

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

Ruiguang 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

Communications in Computer and Information Science

ISBN 978-981-33-4921-6

ISSN 1865-0937 (electronic)

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

Perosonalized 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 Rui¹ , Jingzheng Wu¹ , Yanjie Shao¹, Tianyue Luo¹, Mutian Yang^{1,2},
Yanjun Wu¹, and Bin Wu¹

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

² 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

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.

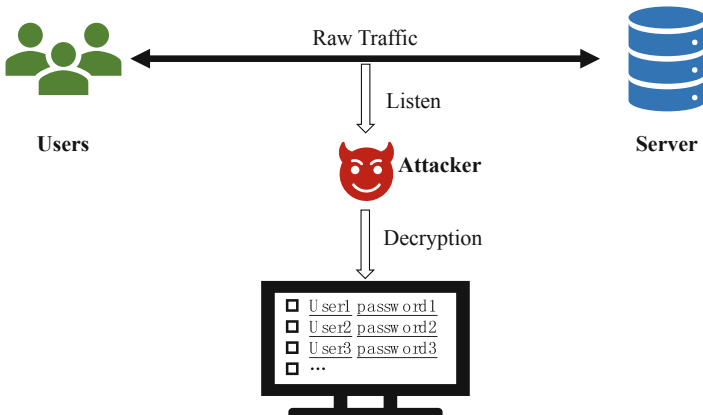


Fig. 1. Traditional password sniffing attack model.

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.

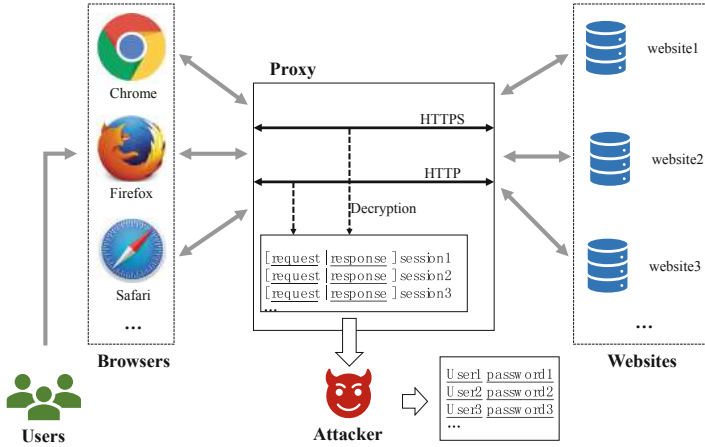


Fig. 2. HTTP session password sniffing (HSPS) attack model.

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.

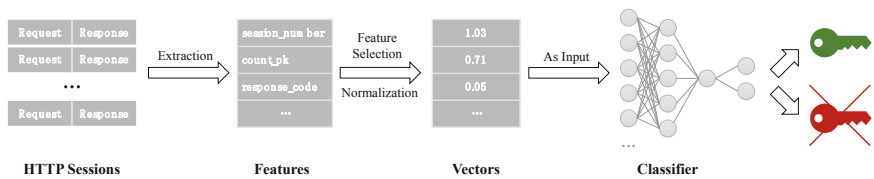


Fig. 3. An overview of the PassEye.

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, μ denotes the arithmetic mean, σ 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.

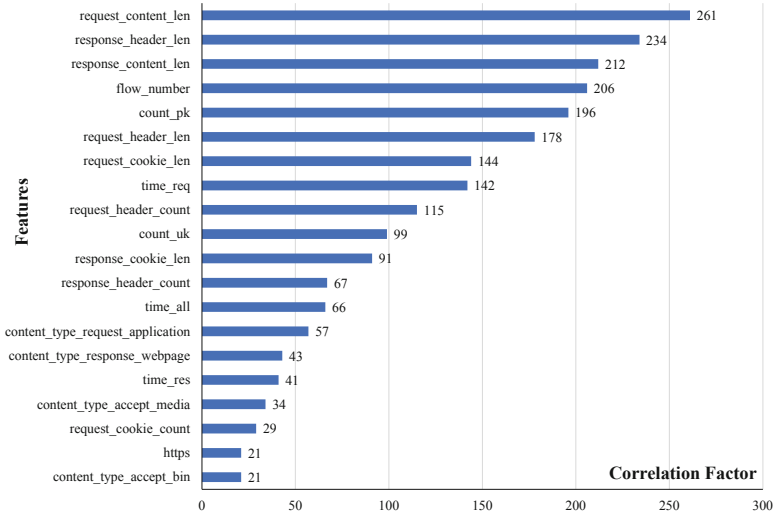


Fig. 4. Correlation of the top 20 features.

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	99.38%	99.46%	99.20%	99.33%

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/utils/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
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/1709.02656) [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^(✉), 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–Macao 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^(✉), 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.

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

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

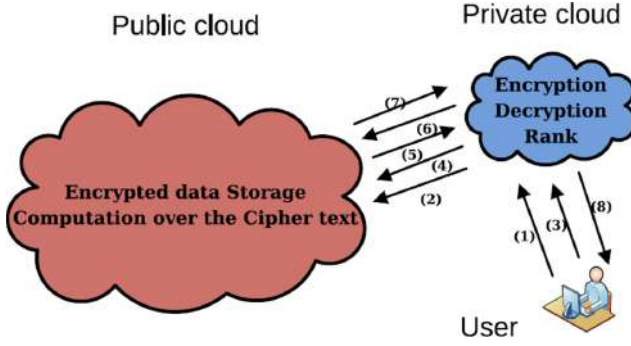


Fig. 1. Schematic diagram of secure ranked search model

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.

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

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

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

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.

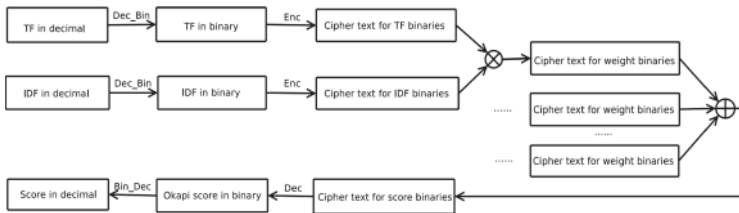


Fig. 2. Implementation of the FHE based ranking

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 Chen¹, Hongbin Ma¹(✉), Pujun Ji¹, Haiting Liu¹, and Yan Liu²

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

² 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.

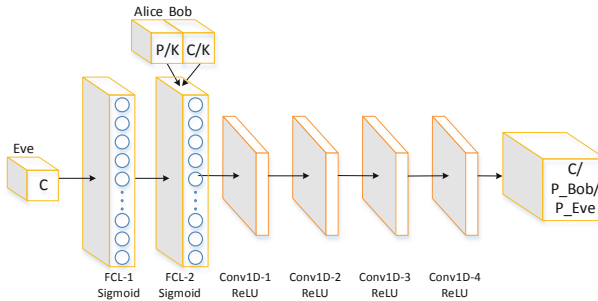


Fig. 1. The network structure of Alice, Bob and Eve.

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 θ_A and θ_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.

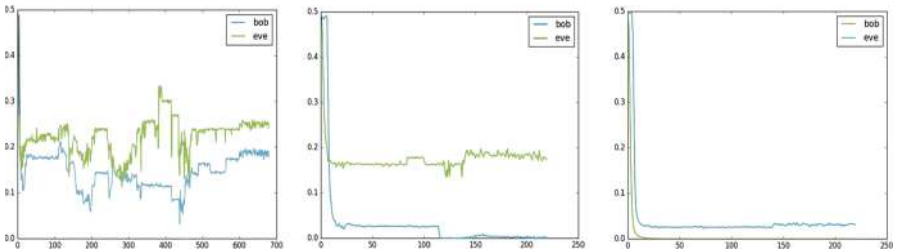


Fig. 2. The decryption of Bob and Eve when the amount of leaked plaintext is 8, 10, and 16 bits respectively.

