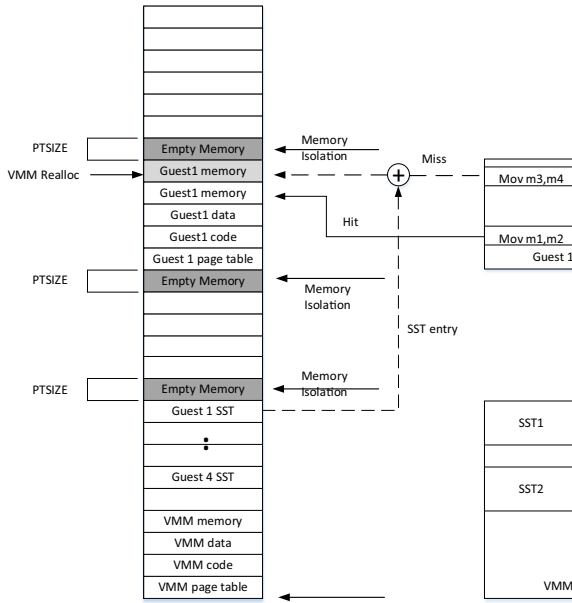


Fig. 7. High-speed memory isolation

In this paper, in the process of initializing the client by VMM, each client is allocated a completely isolated memory area. Each client maintains its own page table. When the client's GPA (Guest Physical Address) to HPA (Host Physical Address) After the query of Address) hits, with the assistance of ARM virtualization technology, there is no need to switch to the host machine. In this way, performance can be improved and the interference between different clients can be isolated.

### 3.3 Test and Analysis

In this vehicle-mounted resource hardware isolation system, it mainly involves the involvement and realization of CPU isolation, peripheral equipment and memory isolation, etc. This article mainly verifies the correctness of several aspects by writing a set of test cases: 1. CPU isolation characteristic test. 2. Peripheral equipment isolation characteristic test. 3. Memory resource isolation characteristic test. 4. Performance test. The tests are in two Linux clients. The test case judges whether the expected function is completed according to the result output by the terminal.

### 3.3.1 Test Environment

At present, the vehicle-mounted platform can already run on the NVIDIA Jetson TK1 kit, but in order to obtain the visualization parameters for testing, this article runs the platform in the Qemu environment. In order to maintain consistency, all parameters are simulated NVIDIA Jetson TK1 parameters (Table 1).

**Table 1.** TestBed

Type	Version
Kernel	Linux 4.2.0-16-generic
CPU	NVIDIA Jetson TK1
AHB Bus Frequency	1000 MHz
Disk	Samsung SSD 850 EVO 120 GB
Memory	DDR3 4096 MB*2
QEMU	qemu-2.3.0

### 3.3.2 CPU Isolation Testing

The CPU isolation of VMM allocates a unique CPU representation structure in the memory for each client when the client is started. Each structure is maintained by VMM. In this way, the isolation is completed and the resource usage of each client will be mapped to this structure, and the allocated resources are represented by the following structure (Fig. 8):

```

struct cell {
    struct kobject kobj;
    struct list_head entry;
    unsigned int id;
    cpumask_t cpus_assigned;
    u32 num_memory_regions;
    struct jailhouse_memory *memory_regions;
#ifdef CONFIG_AHB
    u32 num_AHB_devices;
    struct jailhouse_ahb_device *ahb_devices;
#endif /* CONFIG_AHB */
};

```

**Fig. 8.** Cell Structure

In the configuration of the client linux-arm-demo, there are a total of 4 cores, and the client occupies a total of 3 CPU. After enabling the client, 3 CPU outputs are obtained through lscpu output (Fig. 9).

There are 4 CPUs in total. After the client is started, 3 CPUs 1–3 are occupied.

```

-> ~ lscpu
Architecture:          armv8
CPU op-mode(s):      32-bit, 64-bit
Byte Order:           Little Endian
CPU(s):              4
On-line CPU(s) list: 0
Off-line CPU(s) list: 1-3
Thread(s) per core:  1
Core(s) per socket:  1
Socket(s):           1
NUMA node(s):       1
Vendor ID:           15
CPU family:          6
Model:              15
Model name:          ARMv8 Processor rev 5
Stepping:            1
CPU MHz:             3192.569
BogoMIPS:            6385.13
Virtualization:
L1d cache:          32K
L1i cache:          32K
L2 cache:           4096K
NUMA node0 CPU(s): 0
    
```

