IJNS

# International Journal of Network Security

ISSN 1816-353X (Print) Vol. 20, No. 5 (Sept. 2018)
ISSN 1816-3548 (Online)

# INTERNATIONAL JOURNAL OF NETWORK SECURITY

**PUBLISHING OFFICE**

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

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

**PUBLISHER: Candy C. H. Lin**

23-75, P.O. Box, Taichung, Taiwan 40199, R.O.C.

# International Journal of Network Security

**Vol. 20, No. 5 (Sept. 1, 2018)**

# Impact of IPSec Protocol on the Performance of Network Real-Time Applications: A Review

Abdullah Abdulrahman Al-khatib and Rosilah Hassan
(Corresponding author: Abdullah Abdulrahman Al-khatib)

Research Center of Software Management and Technology (SOFTAM)
Faculty of Information Science and Technology (FTSM), Universiti Kebangsaan Malaysia
43600 UKM Bangi Selangor, Malaysia
(Email: khteb2003@gmail.com)

(Received Feb. 2, 2017; revised and accepted July 13, 2017)

## Abstract

Network real-time applications are gaining rapid prominence on today's Internet, particularly video and audio streaming, conferencing and telephony applications. The original design of those applications and their corresponding protocols did not consider security and privacy, which form a compelling requirement to many of the Internet users and in many of the involved applications. One of the most common solutions to impose security on network traffic is to use the IPSec security protocol, which adds authentication and encryption to network packets, by which securing network applications. In the context of time-sensitive application such as VoIP and similar real-time applications, an important question that arises is how, if any, the overhead of IPSec operations could affect the performance of network applications. This paper surveys a number of the most relevant works that addressed this question mostly within the last decade. The approach, metrics and findings of each study are briefly described and an overall summary of their main characteristics is presented.

*Keywords: IPsec; VOIP; VPN (Virtual Private Network)*

## 1 Introduction

The real time application is one type of applications make specific quality of service (Qos) demands to the communication network such as maximum delay, maximum loss rate, etc.; and the network once it accepts a connection guarantees the requested services quality. Traditional network protocols such as Ethernet are designed to deliver best effort performance. A best offer network strives to achieve good average performance, and makes no attempt to meet the individual deadline of task. These networks are intended for use in applications where long delays and high data loss under heavy load conditions are acceptable. It is needless to say that such network are insufficient for use in real-time applications. Also, the Current advances in communications technology have helped it possible to support applications in many different fields such as include the Internet, mobile/cell phones, land lines, instant messaging (IM), video conferencing, Internet relay chat, robotic telepresence and teleconferencing. Emails, blogs and bulletin boards are example of non-real-time applications, for which the performance metrics of interest are typically average message/packet delay and throughput. These applications also have strict reliability requirements; indeed, much of the complexity of traditional network protocols arises from the need for loss-free communication between non-real-time applications and data-oriented. Also, the properties of real-time applications are very different from those that are in the non-real-time. As in real-time computing, the distinguishing feature of real-time application is the fact that the value of the communication depends upon the times at which messages are successfully delivered to the recipient [16, 18, 38].

Furthermore, the real-time voice and video streams are isochronous in nature, that is, they can be through of as stream of finite size samples which are generated, transmitted and received at fixed time intervals, imposing a set of time constraints witch must never be exceeded [24, 31].

Many receivers located at geographically different places receive multimedia information from multimedia server. The multimedia information is usually in the form of streaming video and audio. Transmission and processing of these information have firm real-time requirements. Additionally, VoIP is all set to revolutionize voice communications. VoIP stands for Voice over IP, it Provide a generic transport capabilities for real-time multimedia applications and Supports both conversational and streaming applications such as Internet telephony, Internet radio, Videoconferencing, Music-on-demand and Video-on-demand [10].

VoIP applications are normally used with a simple microphone and computer speakers, but IP telephones can also be used, providing an experience identical to normal telephoning. VoIP applications and services require firm real-time data transfer support [28].

## 2 IPSec

Most widely used and very important security technology is Internet Protocol Security (IPsec) [6]. It is used in the authentication and encryption in the public internet to provide the secure access. The IPsec is a set of protocols whose function is to secure communications over the Internet Protocol (IP) by authenticating and / or encrypting each IP packet in a data stream. IPsec also includes protocols for establishing encryption keys [11].

IPsec protocols act on the network layer, Layer 3 of the OSI model. Other extended Internet security protocols such as SSL, TLS and SSH operate from the application layer (Layer 7 of the OSI model). This makes IPsec more flexible because it can be used to protect Layer 4 protocols, including TCP and UDP [5, 15].

In addition, the feature of IPSec is its open standard nature. It complements perfectly with the PKI technology and, although it establishes certain common algorithms, for interoperability reasons, allows to integrate more robust algorithms cryptographic that can be designed in the future [14].

Among the benefits provided by IPSec, it should be noted that [12]:

- It enables new applications such as secure and transparent access to a remote IP node.
- It facilitates business-to-business e-commerce by providing a secure infrastructure on which to conduct transactions using any application. Extranets are an example.
- It allows building a secure corporate network over public networks, eliminating the management and cost of dedicated lines.
- It offers the teleworker the same level of confidentiality that would have in the local network of his company, being not necessary the limitation of access to the sensitive information by problems of privacy in transit.

It is important to note that when we cite the word ”safe” we do not refer only to the confidentiality of the communication, we are also referring to the integrity of the data, which for many companies and business environments may be a much more critical requirement than Confidentiality. This integrity is provided by IPSec as a service added to data encryption or as an independent service. Within IPSec the following components are distinguished [35]:

- Two security protocols: IP Authentication Header (AH) and IP Encapsulating Security Payload (ESP) that provide security mechanisms to protect IP traffic.
- An Internet Key Exchange (IKE) key management protocol that allows two nodes to negotiate the keys and all the parameters necessary to establish an AH or ESP connection.

### 2.1 Authentication Header (AH)

The authentication header (AH) provides the data integrity and the authentication to check and replay the protection, but this authentication heard does not [29]. The AH protocol [21] is the procedure provided within IPSec to ensure the integrity and authentication of IP datagrams. That is, it provides a means to the receiver of the IP packets to authenticate the source of the data and to verify that said data has not been altered in transit. However it does not provide any guarantee of confidentiality, that is, the transmitted data can be viewed by third parties [22].

As its name indicates, AH is an authentication header that is inserted between the standard IP header (both IPv4 and IPv6) and the transported data, which can be a TCP, UDP or ICMP message, or even a complete IP datagram as shown in Figure 1.

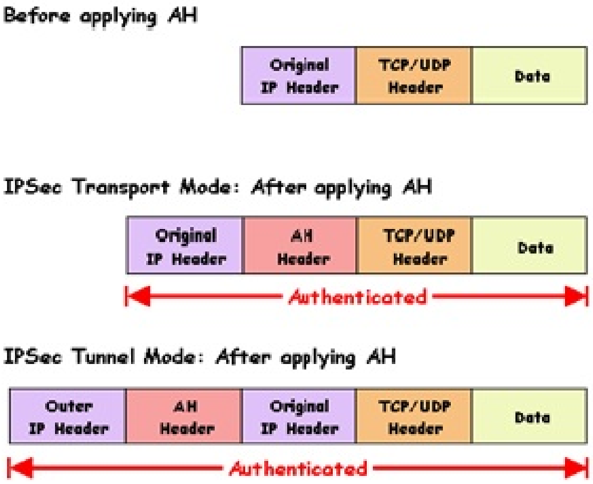


Figure 1: Structure of an AH datagram [27]

AH is actually a new IP protocol, and has assigned it the decimal number 51. This means that the IP header field contains the value 51, instead of the values 6 or 17 that are associated with TCP and UDP respectively. It is inside the AH header where the nature of the upper layer data is indicated. It is important to note that AH ensures the integrity and authenticity of the data transported and the IP header, except the variable fields: TOS, TTL, flags, offset and checksum as shown in Figure 1.

The function of AH is based on an HMAC algorithm [27], that is, a message authentication code. This algorithm consists of applying a hash function to the combination of an input data and a key, the output being a small string of characters that we call extract. This extract has the property that it is like a personal footprint associated with the data and the person who generated it, since it is the only one that knows the key.

### 2.2 Encapsulating Security Payload (ESP)

The main objective of the Encapsulating Security Payload (ESP) protocol [23] is to provide confidentiality by specifying how to encrypt the data that is to be sent and how this encrypted content is included in an IP datagram. In addition, it can offer data integrity and authentication services by incorporating a mechanism similar to AH.

Since ESP provides more functions than AH, the format of the header is more complex; This format consists of a header and a tail that surround the data transported. Such data can be any IP protocol (for example, TCP, UDP or ICMP, or even a complete IP packet). Figure 2 shows the structure of an ESP datagram, which shows how the content or payload travels encrypted [17].

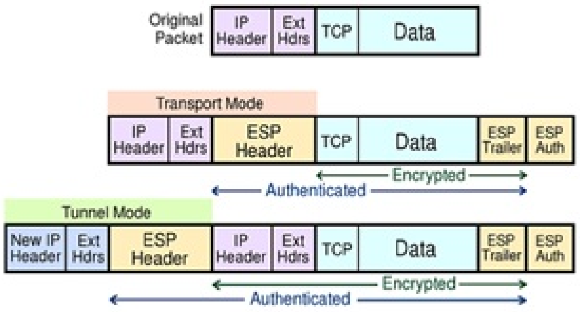


Figure 2: Structure of an ESP [17]

### 2.3 Internet Key Exchange (IKE)

IPsec uses Internet Key Exchange or IKE as the default protocol to control and convey the algorithms, keys, and protocols, and to validate the two parties. It is used to setup security associations [7]. An essential concept in IPSec is that of security association (SA): it is a unidirectional communication channel that connects two nodes, through which protected datagrams flow through previously agreed cryptographic mechanisms. By identifying only one unidirectional channel, an IPSec connection is composed of two SAs, one for each sense of communication. So far it has been assumed that both ends of a security association must be aware of the keys as well as the rest of the information they need to send and receive AH or ESP datagrams. As indicated above, it is necessary for both nodes to agree on both the cryptographic algorithms to be used and the control parameters. This operation can be done by means of a manual configuration, or by some control protocol that is in charge of the automatic negotiation of the necessary parameters; To this operation is called SAs negotiation [19].

The IETF has defined the IKE [20] protocol to perform both this automatic key management function and the establishment of the corresponding SAs. An important feature of IKE is that its utility is not limited to IPSec, but is a standard key management protocol that could be useful in other protocols, such as OSPF or RIPv2. IKE is a hybrid protocol that has resulted from the integration of two complementary protocols: ISAKMP and Oakley. ISAKMP generically defines the communication protocol and syntax of the messages that are used in IKE, while Oakley specifies the logic of how to securely perform the exchange of a key between two parts that are not previously known.

In this paper, we firstly reviewed the overview of IPsec and the main component of it. The other parts of this paper are organized as follows: Section 3 chronological survey of related works, we present some possible methods and techniques to ensure impact IPsec protocol on the performance of network RealTime Applications. Section 4 discussion attention of many researchers note from the works of several years; Section 5 is the conclusion of this paper.

## 3 Surveys of Relevant Works

Internet Protocol Security (IPsec) refers to a set of protocols used to secure communications over IP through the encryption and authentication of all data stream IP packets. It also includes protocols used to establish reciprocal authentication between agents at the start of a session and to facilitate the negotiation of cryptographic keys for use in the session. Several experiments have been conducted to measure IPsec performance and effects on Real-Time Applications.

### 3.1 VoIP Performance

Most of these studies the authors measure IPsec impact on the performance of VoIP with different experiences and network metrics, also the proportion of influence.

The experimental analysis reveals results relative to voice transmittal over secure communication links that employ IPsec [3]. This work reveals the important parameters that characterize real-time voice transmission over an Internet connection secured by IPsec and presents strategies to mitigate some VoIPsec (Voice over IPsec) limitations. The aim is to determine whether available VoIP applications can simply be replaced by VoIPsec. An efficient solution for packet header compression for VoIPsec traffic, which is called cIPsec, is also presented. Effective bandwidth was shown to decrease by up to 50 in terms of VoIP in case of VoIPsec. Meanwhile, voice traffic performance may be degraded by the cryptographic engine because access to this engine in order to prioritize traffic is difficult. Simulation results demonstrated a significant decrease in packet header overhead when the proposed compression scheme is used. This enhances the transmission effective bandwidth to assess several parameters like effective bandwidth usage, crypto-engine throughput, the traffic delay affected by various QoS strategies, and impact of various encryption algorithms on packet delay, a number of tests have been conducted. When using IPsec, voice traffic is affected by two main factors. First, is the increased packet size attributed to the additional headers in the original IP packet. These headers include the ESP header for confidentiality and the new IP header for the tunnel. Second, is the time required to encrypt the payload and headers, as well as to construct new ones [4, 6, 9, 33].

The effects of encryption mechanisms on the quality of voice and speech in widely deployed wireless technologies, namely Bluetooth and 802.11 were experimentally compared [2]. The upper bound is assessed according to the number of simultaneous VoIP calls placed using a single cell of both networks with the application of security. E-Model is used to evaluate service quality. An E-Model-based QoS tool was found to evaluate the effect of the IPsec on VoIP traffic efficiently and objectively. A decrease in average MOS was noted with the increase in the number of simultaneous calls, but IPsec overhead significantly decreased compared with the case in which calls were placed without security [30].

IPsec-based VoIP performance (in terms of throughput, packet loss rate, latency and jitter) was assessed in a 3G-WLAN integration environment. The study was designed to provide guidelines for selecting the appropriate system parameter values for VoIP over WLAN. These works serve as a guide in the choice of system parameter setups that are suitable for VoIP service in a 3G-WLAN integration environment. An IEEE 802.11b access point was found capable of supporting 15 IPsec RTP streams with satisfactory latency, minimal jitter, and no packet loss. IPsec overhead is also reasonable [8, 39].

Experiments in a LAN environment were conducted to identify the influence of 6to4 encapsulation and IPsec on VoIP quality in future IPv6 networks. VoIP performance is assessed with varying background traffic, along with IPv6 and 6to4 encapsulation with and without NAT. Such performance is compared with that of IPv4. To make calls, soft phones are used, and background traffic is generated to simulate link and router congestion. Findings reveal the efficiency of the use of a single Linux box to handle IPsec, 6to4, and NAT processing. Voice quality was verified to remain satisfactory even when the network is operating close to its 100 Mbps capacity. VoIP performance is evaluated based on delta (packet inter-arrival time), packet loss, jitter, throughput, and MOS [13, 40]. The OPNET (Network Simulation Tool) was used to assess the negative effect of VoIP network security at different simulation times. Voice transmission over IPsec was found to increase end-to-end delay, delay variation (jitter), call setup time, and packet loss. Notably, authentication is less expensive than encryption. In the conclusion of their research on VOIPsec traffic, proposed an approach to the QOS issues linked to VOIPsec. This solution addresses the packet size increase associated with the use of IPsec. They use cIPsec, which is an IPsec version that performs internal header compression for a packet. This is done while the data in the internal headers of a packet remain constant or are duplicated in the outer header [34]. The analysis and experimental results are highlighted to facilitate the assessment of the voice traffic QOS. To study the effect of IPSEC VPN, a number of metrics are considered. Three scenarios, namely without firewall, with firewall, and with firewall and VPN, are compared. The results demonstrated that delay variation and packet end-to-end delay for voice traffic rises by using the IPSEC VPN. The key cause behind this is the additional encapsulation time required. MOS was not impacted by the IPSEC VPN. Whether IPsec encryption influences router CPU utilization, voice quality, and required bandwidth was determined. These parameters are functions of the number of calls placed. Integrated Services Cisco 2811 routers, which are suitable for small firms, were used to perform the tests. After IPsec deployment, a steep linear dependence of CPU utilization on the number of processed packets was observed. This could be addressed by adjusting the Voice Payload Size (VPS) to decrease the router packet load [26]. all the previous studies summarized the survey shown in Table 1.

## 3.2 Video Performance

Most of these studies the authors measure IPsec impact on the performance of Video with different experiences and network metrics, also the proportion of influence.

Voice and video communication performance in a LAN are measured. This includes such factors as the wireless hop, given that data transmissions alternately occur over the wireless hop, both through IPsec and plain IP. IPsec is found suitable for use to secure multimedia communications over a wireless link without a notable decrease in perceived quality. This experimental study primarily revealed that IPsec is suitable for using protect real-time communications interactive even when using a wireless link IPsec. When an infrastructure with sufficient bandwidth is used, the effect of IPsec cannot be perceived by users. The small differences in the network metrics become undetectable. The authors evaluated the measurements in terms of network parameters like delay, loss, and jitter and with respect to perceived quality. In this paper, it shows that the Internet security protocol (IPsec) can be used to provide secure multimedia communications over a wireless link without significantly degrading the perceived quality [25].

This work considered such metrics as End to End Packet Delay, Packet Delay Variation, Traffic Received, MOS (Mean Opinion Score), and Traffic Sent to assess the effect of encryption on IP network video transmission. To describe IPsec tunnels, AH header and ESP are used. These components provide confidentiality, safety, integrity, and non-repudiation, with HMAC-SHA1 and 3DES encryption used for confidentiality, and AES is employed in CBC mode. We also determine the effects of an OpenVPN on the transmitted video. The Experiment results identified video compression type as the most important component that influences video quality. The H264 codec achieves better encryption results, but worse packet loss results. In the latter part of measurement, a new network was created using IPsec tunnel. ESP and AH header are employed. Authentication was performed using a shared password (ISAKMP) and a hash function called SHA (HMAC variant). Symmetric ciphers AES and 3DES were used and their topologies are demonstrated to meet encryption standards [36].

An empirical investigation of the parameters influenced by IPsec implementation in IPv6 and 6to4 Tunnelled Migration Networks was conducted to assess the performance decay that ensues after incorporating security. IPsec significantly influenced the network, such that performance was degraded with security was incorporated. This performance decay impacted realtime applications, such as VoIP and video conferencing, which are highly sensitive to delay. The experiment employed various network scenarios: IPv6 with IPsec, IPv6 Only, and 6to4 Tunnelling (IPv6 to IPv4 Migration Technique) test-bed with two computers having the same specifications. One computer was used as the server, whereas the other was the client. Notably, no IPsec design exists for an IPv6-only network. The IPsec Tunnel was placed between Routers R2 and R1. We then verified the effect on the IPv6 to IPv4 migration network (6to4). Under these conditions, the server and PC-1 were running IPv6, while the network was IPv4. Moreover, a 6to4 Tunnel was designated as the IPsec Tunnel between R2 and R1 [37].

The authors studied the performance the real time multimedia with IPSec tunnel implemented with different operation system Windows 2000 and Novell Netware. The results from two platforms were not different between the encryption and unencrypting in this experimented. The implement of encryption tools (hardware and software) are affected in network performance with use different platform like Windows 2000 and Novell Netware, etc. and they focused on client-to-site VPN topology [1]. The authors analysis QoS in videoconference by using the IPsec with different type of encryption like AES are affected fundamentally in latency when is sent through a VPN because encryption and the traffic load. Also the packet lost for voice 1 and video 2 [32]. all the previous studies summarized the survey shown in Table 1.

## 4 Discussion

It is apparent from the related works that the impact of security (IPSec in particular) on the performance of real-time applications has attracted the attention of quite few researchers. We can make a few observations from the surveyed works. First, while the works vary on their results, it seems that more experimental works confirm the efficiency of IPSec. In particular, the perceptual quality of VoIP communication can be maintained with IPSec though at the expense of some bandwidth increase and end-to-end delay, which are deemed less than serious by some researchers. Second, a smaller part of the studies use simulations to estimate the effect of IPSec. This approach might be less reliable than a true test-bed based experimental approach, and it is noted that these works tend to highlight larger impact of IPSec on the network. Among those few works, the OPNET simulator is frequently the simulator of choice. Finally, we note that most of the previous studies are focused on the voice communications through the VoIP application, and did not give much attention to video. This might suggest a potential gap that can be filled by further research on various video conferencing as well as video streaming applications under the function of IPSec protocol.

## 5 Conclusions

This paper surveyed the most relevant works on the overhead caused by security on network performance. In particular, the survey focused on the works that studied the performance of real-time applications when IPSec protocol is employed. Both securing the transmission and ensuring minimum QoS measures are important in current networks, and the study of their interaction is imperative. Quite a few research works experimented with various scenarios of voice as well as video communications with IPSec, either in real test-beds or through OPNET-based simulations. Their work and findings have been summarized and a few observations have been made.

## Acknowledgments

The authors would like to acknowledge the assistance provided by the Network and Communication Technology Research Group, FTSM, and UKM in providing facilities throughout the research. This project is partially supported under the ETP-2014-008.

## References

[1] S. Al-Khayatt, *et al.*, "Performance of multimedia applications with IPSec tunneling," in *Proceedings of International Conference on Information Technology: Coding and Computing*, 2002.

[2] A. A. Al-khatib and W. A. Hammood, "Mobile malware and defending systems: Comparison study," *International Journal of Electronics and Information Engineering*, vol. 6, no. 2, pp. 116–123, 2017.

[3] A. A. Al-khatib, and R. Hassan, "Performance evaluation of AODV, DSDV, and DSR routing protocols in MANET using NS-2 simulator," in *International Conference of Reliable Information and Communication Technology*, pp. 276-284, 2017.