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 μ 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 μ 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

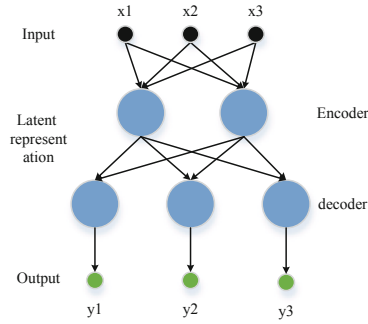


Fig. 3. Basic structure of self-encoder.

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.

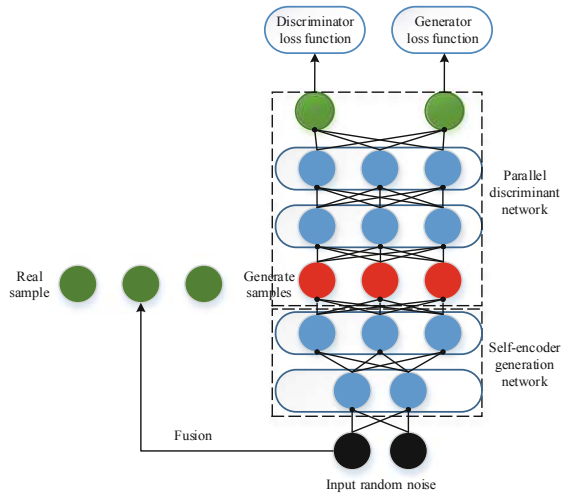


Fig. 4. Parallel training network structure.

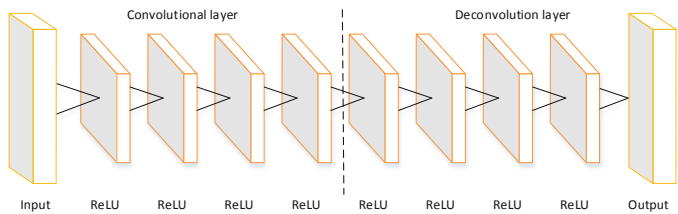


Fig. 5. Generate network structure.

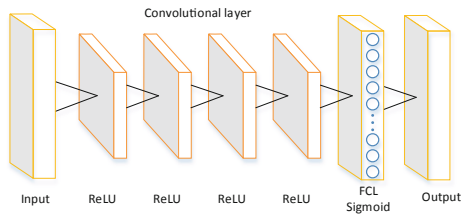


Fig. 6. Discriminate network structure.

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.

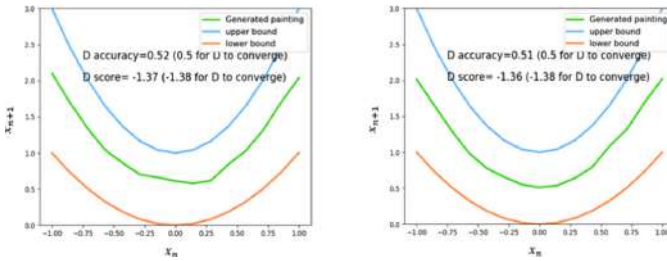


Fig. 7. The number of iterations is 5000 and the discriminator output accuracy is 0.5.

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

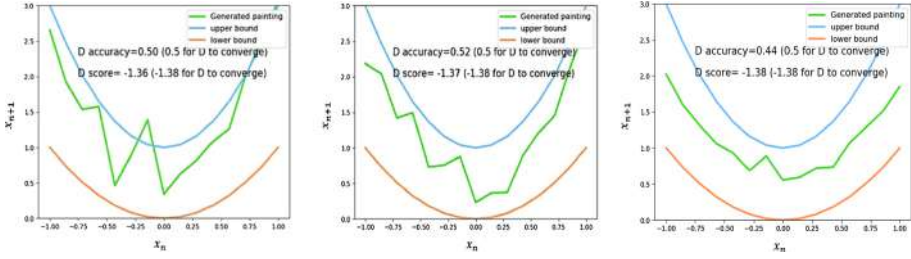


Fig. 8. The number of iterations is less than 500, or the discriminator output accuracy is not selected properly.

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.

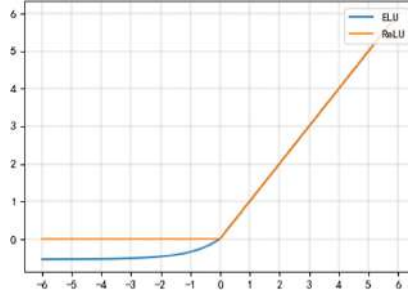


Fig. 9. Comparison of ReLU function and ELU function.

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).

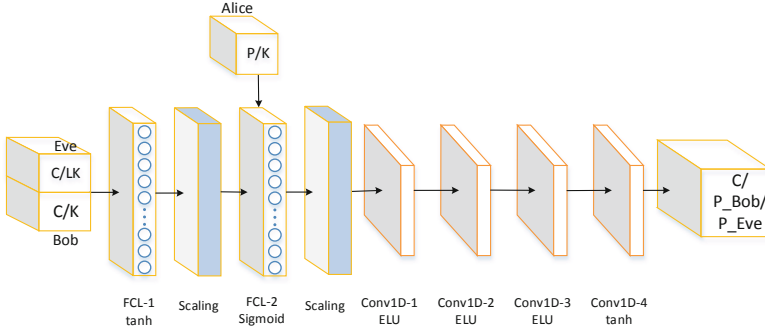


Fig. 10. Improved network model of Alice, Bob and Eve.

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:

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

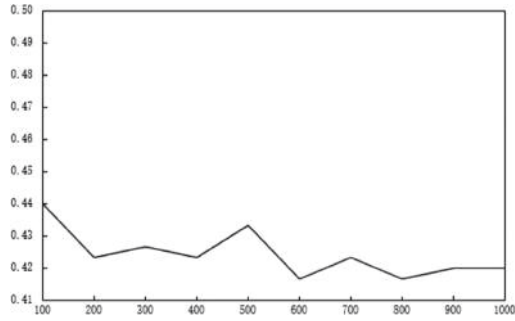
The evaluation process is shown in Table 5.

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.

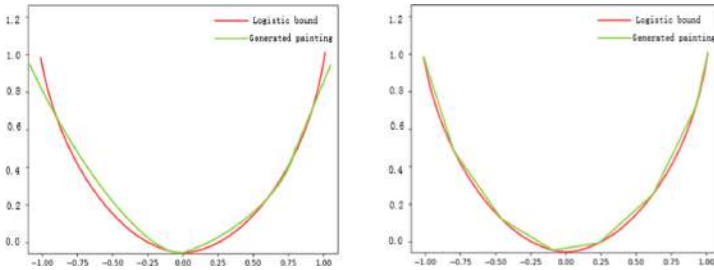
To prevent attackers from attacking through fuzzy keys, the sensitivity of keys must be high enough, and chaotic systems can satisfy this characteristic well [7]. To verify the high sensitivity of the key generated by the key generator, $\mu = 4$ was set in the experiment. The initial value of $x_{01} = 0.256$ 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.

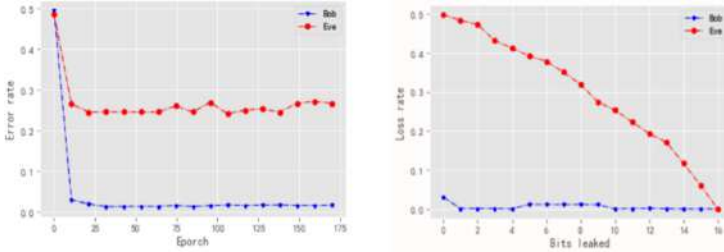


Fig. 13. Results of the security evaluation of the encryption model.

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
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

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 Yu¹, Guangliang Yang², Tongxin Li¹, Xinhui Han¹(✉), Shijie Guan¹, Jialong Zhang³, and Guofei Gu²

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

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

³ 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.