Fig. 9. CPU testing info

### 3.3.3 PCI Device Isolation Testing

After enabling QEMU-VM, all AHB devices in all systems are traversed and output to the console, as shown in Fig. 1. Enable the client PCI-demo. The configuration file of this demo only applies for a PCI device from VMM, as shown in Fig. 2. According to the expected idea, the client successfully applies for a PCI device from QEMI-VM: 00:1b.0. After applying, use the device, and after using it, release the device, as shown in the figure below, the device is successfully returned to QEMU-VM (Figs. 10 and 11).

```

Adding PCI device 00:01.0 to cell "QEMU-VM"
Adding PCI device 00:02.0 to cell "QEMU-VM"
Adding PCI device 00:1b.0 to cell "QEMU-VM"
Adding PCI device 00:1f.0 to cell "QEMU-VM"
Adding PCI device 00:1f.2 to cell "QEMU-VM"
Adding PCI device 00:1f.3 to cell "QEMU-VM"
Adding PCI device 00:1f.7 to cell "QEMU-VM"
Adding virtual PCI device 00:0f.0 to cell "QEMU-VM"
Page pool usage after late setup: mem 179/1498, remap 65606/131072
Activating hypervisor
    
```

Fig. 10. List PCI Devices In different VMs

```

Removing PCI device 00:1b.0 from cell "QEMU-VM"
Adding PCI device 00:1b.0 to cell "pci-demo"
Created cell "pci-demo"
Page pool usage after cell creation: mem 196/1498, remap 65606/131072
Started cell "pci-demo"
    
```

Fig. 11. Remove the PCI Device from a VM

### 3.3.4 Memory Isolation Testing

Enter the address belonging to the client and return the valid result. Enter a physical address that does not belong to the client, and invalid is returned. As shown in the figure, the output result of the vra program is the same as the expected result, indicating that the memory isolation is executed successfully, and different clients can only access their assigned physical address space, but cannot access the address space of other clients (Fig. 12).

```
→ project sudo vra cell list
ID      Name      State      Assigned CPUs      Failed CPUs
0       QEMU-VH  running    0,1,3
1       pci-demo  running    2
→ project sudo vra cell vra pci-demo 0x3f000000
3.4
request of the physical address 0x0a3f00000 of pci-demo cell is invalid
→ project sudo vra cell vra pci-demo 0x3f100000
3.4
request of the physical address 0x0a3f10000 of pci-demo cell is valid
→ project
```

Fig. 12. Memory usage

### 3.3.5 Install a Simple Behavior Detection Engine Testing

First run the linux-arm-demo client in VMM, which is used to run the anti-virus software runtime environment. As shown in Fig. 1, next run a behavior detection engine in the client. As shown in Fig. 2. As shown in Fig. 3, the client runs a centos image, which contains basic shell tools that can perform read and write operations. This also provides a runtime basis for the behavior detection engine that needs to be run in this experiment (Fig. 13).

```
ID      Name      State      Assigned CPUs      Failed CPUs
0       QEMU-VH  running    0
1       linux-arm-demo  running    1,3      2
```

Fig. 13. VMs engine running status

The behavior detection engine can start, run, and detect. After starting the mirror, you can read and write files in the mirror. It can be seen from Fig. 3 that the behavior detection engine can run the program normally, download the centos image and start it, and the image contains the engine that can read and write basic files, reaching expectations.

## 4 Conclusion

In this article, we propose a new mechanism to make the new security technologies works with the existing ones. The experimental results shows it. But the system is a rudimentary form. If the virus detection and killing mechanism want to perfectly operated in modern vehicle systems, many other measures are needed to involved, to improve the reliability of the system, such as the division of resources and priority. For the classification of processes' levels, the system itself also needs to pass the certification of security standards, but this scheme effectively uses the current general virus detection and killing mechanism, which reduces the time cost of this research and makes it easier to use previous successful experience.

**Acknowledgments.** I would like to thank AsiaInfo Technologies (Chengdu), Inc for its support in this research, not only for the sufficient arrangement of time, but also for the support of hardware and personnel. At the same time, I would like to thank my chief Yeli Xu, my boss Liang Yu, and my colleague Shengzhe Kan assisted in the development.