Table 1: Summary of the surveyed works on the overhead of IPSec on network real-time applications

| **Authors** | **Year** | **Description** | **Findings** | **IP Protocol** | **Real-time Application** | **Measurement** | **Parameter** | **Type of study** |
|---|---|---|---|---|---|---|---|---|
| *Barbieri* et al. | 2002 | Analysis reveals results relative to voicetransmittal over secure communication links that employ IPsec. | Results show decrease by up to 50 in terms of VoIP in case of VoIPsec. | IPv4 | VoIP | Effective bandwidth usage | Packet delay, Crypto engine throughput | Experimental testbed |
| *Klaue* et al. | 2005 | Video and Voice communication performance in LAN are measured. | The result IPsec is suitable for using protect real-time communications interactive even when using a wireless link. | IPv4 | VoIP, Video conferencing | Perceived quality | Loss, Delay, Jitter | Experimental testbed |
| *Nascimento* et al. | 2006 | The effects of encryption mechanisms on the quality of voice and speech in widely deployed wireless technologies, namely 802.11 and Bluetooth, were experimentally compared. | Found to evaluate the effect of the IPsec on VoIP traffic efficiently and objectively. | IPv4 | VoIP | Voice speech quality | Packet delay, Number of simultaneous VoIP calls, MOS (Mean Opinion Score) | Experimental testbed |
| *Sung and Lin* | 2008 | IPsec-based VoIP performance (in terms of packet loss rate, jitter, throughput and latency) was assessed in a 3G-WLAN integration environment. | The system parameter setups that are suitable for VoIP service in a 3G-WLAN integration environment. | IPv4 | VoIP | VoIP quality | Throughput, Packet loss, Latency Jitter | Experimental testbed |
| *Yasinovskyy* et al. | 2009 | Identify the influence of 6to4 encapsulation and IPsec on VoIP quality in future IPv6 networks. | Reveal the efficiency of the use of a single Linux box to handle IPsec, 6to4, and NAT processing. | IPv4, IPv6 | VoIP | VoIP quality | Throughput, MOS, Packet interarrival time, Jitter, Packet loss | Experimental testbed |
| *Salama* et al. | 2009 | The OPNET (Network Simulation Tool) was used to assess the negative effect of VoIP network security at different simulation times. | Increase delay variation (jitter), end-to-end delay, call setup time, and packet loss. Notably, authentication is less expensive than encryption. | IPv4 | VoIP | VoIP quality | End to end delay, Jitter, Packet loss, Call setup time | Simulation (OPNET) |

Table 2: Summary of the surveyed works on the overhead of IPSec on network real-time applications (Cont.)

| **Authors** | **Year** | **Description** | **Findings** | **IP Protocol** | **Real-time Application** | **Measurement** | **Parameter** | **Type of study** |
|---|---|---|---|---|---|---|---|---|
| *Babu* | 2012 | Study effect of IPSEC VPN, a number of metrics are considered. Three scenarios, namely without firewall, with firewall, and with firewall and VPN, are compared. | Delay variation and packet end-toend delay for voice traffic rises by using the IPSEC VPN. | IPv4 | VoIP | Voice quality | * Jitter * MOS * End to end packet delay * Traffic sent * Traffic received | Simulation (OPNET) |
| *Sevcik* et al. | 2014 | Study the metrics End to End Packet Delay, Packet Delay Variation, Traffic Received, MOS (Mean Opinion Score), and Traffic Sent to assess the effect of encryption on IP network video transmission. | Identified video compression type as the most important component that influences video quality. | IPv4 | Video transmission | Video quality | * Packet loss | Experimental testbed |
| *Mazalek* et al. | 2015 | IPsec encryption influences router CPU utilization, voice quality, and required bandwidth was determined. | Decrease the routers packet load. | IPv4 | VoIP | Voice quality | * CPU utilization * Bandwidth MOS | Experimental testbed |
| *Shah and Parvez* | 2015 | IPsec significantly influenced the network, such that performance was degraded with security was incorporated. | Performance decay impacted realtime applications, such as VoIP and video conferencing, which are highly sensitive to delay. | * IPv4 * IPv6 | * VoIP * Video conferencing | * VoIP * Video quality | * Throughput * IP end-to-end * Jitter * Delay * Packets drop rate Tunnel delay | Simulation based (OPNET) |

[4] M. Babu, "Performance analysis of IPSec VPN over VoIP networks using OPNET," *International Journal of Advanced Research in Computer Science and Software Engineering*, 2012. DOI: 10.5815/ijcnis.2015.12.01.

[5] M. Balfaqih, *et al.*, "Fast handover solution for network-based distributed mobility management in intelligent transportation systems," *Telecommunication Systems*, vol. 64, no. 2, pp. 325-346, 2017.

[6] R. Barbieri, D. Bruschi, and E. Rosti, "Voice over IPsec: Analysis and solutions," in *Computer Security Applications Conference*, 2002.

[7] B. K. Chawla, O. P. Gupta, and B. K. Sawhney, "A Review on IPsec and SSL VPN," *International Journal of Sientific & Engineering Research*, vol. 5, no. 11, 2014.

[8] W. E. Chen, "A call server integrated approach for QoS provisioning of SIP multimedia services in 802.11 wireless networks," in *Vehicular Technology Conference*, 2010.

[9] A. D. Elbayoumy, and S. J. Shepherd, "QoS control using an end-point CPU capability detector in a secure VoIP system," in *10th IEEE Symposium on Computers and Communications*, 2005.

[10] A. Elnashar, M. A. El-Saidny, and M. R. Sherif, *Design, Deployment and Performance of 4G-LTE Networks: A Practical Approach*, John Wiley and Sons, 2014.

[11] S. Frankel and S. Krishnan, *IP Security (IPsec) and Internet Key Exchange (IKE) Document Roadmap*, RFC 6071, 2011.

[12] S. E. Frankel, *et al.*, "Guide to IPsec VPNs: Recommendations of the national institute of standards and technology," *NIST Special Publication*, Special Publication (NIST SP)-800-77, 2005.

[13] A. J. Ghazali, *et al.*, "Building IPv6 based tunneling mechanisms for VoIP security," in *13th International Multi-Conference on Systems, Signals and Devices (SSD'16)*, 2016.

[14] A. P. Hansen, *Public Key Infrastructure (PKI) Interoperability: A Security Services Approach to Support Transfer to Trust*, Dissertations Monterey, California: Naval Postgraduate School, 1999.

[15] R. Hassan, A. A. Al-Khatib, and W. M. H. W. Hussain, "A framework of Universiti Kebangsaan Malaysia patent: UKM patent," in *19th International Conference on Advanced Communication Technology (ICACT'17)*, 2017.

[16] X. Hei, and D. H. K. Tsang, "The Earliest Deadline First scheduling with active buffer management for real-time traffic in the Internet," in *International Conference on Networking*, pp. 45-54, 2001.

[17] P. Jokela, J. Melen, and R. Moskowitz, *Using the Encapsulating Security Payload (ESP) Transport Format with the Host Identity Protocol (HIP)*, RFC 7402, 2015.

[18] N. S. Kambo, D. Z. Deniz, and T. Iqbal, "Measurement based MMPP modeling of voice traffic in computer networks using moments of packet interarrival times," in *International Conference on Networking*, pp. 570-578, 2001.

[19] C. Kaufman, *et al. Internet Key Exchange Protocol Version 2 (IKEv2)*, RFC 7296, 2014.

[20] C. Kaufman, *Internet Key Exchange (IKEv2) Protocol*, RFC 4306, 2005.

[21] S. Kent, *IP Authentication Header*, RFC 2402, 2005.

[22] S. Kent, R. Atkinson, *IP Authentication Header*, RFC 2402, 1998.

[23] S. Kent, *IP Encapsulating Security Payload (ESP05)*, RFC 4303, 2005.

[24] H. J. Kim, and S. G. Choi, "A study on a QoS/QoE correlation model for QoE evaluation on IPTV service," in *The 12th International Conference on Advanced Communication Technology (ICACT'10)*, vol. 2, 2010.

[25] J. Klaue, and A. Hess, "On the impact of ipsec on interactive communications," in *Proceedings of 19th IEEE International Parallel and Distributed Processing Symposium*, 2005.

[26] A. Mazalek, Z. Vranova, and E. Stankova, "Analysis of the impact of IPSec on performance characteristics of VoIP networks and voice quality," in *International Conference on Military Technologies (ICMT'15)*, 2015.

[27] C. Madson and R. Glenn, *The Use of HMAC-MD5-96 within ESP and AH*, RFC 1321, 1998.

[28] R. Moraes, P. Portugal, F. Vasques, *et al.*, "Limitations of the IEEE 802.11 e EDCA protocol when supporting real-time communication," in *IEEE International Workshop on Factory Communication Systems (WFCS'08)*, 2008.

[29] M. Nakhjiri, and M. Nakhjiri, "Internet security and key exchange basics," *AAA and Network Security for Mobile Access: Radius, Diameter, EAP, PKI and IP Mobility*, 2006.

[30] A. Nascimento, *et al.*, "Can I add a secure VoIP call?" in *Proceedings of the 2006 International Symposium on on World of Wireless, Mobile and Multimedia Networks*, 2006.

[31] K. Pulli, *et al.*, "Real-time computer vision with OpenCV," *Communications of the ACM*, vol. 55, no. 6, pp. 61-69, 2012.

[32] J. A. Perez, V. Zarate, A. Montes, "Quality of Service Analysis of site to site for IPSec VPNs for realtime multimedia traffic," in *Advanced Int'l Conference on Telecommunications and Int'l Conference on Internet and Web Applications and Services (AICT-ICIW'06)*, 2006.

[33] P. Radman, *et al.*, "VoIP: Making secure calls and maintaining high call quality," in *Proceedings of the 8th International Conference on Advances in Mobile Computing and Multimedia*, 2010.

[34] G. I. Salama, *et al.*, "Performance analysis of transmitting voice over communication links implementing IPsec," in *Paper in 13th International Conference on Aerospace Sciences and Aviation Technology (ASAT'09)*, Military Technical College, Cairo, Egypt, 2009.

[35] K. Seo, and S. Kent, *Security Architecture for the Internet Protocol*, RFC 2401, 2005.

[36] L. Sevcik, *et al.*, "The impact of encryption on video transmission in IP network," in *Telecommunications Forum Telfor (TELFOR'14)*, 2014.

[37] J. L. Shah, and J. Parvez, "Impact of IPsec on real time applications in IPv6 and 6to4 tunneled migration network," in *International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS'15)*, 2015.

[38] M. Shiwen, *et al.*, "MRTP: A multiflow real-time transport protocol for ad hoc networks," *IEEE Transactions on Multimedia*, vol. 8, no. 2, Apr. 2006.

[39] Y. C. Sung, and Y. B. Lin, "IPsec-based VoIP performance in WLAN environments," *IEEE Internet Computing*, vol. 12, no. 6, 2008.

[40] R. Yasinovskyy, A. L. Wijesinha, and R. Karne, "Impact of IPsec and 6to4 on VoIP quality over IPv6," in *10th International Conference on Telecommunications*, 2009.

# Biography

**Abdullah Abdulrahman Al-khatib** graduated from Al-Ahgaff College Hathramout-Yemen in year 2009. He is work in community college from 2009. He enrolled to study master in University Kebangsaan Malaysia (UKM) in computer science (Network Technology) in 2015. His research interest in (SDN) and IP Security (IPSec).

**Rosilah Hassan** works as an Engineer with Samsung Electronic Malaysia, Seremban before joining Universiti Kebangsaan Malaysia (UKM) in 1997. She obtains her Master of Electrical Engineering (M.E.E) in computer and communication from UKM in 1999. In late 2003, she went to Glassgow, Scotland for her PhD in Mobile Communication from University of Strathclyde. Her research interest is in mobile communication, networking, 3G and QoS. She is a senior lecturer at UKM for more than 10 years. She had received many award such as Top Ten Student for the Faculty of Engineering on the graduation day(certificated) in 1993. After that, on 1994-1996, she received JPA scholarship for undergrade study for Look East Policy Programmer.

# Improving Network Intrusion Detection Using Geometric Mean LDA

Elkhadir Zyad[1], Chougdali Khalid[2], and Benattou Mohammed[1]

*(Corresponding author: Elkhadir Zyad)*

LASTID Research Laboratory, Ibn Tofail University, Kenitra[1]
PO Box 242, Kenitra, Morocco
(Email: zyad.elkhadir@gmail.com)
GEST Research group National School of Applied Sciences (ENSA), Ibn Tofail University, Kenitra[2]

*(Received May 15, 2017; revised and accepted Aug. 15, 2017)*

## Abstract

Anomaly based Intrusion Detection System (IDS) recognizes intrusion by adapting itself to identify normal behavior of the network. It then raises an alarm whenever any suspicious network behaviors are observed. Nonetheless, this kind of IDS is usually prone to small detection rate and high false positive rate due to difficulties involved in building normal network traffic pattern or a model. To avoid as much as possible this issue, many papers exploited a feature extraction method called linear discriminant analysis (LDA) as an intermediate step before constructing the model. Unfortunately, LDA has an important weakness, the class mean vector employed in this method is always estimated by the class sample average. That is not sufficient to provide an accurate estimate of the class mean, particularly with the presence of outliers. In this paper, to overcome that, we propose to use the geometric mean to estimate the class mean vector in LDA modeling. Many experiment on KDDcup99and NSL-KDD indicate that the proposed approach is more effective than numerous LDA algorithms.

*Keywords: Geometric Mean; KDDcup99; LDA; Network Anomaly Detection; NSL-KDD*

## 1 Introduction

The quick proliferation of various network tools which communicate and interact with each others have extremely increased the complexity of the network security and leads to the birth of sophisticated attacks. The classical security techniques such as user authentication, firewall and data encryption, are not able to fully cover the entire landscape of network security. As a consequence, they miss many damageable attacks. Hence, another type of protection is highly recommended, such as Intrusion Detection System (IDS). The latter takes part in containing the network breach by respecting appropriate preventive measures before any significant damages caused by the attacker.

IDS can be generally classified into two different categories: Signature-based IDS and Anomaly based IDS. In the first one, the IDS relies on a database of known attack signatures and produces an alarm wherever it exists any malicious network activities that correspond to one or more stored signatures. This kind of IDS has high detection rate against known attacks, but it is not able to detect new attacks. To overcome this limitation, frequent and expensive updates to the signature database are required. On the other hand, anomaly based IDS tries to build a normal behavior or model with the help of system and network characteristics. Here, the attack is considered as any deviation of traffic patterns from normal behavior. The main advantage of anomaly based IDS is it ability to identify new attacks.

Nevertheless, the actual network traffic data which are often enormous in size, are considered as a major challenge to anomaly based IDS. These kind of traffic slow down the entire detection process and lead in most of times to a biased classification accuracy. Such a large scale dataset usually contains noisy and redundant features which present critical challenges to knowledge discovery and data modeling.

To alleviate that, many feature reduction and feature selection methods have been successfully employed. For examples, the paper [5] proposed a cuttlefish based feature selection techniques to ensure data quality features and eliminate redundant and noisy features. The authors of [1] use Ant Colony Optimization algorithm to select important features. As a result, the IDS can accurately detect a broader range of attacks using smaller number of features. Following the same philosophy, the work [13] employed a Discrete Differential Evolution to identify the adequate features. The obtained results show a significant improvement in detection accuracy. In [7], the authors suggest to use Principal Component Analysis (PCA) and Kernel Principal Component Analysis (KPCA) as a primary step. After that, they classify network connections using k nearest neighbor (K-NN) and decision tree algorithms. In other publication [6], the same authors proposed an improved feature extraction method called PCA Lp using conjugate gradient. Applying this method on the two well-known datasets namely KDDcup99 and NSL-KDD prove the effectiveness of the proposed approach in terms of network attacks detection, false alarms reduction and CPU time minimization.

However, PCA and it variants offer great weights to features with higher variability whether they are effective or not. This fact may bring out the situation where the features have a lake of discriminating characteristics. To deal with that, the scientific community take the advantage of using linear discriminant analysis (LDA) [9] instead of PCA in many pattern recognition problems [3, 15, 21]. The key procedure behind this method is employing the well-known Fisher criterion to extract a linearly independent discriminant vectors and exploit them as basis by which samples are projected into a new space. These vectors contribute in maximizing the ratio of the between-class distance to within-class distance in the obtained space. Recent papers in network security field such [2, 8, 14] exploited an improved variant of this feature extraction method. Hence, this step provides the IDS with an important discrimination power. Meanwhile, it leads to a better attack identification.

In PCA and LDA mathematical formulations, class mean vectors take a significant part. For the first feature extraction technique, the class mean vector contributes in defining the covariance matrix. For the second one, the mean vectors take part in creating the between-class and within-class scatter matrices. However, such vectors are estimated by the class sample averages. Since there are many outliers and some abnormal classes that contain only a few training samples, it becomes difficult to give an accurate estimate of the class mean vectors using the class sample average.

In order to solve the mean calculation issue, a numerous papers had proposed different approaches. For instance, the authors of [12] suggest an algorithm which automatically removes the correct data mean with proved convergence. Experiments on face image datasets show that the approach consistently outperforms many PCA methods. The paper [11] uses a within-class maximum - minimum - median - average vector to construct within-class scatter matrix and between-class scatter matrix instead of within-class mean vector. Recently, the work [23] proposes a harmonic mean based LDA which makes use of weighted harmonic mean of pairwise between-class distance and gives higher priority to maximize small between-class distances. This approach shows a good results when it is applied to many multi-label data sets.

To overcome this weakness in context of intrusion detection, this paper proposes to use the class geometric mean vector [16] to approximate the class mean vector. The class geometric mean vector is less sensitive to outliers. Thus, the geometric mean LDA model should be more robust than the current sample-average based LDA. We will prove this by numerous experiments using two popular data sets namely KDDcup99 and NSL-KDD.

The rest of this paper is organized as follows. In Section 2, we outline LDA. Section 3 presents in details the proposed method. Section 4 introduces the two well known network datasets KDDcup99 and NSL-KDD. Section 5 provides the experimental results and illustrates the effectiveness of the algorithm by comparing it to some LDA approaches. Finally, Section 6 offers our conclusions.

## 2 Linear Discriminant Analysis

LDA [9] seeks to find a projection matrix $G$ such that the Fisher criterion is maximized after the projection of samples. Suppose $X$ is composed of $k$ classes, $[X_1, .., X_k]$. Every $X_i$ contains $n_i$ samples. The between-class and within-class scatter matrices $S_b$ and $S_w$, are defined by

$$S_w = (1/n)\sum_{i=1}^{k}\sum_{x\in X_i}(x-m_i)(x-m_i)^T \qquad (1)$$

$$S_b = (1/n)\sum_{i=1}^{k}(m_i-m)(m_i-m)^T. \qquad (2)$$

$m_i$ is the mean of the $i$th class, and $m$ is the general mean. They are defined as follow:

$$m_i = \frac{1}{n_i}\sum_{x\in X_i}(x) \qquad (3)$$

and

$$m = \frac{1}{n}\sum_{i=1}^{k}\sum_{x\in X_i}(x). \qquad (4)$$

The Fisher criterion is defined by

$$G = \arg\max \frac{G^T S_b G}{G^T S_w G}. \qquad (5)$$

When $S_w$ is invertible, the solutions to Equation (5) can be obtained by performing the following generalized eigenvalue decomposition:

$$S_w^{-1} S_b g_i = \lambda_i g_i. \qquad (6)$$

Where $G = [g_1, \ldots, g_l]$ and $l$ is the number of eigenvectors $g_i$ that correspond to the largest eigenvalues $\lambda_i$.

From Equations (1) and (2), it is clear that the class mean vector contributes significantly in formulation of the between-class and within-class scatter matrices. Thus, it precision must have crucial effect on the resulting linear discriminant vectors $G$. However, the class sample average vector may not approximate precisely the class mean vector (Equations (3) and (4)), when there are only a few samples available for training per class. Furthermore, there are numerous studies which affirms that sample average may not be representative of the true central region for skewed data or data with outliers. In network intrusion case, since there are some classes such as U2R and R2L attacks which provides a few training samples, the resulting matrix $G$ will be seriously blurred.

# 3 Geometric Mean LDA Formulation

## 3.1 Geometric Mean Vector

In probability theory and statistics, the geometric mean $m_g$ of a set of $n$ positive numbers $x_1, x_2, ..., x_n$ is defined as:

$$m_g = (x_1 \times x_2 \times ... \times x_n)^{\frac{1}{n}}. \tag{7}$$

As sample average, the geometric mean [16] can also be used to estimate the central tendency. Furthermore, it is generally considered that this measure is more resistant to outliers (or skewed data). That what we can see in the following example: Suppose we have Data=[3.3, 3.0, 10, 3.1, 1, 3.2, 3.4] with the outliers "1" and "10" then $m$= 3.857 and $m_g$=3.186. We observe that 3.186 is more closer to the central tendency ($\frac{3+3.1+3.2+3.3+3.4}{5} = 3.2$) than 3.857. The geometric mean of a non negative matrix:

$$Z = [Z_1, Z_2, .., Z_n] = \begin{bmatrix} Z_{11} & Z_{21} & Z_{31} & \dots & Z_{n1} \\ Z_{12} & Z_{22} & Z_{32} & \dots & Z_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Z_{1d} & Z_{2d} & Z_{3d} & \dots & Z_{nd} \end{bmatrix}$$

is given by $m_g = (m_{g1}, m_{g2}, .., m_{gn})$ where $m_{gi}$ is the geometric mean of elements on the i-th column of the data matrix $Z$.

## 3.2 Geometric Mean LDA

In small sample cases, the class geometric mean vector $m_{gi}$ generally ensures a better representation of true central tendency, in particular when outliers exist in the training samples. Additionally, it is worthwhile to highlight another merit of geometric operator for dealing with outliers. Differing from many outlier-removing methods that just eliminate outliers from the training sample set, the geometric mean operator is able to derive useful information from it.

Based on all the geometric mean merits, $m_{gi}$ and $m_g$ will be used as estimators of the class mean vector $m_i$ and general mean $m$. To avoid the singularity of the within-class scatter matrix, we would apply PCA [10] on $X$ and get the PCA-projected matrix $X_{PCA}$ with the help of the equation:

$$X_{PCA} = W^T X.$$

Where $W$ is the projection matrix that contains the principal components (PCs). Then, instead of working with $X$ we operate on $|X_{PCA}|$. We apply the absolute value on $X_{PCA}$ in order to make possible the calculation of geometric mean given by Equation (7). After that we compute the new $S_w^g$ and $S_b^g$ with the formulas:

$$\begin{aligned} S_w^g &= (1/n)\sum_{i=1}^{k}\sum_{x\in X_i}(x - m_{gi})(x - m_{gi})^T \\ S_b^g &= (1/n)\sum_{i=1}^{k}(m_{gi} - m_g)(m_{gi} - m_g)^T. \end{aligned}$$

The new proposed Fisher criterion will be defined by

$$G' = \arg\max \frac{G'^T S_b^g G'}{G'^T S_w^g G'}.$$

The solutions to the above problem is reached by:

$$(S_w^g)^{-1}(S_b^g)g'_i = \lambda'_i g'_i.$$

Where $G' = [g'_1, \dots, g'_l]$. The projection of a new vector $x_{new}$ on the space constructed by our approach is obtained by:

$$t_i = (G')^T x_{new}.$$

Hereafter the algorithm is called geomean LDA.