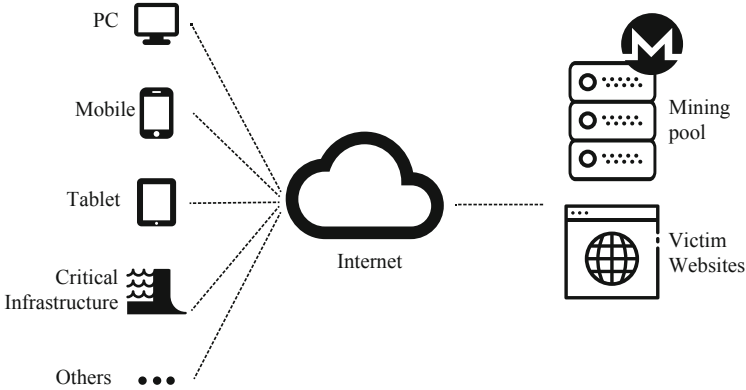


Fig. 1. Example attack scenarios

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 predefined 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.

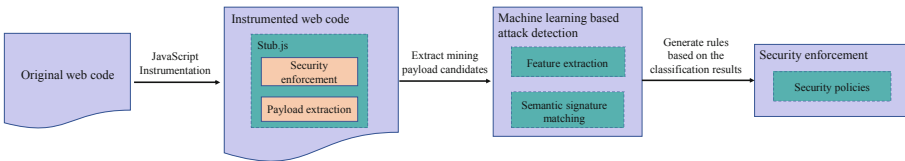


Fig. 2. MinerGate’s workflow

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 pronumeral, 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 global 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.

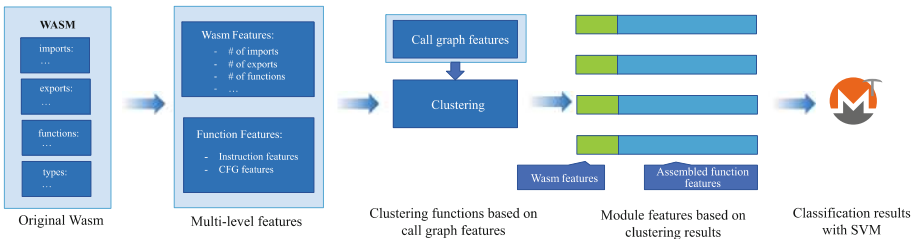


Fig. 3. The classification flow of a WebAssembly module.

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 χ^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: ϵ -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 ϵ -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 $\text{MinPts} = 4$, 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.

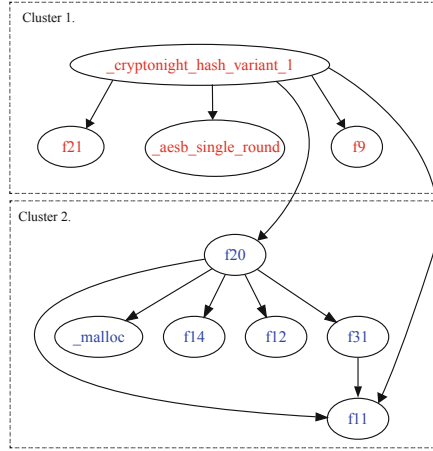


Fig. 4. The clustering result on CoinHive with $MinPts = 4$, only includes functions that are related to `cryptonight_hash_variant_1`.

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.

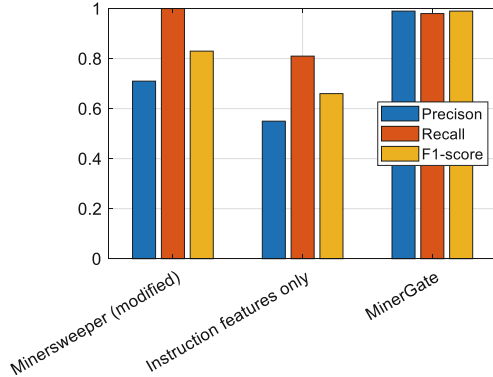


Fig. 5. Results of Cryptojacking discrimination and comparison to other approaches.

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.

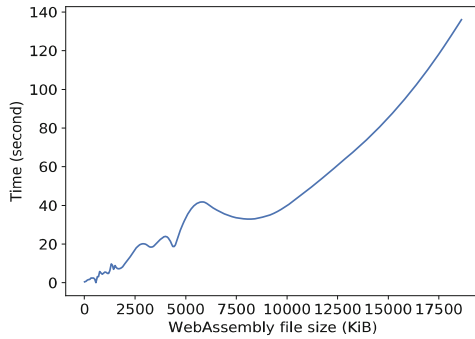


Fig. 6. The time for processing different sizes of WebAssembly in the background.

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. 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
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.: Nocoins (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
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^{1,2} , Yuxi Gao² , Min Yan³ , and Xiaofan Zhang²

¹ Beijing University of Posts and Telecommunications, Beijing 100876, China
poehuang1@163.com

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

³ 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

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.

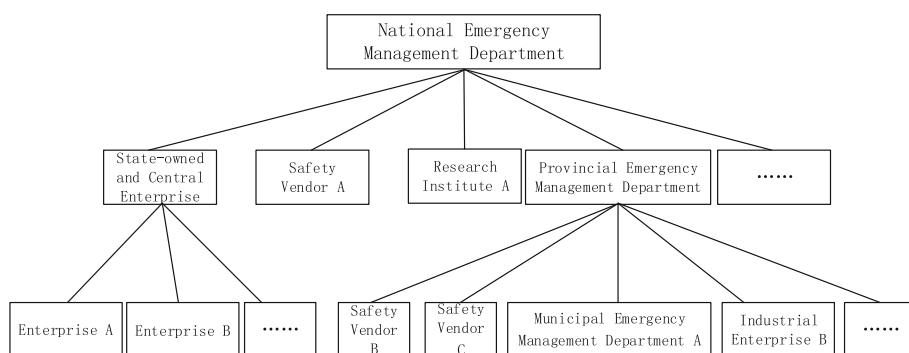


Fig. 1. Management structure of China's ISEMS

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

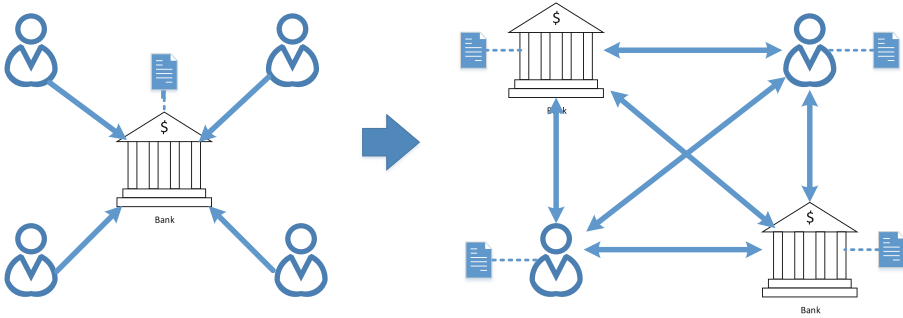


Fig. 2. Centralization and decentralization

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 peers 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