## References

1. Zhou, S.: On embedded network security technology. *Network Security Technology and Application* (in Chinese)
2. Wang, Z.: Research on network security technology based on embedded operating system. *Network Security Technology and Application* (in Chinese)
3. Du, G., Wang, L.: About the security analysis of embedded network firewall. *Network Security Technology and Application* (in Chinese)
4. Lutz, R.: Analyzing software requirements errors in safety-critical. *Embedded Syst.*, 126–133 (1993). <https://doi.org/10.1109/isre.1993.324825>
5. Kane, A.: *Runtime Monitoring for Safety-Critical Embedded Systems* (2015)
6. Nejati, S., Alesio, S.D., Sabetzadeh, M., et al.: Modeling and analysis of CPU usage in safety-critical embedded systems to support stress testing. In: *International Conference on Model Driven Engineering Languages and Systems*. Springer, Heidelberg (2012)
7. Popek, G.J., Goldberg, R.P.: Formal requirements for virtualizable third generation architectures. *Commun. ACM* **17**(7), 412–421 (1974)
8. Krammer, M., Martin, H., Karner, M., Watzenig, D., Fuchs, A.: System Modeling for Integration and Test of Safety-Critical Automotive Embedded Systems, 2 (2013). <https://doi.org/10.4271/2013-01-0189>
9. Ding, Y.: Thinking of virus killing mechanism of embedded devices. *Information Security and Communication Secrecy* (in Chinese)
10. Luo, J., Hubaux, J.-P.: Embedded Security in Cars. *Embedded Security in Cars – Securing Current and Future Automotive IT Applications* (2005). [https://doi.org/10.1007/3-540-28428-1\\_7](https://doi.org/10.1007/3-540-28428-1_7)
11. Babar, S., Stango, A., Prasad, N., et al.: Proposed embedded security framework for internet of things (IoT). In: *IEEE International Conference on Wireless Communication*. IEEE (2011)
12. Othman, N.A., Aydin, I., Karakose, M.: An efficient embedded security system for reduce car accident to build safer world based on IoT. In: *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)* (2019)
13. Ali, S., Al Balushi, T., Nadir, Z.: *Embedded Systems Security for Cyber-Physical Systems* (2018)
14. Kevan, T.: Facing the embedded security crisis. *Desktop Eng.* **23**(9), 16–19 (2018)
15. Nathi, R.A., Sutar, D.S.: Embedded payload security scheme using CoAP for IoT device. In: *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)* (2019)
16. Ye, H.: Security protection technology of cyber-physical systems. *Int. J. Security Appl.* **9**, 159–168 (2015)
17. Dong, P., Han, Y., Guo, X., Xie, F.: A systematic review of studies on cyber physical system security. *Int. J. Security Appl.* **9**, 155–164 (2015)

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



# Author Index

- Bai, Jinghua 97
- Cao, Huan 193
- Chen, Xiaoming 16
- Chen, Xuguang 37
- Chen, Yue 193
- Dai, Guangchong 117
- Duan, Pengyu 206
- Gao, Junshi 16, 117
- Gao, Yuxi 71
- Gu, Guofei 50
- Gu, Jie 133
- Guan, Hongchao 141
- Guan, Shijie 50
- Guo, Yanchun 16
- Han, Xinhui 50
- Huang, Daochao 29
- Huang, Haibo 71
- Ji, Pujun 37
- Kan, Shengzhe 222
- Lei, Zhengchao 193
- Li, Chao 206
- Li, Ruiguang 86, 206
- Li, Tongxin 50
- Liu, Haiting 37
- Liu, Huichuan 175
- Liu, Jiale 175
- Liu, Weixin 97
- Liu, Yan 37
- Liu, Zhihong 175
- Luo, Tianyue 3
- Ma, Hongbin 37
- Ma, Huan 165
- Ma, Jianfeng 175
- Ma, Yuan 141
- Rao, Yu 97
- Rui, Zhiqing 3
- Shao, Yanjie 3
- Shen, Meng 206
- Shen, Shijun 29
- Su, Yongtao 193
- Sun, Shuo 16
- Wang, Haitao 117
- Wang, Wei 165
- Wang, Weiqi 86
- Wang, Xuguo 222
- Wang, Yumo 215
- Wu, Bin 3
- Wu, Jingzheng 3
- Wu, Lili 193
- Wu, Yanjun 3
- Xu, Dawei 86
- Xu, Yeli 222
- Yan, Hanbin 97
- Yan, Hanbing 141
- Yan, Min 71
- Yang, Guangliang 50
- Yang, Mutian 3
- Yang, Tianpu 16, 117
- Yu, Guorui 50
- Yu, Hao 206
- Yu, Yuqi 141
- Zeng, Yong 175
- Zhai, Rui 117
- Zhang, Jialong 50
- Zhang, Jiuling 29
- Zhang, Qinghua 215
- Zhang, Xiaofan 71
- Zhao, Chunping 193
- Zhou, Hao 97, 141
- Zhu, Liehuang 86
- Zhu, Lihuang 206
- Zhu, Tian 97
- Zhu, Xiaoyan 175