# 4 The Simulated Databases

## 4.1 KDDcup99

The objective of 1999 KDD intrusion detection contest is to create a standard dataset [18] to evaluate research in intrusion detection. The dataset is prepared and managed by DARPA Intrusion Detection Evaluation Program. It is composed of many TCPdump raws, captured during nine weeks.

The first seven weeks were devoted to create training data. The latter represents four gigabytes of compressed binary TCP dump data, equivalent to five million connection records. Similarly, in last two weeks, the program captured around two million connection records and considered it as testing data. The KDD dataset was employed in the UCI KDD1999 competition whose goal is developing intrusion detection system models. the attacks simulated in this competition fall into four main categories: DOS, R2L, U2R, PROBE. In the first category an attacker tries to prevent legitimate users accessing or consume a service via back, land, Neptune, pod Smurf and teardrop. In R2L, the attacker tries to gain access to the victim system by compromising the security via password guessing or breaking. To perform U2R, the intruder tries to access super users (administrators) privileges via Buffer overflow attack. The last type of attack consists in gaining information about the victim machine by checking vulnerability on the victim machine. *e.g.*, Port scanning.

The KDD Cup99 dataset is available in three different files such as KDD Full Dataset which contains 4898431 instances, KDD Cup 10% dataset which contains 494021 instances, KDD Corrected dataset which contains 311029 instances. In this paper, training data are taken from KDD Cup 10% and testing data from KDD Corrected dataset.

Each sample of the dataset is a connection between two network hosts according to network protocols. It is described by 41 attributes. 38 of them are continuous or discrete numerical attributes, the other are categorical attributes. Each sample is labeled as either normal or one specific attack. The dataset contains 23 class labels out of which 1 is normal and remaining 22 are different attacks. The total 22 attacks fall into four categories as forth-mentioned attacks.

KDD Cup 99 features can be classified into three groups:

1) Basic features: This category represents all the attributes that can be extracted from a TCP/IP connection. Most of these features leading to an implicit delay in detection.

2) Traffic features: This category contains features that are computed with respect to a window interval.

3) Content features: The majority of DoS and Probing attacks have many intrusion frequent sequential patterns, this is due to the fact that these attacks establish many connections to the host(s) in a very short period of time. Unlike these attacks, the R2L and U2R attacks do not have any intrusion frequent sequential patterns. The R2L and U2R attacks are embedded in the payload of the packets, and normally include only a single connection. To identify these kinds of attacks, some relevant features are needed to identify suspicious behavior in the packet payload. These features are called content features.

### 4.2 NSL-KDD

NSL-KDD [19] is a data set proposed to solve some of the shortcomings of the KDD'99 data set discussed in [17]. To summarize, the new dataset proposes a reasonable number of train records (125973 samples) and test sets (22544 samples). This advantage makes it affordable to run the experiments on the complete set without the need to randomly select a small portion. Consequently, evaluation results of different research work will be consistent and comparable. In addition, there is no redundancy sample present in the dataset and testing set contains some attack which are not present in the training set.

## 5 Experiments and Discussion

In this section, several experiments were designed to demonstrate the effectiveness of our proposed method. In order to show it high accuracy in an all-round way, we compare geomean LDA with other popular methods such as LDA [9], Direct LDA [22], median LDA [20], null space LDA [4]. KDDcup99 and NSL-KDD were selected for evaluation.

To estimate the accuracy of these methods we employ two factors:

$$\begin{aligned} DR &= \frac{TP}{TP+FN} \times 100 \\ FPR &= \frac{FP}{FP+TN} \times 100. \end{aligned}$$

(DR) and (FPR) mean Detection Rate and False Positive Rate. True positives (TP) refer to attacks correctly predicted. False negatives (FN) represent intrusions classified as normal instances, false positive (FP) are normal instances wrongly classified, and true negatives (TN) are normal instances classified as normal. Based on the above measures, the most reliable feature extraction method will be the one which improves DR as much as possible and tries to minimize FPR.

Concerning the experiments settings, we decide to vary the number of training samples and keep test dataset unchanged with the following composition (100 normal data, 100 DOS data, 50 U2R data, 100 R2L data, and 100 PROBE). The way we modify training samples consists in increasing the number of DOS and PROBE attacks on the one hand, on the other hand, we set normal training data at 1000 samples. U2R and R2L samples are fixed at 100. In order to get a realistic detection rate (DR) and FPR, the operation of sample selection was done randomly for thirty times. Then DR and FPR took the average. Since our goal is evaluating the efficacy of feature extraction method, we use a simple classifier, the nearest neighbor classifier.

Table 1: Detection rate (%) of geomean LDA, LDA and median LDA in different PCA and LDA space dimensions

| The method | PCs | LDs | KDDcup99 | NSL-KDD |
|---|---|---|---|---|
| Geomean LDA | 3 | 3 | 60.60 | 60.38 |
| | 3 | 2 | 58.26 | 59.09 |
| | 3 | 1 | 48.52 | 52.39 |
| LDA | 3 | 3 | 61.68 | 54.35 |
| | 3 | 2 | 59.61 | 58.43 |
| | 3 | 1 | 50.45 | 42.70 |
| Median LDA | 3 | 3 | 60.63 | 58.91 |
| | 3 | 2 | 59.71 | 56.44 |
| | 3 | 1 | 40.88 | 42.12 |

To avoid the singularity of the within-class scatter matrix in LDA and median LDA, we employ PCA as first dimension reduction, then the algorithms are performed in the PCA-transformed space. The aim of the first experiment is to find the adequate dimension of the subspace transformed by PCA, such that the algorithms can be applied and give optimal results (high DR and less FPR). One proposition to do that is fixing training data at 1000 normal, 100 DOS, 50 U2R, 100 R2L, 100 PROBE. Then we apply LDA in different PCA dimension spaces and pick up the values which ensure good DR. Table 1 shows this manipulation on the two databases, number of PCs

Table 2: Detection rate (%) Direct LDA and Null space LDA in different LDA space dimensions

| Database | The method | 5 LDs | 4 LDs | 3 LDs |
|---|---|---|---|---|
| KDDcup99 | Null space LDA | 59.48 | 60.86 | 59 |
| | Direct LDA | 42.85 | 46.28 | 60.28 |
| NSL-KDD | Null space LDA | 57.58 | 58.84 | 58.37 |
| | Direct LDA | 47.14 | 46.85 | 50.28 |

means dimension kept after applying PCA, and LDs refer to the number of top discriminant vectors. We observe that choosing three PCs and three LDs contribute significantly to get a higher DR for both median LDA and the proposed approach. LDA excels with three PCs and two LDs on NSL-KDD. In the same frame of mind, we look for the number of LDs that improve the rest of LDA models efficiency. According to TABLE II, we note that three discriminant vectors ensures good DR for Direct LDA. Null space LDA needs four discriminant vectors. Thus, all the stated above LDA models will use these parameters in the next experiments.

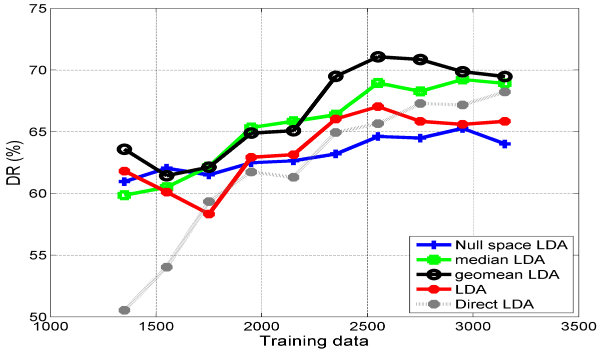


Figure 1: DR of geomean LDA, LDA, Null space LDA and median LDA on KDDcup99

Figures 1 and 2 exhibit the results we found when we compare our approach to the aforementioned LDA models for KDDcup99. According to Figure 1, we observe that geomean LDA overcomes all LDA models once the training data surpasses 2000. The reason behind this phenomenon seems to be that more there are training samples more the effect of outliers is visible. Since the other algorithms except median LDA work with the class sample average vector, they will be more sensitive to outliers, that what decrease their efficiency. Median LDA is also resistant to the skewed data because it estimates the class mean vector by a robust measure which is the median vector. However it still inferior to geomean LDA in this case. A possible explanation of this fact may reside in the distribution nature of data. We expect that the data follows a log-normal distribution which gives an important advantage to geometric mean. Concerning FPR, Figure 2 asserts that the proposed method produces a false positive rate lower than 2.7%. This means that the method is very able to distinguish normal instances from attacks.

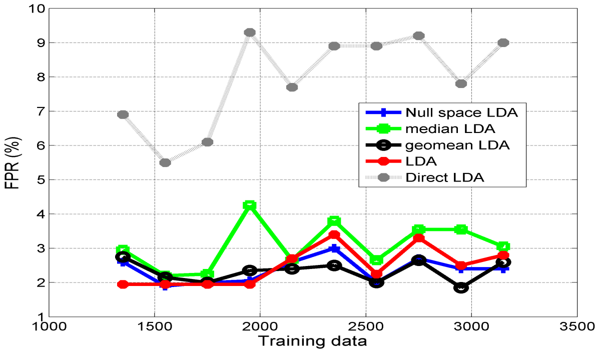


Figure 2: FPR of geomean LDA, LDA, Null space LDA and median LDA on KDDcup99

On NSL-KDD, we see from Figures 3 and 4 that geomean LDA achieves at least 3% improvement over LDA and Null space LDA, 1% over median LDA. The approach takes the lead permanently over Direct LDA. In term of FPR, Figure 4 shows that geomean LDA still gives the fewest values in the company of median LDA and LDA.

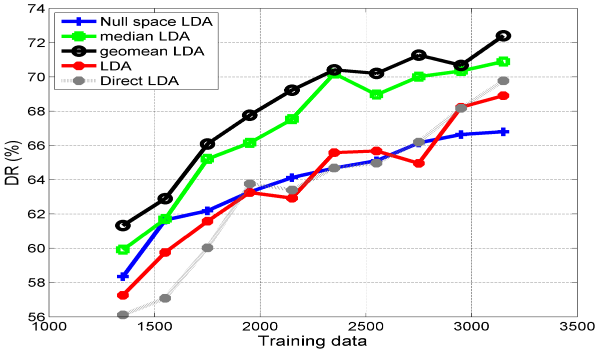


Figure 3: DR of geomean LDA, LDA, Null space LDA and median LDA on NSL-KDD

In the next experiment, we evaluated the proposed method when changing the parameter $K$ of K-NN classifier. In order to make the manipulation possible, we fixed a number of training data having the following settings: 1000 normal, 100 DOS,50 U2R, 100 R2L and 100 PROBE instances. Then, we increased $K$ and show it effect on DR and FPR.

From Figure 5 we can see that geomean LDA preserves it superiority in giving high DR. It produces at least 62% and achieves 65% as a maximal detection rate when $K = 5$. Moreover, the figure asserts that the proposed method overcomes all the other aforementioned LDA variants. In term of false positive rate, we observe from Figure 6 that geomean LDA has the fewest false positive rate in the

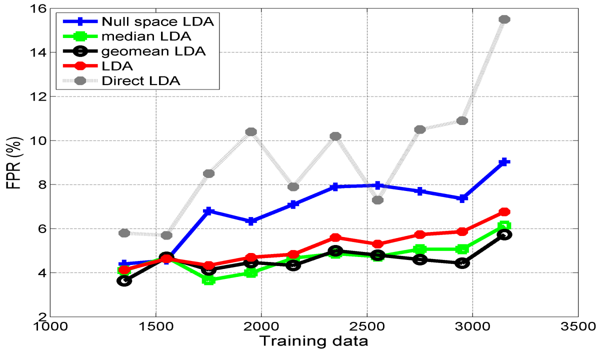


Figure 4: FPR of geomean LDA, LDA, Null space LDA and median LDA on NSL-KDD

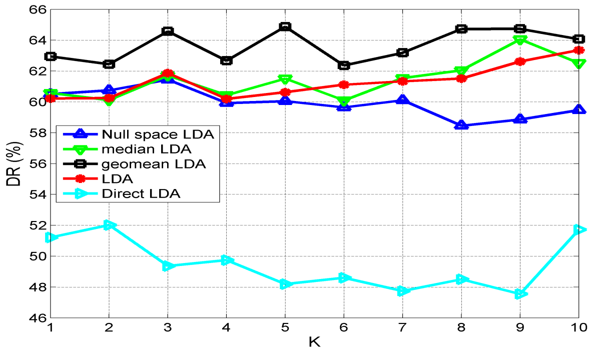


Figure 5: K vs. DR(%) for KDDcup99

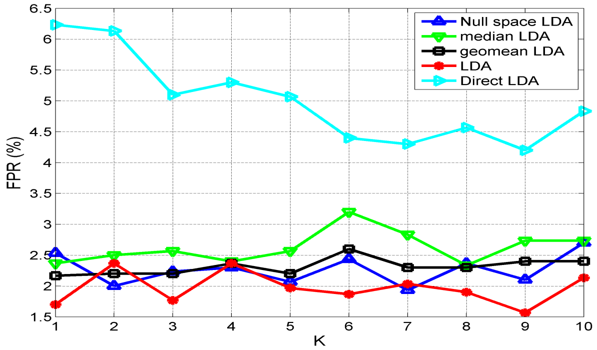


Figure 6: K vs. FPR(%) for KDDcup99

company of Null space LDA and LDA.

When we reproduce the same experiment on NSL-KDD, we get the following results: From Figure 7, it is clear that geomean LDA and median LDA are the LDA variants which ensure a better DR. The false positive rate of the proposed method is acceptable. In fact, it produces a less FPR than Direct LDA and median LDA. However, the other LDA methods take the advantage once $K$ surpasses 4.

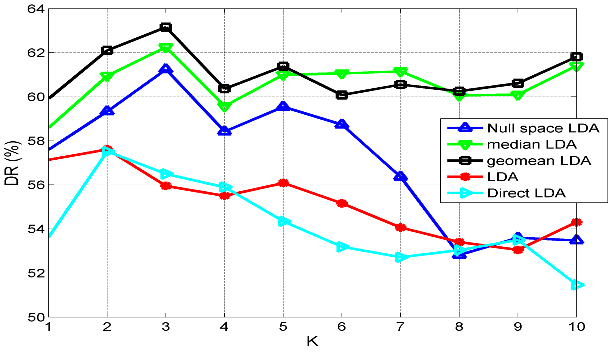


Figure 7: K vs. DR(%) for NSL-KDD

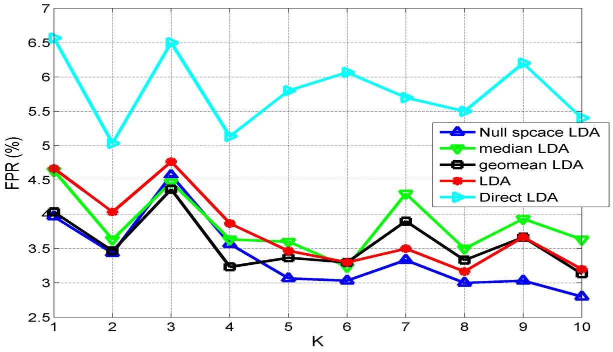


Figure 8: K vs. FPR(%) for NSL-KDD

## 6 Conclusion

In this paper, we improve the robustness of LDA in detecting network intrusions, by using class geometric mean vector, rather than the class sample average. Therefore, the proposed method called geomean LDA is more robust to outliers and preserve useful discriminant information. Experiments on KDDcup99 and NSL-KDD demonstrate the effectiveness of the proposed model, meanwhile they show it superiority over some Fisher's LDA-based algorithms such as classical LDA, null space LDA, median LDA and Direct LDA.

## References

[1] Mehdi Hosseinzadeh Aghdam and Peyman Kabiri, "Feature selection for intrusion detection system using ant colony optimization.," *IJ Network Security*, vol. 18, no. 3, pp. 420–432, 2016.

[2] Rana Aamir Raza Ashfaq, Xi-Zhao Wang, Joshua Zhexue Huang, Haider Abbas, and Yu-Lin He, "Fuzziness based semi-supervised learning approach for intrusion detection system," *Information Sciences*, vol. 378, pp. 484–497, 2017.

[3] Juan Bekios-Calfa, José M Buenaposada, and Luis Baumela, "Robust gender recognition by exploiting

facial attributes dependencies," *Pattern Recognition Letters*, vol. 36, pp. 228–234, 2014.

[4] Li-Fen Chen, Hong-Yuan Mark Liao, Ming-Tat Ko, Ja-Chen Lin, and Gwo-Jong Yu, "A new lda-based face recognition system which can solve the small sample size problem," *Pattern recognition*, vol. 33, no. 10, pp. 1713–1726, 2000.

[5] Adel Sabry Eesa, Zeynep Orman, and Adnan Mohsin Abdulazeez Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems," *Expert Systems with Applications*, vol. 42, no. 5, pp. 2670–2679, 2015.

[6] Zyad Elkhadir, Khalid Chougdali, and Mohamed Benattou, "Network intrusion detection system using pca by lp-norm maximization based on conjugate gradient," *International Review on Computers and Software (IRECOS)*, vol. 11, no. 1, pp. 64–71, 2016.

[7] Zyad Elkhadir, Khalid Chougdali, and Mohammed Benattou, "Intrusion detection system using pca and kernel pca methods," in *Proceedings of the Mediterranean Conference on Information & Communication Technologies 2015*, pp. 489–497. Springer, 2016.

[8] Zyad Elkhadir, Khalid Chougdali, and Mohammed Benattou, "A median nearest neighbors lda for anomaly network detection," in *International Conference on Codes, Cryptology, and Information Security*, pp. 128–141. Springer, 2017.

[9] Keinosuke Fukunaga, *Introduction to statistical pattern recognition*. Academic press, 2013.

[10] Ian Jolliffe, *Principal component analysis*. Wiley Online Library, 2002.

[11] Li Li, Hongwei Ge, and Jianqiang Gao, "Maximum–minimum–median average msd-based approach for face recognition," *AEU-International Journal of Electronics and Communications*, vol. 70, no. 7, pp. 920–927, 2016.

[12] Feiping Nie, Jianjun Yuan, and Heng Huang, "Optimal mean robust principal component analysis," in *International Conference on Machine Learning*, pp. 1062–1070, 2014.

[13] Ebenezer Popoola and Aderemi Oluyinka Adewumi, "Efficient feature selection technique for network intrusion detection system using discrete differential evolution and decision.," *IJ Network Security*, vol. 19, no. 5, pp. 660–669, 2017.

[14] Alaoui-Adib Saad, Chougdali Khalid, and Jedra Mohamed, "Network intrusion detection system based on direct lda," in *Complex Systems (WCCS), 2015 Third World Conference on*, pp. 1–6. IEEE, 2015.

[15] Carolina Santos Silva, Flávia de Souza Lins Borba, Maria Fernanda Pimentel, Marcio José Coelho Pontes, Ricardo Saldanha Honorato, and Celio Pasquini, "Classification of blue pen ink using infrared spectroscopy and linear discriminant analysis," *Microchemical Journal*, vol. 109, pp. 122–127, 2013.

[16] Stevan Stević. "Geometric mean,". in *International Encyclopedia of Statistical Science*, pp. 608–609. Springer, 2011.

[17] Mahbod Tavallaee, Ebrahim Bagheri, Wei Lu, and Ali-A Ghorbani, "A detailed analysis of the kdd cup 99 data set," in *Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defence Applications 2009*, 2009.

[18] UCI. *KDD Cup 1999 Data*, The UCI KDD Archive Information and Computer Science University of California, Irvine, Oct. 2014. (`http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html`).

[19] UNB. *NSL-KDD Data Set for Network-based Intrusion Detection Systems*, Mar. 2014. (`http://nsl.cs.unb.ca/NSL-KDD/`).

[20] Jian Yang, David Zhang, and Jing-yu Yang, "Median lda: a robust feature extraction method for face recognition," in *2006 IEEE International Conference on Systems, Man and Cybernetics*, vol. 5, pp. 4208–4213. IEEE, 2006.

[21] Mo Yang and Shiliang Sun, "Multi-view uncorrelated linear discriminant analysis with applications to handwritten digit recognition," in *Neural Networks (IJCNN), 2014 International Joint Conference on*, pp. 4175–4181. IEEE, 2014.

[22] Hua Yu and Jie Yang, "A direct lda algorithm for high-dimensional datawith application to face recognition," *Pattern recognition*, vol. 34, no. 10, pp. 2067–2070, 2001.

[23] Shuai Zheng, Feiping Nie, Chris Ding, and Heng Huang, "A harmonic mean linear discriminant analysis for robust image classification," in *Tools with Artificial Intelligence (ICTAI), 2016 IEEE 28th International Conference on*, pp. 402–409. IEEE, 2016.

# Biography

**Elkhadir Zyad** is a PhD student in Faculty of science, Ibn Tofail University, Kenitra, Morocco. He obtained his Master degree in computer science in 2013 from the same Faculty. He is an IEEE member. His main research interest is to develop new feature extraction algorithms for pattern recognition problem such as network intrusion detection.

**Chougdali Khalid** is an associate Professor of Computer Science at the National School of Applied Sciences, Knitra. In 2010 he obtained his PhD degree from Mohamed V-Agdal university in computer science. His main research interest are network security, pattern recognition and biometrics.

**Mohamed Benattou** is a Professor of Computer Science at the IBN TOFAIL University KNITRA where he directs the Computer Science and Telecommunication Laboratory. He has also held several positions in his French academic career: University of PAU, University of ORSAY Paris XI, 3IL and Xlim Laboratory. His research interests include distributed testing, secure testing, and software testing.