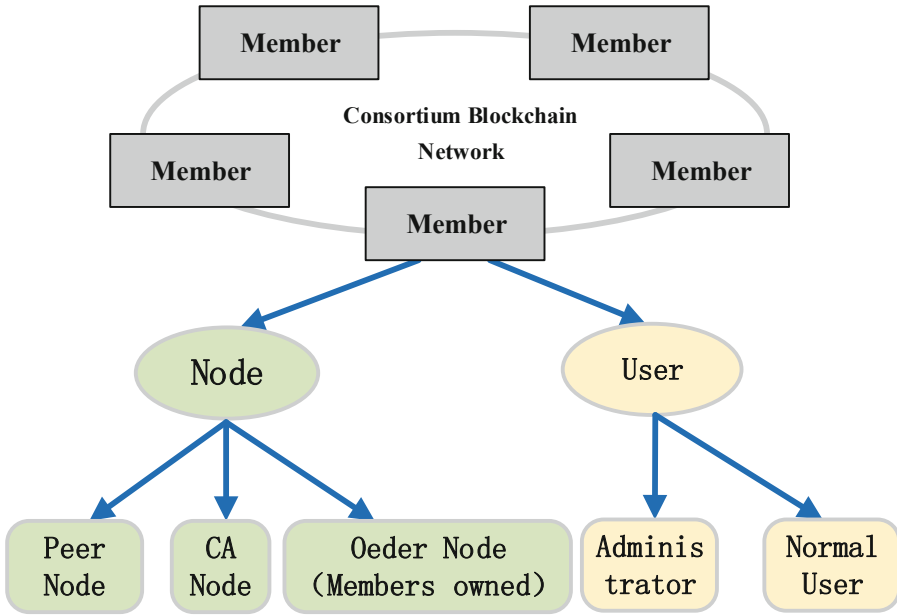


Fig. 3. Consortium blockchain network architecture model

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

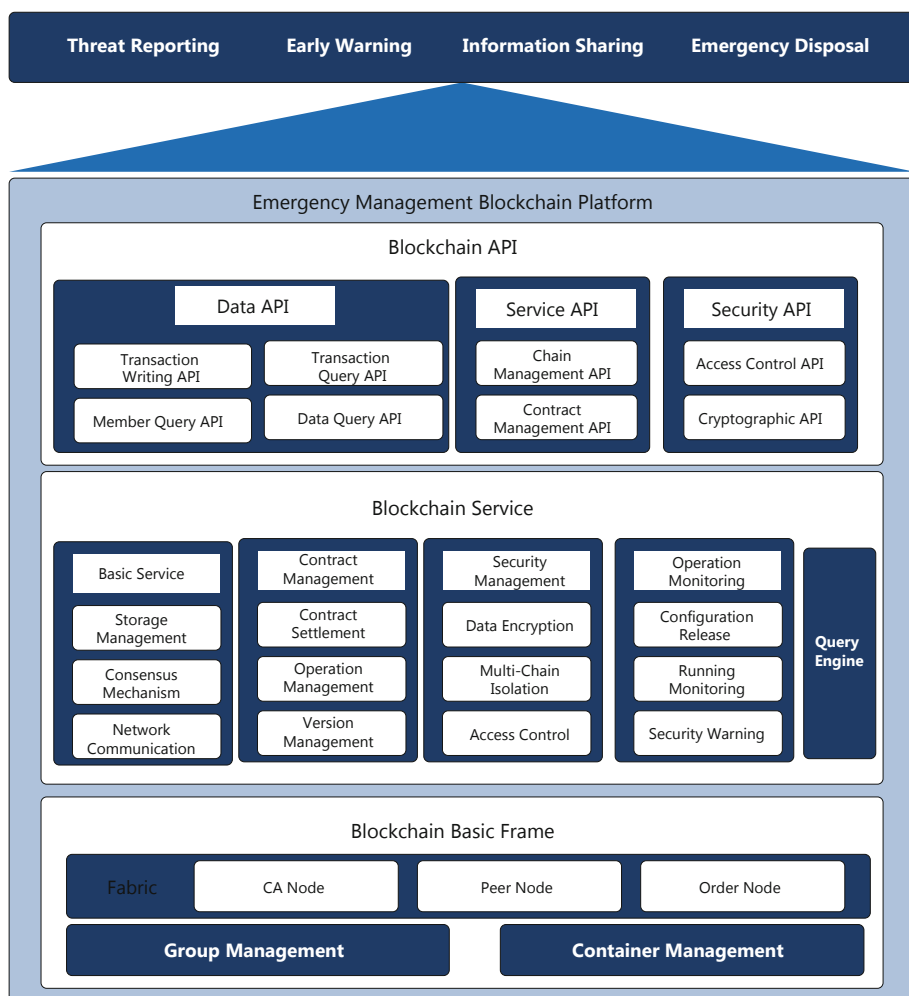


Fig. 4. Technical framework of ISEMS based on consortium blockchain

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.

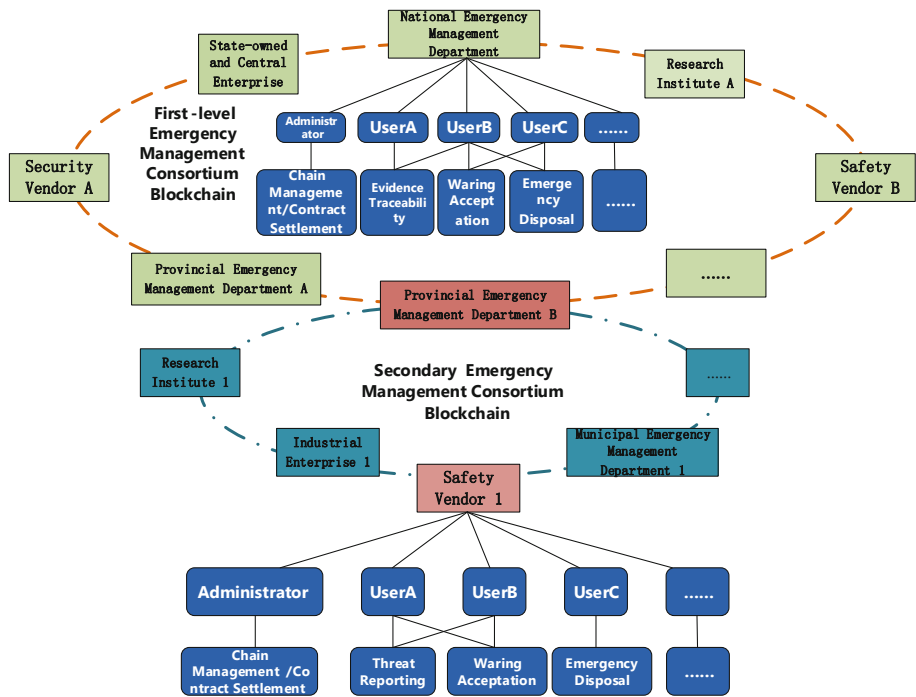


Fig. 5. Organizational structure of ISEMS based on consortium blockchain

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.