# Scalable Approach Towards Discovery of Unknown Vulnerabilities

Umesh Kumar Singh[1], and Chanchala Joshi[2]
(Corresponding author: Chanchala Joshi)

School of Engineering and Technology, Vikram University Ujjain[1]
Institute of Computer Science, Vikram University Ujjain[2]
University Road, Madhav Bhavan, Near Vikram Vatika, Ujjain, Madhya Pradesh 456010, India
(Email: chanchala.joshi@gmail.com)

(Received Apr. 8, 2017; revised and accepted July 2, 2017)

## Abstract

Of all the hazards confronting enterprise IT systems, zero-day vulnerabilities are among the most harmful. Zero-day vulnerabilities are flaws that leave users exposed to network attacks before a patch or work around is available. Every day an exploit remains unpatched, our risk of a data breach increases dramatically. Only a multi-layered approach that fully integrates with organization's IT defense stands a chance of stopping them. This paper presented a novel hybrid three layer architecture framework for zero-day attack detection and risk level assessment with respect to likelihood of exploits. The first layer of the proposed framework is liable to detect the unknown vulnerability which is based on statistical, signature and behavior based techniques; the second layer focuses on risk measurement; and the third physical layer contains centralized database and centralized server that are used during processing of first two layers. The proposed framework is analyzed in network environment of Vikram University Ujjain, India in order to evaluate the performance; experimental results show detection rate of 89% with 3% false positive rate.

*Keywords: Attack Graphs; Attackrank; Intrusion Detection; Vulnerability Analysis; Zero-day Attacks*

## 1 Introduction

In today's network system, organizations have taken great care to secure their network but even with responsible and sustained investment in defenses, they're still at risk. Attackers can bypass organization security through unknown vulnerabilities that are not listed by security persons. In a well-guarded network, a loophole may reveal by the persistent probing of a determined hacker. Attackers can leverage vulnerabilities present in network configuration to penetrate the target network [**?**, **?**]. Besides known vulnerabilities, attackers find a zero-day through hours, weeks or months of painstaking effort through lines of code, to find some weakness, some flaw that methodically barrages the target application, for which even developers are not aware of. Attackers can force the network to reveal a small crack in the defense that provides them access to secretly execute their code. This is how a network is breached trough zero-day.

Zero-day is a vulnerability that has previously unknown and unpatched and therefore can be exploited by a threat actor to gain entry to a target network. Cyber criminals are increasing the success rate of attacks by finding and exploiting Zero-day vulnerabilities. In most of the cases, information about vulnerability is not available until attacks have already taken place. As a result, attacks using zero-day exploits are hard to identify and analyze. With zero-day vulnerability in hand, the hacker has a choice that, he may either help the software vendor by providing them information about the discovered vulnerability or sell it to the black market broker who may further sell the invented exploit at highest rate. Zero-day exploits have an element of surprise as they are previously unrevealed; an attacker incorporates the zero-day exploit into their charted list of vulnerabilities and once the penetration program process and payload is concocted, attack is launched.

There is actually no protection against zero-day when the attacks were first observed. Traditional security approaches discover the vulnerabilities by generating signatures, but in case of zero-day this information is unknown, so it is extremely difficult to detect zero-day with traditional defenses [18]. Attackers are highly skilled therefore their discovered vulnerability remains unknown to public for months or even years, which provides plenty of time to attacker to cause irreparable harm [2, 20]. According to FireEye [1], a typical zero-day attack may last for 310 days on average. Therefore, dealing with zero-day is clearly a challenging task.

This paper presents a three-layer, two-phase architecture framework for zero-day attack detection and analysis. The framework consists of three layers: zero-day attack path generator, risk analyzer and physical layer. The first layer is liable to detect the unknown vulnerability; the second layer is an analyzer layer, which is assigned to analyze the generated attack and the third layer is physical layer which consists of centralized database and centralized server that are used during processing of first two layers.

The proposed framework performs two-phase working and follows a probabilistic approach for identification of the zero-day attack path and further to rank the severity of identified zero-day vulnerability. During first phase an attack graph is built from captured network scenario at any time stamp by levering the favorable attack conditions collected from various information sources. These conditions represent the abnormal system and network activities that are noticed by security persons or security sensors such as Intrusion Detection Systems. Based on the generated attack graph during first phase, second phase discovered the most probabilistic hosts by the proposed AttackRank algorithm to rank the severity of discovered vulnerability. In the proposed layered approach, once the basic network graph is generated, layers are supposed to execute dedicated functionality in parallel. Parallel work of each layer improves the performance of our proposed approach.

## 2 Related Work

Many research groups have proposed various methodologies to protect against Zero-day attacks. These methods are classified as statistics, signatures and behavior techniques [19]. Statistical based zero-day detection approaches [7] cannot be applied for real-time instantaneous detection and protection. They relying on static attack profiles therefore require a manual modification of detection settings. Signatures based techniques are broadly used yet they need an improvement to generate high class signatures. Kaur and Singh [8] proposed a hybrid approach for identification of zero-day although it is applicable only for polymorphic warm detection. In this paper we proposed a probabilistic approach for detection of zero-day attacks. The proposed approach integrates the signature based and behavior based methods of zero-day identification. The next part of our work focuses on measuring the risk level of identified malicious activity.

In the field of vulnerabilities's risk level calculation many researcher have made attempts to measure security risks of vulnerabilities [6, 17]. In our previous work we provide an approach for measuring the risk level of vulnerabilities using Hazard metric [15] with the involvement of frequency [16] and impact [14] factors. However, zero-day attacks risks level measurement is like measuring an immeasurable. We cannot measure the severity of vulnerability while it is not known. Therefore, we are considering the degree of exploitability, while measuring the risks of zero-day. The approach is based on link analysis algorithm [3] used for personalized web. We made an assumption that, an attacker can only advance his attack position to a node that has connectivity and vulnerability to be exploited. The proposed framework provides a method for zero-day detection and estimates the likelihood of system being intruded by attacker.

## 3 Proposed Approach

A three-phase architecture of Zero-day attack analysis is proposed in the paper which is shown in Figure 1. The architecture consists of three layers: zero-day attack path generator and risk analyzer. The first layer is liable to detect the unknown vulnerability, the second layer is an analyzer layer, which is assigned to analyze the generated attack graph in order to measure the risk of vulnerability and the third layer is physical layer, consists of centralized database and centralized server.

In the proposed three-layer architecture layers are supposed to execute dedicated functionality in parallel. Parallel work of each layer improves the performance of proposed approach.

### 3.1 Zero-day Attack Path Generator Layer

The first layer of proposed framework is liable to detect the unknown vulnerability. The main components of first layer are: Snort anomaly detector, attack-graph generator, detection engine and zero-day attack path generator. To capture intrusion propagation, we first build the attack graph by capturing the network scenario at specific timestamp. The attack graph is generated by sensing the anomalous behavior or abnormal activities of network that are noticed by security persons or security sensors such as IDS. These anomalies represent the probability of host being infected in an attack graph. Detection engine analyses the mysterious anomalous activities in parsed attack-graph that could be an attack and suspicious activities are preserved as zero-day exploit.Zero-day attack path generator layer consist of four modules:

#### 3.1.1 Parsing and Dependency Extraction trough Snort

The purpose of this module is to detect and filter known attacks from the captured network scenario, which is done through parsing by defining a set of malicious behavior rules that are set up or configured by an administrator. By establishing a good network profile, it is easier to identify anonymous bad behavior. For this purpose, Snort 2.9.7.6 is programmed as an anomaly detector. Snort is used to detect and filter the known attacks, by implementing a good network setup [10,13]. The packets that match with the Snort profile are known attacks and after storing their information in centralized database these packets are discarded. The packets that are partially matched or not matched are forwarded to next step for further analysis.

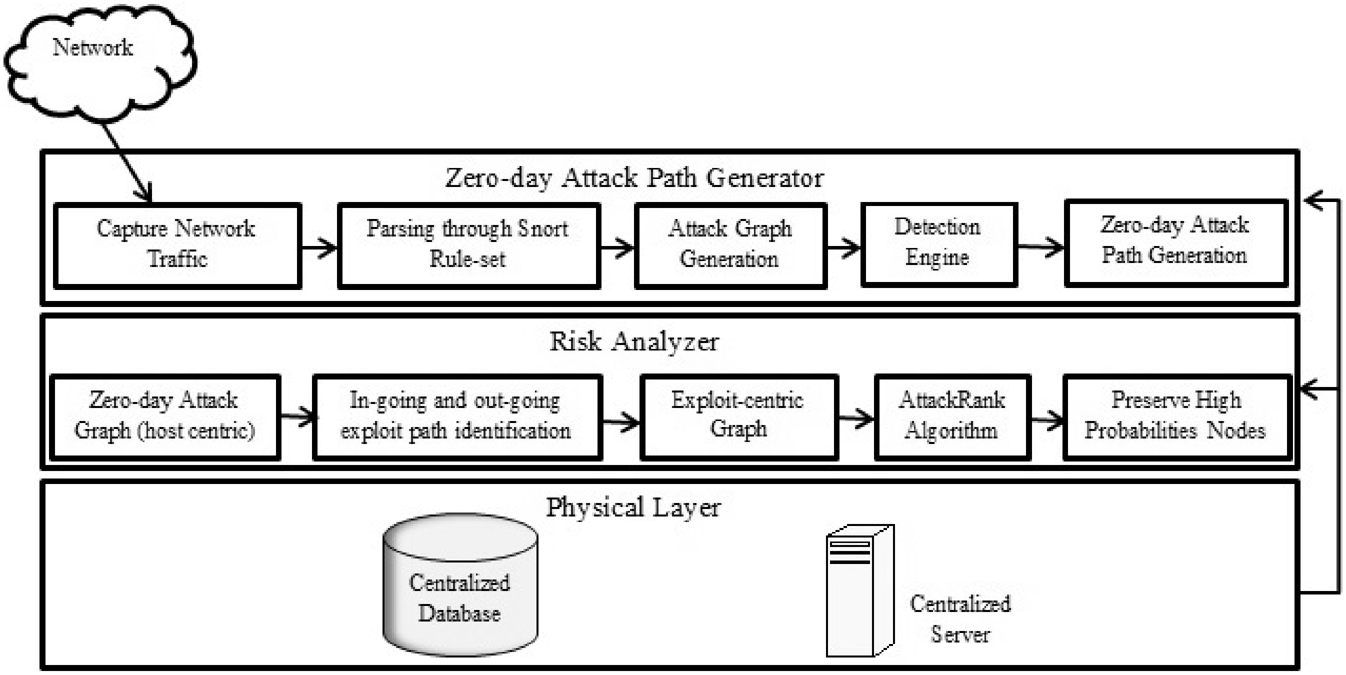

Network
Zero-day Attack Path Generator
Capture Network Traffic
Parsing through Snort Rule-set
Attack Graph Generation
Detection Engine
Zero-day Attack Path Generation
Risk Analyzer
Zero-day Attack Graph (host centric)
In-going and out-going exploit path identification
Exploit-centric Graph
AttackRank Algorithm
Preserve High Probabilities Nodes
Physical Layer
Centralized Database
Centralized Server

Figure 1: Proposed three layers architecture for zero-day attack detection and analysis

### 3.1.2 Attack Graph Generation

After filtering the known attacks through parsing, further analysis is done by Snort tagger on extracted mysterious packets which didn't trigger any alert. Tagging packets is a way to continue logging packets from a session or host that generated an event in Snort [12]. Tagged traffic is logged to allow analysis of response codes and post-attack traffic. The function of tagger is to monitor the traffic, tag the packets and send it to detection engine for further analysis. The tagger creates a new identifier based on 16-bit hash of a packet. The tag value and label for the filtered packet is stored in a table $< Tag, Label >$ for later use. The Tag value is calculated for the 6 attributes (arvl_time, source_ip, destination_ip, source_port, destination_port, protocol). The tag value is stored for later use and an attack graph of extracted nodes and mysterious conditions is generated in this module.

### 3.1.3 Detection Engine

The parsing module is not able to respond to an unknown attack; therefore run-time analysis is performed by Snort NIDS (Network Intrusion Detection System) to monitor network traffic in order to detect suspicious activity, which could be an attack or unauthorized activity. Detection engine receives the parsed packets, compares them with existing good traffic and detect unknown observations. The good traffic is the collection of safe machines on which all possible security mechanisms are applied. Security privileges and policies are defined for these safe systems and they do not participate in any malicious activity. A trust value has been assigned to these machines based upon the past experience. Algorithm 1 explains the procedure of detection engine.

**Algorithm 1** zero_day_detection

```
Begin
Capture network_scenario.
while Not end of packet in network_scenario do
    if (equals(packet_content,snort_rules)) then
        drop current_packet;
    else
        preserve filtered_packet ⇐ current_packet;
    end if
    tag ⇐ hash(preserve filtered packet (arvl_time, source_ip, destination_ip, source_port, destination_port, protocol))
    update_database(tag);
    tagged_packet ⇐ preserve filtered_packet
    if ( NOT ( is Malicious (tagges_packet)) ) then
        Capture good traffic network scenario from safe systems
        Extract features and update Snort NIDS database
    else
        unknown ⇐ tagges_packet;
        insert unknown;
        update zero_day_database(unknown);
    end if
end while
End
```

### 3.1.4 Zero-day Attack Graph Generator

In an attack graph, each node represents a host behavior at specific timestamp. For example, let us assume that the network scenario, captured in a time window T[$tbegin, tend$] is denoted as $\sum T$ and the set of hosts involved in $\sum T$ is denoted as $O_T$ , then the attack graph is a directed graph G(V, E),where:

- $V$ is the set of nodes that represent hosts in a given time window.
- $E$ is the set of directed edges that represent conditions or dependencies.
- If an attacker at given timestamp navigates in between two hosts $out_i, in_j, i, j >= 1$, and a dependency relation $dep_c : out_i - > in_j$ , where $out_i$ is the $i^{th}$ instance of host $out \in O_T$ , and $in_j$ is the $j^{th}$ instance of host $in \in O_T$ , then $V = V \cup \{out_i, in\}, E = E \cup \{dep_c\}$.

Figure 2 depicts an attack graph.

## 3.2 Risk Analyzer Layer

The first phase of our proposed methodology built an attack graph as chains of possible vulnerability exploits, which can help security persons to locate security flaws. The second phase of the proposed framework focuses to rank the nodes of attack graph based on likelihood of an attacker reaching these states. The ranking determines more vulnerable attack paths which require more immediate attention for network security. The proposed approach is based on link analysis algorithm used for personalized web [3]. We are considering three prominent attributes: Attack Vector, Attack Complexity and Authentication, of severity matrix [11] while determining the exploitability of vulnerability.

**Attack Vector:** Access vector represents difficulty from the access location (e.g. local, network accessible or remote) required to exploit the vulnerability. The more remotely an attacker can exploit the vulnerability, the greater the exploitability value will be.

**Attack Complexity:** It indicates the level of effort required to exploit the vulnerability after an access to the target point is gained. It's range in between low, medium and high. For example, a Denial of Service in a network has low complexity since the vulnerability can be exploited once an attacker gains access of the network. The lower the complexity is, the higher the exploitability will be.

**Authentication:** Authentication is defined to measure the number of authentications required (e.g., multiple instances, single instance or no instance) before network vulnerability can be exploited.

The assumption behind measurement of likelihood is that a highly exploitable vulnerability is more likely to be misused by attackers and consequently should have a higher frequency.

### 3.2.1 AttackRank Algorithm

To reveal attack paths of the higher risks zero-day vulnerabilities from the massive attack graph, the nodes with high probabilities are to be preserved, while the link between them should not be broken. We implemented an AttackRank algorithm based on the PageRank algorithm [9] to tag each node in the attack graph as either possessing high probability itself, or having both an ancestor and a descendant with high probabilities. The tagged nodes are the ones that actually propagate the infection through the network, and thus should be preserved in the final graph. Proposed AttackRank algorithm is based on PageRank and measures the likelihood of exploit in an attack graph. However network attacker behavior is different than web surfer behavior in a manner as during an attack, an attacker has options to continue or quit attacking on a current path (because of security privileges and policies it is too hard to lead to his goal). AttackRank algorithm made an assumption as if the attacker dump the current attack path then he will find an alternative path by backtracking (from one of the set of previous states) and if he continues attacking then he will attempt to each of the possible navigational states with a probability based on how hard its vulnerabilities can be exploited. With this assumption we proposed an AttackRank algorithm to find frequency of exploit.

**Algorithm 2** AttackRank algorithm for likelihood detection of exploit

```
Begin
Initialize the Graph AttackRank G(V,E)
I: set of initial states  V
while $u, v \in V$ do
  if $u \in in_{l}ink(v)$ then
    attackrank(v)⇐ attackrank(in(u))/ out(v)
  else
    if $v \in out(v)$ then
      attackrank(v)⇐ 1 - (attackrankin((u))/ out
      (v))
    else
      attackrank(v)⇐ 1
    end if
  end if
end while
End
```

## 3.3 Physical Layer

The third layer of the proposed framework, the physical layer contains centralized database and centralized server that are used during processing of first two layers. All of the information (malicious or non-malicious activities, known or unknown exploits) is stored in the database server of physical layer. Database is continuously updated by the records in the audit network data repository that do not yet have any sort of context profile.

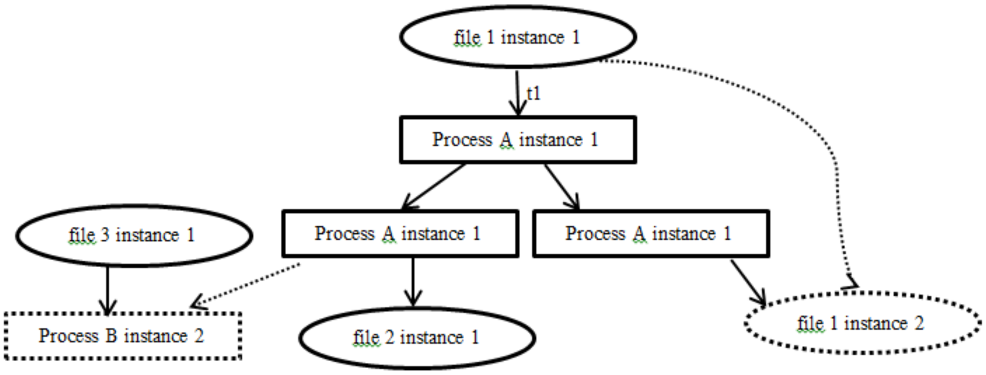


Figure 2: An attack graph

## 4 Experimental Setup

The experimental analysis of the proposed model has been conducted in the network of Vikram University, India campus. Vikram University campus consists of diverse multi-disciplinary departments, has a network of more than 500 computers, providing connectivity to different users in various institutes and hostels. Overall structure of the campus network is shown in Figure 3,

To test performance of proposed framework, a group of 7 hosts, playing miscellaneous roles are selected from the campus network that forms the testbed for this case study. Testbed is comprised of hosts in diverse physical locations (as shown in Figure 4 that includes network server located at academic block within the contact range of firewall (208.91.191.121), server located at School of Engineering and Technology (128.168.1.4), and other machines. The structure of testbed is shown in Figure 4.

The External scan is done through a router or firewall by the means of Nessus [4] vulnerability scanner. The internal scan took place at the School of Engineering and Technology (SoET) location, and was plugged into a server that resides inside Vikram University network. In Figure 4 the placement of the blue scanner is inside the firewall, so it can scan internal vulnerabilities and the red scanner is used for external vulnerabilities scan. These internal and external vulnerability scans are used to collect data to assess the effectiveness of current security measures taken at the Vikram Universitys network. The internal scan took place at the School of Engineering and Technology (SoET) location, and was plugged into a server that resides inside Vikram University network. The objective is to avoid external security counter measures to get a detailed view at system configurations. The external scan is for determining the security posture through Internet users view. The point behind external scanning is to identify what a hacker would see if he were trying to probe Vikram University network.

To measure the performance of proposed work four standard terminologies TP, FP, FN and TN were used, True Positive (TP) means the number of correctly identified malicious codes; True Negative (TN) refers to the number of correctly identified benign codes, means the non-malicious code that is classified as genuine code. FN and FP refers to misjudgments: False Positive (FP) shows that the alarm is generated when there is no actual attack, means the number of incorrectly identified trusted code as malicious code. False Negative (FN) is when the system fails to detect the malware activity due to it being similar to expected activity or no signature being available in the database. Table 1 summarizes the possible cases of classification scheme.

The rates of TP, FP, FN and TN will be computed by using four standard metrics to evaluate the performance of our technique: *True Positive Rate (TPR):* This is the percentage of correctly identified malicious codes which is measured as the ratio between the number of events that have accurately classified as positive and the total number of events that can be classified as positive, which is given by:

$$TPR = (|TP|)/(|TP| + |FN|)$$

**False Positive Rate (FPR):** This is the percentage of wrongly identified benign codes, measured as the ratio between the numbers of events that were considered positive on the number of events that should have been negative, which is given by:

$$FPR = (|FP|)/(|FP| + |TN|)$$

**False Negative Rate (FNR):** This is the rate of incorrectly rejected malicious code.

$$FNR = (|FN|)/(|TP| + |FN|)$$

**True Negative Rate (TNR):** This is the percentage of correctly identified benign codes.

$$TNR = (|TN|)/(|FP| + |TN|)$$

Accuracy, Precision and Recall will be used to evaluate the filtering accuracy of zero-day attack path. Accuracy is used to evaluate the accuracy of the classification results,

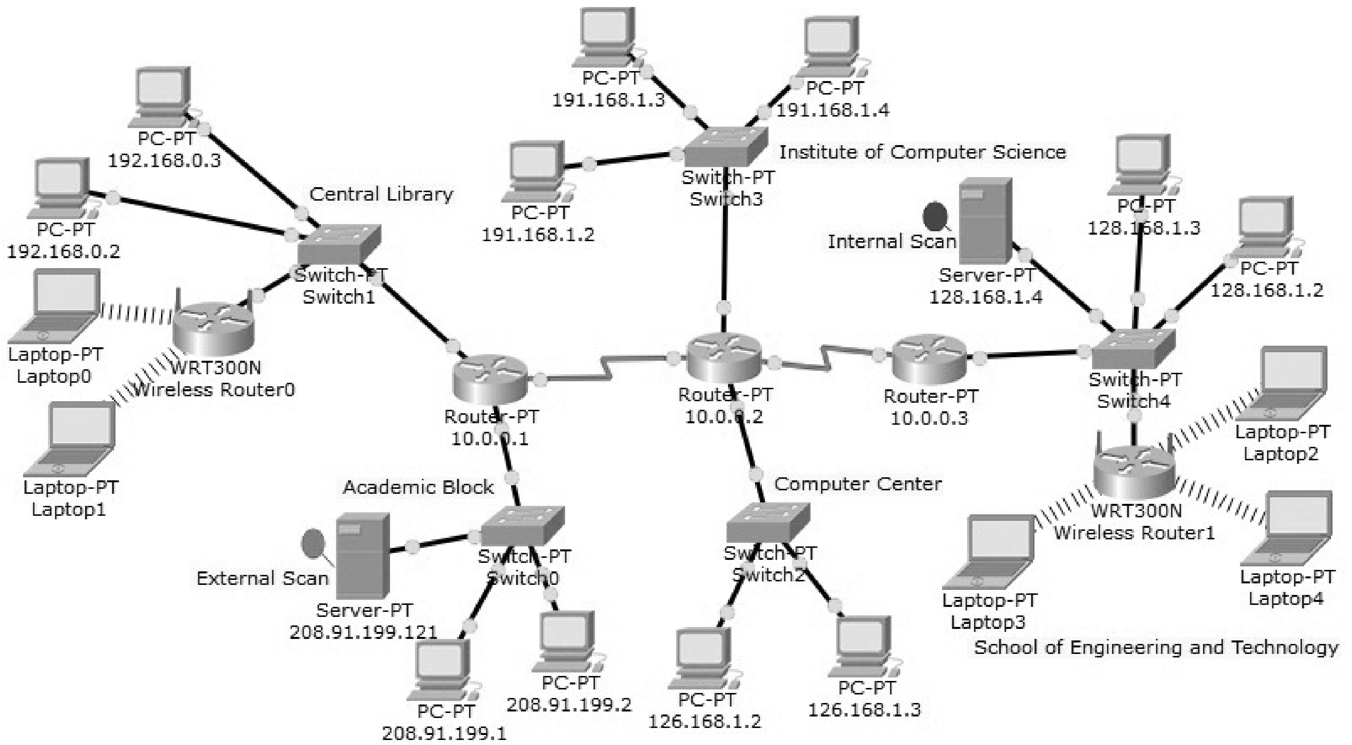


Figure 3: Network architecture of Vikram University, India campus

Table 1: Judgment cases

| **Classification** | Malicious Code | Non-Malicious Code |
|---|---|---|
| Malicious Code | True Positive (TP) | False Positive (FP) |
| Non-Malicious Code | True Negative (TN) | False Negative (FN) |

namely, the proportion of the malicious codes that are accurately classified to their own categories. It is calculated through the following formula:

$$Accuracy = (|TP| + |TN|)/(|TP| + |FP| + |FN| + |TN|)$$

Precision is the positive detection value which measures the effectiveness of zero-day detection system. Precision is used to evaluate the proportion of malicious codes among all the network activities that are judged to be malicious in nature. It is calculated through the following formula:

$$Precision = (|TP|)/(|TP| + |FP|)$$

Recall is used to evaluate the proportion of the malicious code that are accurately classified as malicious

$$Recall = (|TP|)/(|TP| + |FN|)$$

F-measure, is the harmonic mean of precision and recall, is adopted as one of the measuring indexes of the filtering mechanism, and calculated as:

$$Fmeasure = 2\times(Precision\times Recall)/(Precision+Recall)$$

F-measure is thus a means of evaluation that could combine precision and recall effectively.

Receiver Operating Characteristics (ROC) Curve is a very popular technique for measuring the relationship between the TP and FP rates of the anomaly detection system. The ROC curve uses a function of the FP rate on which the TP rate is plotted for different points. Closer the value off ROC area is to 1, it is good and when it is closer to 0.0, it is poor.

# 5 Evaluation of Proposed Model

In this section, we designed performed experiments to confirm the accuracy and efficiency of the proposed method. We built a test-bed network and launched an attack towards it. Since zero-day exploits are not readily available, we emulate zero-day vulnerabilities with known vulnerabilities. E.g., we treat CVE-2016-5387 as zero-day vulnerabilities by assuming the current time is Dec 31, 2015. The strategy of emulation also brings another benefit. The information for these known zero-days vulnerabilities can be available to verify the correctness of our experiment results.

The basic components of the testbed (Figure 4) are 2 servers for network vulnerabilities scan. 208.91.199.121 performs the external scanning through a router or fire-

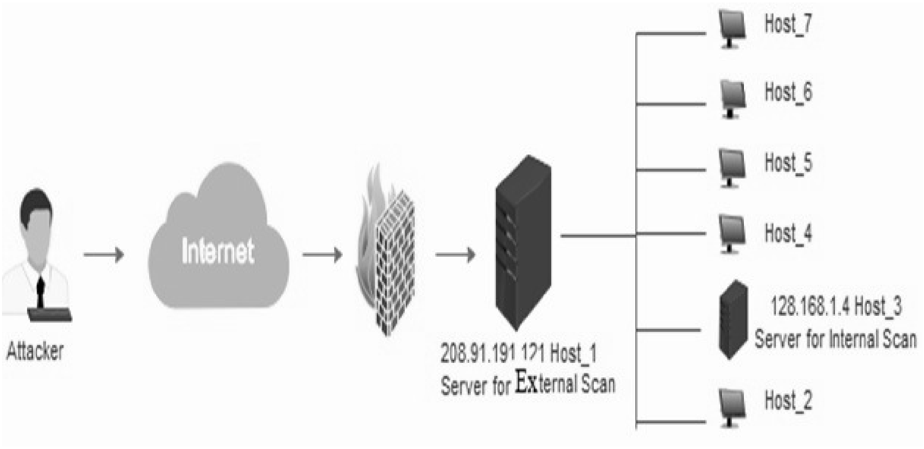


Figure 4: Experimental setup of the testbed

wall, by the means of the Nessus vulnerability scanner. Nessus placed within contact range of University, and generates details about active services, credentials and successful attacks. Scanning activities result that the server 208.91.199.121 has two open ports, tcp80 listening to HTTP traffic and tcp22 listening to SSH traffic. The SSH connection allows system administrators to do maintenance work remotely from within the subnet administration. The SSH service has vulnerabilities CVE-2012-5975, CVE-2014-6271 and CVE-2015-5600. CVE-2012-5975 allows remote attackers to bypass authentication via a crafted session involving entry of blank passwords; CVE-2015-5600 does not properly restrict the processing of keyboard-interactive devices within a single connection, which makes it easier for remote attackers to conduct brute-force attacks or cause a denial of service; and CVE-2014-6271 allows remote attackers to execute arbitrary code via a crafted environment. HTTP service has vulnerabilities CVE-2016-5387 and CVE-2015-3183. CVE-2016-5387 allows remote attackers to redirect an application's outbound HTTP traffic to an arbitrary proxy server via a crafted Proxy header in an HTTP request; and CVE-2015-3183 allows remote attackers to conduct HTTP request smuggling attacks via a crafted request. Both of these HTTP service vulnerabilities are present in the Apache HTTP Server.

We have also generated an attack environment that contains real exploit code as well as normal network traffic to web servers. Strong efforts were undertaken to make environment as realistic as possible. Acunetix, Nexpose and Metasploit [5] were used for internal scanning of vulnerabilities. Polymorphic engines ADMmutate, clet, Alpha2, CountDown, JumpCallAdditive and Pex were applied to the unencrypted exploits. True Positive Rate (TPR), False Positive Rate (FPR) and Receiver Operating Characteristics (ROC) Curve parameters are used to evaluate performance and accuracy of proposed layered architecture. Figure 5 and Figure 6 represent true detection rate and false positive rate of Zero-day attack.

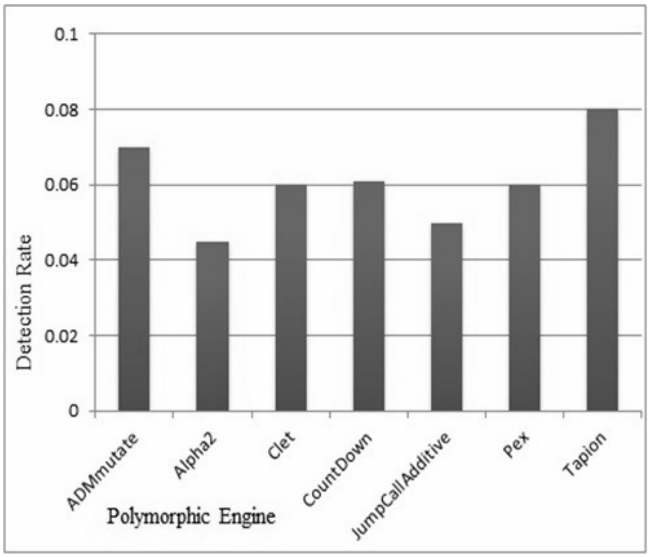


Figure 5: True positive rate

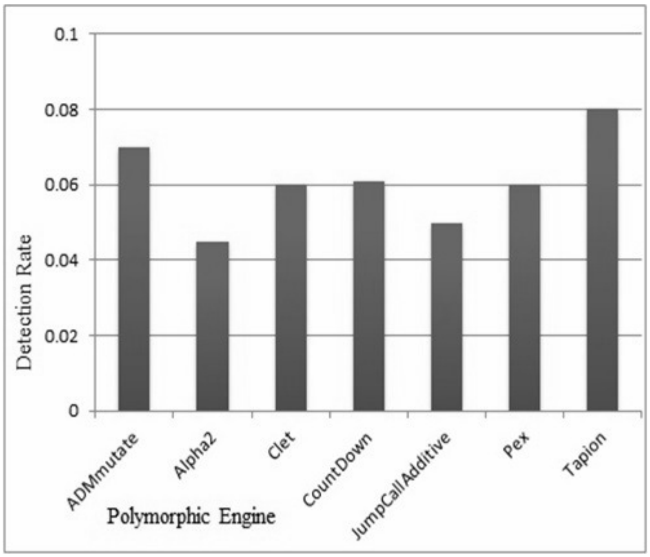


Figure 6: False positive rate

Figure 7 represents ROC curve. It is drawn by taking the average value of TPR, and in the figure, it is clearly shown that ROC is closer to 1 which proves the efficiency of our proposed approach.

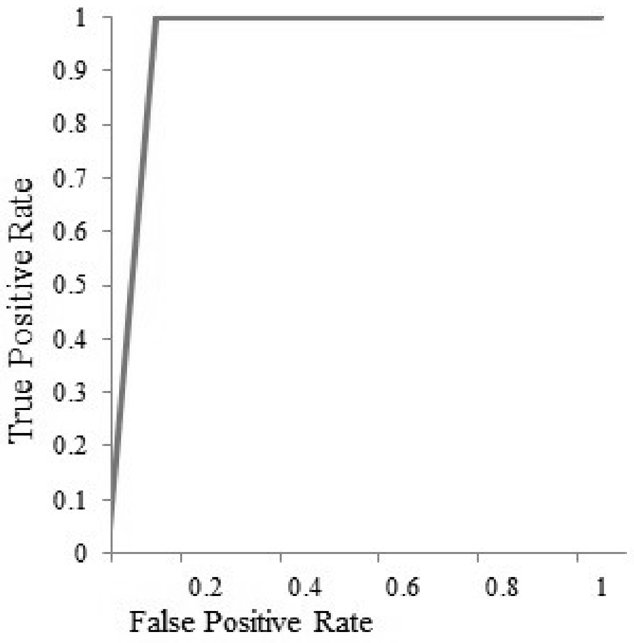


Figure 7: Average value of receiver operating characteristics curve

## 6 Conclusions

Zero-day vulnerability refers to the hole in the software or network system that is unknown to user and has no patch available. An attacker may take advantage of zero-day vulnerability by creating an exploit to gain access to the target network for stealing sensitive data, legal documents and other crucial information. This paper analyzed the lifecycle of zero-day vulnerabilities and proposed a three layer architecture framework for zero-day exploit identification and assessment; all three layers of proposed architecture have assigned specific functionalities to operate and execute in parallel to increase the performance. The first layer identifies the malicious activity that is not previously known, the second layer rank the identified vulnerability with respect to frequency of exploit, the third layer consists of database server and the centralized server. We designed our experiments to verify the efficiency of our proposed approach by using various standard parameters. In our experiments, it was observed that the best or truest detection rate was 89% and the false positive rate was 3%.

## Acknowledgments

Authors are thankful to MP Council of Science and Technology, Bhopal, for providing support and financial grant for the research work.

## References

[1] FireEye. "Zero-day danger: A survey of zero-day attacks and what they say about the traditional security model,". tech. rep., 2015.

[2] D. Song J. Caballero, T. Kampouris and J. Wang, "Would diversity really increase the robustness of the routing infrastructure against software defects?," in *Proceedings of the Network and Distributed System Security Symposium*, 2008.

[3] C. Joshi and U. K. Singh, "A novel approach towards integration of semantic web mining with link analysis to improve the effectiveness of the personalized web," *International Journal of Computer Application(IJCA, 0975 8887)*, vol. 128, pp. 1–5, Oct. 2015.

[4] C. Joshi and U. K. Singh, "Analysis of vulnerability scanners in quest of current information security landscape," *International Journal of Computer Application(IJCA, 0975 8887)*, vol. 145, pp. 1–7, Jul. 2016.

[5] C. Joshi and U. K. Singh, "Performance evaluation of web application security scanners for more effective defense," *International Journal of Scientific and Research Publications, ISSN 2250-3153*, vol. 6, pp. 660–667, Jun. 2016.

[6] C. Joshi and U. K. Singh, "Information security risks management framework- a step towards mitigating security risks in university network," *Journal of Information Security and Applications, Elsevier*, vol. 35, p. 128137, Jun. 2017.

[7] R. Kaur and M. Singh, "Automatic evaluation and signature generation technique for thwarting zero-day attacks," in *Proceedings of the Second International Conference, SNDS 2014*, pp. 298–309, India, Mar. 2014.

[8] R. Kaur and M. Singh, "Efficient hybrid technique for detecting zero-day polymorphic worms," in *Proceedings of IEEE International Advance Computing Conference (IACC)*, pp. 95–100, India, Feb. 2014.

[9] S. Brin L. Page and R. Motwani. "The pagerank citation ranking: bringing order to the web,". tech. rep., 1998.

[10] X. Liu and Y. Ye, "Intrusion detection system based on snort," in *Proceedings of the 9th International Symposium on Linear Drives for Industry Applications*, vol. 3, Springer-Verlag Berlin Heidelberg, 2014. Lecture Notes in Electrical Engineering 272.

[11] P. Mell and K. Scarfone. "Cvss: A complete guide to the common vulnerability scoring system version 2.0,". tech. rep., 2007.

[12] V. Shah N. Patel and K. Pancholi, "An analysis of network intrusion detection system using snort," *International Journal for Scientific Research and Development*, vol. 1, no. 3, pp. 410–412, 2013.

[13] M. Roesch, "Snort- lightweight intrusion detection for networks," in *Proceedings of LISA 99: 13th Systems Administration Conference Seattle*, pp. 229–238, Washington, USA, Nov. 1999.

[14] U. K. Singh and C. Joshi, "Information security assessment by quantifying risk level of network vulnerabilities," *International Journal of Computer Application(IJCA, 0975 8887)*, vol. 156, pp. 37–44, Dec 2016.

[15] U. K. Singh and C. Joshi, "Quantifying security risk by critical network vulnerabilities assessment," *International Journal of Computer Application(IJCA, 0975 8887)*, vol. 156, pp. 26–33, Dec 2016.

[16] U. K. Singh and C. Joshi, "Quantitative security risk evaluation using cvss metrics by estimation of frequency and maturity of exploit," in *Proceedings of the World Congress on Engineering and Computer Science, Vol I (WCECS2016)*, San Francisco, USA, Oct. 2016.

[17] U. K. Singh and C. Joshi, "Information security risk management framework for university computing environment," *International Journal of Network Security*, vol. 19, pp. 742–751, Sep. 2017.

[18] U. K. Singh, C. Joshi, and S. K. Singh, "Zdar system: Defending against the unknown," *International Journal of Computer Science and Mobile Computing*, vol. 5, pp. 143–149, Dec. 2016.

[19] U. K. Singh, C. Joshi, and S. K. Singh, "Zero day attacks defense technique for protecting system against unknown vulnerabilities," *International Journal of Scientific Research in Computer Science and Engineering*, vol. 5, pp. 13–18, Feb. 2017.

[20] S. Zhu Y. Yang and G. Cao, "Improving sensor network immunity under worm attacks: a software diversity approach," in *ACM Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*, pp. 149–15, 2008.

# Biography

**Umesh Kumar Singh** received his Doctor of Philosophy (Ph.D.) in Computer Science from Devi Ahilya University, Indore(MP)-India. He is currently Associate Professor in Computer Science and Director in School of Engineering & Technology, Vikram University, Ujjain(MP)-India. He has authored 6 books and his about 150 research papers are published in national and international journals of repute. He was awarded Young Scientist Award by M.P. council of Science and Technology, Bhopal in 1997. He is reviewer of various International Journals and member of various conference committees. His research interest includes Computer Networks, Network Security, Internet & Web Technology, Client-Server Computing and IT based education.

**Chanchala Joshi** received her Master of Science in Computer Science and Master of Philosophy in Computer Science from Vikram University, Ujjain(MP)-India. Currently, she is Junior Research Fellow and doctoral student in Institute of Computer Science, Vikram University, Ujjain(MP)-India. Her research interest includes network security, security measurement and risk analysis.

# Detection and Analysis Cerber Ransomware Based on Network Forensics Behavior

Ade Kurniawan[1] and Imam Riadi[2]
*(Corresponding author: Ade Kurniawan)*

Department of Informatics Engineering, Universal University[1]
Kompleks Maha Vihara Duta Maitreya, Sungai Panas, Batam 29456, Kepulauan Riau, Indonesia
(Email: ade.kurniawan@uvers.ac.id)
Department of Information System, Ahmad Dahlan University[2]

*(Received Apr. 3, 2017; revised and accepted July 2, 2017)*

## Abstract

Kaspersky and other information security firms mentioned 2016 as the year of Ransomware. The impact of attacks has allowed financial damage on the business or individual. The FBI estimates that losses incurred in 2016 will top US$ 3 billion. Meanwhile, cyber criminals use malware: Trojans, Spyware, and Keyloggers, all of which require long tremendous effort to transfer benefits into their bank accounts; while Ransomware makes the process automatic and easy by using a business model of Ransomware as a Service (RaaS). Therefore, Ransomware are made more sophisticated and more effective as to avoid detection and analysis. In this paper, we present a new insight into detection by analyzing Cerber Ransomware using Network-Forensic-Behavioral-Based. This paper is aimed to reconstruct the attack of timestamp, to identify the infected host and malware, to compromise websites involved in the chain of infection, to find campaigns scripts, and to exploit kits and payload Ransomware.

*Keywords: Cerber; Detection; Malware; Network Forensic; Ransomware*

## 1 Introduction

A hospital in Los Angeles in 2016 occurred "network infiltration" by disabling the network and computers with Ransomware, cyber criminals demanded a ransom of $ 17,000 to restore the network and computer full of important and confidential information of patients [1]. Ransomware is a type of Malware that restrict access to information by encrypting files and folders with a key is impossible to resolve and the cybercriminal will ask a ransom to unlock access to files and folders [2, 1, 3].

Ransomware is becoming popular among cyber criminals to make money in an easy way [4]. Ransomware has an impact of damage and anxiety to the business characterized by an increased the number of attacks Ransomware statistical average 100-300 percent in 2016 [5, 6, 7] with the report number of incidents increased up to 4000 percent [8]. In 2016 and is estimated in 2017 there was three Ransomware is TeslaCrypt, Locky, and CERBER who rules the world of Ransomware [9, 6]. Now Malware authors create Ransomware more sophisticated, more effective, and using anti forensic to avoid detection and analysis of each commit crimes [10, 11, 12].

Ransomware detection method generally divided into three approaches; Static feature- based, host-based and behavior-based Network Behavior Analysis [13, 14, 15]. Static feature- based widely used by antivirus software and easily avoided by attackers, such as an attacker using packaging techniques or structural change their malware code [16]. Host-behavior-based methods or dynamic analysis where artifacts malware is executed in an environment VM (virtual machine) which also has limitations due to the current Malware can detect a VM environment or host computer [17, 18] and also less capable of detecting new malware samples, and tends to produce false warnings or generate misclassification [14].

Cerber ransomware can infect via several different methods with the impact more damaging and more expensive. General scheme of distribution, spread and infections of ransomware through Network-based such as downloading a file, e-mail phishing, drive-by download or compromised website and others [11, 3] and therefore in this paper, we offer an approach to the detection and analysis Cerber Ransomware with Network Forensics Based behavior Method of because this approach has the ability to identify abnormal traffic patterns during the operation of the network. [19, 20]

Use of the approaches Network Forensic Behavior Based could reconstruct the events of the beginning of a spread, starting with the first infection of CERBER Ransomware on the host computer named STIWIE PC, find the Trojan Godzilla, pseudoDarkleech script as the Campaign to redirect network traffic victim to the server of exploit kit (EK) and a payload Ransomware used by cyber criminals.

This paper is structured as follows, in Section 2 we describe Ransomware, Cerber, and Network Forensics. In Section 3 we describe Methodology, the hardware, and software used to analyze Cerber Ransomware by using Network Forensics Behavior Based. In Section 4 which is the result of an analysis of the findings of this paper. Section 5 is part of the Conclusion and Future Work.

## 2 Basic Theory

### 2.1 Ransomware

Ransomware is a type of malware that restricts access to important information an individual or company with a way to encrypt files and will ask for a ransom payment in exchange for the decryption key to restoring encrypted files [21, 11]. The embryo of ransomware called PC Cyborg started in 1989 by Dr. Joseph Popp [22].