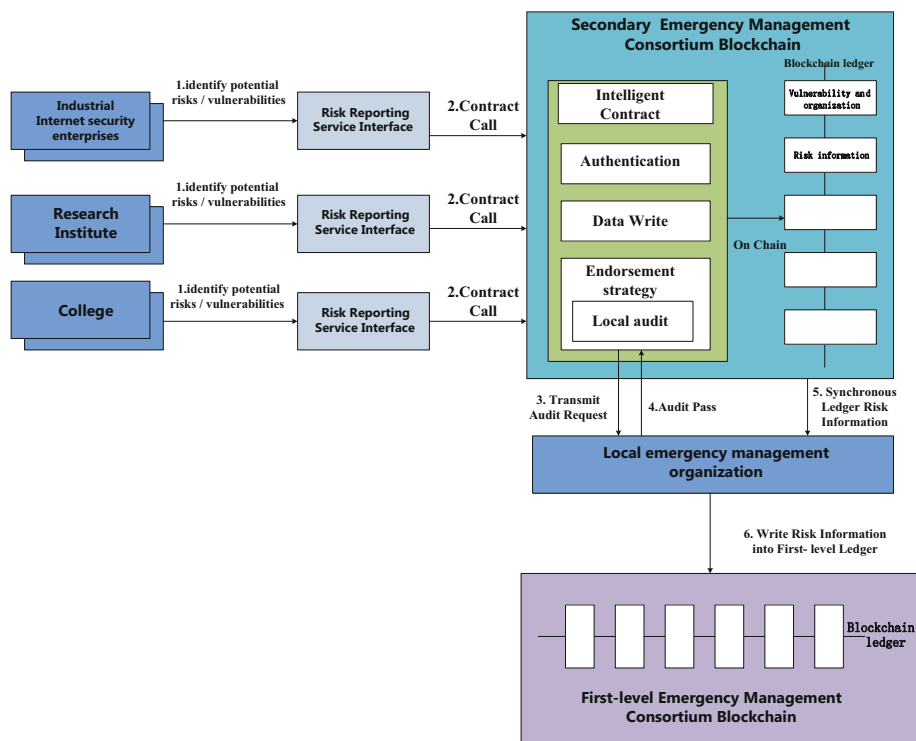


Fig. 6. Threat reporting process based on blockchain

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

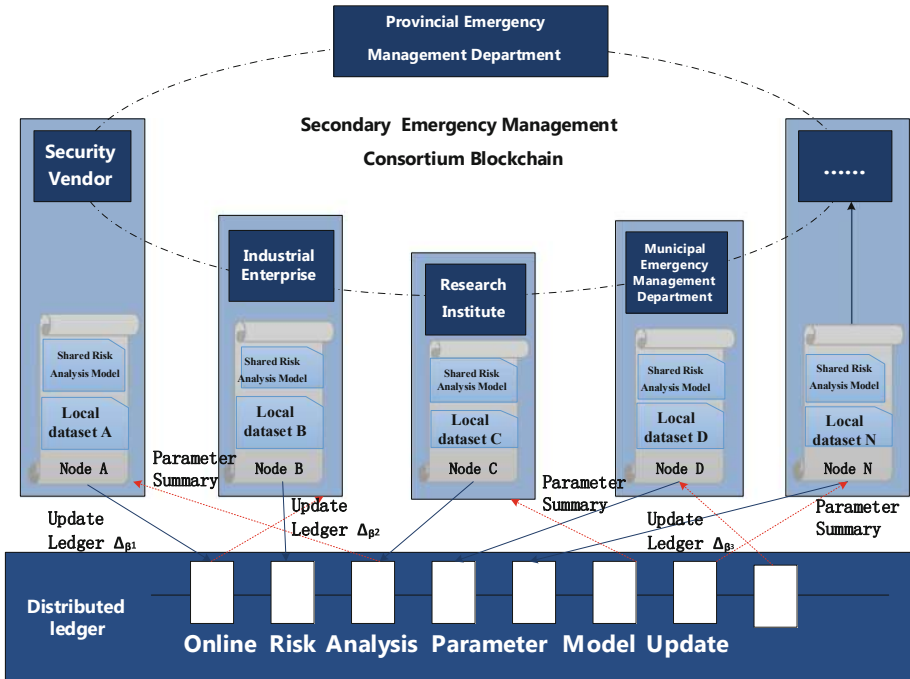


Fig. 7. Risk analysis process based on blockchain

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.

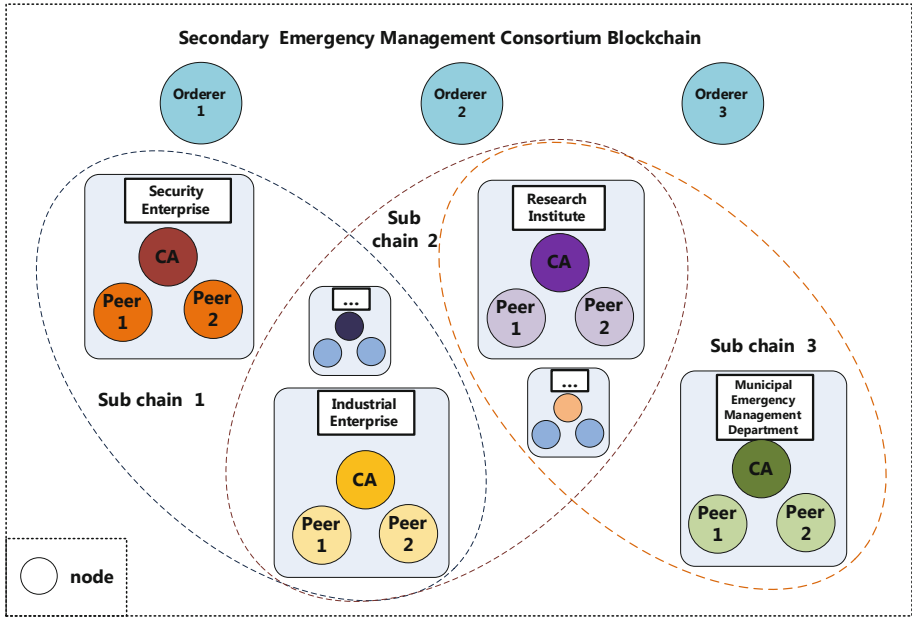


Fig. 8. Multi-organization sub chain division structure

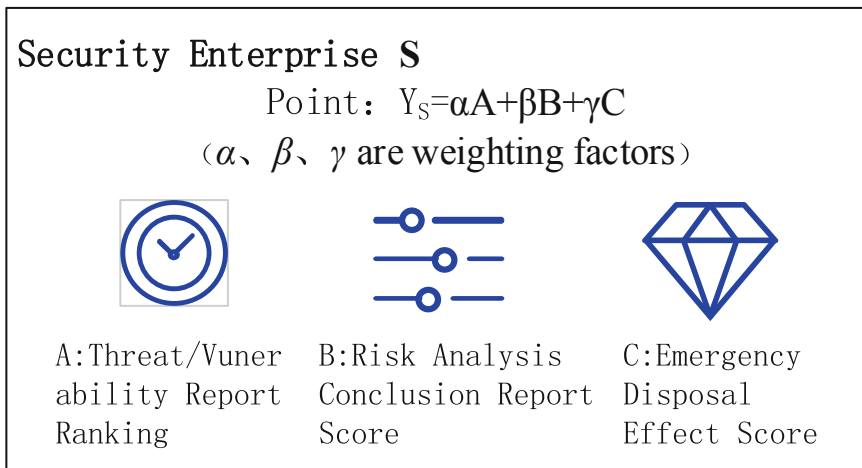


Fig. 9. User bonus point model

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., Glauco, 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
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^{1,2(✉)}, Weiqi Wang², Liehuang Zhu¹, and Ruiguang Li^{2,3}

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

² College of Cybersecurity, Changchun University, Changchun, China

³ 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.

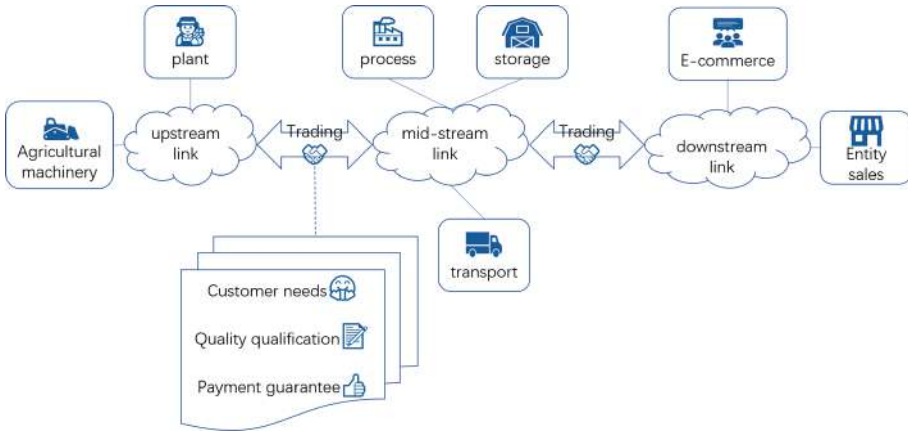


Fig. 1. Agricultural supply chain model

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.