After infection, the PC Cyborg will hide all the file folders and encrypt files on the C: drive. A script message asked for a ransom of $ 189 directed to the PC Cyborg Corporation [11]. The first attack Ransomware uses public key cryptography to incorporate a combination of viruses and Trojan horses called cryptovirus and they called "cryptovirological attacks" [23]. The five phases of ransomware [11] shown in Figure 1:

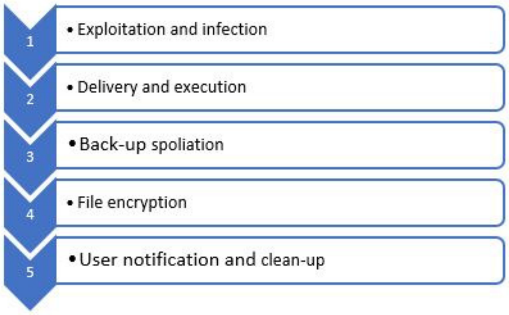


Figure 1: Five phases of ransomware

The following explanation of the five-phase mentioned above:

**Exploitation and infection:** Ransomware file needs to be executed on a computer. The spreading process and infection are often carried out through phishing emails or exploit security holes in software applications, for example, Adobe Flash and Internet Explorer.

**Delivery and execution:** After Exploitation and infection processes, Ransomware executable will be sent back to the victim's system. After executing, the mechanism of this process can take several seconds, depending on network latency. Ransomware is most often executable network deployment through strong encrypted and placed in the folder % APPDATA% or% TEMP% in the user profile.

**Back-up spoliation:** shortly after Delivery and execution processes, Ransomware will search for a file and folder backup and delete all the files for avoiding the victim that will restore files and folders that have been encrypted. In a Windows system, vssadmin tool delete volume shadow copy of the system, such as cryptolocker Ransomware and Locky will run the command to remove all shadow copies of the system. File encryption; once the file, folder and shadow copy back-up were completely removed, the malware will perform the secure key exchange with the command and control (C2) server, build an encryption key that will only be used on the local system. Ransomware will identify uniquely to each local system to distinguish the strong encryption keys among them using the AES 256 algorithm the encryption process can take anywhere from several minutes to hours depending on network latency, number and size of documents and the number of connected devices.

**User notification and clean-up:** in this phase extortion requests and payment instructions are presented to the victim. Instructions extortion requests and saved to the hard drive, sometimes the instruction file in the same folder with the encrypted files as an example of CryptoWall version 3 with the file name HELP_DECRYPT.

### 2.2 Cerber Ransomware

Cerber is one kind of sophisticated malware, with a business model Ransomware as a Service. Emerging Cerber Ransomware about 4 March 2016 in Russia and the spread is usually through botnets, spam emails and drive-by downloads [24]. When it infected, the victim data files are encrypted using AES encryption algorithm and will be notified to the victim must pay a ransom of its ordinary in the form of digital currency such as Bitcoins to receive and access their files get back [25].

Cerber will identify each victim by country, by checking the IP Geolocation country of origin of the victim, if the computer of one of the following countries (Armenia, Azerbaijan, Belarus, Georgia, Kyrgyzstan, Kazakhstan, Moldova, Russia, Turkmenistan, Tajikistan, Ukraine, Uzbekistan) will end itself and does not encrypt the computer [26].

After the executed, CERBER will install itself in the folder% AppData% {2ED2A2FE-872C-D4A0-17ACE301404F1CBA}. Windows configures automatically boot into Safe Mode and the next reboot the network mode Cerber start automatically when the user logs into Windows, to run the screensaver when the system is idle for execute itself every minute and display false alert system until the computer is restarted [27]. To make sure the victim will be begging ransom, Cerber left three notes (# decrypt MY FILES # .html, # decrypt MY FILES # .txt, and # decrypt MY FILES # .vbs) in each folder that has been encrypted.

### 2.3 Network Forensic

Network Forensics is a branch of Digital Forensics that use proven scientific techniques to collect, to use, to identify, examine, linking, to analyze, and documenting digital evidence from several sources of digital evidence and electronic evidence [28, 29, 30]. Network Forensic very reliable to capture the network traffic to and from one or multiple hosts that can later be revealed channels, methods, and the spread of malicious code [31, 30].

Obstacles often faced by the Network of Forensic investigators are gathering evidence and acceptance are often vague, poorly understood, or lack of evidence. When performing network forensics, investigators often work with a live system (online) that cannot be taken offline. This may include routers, switches, servers and other types of network devices [20].

Forensic evidence gathering for network similar to the collection of digital forensic investigation [32] but digital evidence network-based often highly volatile and should be collected through active ways inherent of evidence gathering system [33, 20].

## 3 Methodology

Preparation stage starts with the setup of hardware and software that will be used in this study. Hardware used in this study is a Notebook Processor: Intel (R) Core (TM) i7-6500U CPU @ 2.30GHz, 8GB RAM, 250GB SSD, Intel 530 Graphics Card. Software used in this study is Wireshark Version 2.2.5 and dataset from `http://www.malware-traffic-analysis.net/`.

In general, there are three methods for detecting malware: static feature, host behavior, and network-based behavior [21, 20, 12]. Detection methods used in the study Cerber Ransomware is a network-based behavior. Network behavior is to identify traffic patterns that did not occur during normal operation of the network by checking Packet Inspection: checking header, protocol, viruses, spam. Signature Detection: It monitors the content of packets in the network and comparing the pattern of attacks before configuration [30].

Inside the Network Forensics Investigation research, we use OSCAR Methodology (Obtain Information, Strategies, Collect Evidence, Analyze and Report) [20]. Illustration of the Methodology is shown in Figure 2.

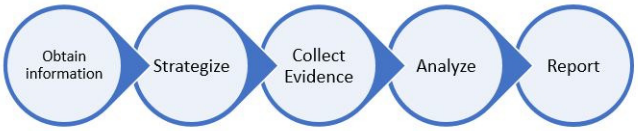


Figure 2: OSCAR methodology

### 3.1 Obtain Information

Two important things that need to be done network forensic Investigator at the beginning of the investigation: to obtain information about the incident itself and get information about the environment. Important points to note regarding the incident is a description of what happened, the timestamp (date, time, and method of the invention of the incident), people involved, systems and data involved, the manager Incident and processes, legal issues, time for investigation/recovery/resolution and Goals.

Social and political dynamics could change during the incident, investigators need to spend some time understand and respond to specific events. The following things about the environment: Business models, Legal issues, network topology, network available sources of evidence, Organizational structure, Incident Response Management Process, Resources available (staff, equipment, funding, and time).

### 3.2 Strategies

Network Forensics Investigator must work efficiently [19], because of network forensics keeps potential sources of very important evidence, some of which are also very volatile. Strategies points to consider in network forensic is Understanding the purpose the period of the investigation, a list of resources (personnel, time, and equipment), identify possible sources of evidence, to estimate the value and the cost of obtaining the evidence, list prioritizes the acquisition and plans initial acquisition/analysis.

### 3.3 Collect Evidence

Three essential components that must be done each time the Network Forensic Investigator to obtain evidence: Document, Capture and Store/Transport. Make sure the document keeps a log of all the systems accessible and all actions taken during the collection of evidence, as well as noting the date, time, sources of, methods of acquisition, and the name of the investigator and the chain of custody [30, 20].

### 3.4 Analysis

The important elements to consider in the Analyze phase is:

- Correlation: The advantages of network forensics involves multiple sources of evidence such as time stamp and other sources of evidence that can be correlated that would become sources of new evidence.
- Timeline: Once a data source is some evidence has been collected and correlated, we are building a timeline of activities, recount comprehension who is doing what, when, and how the basis of the case.
- Events of Interest: Certain events will stand out, potentially more relevant than another event. Network

forensics investigator must isolate events of interest and search for to understand how it happened.

- Interpretation: The necessary expertise to identified potential sources of additional evidence and build a theory of possible events. It is most important that you separate your interpretation of the evidence of the fact. Your interpretation of evidence always hypotheses, which can be proved or disproved.

### 3.5 Report

Reporting the most important aspect of the investigation, any Network Forensic report must be attention to the following points:

- Conceived by non-technical layman: Legal Team, Manager, Human Resources Personnel;
- Delivered detailed and structured;
- Factual.

In short, should be able to explain the results of an investigation is unreasonable to non-technical people, while retaining scientific principles.

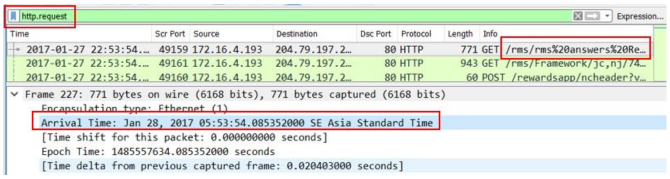
## 4 Result

In our research focus is on the side of the detection and analyze. Obtain Information from this study suspected of infection and spread of Ransomware in a corporate environment via a Network. Phase Strategies the company before infected with malware is to installing packet capture tools to capture every traffic if an illegal act when there is either an attack from the inside or from the outside that later can become digital evidence to support the forensic measures if there is a violation of the law. The dataset for research using sample data from `http://malware-traffic-analysis.net/index.html`, file format packet capture (PCAP) with a filename 2017-01-28 traffic-analysis.pcap file size 3,173 KB.

### 4.1 Analysis

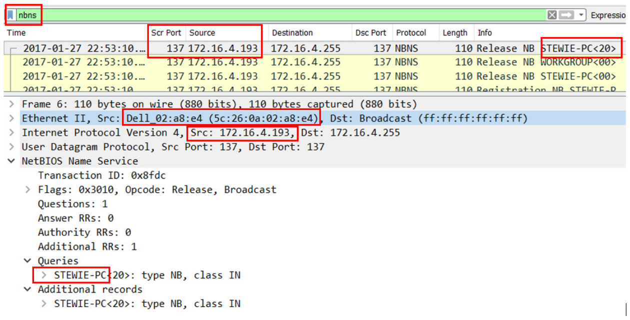
Timestamp in the digital forensic very important role because it contains information related to the show in a condition when or time [28]. Detection and forensic analysis using Wireshark Network with filter HTTP.request first thing to do is to determine when the first time the host computer is infected, show in Figure 3 shows the first time an infected computer is on time 2017-01-27 22:53:54 UTC or January 28, 2017 05:53:54 SE Asia Standard Time.

After knowing the date and time of the infection the next phase is to detect and analyze the IP and the hostname of the computer has been infected. IP detection, MAC Address Hostname and NetBIOS analysis performed by using filter NBNS. NetBIOS is an application which allows a computer to communicate with computers on the Local Area Network (LAN). Analysis of IP and MAC address of who the victims were first the infected show in Figure 4, IP Host Computers infected victims is 172.16.4.193 with MAC Address 5c: 26: OA: 02: a8: e4 the network card from hardware vendors Dell and with Host Stewie name PC.



```
Time                      Scr Port Source        Destination     Dsc Port Protocol  Length Info
2017-01-27 22:53:54...    49159 172.16.4.193  204.79.197.2...    80 HTTP       771 GET /rms/rms%20answers%20Re...
2017-01-27 22:53:54...    49161 172.16.4.193  204.79.197.2...    80 HTTP       943 GET /rms/Framework/jc,nj/74...
2017-01-27 22:53:54...    49160 172.16.4.193  204.79.197.2...    80 HTTP        60 POST /rewardsapp/ncheader?v...
Frame 227: 771 bytes on wire (6168 bits), 771 bytes captured (6168 bits)
    Encapsulation type: Ethernet (1)
    Arrival Time: Jan 28, 2017 05:53:54.085352000 SE Asia Standard Time
    [Time shift for this packet: 0.000000000 seconds]
    Epoch Time: 1485557634.085352000 seconds
    [Time delta from previous captured frame: 0.020403000 seconds]
```

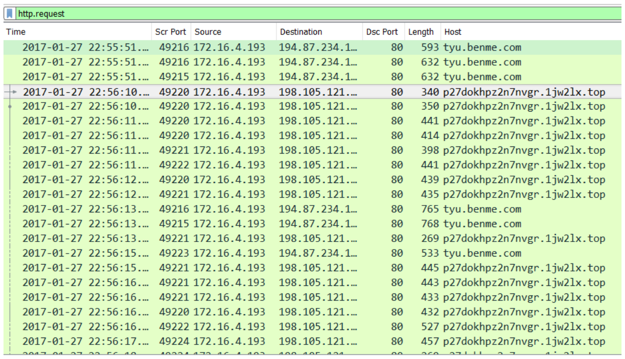
Figure 3: Date and time of the infection



```
Time                    Scr Port Source        Destination   Dsc Port Protocol  Length Info
2017-01-27 22:53:10...    137 172.16.4.193  172.16.4.255    137 NBNS       110 Release NB STEWIE-PC<20>
2017-01-27 22:53:10...    137 172.16.4.193  172.16.4.255    137 NBNS       110 Release NB WORKGROUP<00>
2017-01-27 22:53:10...    137 172.16.4.193  172.16.4.255    137 NBNS       110 Release NB STEWIE-PC<00>
2017-01-27 22:53:10       137 172.16.4.193  172.16.4.255    137 NBNS       110 Registration NB STEWIE-P
> Frame 6: 110 bytes on wire (880 bits), 110 bytes captured (880 bits)
> Ethernet II, Src: Dell_02:a8:e4 (5c:26:0a:02:a8:e4), Dst: Broadcast (ff:ff:ff:ff:ff:ff)
> Internet Protocol Version 4, Src: 172.16.4.193, Dst: 172.16.4.255
> User Datagram Protocol, Src Port: 137, Dst Port: 137
v NetBIOS Name Service
     Transaction ID: 0x8fdc
   > Flags: 0x3010, Opcode: Release, Broadcast
     Questions: 1
     Answer RRs: 0
     Authority RRs: 0
     Additional RRs: 1
   v Queries
     > STEWIE-PC<20>: type NB, class IN
   v Additional records
     > STEWIE-PC<20>: type NB, class IN
```

Figure 4: NBNS traffic analysis in wireshark

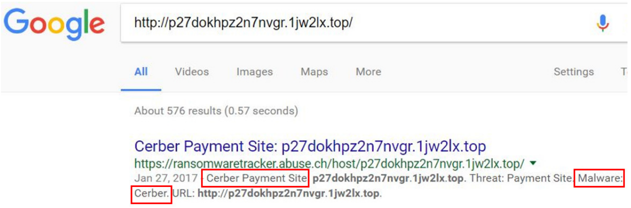
IP, MAC Address and hostname we already know, the next phase determine malware which infects the host name of the Stewie PC. After deep analysis of several packet shown in Figure 5 traffic to the domain.top, usually malware author used domain.top in conducting criminal activities. List Domain [34] which is generally used is Domains lclebb6kvohlkcml.onion [.] link lclebb6kvohlkcml.onion [.] nu bmacyzmea723xyaz.onion [.] link bmacyzmea723xyaz.onion [.] nu nejdtkok7oz5kjoc.onion [.] link nejdtkok7oz5kjoc.onion [.] nu.



```
Time                    Scr Port Source        Destination     Dsc Port Length Host
2017-01-27 22:55:51...  49216 172.16.4.193  194.87.234.1...    80    593 tyu.benme.com
2017-01-27 22:55:51...  49216 172.16.4.193  194.87.234.1...    80    632 tyu.benme.com
2017-01-27 22:55:51...  49215 172.16.4.193  194.87.234.1...    80    632 tyu.benme.com
2017-01-27 22:56:10...  49220 172.16.4.193  198.105.121....    80    340 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:10...  49220 172.16.4.193  198.105.121....    80    350 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:11...  49220 172.16.4.193  198.105.121....    80    441 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:11...  49220 172.16.4.193  198.105.121....    80    414 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:11...  49221 172.16.4.193  198.105.121....    80    398 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:11...  49222 172.16.4.193  198.105.121....    80    441 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:12...  49220 172.16.4.193  198.105.121....    80    439 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:12...  49221 172.16.4.193  198.105.121....    80    435 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:13...  49216 172.16.4.193  194.87.234.1...    80    765 tyu.benme.com
2017-01-27 22:56:13...  49215 172.16.4.193  194.87.234.1...    80    768 tyu.benme.com
2017-01-27 22:56:13...  49221 172.16.4.193  198.105.121....    80    269 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:15...  49223 172.16.4.193  194.87.234.1...    80    533 tyu.benme.com
2017-01-27 22:56:15...  49221 172.16.4.193  198.105.121....    80    445 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:16...  49221 172.16.4.193  198.105.121....    80    443 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:16...  49221 172.16.4.193  198.105.121....    80    433 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:16...  49220 172.16.4.193  198.105.121....    80    432 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:16...  49222 172.16.4.193  198.105.121....    80    527 p27dokhpz2n7nvgr.1jw2lx.top
2017-01-27 22:56:17...  49224 172.16.4.193  198.105.121....    80    457 p27dokhpz2n7nvgr.1jw2lx.top
```

Figure 5: Information gathering

From result analyze was we found that the domain used by cyber criminals, with the aid the Google search engine with the keywords p27dokhpz2n7nvgr.1jw2lx.top. Google.com search results show in Figure 6 describes found malware which infects PCs Stewie is CERBER Ransomware.



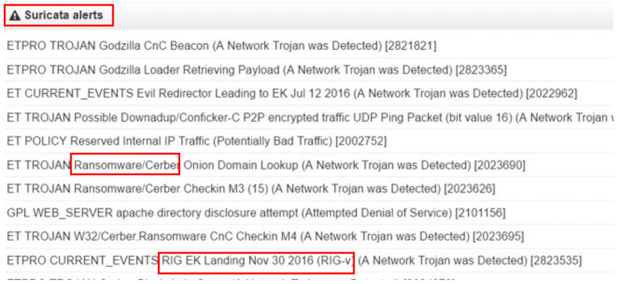
http://p27dokhpz2n7nvgr.1jw2lx.top/

About 576 results (0.57 seconds)

Cerber Payment Site: p27dokhpz2n7nvgr.1jw2lx.top
https://ransomwaretracker.abuse.ch/host/p27dokhpz2n7nvgr.1jw2lx.top/
Jan 27, 2017 - Cerber Payment Site **p27dokhpz2n7nvgr.1jw2lx.top**. Threat: Payment Site, Malware: Cerber. URL: **http://p27dokhpz2n7nvgr.1jw2lx.top**.

Figure 6: Result p27dokhpz2n7nvgr.1jw2lx.top

In Figure 7, we show the result of PCAP that has been uploaded to https://www.virustotal.com alert shows the results of Suricata that display found an actor/cybercriminal Cerber used RIG EK (Exploit Kit).

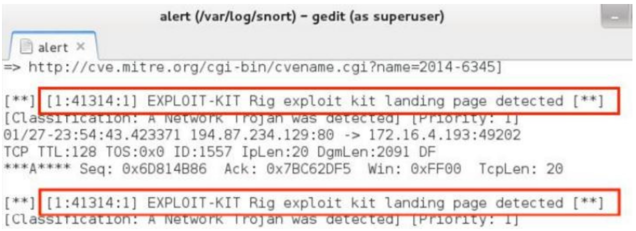


**Suricata alerts**

ETPRO TROJAN Godzilla CnC Beacon (A Network Trojan was Detected) [2821821]
ETPRO TROJAN Godzilla Loader Retrieving Payload (A Network Trojan was Detected) [2823365]
ET CURRENT_EVENTS Evil Redirector Leading to EK Jul 12 2016 (A Network Trojan was Detected) [2022962]
ET TROJAN Possible Downadup/Conficker-C P2P encrypted traffic UDP Ping Packet (bit value 16) (A Network Trojan
ET POLICY Reserved Internal IP Traffic (Potentially Bad Traffic) [2002752]
ET TROJAN Ransomware/Cerber Onion Domain Lookup (A Network Trojan was Detected) [2023690]
ET TROJAN Ransomware/Cerber Checkin M3 (15) (A Network Trojan was Detected) [2023626]
GPL WEB_SERVER apache directory disclosure attempt (Attempted Denial of Service) [2101156]
ET TROJAN W32/Cerber.Ransomware CnC Checkin M4 (A Network Trojan was Detected) [2023695]
ETPRO CURRENT_EVENTS RIG EK Landing Nov 30 2016 (RIG-v) (A Network Trojan was Detected) [2823535]

Figure 7: RIG exploit kit landing

Another scenario when the pcap file is ran on Snort as shown in Figure 8 RIG exploit kit landing page has detected. Exploit Kit (EK) is a server-based framework, exploitation by taking advantage of vulnerabilities in a software application that usually associated web browser and infects the victim without realizing have been infected. RIG EK is a gateway delivery and distribution of malware that functions direct the victim to execute a malware payload.



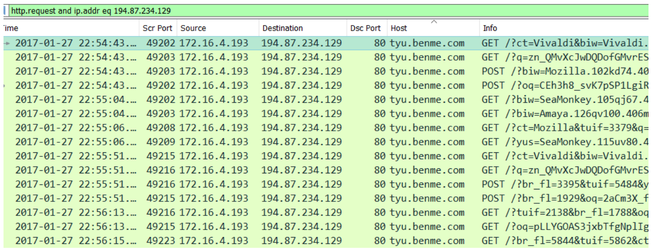
```
=> http://cve.mitre.org/cgi-bin/cvename.cgi?name=2014-6345]

[**] [1:41314:1] EXPLOIT-KIT Rig exploit kit landing page detected [**]
[Classification: A Network Trojan was detected] [Priority: 1]
01/27-23:54:43.423371 194.87.234.129:80 -> 172.16.4.193:49202
TCP TTL:128 TOS:0x0 ID:1557 IpLen:20 DgmLen:2091 DF
***A**** Seq: 0x6D814B86  Ack: 0x7BC62DF5  Win: 0xFF00  TcpLen: 20

[**] [1:41314:1] EXPLOIT-KIT Rig exploit kit landing page detected [**]
[Classification: A Network Trojan was detected] [Priority: 1]
```

Figure 8: Snort result

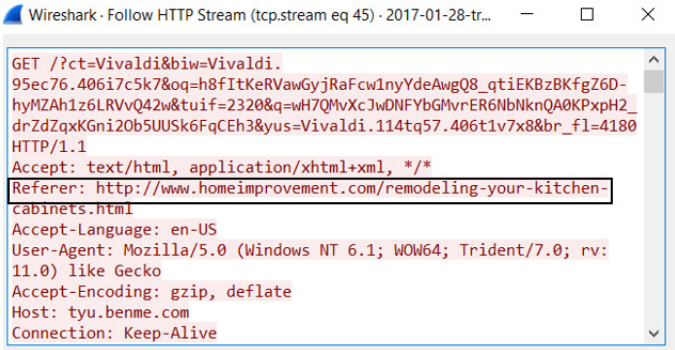
In Figure 9 shows the result of the filtering http.request and ip.addr eq 194.87.234.129 that shows the IP address associated with Rig EK. In general, the spread of Ransomware using two methods: first through malicious spam (mail spam) and Exploit Kit. Malicious spam (mail spam) is a way of spreading and distribution directly to a ransomware victims to enter the link that has been infected with malware and takes an active part on the victim to click a link or attachment files that have been injected malware. The second method is to use exploits Kit. Exploits Kit (EK) is designed to work behind the scenes, which is used by cyber criminals to automate the exploitation of security holes in the victim's machine when it is active browsing [35]. EK does not require such active actions of the victim clicks on a link or attachment.



```
http.request and ip.addr eq 194.87.234.129
ime                      Scr Port Source       Destination    Dsc Port Host          Info
2017-01-27 22:54:43.…   49202 172.16.4.193 194.87.234.129       80 tyu.benme.com GET /?ct=Vivaldi&biw=Vivaldi.
2017-01-27 22:54:43.…   49203 172.16.4.193 194.87.234.129       80 tyu.benme.com GET /?q=zn_QMvXcJwDQDofGMvrES
2017-01-27 22:54:43.…   49203 172.16.4.193 194.87.234.129       80 tyu.benme.com POST /?biw=Mozilla.102kd74.40
2017-01-27 22:54:43.…   49202 172.16.4.193 194.87.234.129       80 tyu.benme.com POST /?oq=CEh3h8_svK7pSP1LgiR
2017-01-27 22:55:04.…   49202 172.16.4.193 194.87.234.129       80 tyu.benme.com GET /?biw=SeaMonkey.105qj67.4
2017-01-27 22:55:04.…   49203 172.16.4.193 194.87.234.129       80 tyu.benme.com GET /?biw=Amaya.126qv100.406m
2017-01-27 22:55:06.…   49208 172.16.4.193 194.87.234.129       80 tyu.benme.com GET /?ct=Mozilla&tuif=3379&q=
2017-01-27 22:55:06.…   49209 172.16.4.193 194.87.234.129       80 tyu.benme.com GET /?yus=SeaMonkey.115uv80.4
2017-01-27 22:55:51.…   49215 172.16.4.193 194.87.234.129       80 tyu.benme.com GET /?ct=Vivaldi&biw=Vivaldi.
2017-01-27 22:55:51.…   49216 172.16.4.193 194.87.234.129       80 tyu.benme.com GET /?q=zn_QMvXcJwDQDofGMvrES
2017-01-27 22:55:51.…   49216 172.16.4.193 194.87.234.129       80 tyu.benme.com POST /?br_fl=3395&tuif=5484&y
2017-01-27 22:55:51.…   49215 172.16.4.193 194.87.234.129       80 tyu.benme.com POST /?br_fl=1929&oq=2aCm3X_f
2017-01-27 22:56:13.…   49216 172.16.4.193 194.87.234.129       80 tyu.benme.com GET /?tuif=2138&br_fl=1788&oq
2017-01-27 22:56:13.…   49215 172.16.4.193 194.87.234.129       80 tyu.benme.com GET /?oq=pLLYGOAS3jxbTfgNplIg
2017-01-27 22:56:15.…   49223 172.16.4.193 194.87.234.129       80 tyu.benme.com GET /?br_fl=5844&tuif=5862&ct
```

Figure 9: HTTP requests to the rig exploit kit internet protocol address

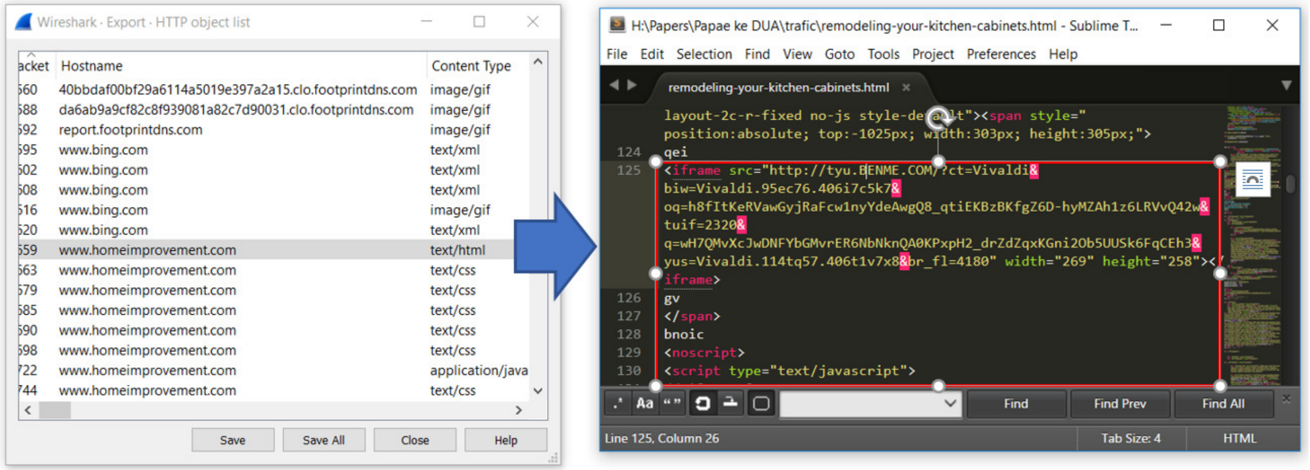
Filtering of HTTP requests on all IP addresses EK Rig in Wireshark, phase detects and analyze RIG EK and the website domain that mediates the spread of and infection of the host computer by way of the Following TCP Stream the packet as shown in figure 10. Following the results of the TCP stream shows the result found host computer is a www.homeimprovement.com address. From analysis of known victims access to bing.com is doing a search with keywords "remodeling your kitchen cabinets" in the address Referrer: http://www.bing.com/search?q=home+improvement+remodeling+your+kitchen\&qs=n\&sp=-1\&PQ=homeimprovement+++yourremodeling


Wireshark · Follow HTTP Stream (tcp.stream eq 45) · 2017-01-28-tr...

```
GET /?ct=Vivaldi&biw=Vivaldi.
95ec76.406i7c5k7&oq=h8fItKeRVawGyjRaFcw1nyYdeAwgQ8_qtiEKBzBKfgZ6D-
hyMZAh1z6LRVvQ42w&tuif=2320&q=wH7QMvXcJwDNFYbGMvrER6NbNknQA0KPxpH2_
drZdZqxKGni2Ob5UUSk6FqCEh3&yus=Vivaldi.114tq57.406t1v7x8&br_fl=4180
HTTP/1.1
Accept: text/html, application/xhtml+xml, */*
Referer: http://www.homeimprovement.com/remodeling-your-kitchen-
cabinets.html
Accept-Language: en-US
User-Agent: Mozilla/5.0 (Windows NT 6.1; WOW64; Trident/7.0; rv:
11.0) like Gecko
Accept-Encoding: gzip, deflate
Host: tyu.benme.com
Connection: Keep-Alive
```

Figure 10: Follow HTTP stream to find referrer

From result analyze www.homeimprovement.com been compromised website in spreading RIG EK. RIG EK is a sophisticated delivery method, the system for distributing malware via EK involves many other components in the chain of events malware infection. Basically, RIG EK with various tricks to direct traffic to the server EK users before sending malware. Actors used campaigns to guide


Wireshark · Export · HTTP object list
Save
Save All
Close
Help
H:\Papers\Papae ke DUA\trafic\remodeling-your-kitchen-cabinets.html - Sublime T...
File
Edit
Selection
Find
View
Goto
Tools
Project
Preferences
Help
remodeling-your-kitchen-cabinets.html
Aa
Find
Find Prev
Find All
Line 125, Column 26
Tab Size: 4
HTML

| acket | Hostname | Content Type |
|---|---|---|
| 560 | 40bbdaf00bf29a6114a5019e397a2a15.clo.footprintdns.com | image/gif |
| 588 | da6ab9a9cf82c8f939081a82c7d90031.clo.footprintdns.com | image/gif |
| 592 | report.footprintdns.com | image/gif |
| 595 | www.bing.com | text/xml |
| 502 | www.bing.com | text/xml |
| 508 | www.bing.com | text/xml |
| 516 | www.bing.com | image/gif |
| 520 | www.bing.com | text/xml |
| 559 | www.homeimprovement.com | text/html |
| 563 | www.homeimprovement.com | text/css |
| 579 | www.homeimprovement.com | text/css |
| 585 | www.homeimprovement.com | text/css |
| 590 | www.homeimprovement.com | text/css |
| 598 | www.homeimprovement.com | text/css |
| 722 | www.homeimprovement.com | application/java |
| 744 | www.homeimprovement.com | text/css |

```
layout-2c-r-fixed no-js style-default"><span style="
position:absolute; top:-1025px; width:303px; height:305px;">
qei
<iframe src="http://tyu.BENME.COM/?ct=Vivaldi&
biw=Vivaldi.95ec76.406i7c5k7&
oq=h8fItKeRVawGyjRaFcw1nyYdeAwgQ8_qtiEKBzBKfgZ6D-hyMZAh1z6LRVvQ42w&
tuif=2320&
q=wH7QMvXcJwDNFYbGMvrER6NbNknQA0KPxpH2_drZdZqxKGni2Ob5UUSk6FqCEh3&
yus=Vivaldi.114tq57.406t1v7x8&br_fl=4180" width="269" height="258"></
iframe>
gv
</span>
bnoic
<noscript>
<script type="text/javascript">
```

Figure 11: Export object list and pseudoDarkleech script

traffic to the victim server EK. Actors and campaigns two different terms, an actor may use one or several campaigns to distribute malware. One actor may have used the same campaigns to distribute various types of malware. The next stage was to determine the campaign's script used to deliver Cerber is a way to export object in the packet capture as shown in Figure 11.

PseudoDarkleech is a commonly used campaigns Cerber author, function to redirect traffic from the victim to Exploit Kit server with a stealth mode. The pseudoDarkleech script has the task of injecting web pages and a web server through on the root level.

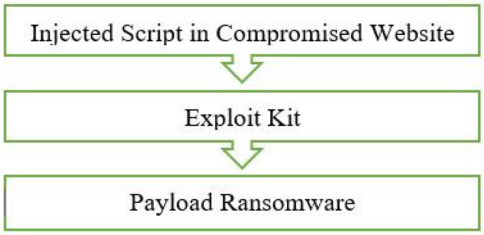


Figure 12: Chain of events pseudoDarkleech campaign

Explanation chain of events is as follows pseudoDarkleech campaign shown in Figure 12:

1) The first victim visits a website (compromised website) that have been compromised or malicious scripts injected and malicious script from compromised websites to make an HTTP request on Exploit Kit Landing Page.

2) Landing page EK finding and determine whether the computer has vulnerability are usually browser-based applications and Adobe flash player and furthermore sending EK Exploit to take advantage of the vulnerable application.

3) If the exploit is successful, EK sending payload Ransomware and carry out activities to access and encryption of files and folders unnoticed, the victim completely have been infected by the payload Ransomware.

## 5 Conlusion and Future Work

The use of Network Forensic Behavior Based successfully detect and analyze Cerber Ransomware as through the reconstruction Cerber Ransomware chain of events as shown in Figure 13.

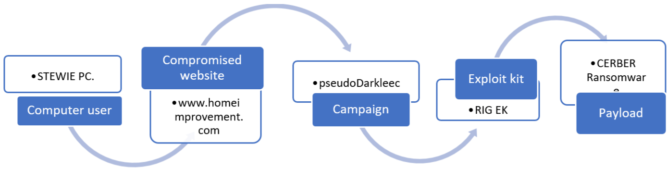


Figure 13: Chains of event

Started from the host computer named STIWIE PC, the victim then performs a search on a search engine bing.com for the referral advice from search engine bing.com STIWIE visits www.homeimprovement.com PC. Website analysis shows the results found have been injected by cyber criminals/actors of making the site into a Compromised Websites for the Campaign. The analysis phase detected Campaign successfully used pseudoDarkleec script to redirect a victim to the server by using RIG Exploit Kit EK to download a malware payload that named CERBER Ransomware for future work required Network Forensic deep on the side of compromised websites and Exploits Kit server. Exploit Kit is currently in delivery has encrypted binary code that has made it harder to be detected and analyzed.

Furthermore, the suggestion to users to stay updated browser application and patch vulnerability because the weakest point in the security chain is the human being, the solution is to strengthen the end point in a human side to build "Human Firewall".

## References

[1] M. Labs, "Understanding ransomware and strategies to defeat it," Tech. Rep. 1, Desember 2016.

[2] N. Khoa, T. Dat, M. Wanli, and S. Dharmendra, "An approach to detect network attacks applied for network forensics," in *2014 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, Xiamen University - China, Augustus 2014, pp. 655–660.

[3] E. Tim, "Ransomware: threat and response," *Network Security*, vol. 2016, no. 10, pp. 17–19, 2016.

[4] P. A. NETWORKS, "Exploit kit getting in by any means neccasary," Tech. Rep. 2, June. 2017.

[5] MalwareBytes, "State of malware report," Tech. Rep. 1, January 2017.

[6] A. Rab, A. Neville, A. Anand, C. Wueest, D. Tan, H. Lau, J. DiMaggio, J. Graziano, L. O'Brien, O. Cox, P. Coogan, S. Meckl, and Y. L. Chong, "Ransomware and businesses 2016," Tech. Rep. 1, July 2016.

[7] M. I. Trend, "Trendlabs sm 2016 1h security roundup," Tech. Rep. 1, January. 2017.

[8] M. Labs, "Mcafee labs threats report," Tech. Rep. 1, Jan 2017.

[9] S. Gordon, "Malvertising hits dating websites," *Network Security*, vol. 2015, no. 9, p. 2, 2015.

[10] A. A. Ali and Z. N. A. Kamarul, "Attack intention recognition: A review," *International Journal of Network Security*, vol. 19, no. 2, pp. 244–250, 2017.

[11] B. Ross, "Ransomware attacks: detection. prevention and cure," *Network Security*, vol. 2016, no. 9, pp. 5–9, 2016.

[12] S. Toshiki, Y. Takeshi, A. Mitsuaki, C. Daiki, and Y. Takeshi, "Efficient dynamic malware analysis based on network behavior using deep learning," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Washington. DC USA, December 2016, pp. 1–7.

[13] K. Clemens, C. PaoloMilani, and K. C. and, "Effective and efficient malware detection at the end host clemens," in *Proceedings of the 18th conference on USENIX security symposium*, Montreal. Canada, August 2011, pp. 70–82.

[14] Z. Mohd, S. Shahrin, A. M. Faizal, S. S. Rahayu, and H. C. Yun, "A comparative study on feature selection method for n-gram mobile malware detection," *International Journal of Network Security*, vol. 19, pp. 1–7, 2017.

[15] A. Saeed and S. Paul, "Optimised malware detection in digital forensics," *International Journal of Network Security*, vol. 6, no. 1, pp. 01–15, 2014.

[16] L. Andy, C. Armour, and B. pearce. Jack, "Ransomware becomes the most prevalent form of malware and hits an ever-wider range of victims," *Network Security*, vol. 2017, no. 2, pp. 1–2, 2017.

[17] P. Ebenezer and A. Aderemi, "Efficient feature selection technique for network intrusion detection system using discrete differential evolution and decision tree," *International Journal of Network Security*, vol. 19, no. 5, pp. 660–669, 2017.

[18] D. Sanjeev, L. Yang, Z. Wei, and C. Mahintham, "Semantics-based online malware detection: Towards efficient real-time protection against malware," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 2, 2016.

[19] M. M. H. and K. S. R., "Network forensic tool - concept and architecture," in *Fifth International Conference on Communication Systems and Network Technologies*, Gwalior. MP. India, April 2015.

[20] D. Sherri and H. Jonathan, *Network Forensics: Tracking Hackers Through Cyberspace.* Cambridge: Prentice Hall, 2012.

[21] K. Cabaj and W. Mazurczyk, "Using software-defined networking for ransomware mitigation: the case of cryptowall," *IEEE Network*, vol. 30, no. 6, pp. 14–20, 2016.

[22] Monika, Z. Pavol, and L. Dale, "Experimental analysis of ransomware on windows and android platforms: Evolution and characterization," *Procedia Computer Science*, vol. 94, pp. 465–472, 2016.

[23] Y. A. and M. Yung, "Cryptovirology: extortion-based security threats and countermeasures," in *IEEE Symposium on Security and Privacy*, vol. 5111, Oakland. CA. USA, May 1996, pp. 129–140.

[24] G. S. Sanjay and K. Kamalanathan, "Understanding and defending crypto-ransomware," *ARPN Journal of Engineering and Applied Sciences*, vol. 12, no. 12, pp. 3920–3925, 2017.

[25] P. Athina and K. Vasilios, "Differential malware forensics," *Digital Investigation*, vol. 10, no. 4, pp. 311–322, 2013.

[26] R. Pratyush and K. Prabhakar, "Network detection of ransomware delivered by exploit kit," *ARPN Journal of Engineering and Applied Sciences*, vol. 12, no. 12, pp. 3885–3889, 2017.

[27] A. Alexander and C. Anders, "The state of ransomware. trends and mitigation techniques," in *2017 IEEE East-West Design.*

[28] K. Ade, R. Imam, and L. Ahmad, "Forensic analysis and prevent of cross site scripting in single victim attack using open web application security project (owasp) framework," *Journal of Theoretical and Applied Information Technology*, vol. 95, no. 6, pp. 1363–1371, 2017.

[29] B. Nadia, K. Mohamed, Z. Khaled, and B. Chafika, "Iwnetfaf: An integrated wireless network forensic analysis framework," in *2016 Cybersecurity and Cyberforensics Conference (CCC)*, Amman. Jordan, Augustus 2016, pp. 35–40.

[30] J. R.C. and P. E. S, *Fundamentals of Network Forensics*. London: Springer, 2016.

[31] R. Imam, E. Jazi, A. Ahmad, and Subanar, "Log analysis techniques using clustering in network forensics," *International Journal of Computer Science and Information Security*, vol. Vol. 10, 2013.

[32] C. Eoghan, *Digital Evidence and Computer Crime*. Maryland: Elsevier Academic Press, 2011.

[33] B. M, C. J, and C. Z, "Forensic analysis of social networks (case study)," in *IProceedings of the ITI 2011. 33rd International Conference on Information Technology Interfaces (ITI 2011*, no. Augustust, Cavtat/Dubrovnik. Croatia, June 2013, pp. 219–223.

[34] S. James, S. Drew, and S. Visiting, *Cerber & KeRanger : The Latest Examples of Weaponized Encryption*. Pittsburgh: Institute for Critical Infrastructure Technology, 2016.

[35] L. Yassine and S. E. Mamouna, "An approach to detect network attacks applied for network forensics," in *2017 International Conference on Cyber Security And Protection Of Digital Services*, London, United Kingdom, Junes 2017, pp. 1–10.

# Biography

**Ade Kurniawan** received his Masters degree in Digital Forensic in 2014 from Universitas Islam Indonesia. He is currently lcturer Department of Informatics Engineering of Universal University. His research interests include Computer, Network Security, and Digital Forensics.

**Imam Riadi** is an Associate Professor, Department of Information System, Ahmad Dahlan University. Received his PhD Degree in Faculty of Sciences from the Universitas Gadjah Mada. His research interest includes computer security, Network Forensic, and Data Minning.

# A Reversible Watermarking Algorithm in the Lossless Mode of HEVC

KwangHyok Jo, Weimin Lei, Zhaozheng Li, and Xiaoshi Song

*(Corresponding author: Weimin Lei)*

School of Computer Science and Engineering, Northeastern University
Shenyang 110169, China
(Email: leiweimin@ise.neu.edu.cn)

*(Received May 5, 2017; revised and accepted Aug. 26, 2017)*

## Abstract

Reversible watermarking is considered to be an effective approach for copyright protection. However, it requires much more embedding capacity. To solve the problem, we consider reversible watermarking in the HEVC lossless mode which bypasses all nonreversible operations to improve the embedding capacity. Further, to prevent the performance loss of compression efficiency, the residual sample-based prediction method is applied together with the reversible watermarking. It is shown that the proposed algorithm not only provides enough space for embedding but also has the controlling capability. In addition, it does not degrade the performance of compression efficiency.

*Keywords: HEVC; Information Hiding; Reversible Watermarking; Watermarking*

## 1 Introduction

With advances in digital video service and digital video processing technology, legitimate video propagation is at great risk due to the ease of manipulation, tampering, redistribution and illegal copying of digital media. As such, the protection and enforcement of intellectual property rights for digital media has become an important issue. Watermarking technology plays an important role in securing multimedia data against illegal recording and retransmission [13]. Particularly, over recent years, a special kind of digital watermarking, namely the reversible watermarking, has been extensively studied, which not only provides the protection of copyright by embedding the assigned watermark into the original media but also enables the recovery of the original media from the suspected media [**?**, 11]. Therefore, the reversible watermark can be used to determine the ownership of the digital media by comparing the retrieved watermark with the assigned one. It is worth noting that despite its advantages, the reversible watermarking schemes may suffer certain performance loss due to the fact that additional recovery information has to be embedded into the original media. As such, how to guarantee the accuracy of the retrieved watermark under the constrained embedding capacity is considered to be the main challenge in the design of reversible watermarking.

High efficiency video coding (HEVC) [14], also known as H.265 and MPEG-H Part 2, is a loss video compression standard developed by Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG). Particularly, with almost twice the compression efficiency of the previous standard, *i.e.* H.264/AVC, HEVC has been widely applied in recording, compression, and distribution of high-resolution video contents [4, 8, 9]. It is worth noting that for applications such as medical imaging, preservation of artwork, image archiving, remote sensing, and image analysis, lossless compression is required. As such, with growing demand for these applications, the HEVC standard has specified the HEVC lossless mode to enable the lossless compression. It is also worth noting that HEVC lossless mode is achieved by skipping the transform, quantization, and in-loop filtering operations [1, 10].