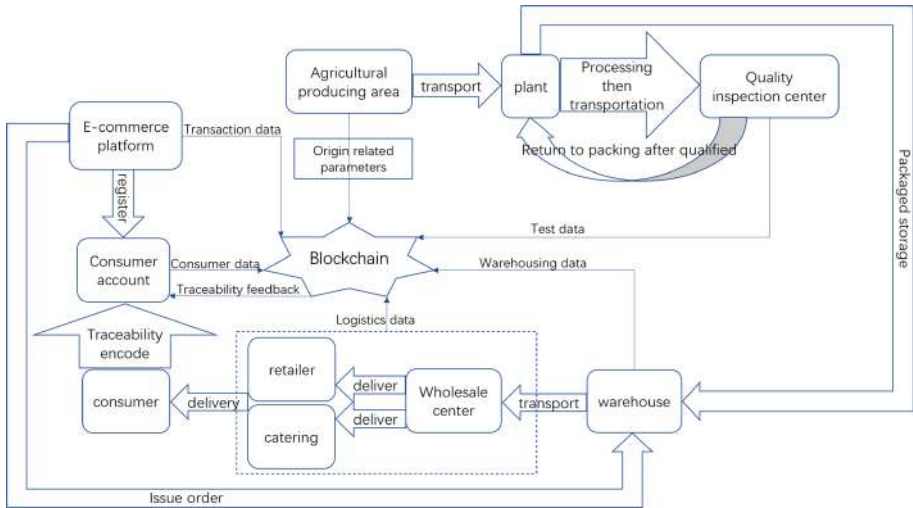


Fig. 2. Agricultural product traceability framework

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
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 Rao¹, Weixin Liu²(✉), Tian Zhu¹, Hanbin Yan¹, Hao Zhou¹, and Jinghua Bai²

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

² 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.Xorrdos. 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.

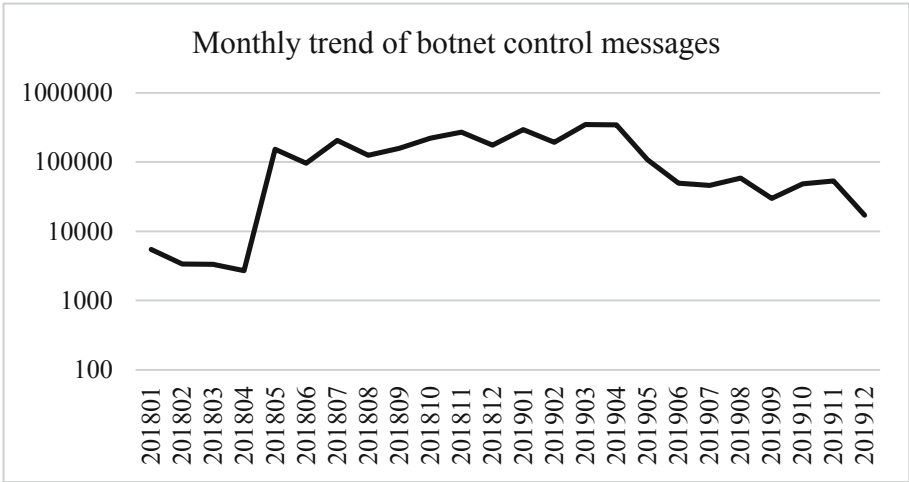


Fig. 1. Monthly trend of botnet control messages

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.

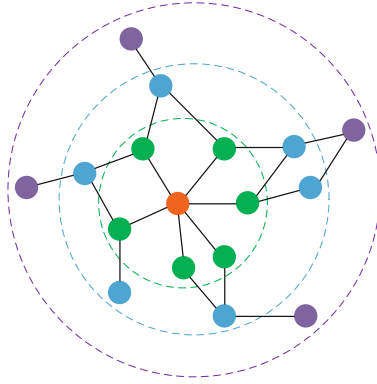


Fig. 2. Multi-hop in graph (Color figure online)

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.

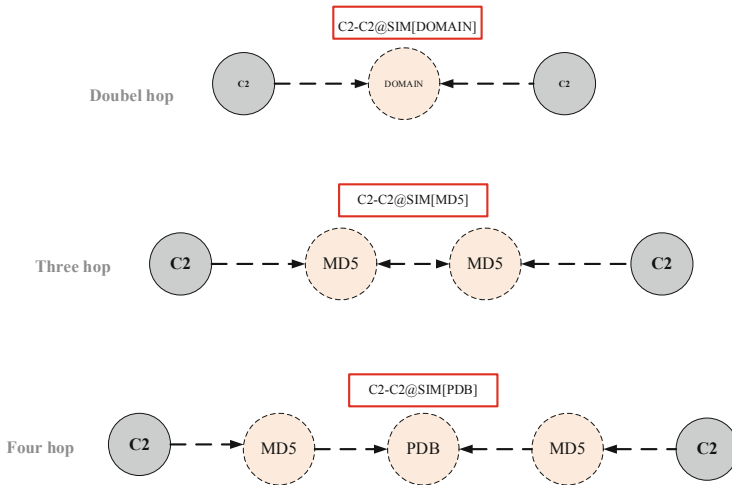


Fig. 3. Multi-hop in DDoS attack

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.

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

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

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, ω_e is weight of C2-C2@SIM[A_e], 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.

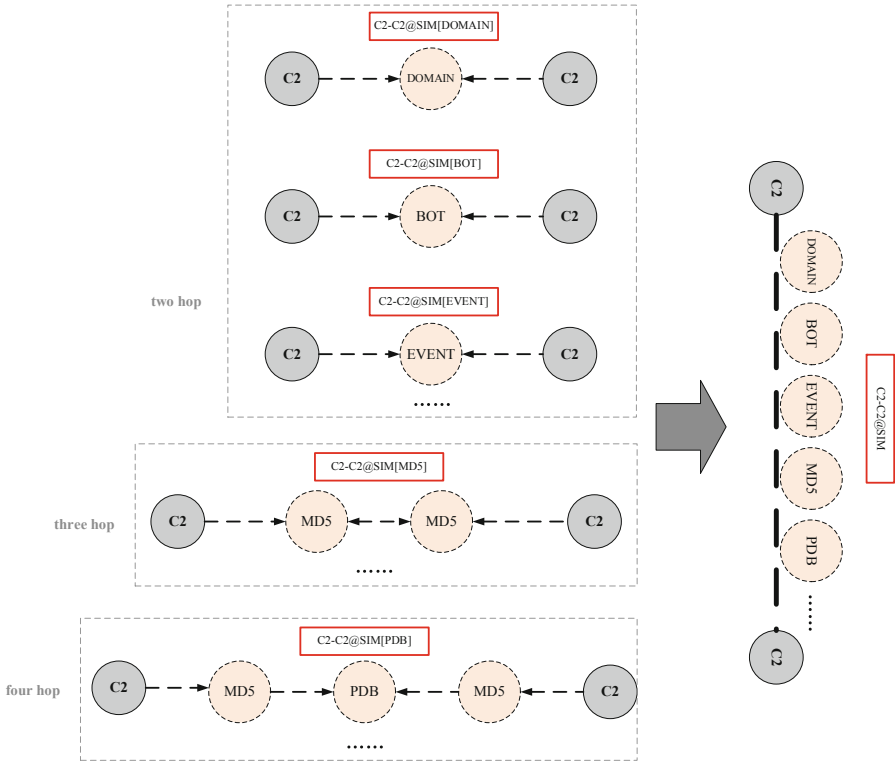


Fig. 4. Similarity calculation of C2s

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.

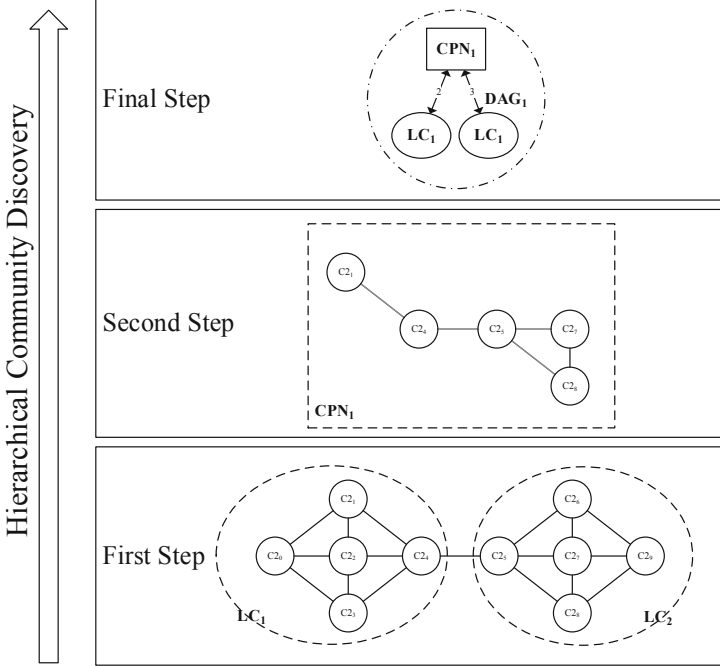


Fig. 5. Hierarchical community discovery workflow.

Hence, we need to establish a community detection workflow to capture closeness from behavior and evidence. Meanwhile, similarity connections between C2 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 C2 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 C2 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 C2 vertices. Each final DAG group consists of a set of C2 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 δ -function $\delta(u, v)$ is 1 if $u = v$ and 0 otherwise and $m = \frac{1}{2} \sum_{i,j} A_{ij}$.

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.

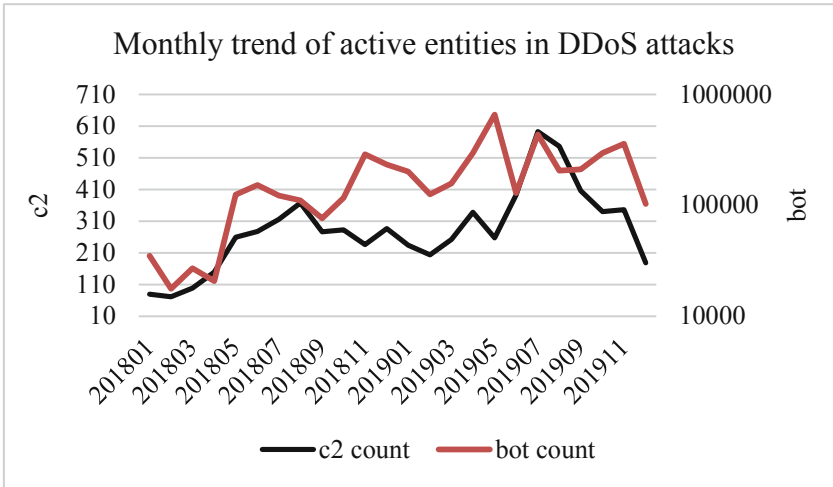


Fig. 6. Monthly trend of active entities in DDoS attacks

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.

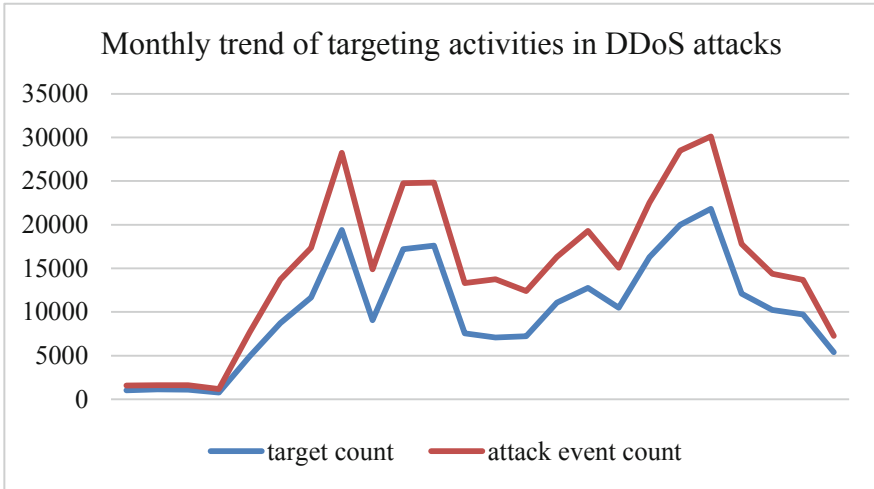


Fig. 7. Monthly trend of targeting activities in DDoS attacks

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.

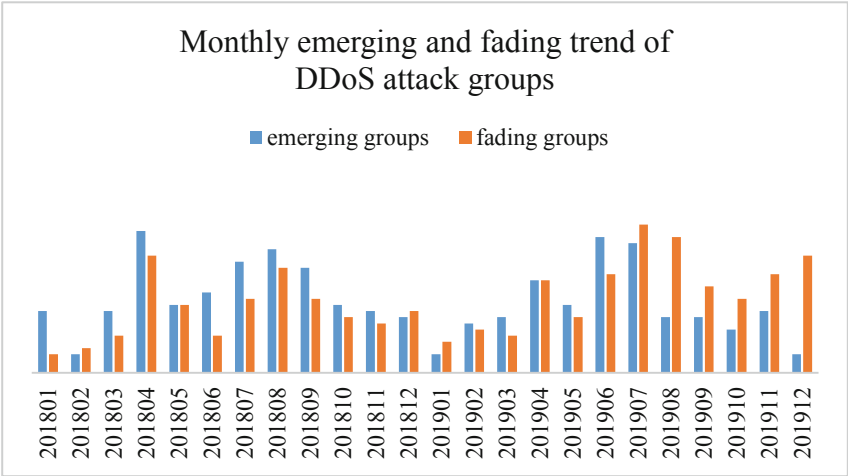


Fig. 8. Monthly emerging and fading trend of DDoS attack groups

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.

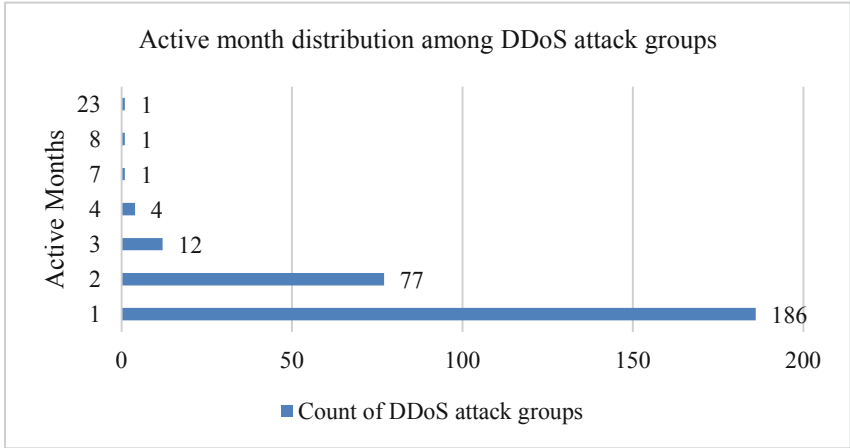


Fig. 9. Active month distribution of DDoS attack groups

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].

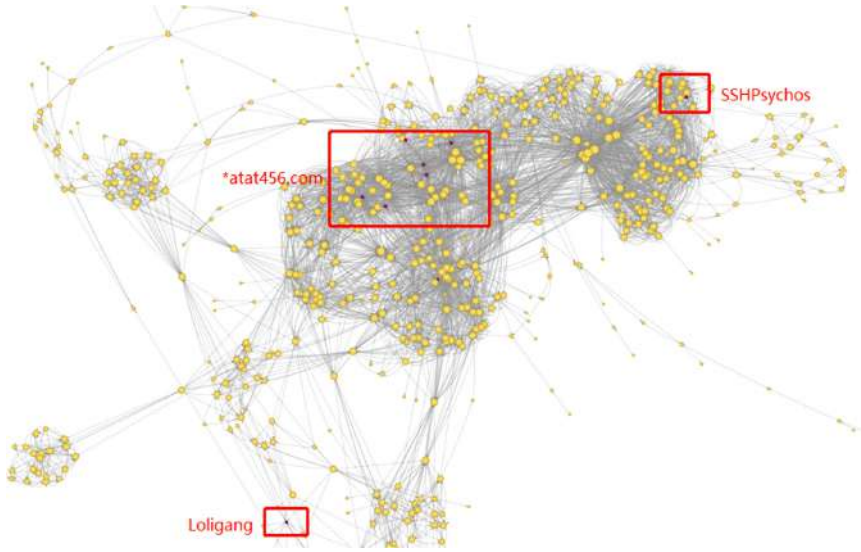


Fig. 10. Subgraphs related to DDoS attack group uncovered by external security researchers

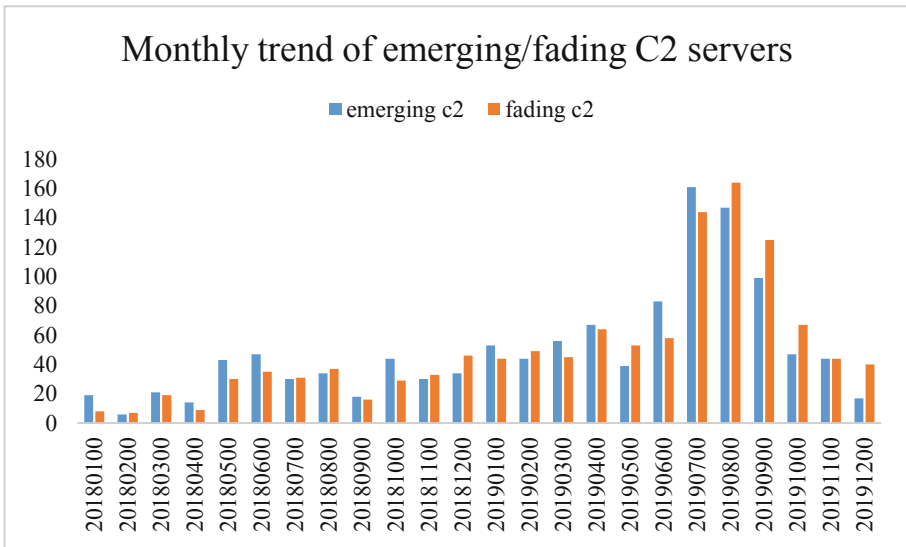


Fig. 11. Monthly trend of emerging/fading C2 servers of DDoS Attack group G769

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
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
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^(✉), 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.

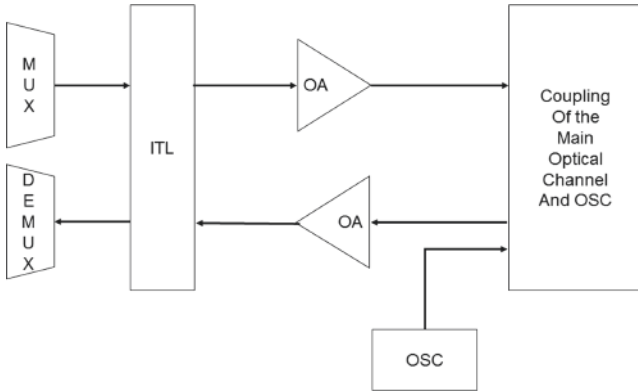


Fig. 1. Main optical-layer architecture of AWG multiplexers/demultiplexers

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).

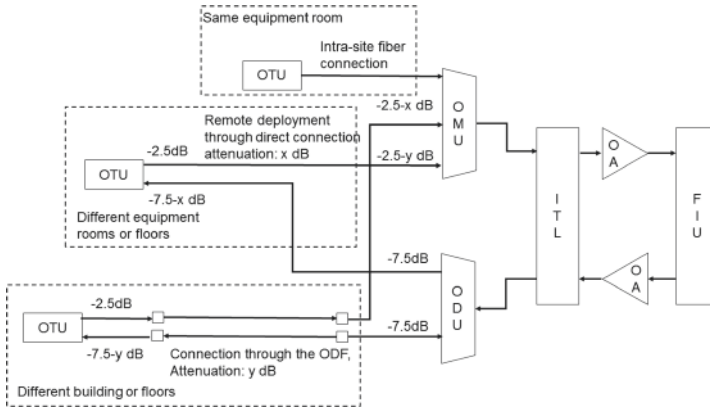


Fig. 2. Direct connection between OTU boards and multiplexers/demultiplexers using optical fibers

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.

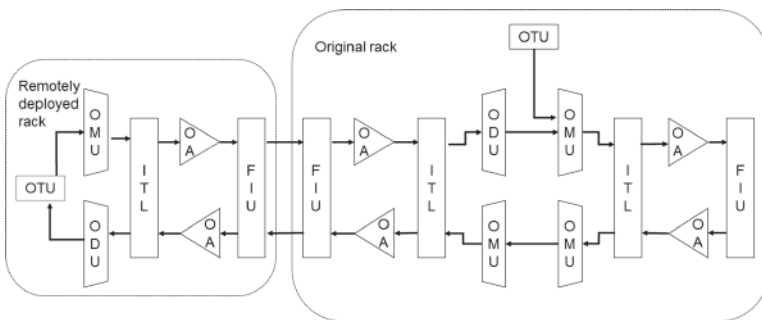


Fig. 3. Adding OMSs

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.

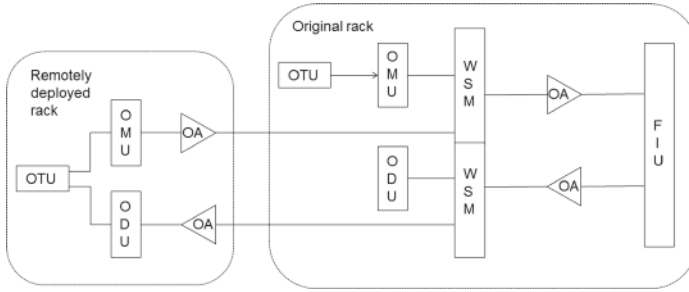


Fig. 4. Adding WSS boards for demultiplexing/multiplexing

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.

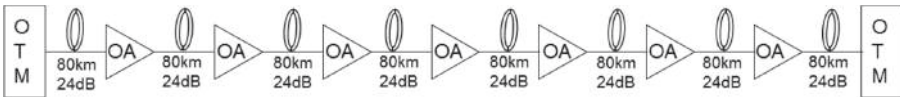
The OSNR margin is small, and OMSs cannot be added for remote deployment. In this scenario, the direct fiber connection mode is recommended so that the attenuation is controlled within the required range.

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 (0.5×4), 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 (0.5×2), 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
8×24 dB	1 dBm	20.6 dB
8×28 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
8×24 dB	1 dBm	20.6 dB
8×28 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
8×24 dB	1 dBm	20.4 dB
8×28 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
8×24 dB	1 dBm	20.2 dB
8×28 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.

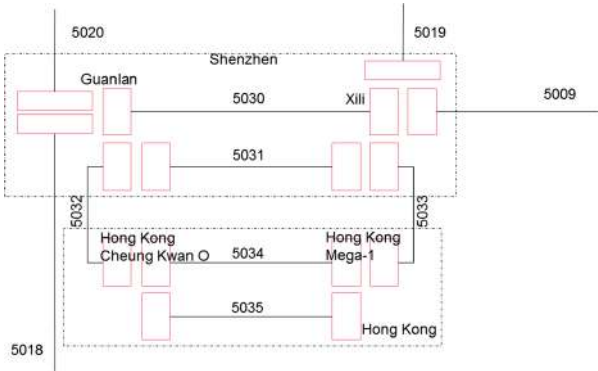


Fig. 5. Network architecture of the Xili equipment room in Shenzhen

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.