In this paper, under the lossless mode, we propose a reversible watermarking algorithm for HEVC in lossless mode. It is worth noting that HEVC lossless mode bypasses all nonreversible procedures such as transform, quantization, the inverse operations, and all in-loop filtering operations including de-blocking filter, sample adaptive offset (SAO), and adaptive loop filter (ALF) in the encoder and decode. As such, under the HEVC lossless mode, the original video can be successfully retrieved from the watermarked video without any loss. It is shown that the proposed algorithm not only provides enough space for embedding of desired payload but also has the capability of controlling embedding capacity. In addition, compared with the HEVC lossless mode, the proposed algorithm does not degrade the performance of compression efficiency.

The remainder of this paper is organized as follows. In

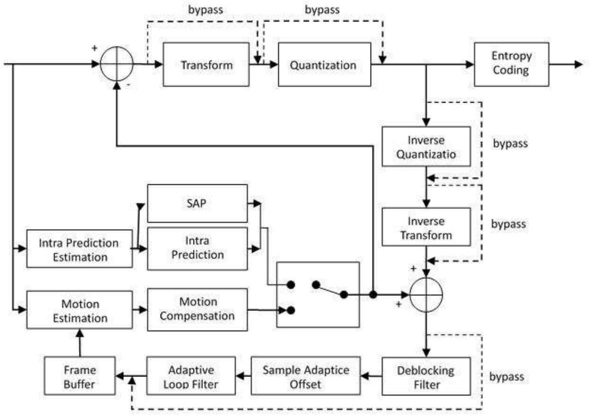


Figure 1: The framework of the HEVC lossless mode

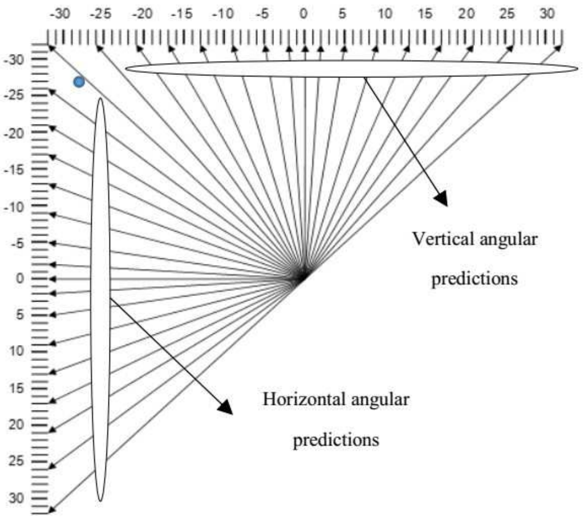


Figure 2: Angular intra prediction angle definition in HEVC

Section 2, a briefly overview of the HEVC lossless mode including its structure and base coding modules is presented. In Section 3, we analyze the distribution characteristics of prediction residual in the HEVC loss-less mode and the reversible watermarking algorithms for spatial domain. In Section 4, a reversible watermarking algorithm in the HEVC lossless mode is proposed. In Section 5, the feasibility and the performance of the proposed algorithm are verified through experiments. Finally, conclusions are presented in Section 6.

# 2 HEVC Lossless Coding Mode

## 2.1 Overview of the HEVC Lossless Mode

The overall structure of the HEVC lossless mode is shown in Figure 1. In Figure 1, dashed lines represent the bypass. It is assumed that all bypass operations are activated in the HEVC lossless mode [2].

In lossless coding, no distortion is allowed in reconstructed frames. Therefore, to achieve lossless coding, transform, quantization, their inverse operations, and all in-loop filtering operations including de-blocking filter, sample adaptive offset (SAO), and adaptive loop filter (ALF) are bypassed in the encoder and decoder since they are not reversible in general. Also, sample-based angular prediction (SAP) is applied to re-place the existing intra prediction method.

The design principle of HEVC lossless mode is to leverage the existing HEVC loss coding structure and to maximize logic sharing with the existing HEVC coding tools. As such, the implementation of HEVC lossless coding are subject to this special constraint in addition to common complexity and coding efficiency tradeoff consideration. The main coding modules of the HEVC lossless mode is described in detail in the following subsection.

## 2.2 Main Coding Modules of the HEVC Lossless Mode

### 2.2.1 SAP: Sample-Based Angular Prediction

As was noted previously, the transform and quantization operations are bypassed in lossless coding. Thus, for a pixel sample within a prediction unit (PU), the samples in its neighboring blocks as well as the immediate neighboring samples within the PU can be utilized for prediction. In order to explore spatial sample redundancy in intra-coded frame, SAP is employed instead of general HEVC intra prediction. As shown in Figure 2, 33 angles are defined and these angles are categorized into two classes: vertical and horizontal angular prediction.

The main idea of SAP is to use two neighboring samples in its top row as its reference samples to predict the current pixel when the prediction direction is vertical angular prediction and use two neighboring samples in its left column as its reference samples to predict the current pixel when the prediction angle is horizontal angular prediction. The SAP method can improve coding performance significantly.

### 2.2.2 Residual Sample-Based Prediction

The residual sample-based prediction was proposed as a method for enhancement of compression performance of HEVC lossless mode [3]. The flowchart of the residual sample-based prediction is shown Figure 3.

As shown in Figure 3, the residual sample-based prediction is implemented as a separate module. In this method, sample-based prediction is performed on the prediction residues without any modification to the existing intra prediction process. Also, the sample-based prediction process does not depend on the intra prediction mode and color component, the same prediction process is applied to all intra residual samples.

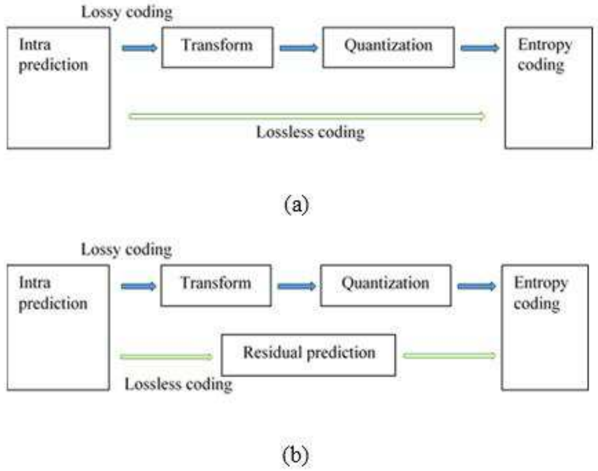


Figure 3: The framework of the residual sample-based prediction

| | |
|---|---|
| *c* | *b* |
| *a* | *x* |

Figure 4: Residual pixel and its neighbors

Figure 4 illustrates the neighboring residuals used in predictor. The prediction of the residual x is denoted as P(x) and is computed as Equation (1).

$$P(x) \quad = \quad \left\{ \begin{array}{l} \min(\mathrm{a,b}) \text{ if } c \geq \max(\mathrm{a,b}) \\ \max(\mathrm{a,b}) \text{ if } c \leq \min(\mathrm{a,b}) \\ a+b-c \text{ otherwise} \end{array} \right\} \qquad (1)$$

where min (a,b) and max (a,b) represent the larger and smaller function in calculating a and b, respectively.

If any neighboring samples are not available due to x being in the first row or column of a block, the same process is carried out with the missing values set to zero. It is reported that the proposed scheme improves the lossless intra coding performance by average 6.5%.

# 3 Analysis of Reversible Watermarking Algorithm

## 3.1 Distribution Characteristics of Residual Data in the HEVC Lossless Mode

In lossless coding, the residual data is not quantized transform coefficients but differential pixel after pre-diction. As a result, the residual data in lossless coding has different characteristics than that in the loss coding. Considering the Slide Editing test sequence as an example, the histogram of its intra luma prediction residuals in lossless coding and the histogram of its quantized transform coefficients when coded using QP=27 are shown in Figure 5 and Figure 6, respectively [6]. From these results, we observe that large luma prediction residuals appear with relatively higher frequency in the lossless case, with respect to the transform coefficients. In lossless coding, transform coefficients are usually concentrated at the upper left corner of the transform unit (TU) while it is not the case for quantized residuals. Prediction residuals may often appear in the bottom of a residual block. And it is observed that for relatively small TU, *e.g.* an 8 × 8 or a 4×4 TU, when intra prediction is in vertical direction, the residual will often appear in vertical direction. Thus, a vertical scan will often result in better performance. Similarly, when the intra prediction is in horizontal direction, a horizontal scan will often be better. The energy distribution of prediction residuals is often the inverse of the transform coefficients. For example, when diagonal scan is used, the energy in transform coefficients often decreases from the top-left corner to the bottom-right corner while the energy in the prediction residuals often increase from the top-left corner to the bottom-right corner.

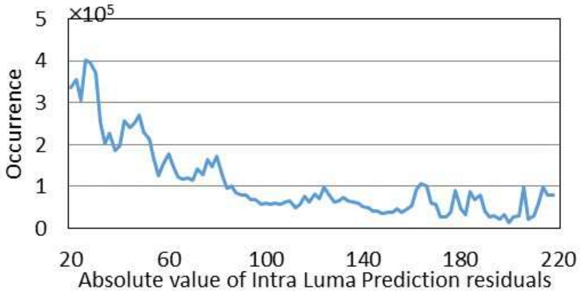


Figure 5: The histogram of intra luma prediction residuals

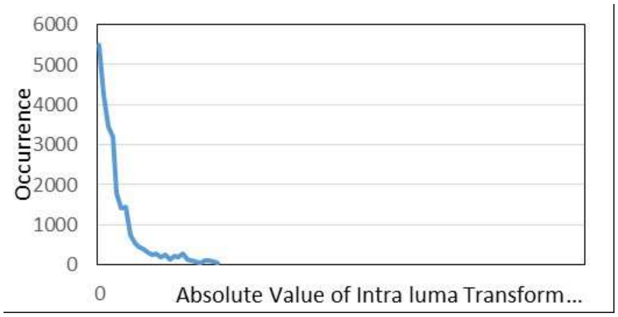


Figure 6: The histogram of intra coded luma transform coefficient levels

## 3.2 Analysis of Reversible Watermarking Algorithms

Prediction residuals are spatial domain data. For data over spatial domain three classic methods, the difference-expansion (DE) method, the histogram shifting method, and the prediction error based method, are widely studied for typical reversible watermarking algorithms, which are specified as follows.

A common feature of DE methods [12, 17] is using a decorrelation operator to create features with small magnitudes. Data can be embedded by expanding these features to create vacant space into which message bits are embedded. DE methods usually suffer from undesirable distortion when the values of features are large. It is worth noting that HEVC lossless mode is designed to support high video quality. Thus, DE method might not be suitable for applications where higher image quality is demanded.

In histogram-shifting based techniques [15, 16], a histogram of feature elements is created. As such, the data can be embedded by shifting histogram bins. Unfortunately, the capacity of histogram-shifting based techniques is low and highly depends on the histogram distribution of the cover image. In general, the higher the peak of image histogram, the more the embedding capacity is. But, as observed from Figure 5 and Figure 6, compared with H.264/AVC, the HEVC lossless mode does not concentrate on certain values and instead has several peak points in histogram distribution of prediction residual. This fact indicates that the histogram-shifting based techniques is not suitable for the prediction residual of HEVC.

The prediction error based method is considered as another type of reversible watermarking method [7, 18]. Particularly, Hong *et al.* has proposed a reversible data hiding technique based on modification of prediction errors (MPE). In MPE, since the histograms in the domain of prediction errors are sharply distributed, the embedding capacity is higher than that of traditional histogram-shifting method for the same image quality. The embedding process of modification of prediction errors (MPE) involves calculating the prediction errors from the neighborhood of a given pixel and then embedding the message bits in the modified prediction errors. The median edge detection (MED) predictor is used in MPE to predict pixel values, which is same as that used in the prediction residual. Particularly, based on the values of context pixels a, b, and c, MED predicator is able to predict current pixel x by applying raster scan order and edge rule, which is as shown in Figure 4. The output of the predicted value $\hat{x}$ is given by Equation (1). Assuming that the predicted result of pixel x is , the prediction error e can be then obtained by subtracting the prediction from x, *i.e.*, $e = x - \hat{x}$.

MPE not only has the capability to control the capacity-PSNR, where fewer data bits need less error modification, but also can be applied to images with flat histogram, which is in accordance with the distribution characteristics of prediction residual of HEVC. Also, the calculation method in the prediction errors is consistent with the prediction residual method for improvement of performance in HEVC lossless mode, which means that MPE algorithm can be as an efficient solution for implement of reversible watermarking in HEVC lossless mode.

# 4 Proposed Algorithm

## 4.1 Embedding Algorithm

According to the structure of HEVC encoder, it can be classified into in-feedback loop and out-feedback loop for information embedding position. It is worth noting that if the watermarking is performed in the in-feedback loop, it is difficult to retrieve original video. This is because that the embed information of the current block can only be used for reconstruction and prediction of next block. Therefore, in order to implement reversible watermarking, it is favorable to embed information in out-feedback loop.

In this paper, the MPE reversible watermarking algorithm is implemented during the residual prediction progress in HEVC lossless mode. As mentioned above, if the residual sample-based prediction is applied after intra prediction in HEVC lossless mode, the compression efficiency can be improved by 6.5% on average. The residual sample-based prediction operation in HEVC lossless mode is same as the calculation operation of prediction error in MPE. Thus, to implement MPE during the residual sample-based prediction progress in HEVC lossless mode, we need to add the watermark embedding operation into the residual sample-based prediction progress.

In HEVC, there are four effective intra prediction block, in which the sizes of samples ranges from $4 \times 4$ to $32 \times 32$. Due to the characteristics of human visual system, the video visual quality distortion caused by watermarking embedding is more visible in smooth regions than that in complex regions. As such, it is more beneficial to use relatively large size blocks for prediction in smooth regions. While, in complex regions, the small size blocks are more preferable. It is worth noting that for small size blocks, the change of the block size may have only trivial impact on visual quality of the video. Therefore, we selected $4 \times 4$ luma block of intra predicted I frame as embedding region.

The proposed algorithm is performed independently in each transform block such that the change of the currant block due to the watermarking does not impact other blocks. Also, since we select block of size $4 \times 4$ as the watermark embedding block, the value of M and N is 3 in MPE.

In addition, the proposed algorithm is implemented in a manner which is independent of intra prediction modes and color components, the proposed algorithm is applied to all $4 \times 4$ intra residual samples. The frame-work of the proposed algorithm is shown in Figure 7, in which the prediction residual block in Figure 3 (b) is substituted to watermarking block.

The detailed embedding steps are listed as (Algorithm 1).

## 4.2 Extraction Algorithm

Through entropy decoding, the PU residual of size $4 \times 4$ can be obtained. For all PU of size $4 \times 4$, we use the same

**Algorithm 1** The Embedding algorithm

```
Input: An intra predicted residual block I, and the watermark bit stream W.
Output: Watermarked block $I'$ of size $4 \times 4$, the end of embedding position L and the auxiliary information A.
1:  if block size is $4 \times 4$ then
2:      go to STEP 6,
3:  else
4:      sent the residual sample-based prediction block
5:  end if
6:  Prepare an empty matrix $I'$. Then, initialize the pixel values in the first row and first column of $I'$, respectively, to the pixel values in the corresponding position of I.
7:  for $1 \leq i \leq 3$, $1 \leq j \leq 3$ do
8:      scan each pixel $I'_{i.j}$ by using raster scan order.
9:  end for
10: if $I'_{i,j} = 0$ or $I'_{i,j} = 255$ then
11:     set $I'_{i,j} = I_{i,j}$ and record the position (i, j) in array A as the auxiliary information for recovering, then proceed to next pixel.
12: end if
13: Use Equation (1) to predict the value of $I'_{i.j}$ by setting $a = I'_{i,j-1}$, $b = I'_{i-1,j}$ and $c = I'_{i-1,j-1}$. Let the predicted result to be $\hat{I}'_{i,j}$.
14: Calculate prediction error e. Prediction error e is the difference between pixel $I_{i,j}$ and its predicted result $\hat{I}'_{i,j}$, $i.e.$, $e = I_{i,j} - \hat{I}_{i,j}$.
15: if all the watermark bits in W have been embedded then
16:     set L = (i, j), and go to STEP 33.
17: end if
18: if e = 0 or e = -1 then
19:     go to STEP 23 for data embedding.
20: else
21:     go to Step 30.
22: end if
23: if the to-be-embedded bit is 0 then
24:     the prediction error e remains unchanged.
25: else if the to-be-embedded bit is 1 and e = 0 then
26:     modify the value of e to e+ 1
27: else if if e = -1 then
28:     modify the value of e to e - 1. After embedding, go to Step 33.
29: end if
30: if e ¿ 0 then
31:     modify the value of e to e + 1, if e ¡ -1, then modify the value of e to e-1.
32: end if
33: Set $I'_{i,j} = \hat{I}'_{i,j} + e$.
34: if $i \neq M - 1$ or $j \neq N - 1$ then
35:     update the index i and j, go to STEP 7.
36: else
37:     the embedding procedure is completed.
38: end if
39: End
```

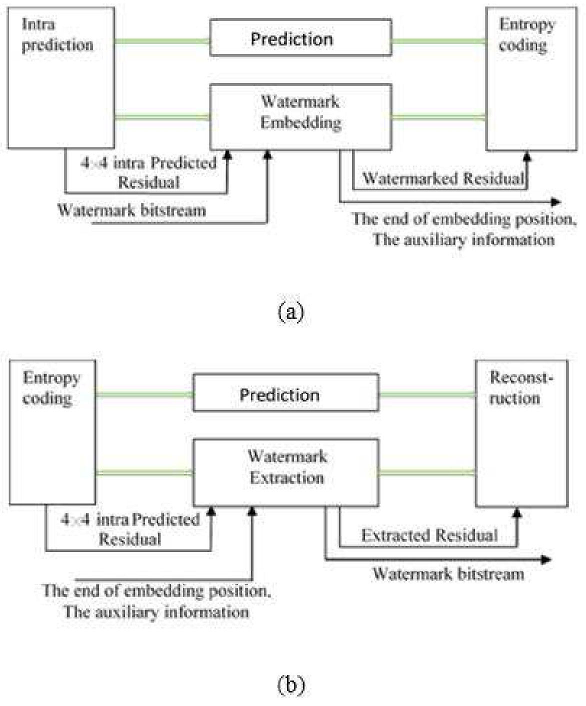


Figure 7: The framework of proposed algorithm within the HEVC codec

scan order as in the embedding phase to predict pixel values, and then calculate the prediction error e. It is worth noting that, if the value e is 0 or -1, the embedded watermark bit is 0, while if the value e is 1 or -2, the embedded watermark bit is 1, else if the value e is not one of the four numbers -2, -1, 0 and 1, then there is no bit embedded. Since we have modified the prediction errors during embedding, the original video can be restored by modifying the prediction errors back to their original value. The detailed steps for extracting watermark and recovering the original video are listed as (Algorithm 2).

## 5 Experimental Results and Analysis

In this section, we evaluate the feasibility and the performance of the proposed algorithm in terms of the embedding capacity and the peak signal to noise ratio (PSNR) difference. We also compare the compression performance of the proposed algorithm with the HEVC lossless mode and the residual sample-based prediction mode, respectively.

The proposed algorithm is simulated in the HM-16.9 model of the HEVC reference software, and the experiments are performed by embedding and extracting random generated bit streams. In the experiments, RDOQ, de-blocking filter, SAO and ALF are disabled. Also, for HE configurations, InternalBitDepth is set to equal to InputBitDepth in the configuration files. HEVC lossless mode is useful in coding video sequences with mixed contents, *e.g.* natural video with overlaid text and graph-

**Algorithm 2** The Extraction algorithm

```
Input: Stego frame $I'$, the end of embedding position L and auxiliary information A.
Output: The bit stream W and the original frame $I''$.
 1: Prepare a matrix $I''$ to store the recovered block. The size of $I''$ is the same as stego frame $I'$. Initialize the pixel values in the first row and the first column of $I''$ to the pixel values in the corresponding positions of $I'$.
 2: for $1 \leq i \leq 3$; $1 \leq j \leq 3$ do
 3:     scan each pixel $I'_{i,j}$ in the stego frame by using the raster scan order.
 4: end for
 5: if (i, j) was recorded as the auxiliary information in A then
 6:     set $I''_{i,j} = I'_{i,j}$ and proceed to next pixel.
 7: end if
 8: Predict the value $\hat{I}'_{i,j}$ using Equation (1), by setting $a = I'_{i,j-1}$, $b = I'_{i-1,j}$ and $c = I'_{i-1,j-1}$. Suppose the predicted value is $\hat{I}'_{i,j}$.
 9: Calculate the prediction error $e = I_{i,j} - \hat{I}_{i,j}$.
10: According to the parameter L, decide whether all the embedded information has been extracted or not. If they are, then go to STEP 24.
11: if e = 0 then
12:     the embedded watermark bit is 0, and the prediction error e remains unchanged.
13: else if e = 1 then
14:     the embedded watermark bit is 1, and the prediction error e is modified to e - 1.
15: else if e = -1 then
16:     the embedded watermark bit is 0, and the prediction error e remains unchanged.
17: else if e = -2 then
18:     the embedded watermark bit is 1, and the prediction error e is modified to e + 1.
19: else if e ¿ 1 then
20:     prediction error e is modified to e - 1.
21: else if e ¡ -2 then
22:     the prediction error e is modified to e + 1.
23: end if
24: Set $I''_{i,j} = \hat{I}'_{i,j} + e$.
25: if $i \neq 3$ or $j \neq 3$ then
26:     update the value of i and j. After that, we first go to the next pixel, and then go to STEP 5.
27: else
28:     we have finished the watermark extraction and original video recovery. It is worth not-ing that the obtained result $I''$ is same as original block I.
29: end if
30: End
```

ics, thus we selected F class sequences (BasketballDrillText, ChinaSpeed, SlideEditing, SlideShow) for testing and each sequence is shown in Figure 8. Parameters of test sequences are shown in Table 1 [5].

First, we evaluate the embedding capacity of the proposed algorithm. Table 2 shows distribution of $4 \times 4$ residual in all size of residual for HEVC lossless mode. As shown in Table 2, $4 \times 4$ prediction residual occupies most among all prediction blocks.

Table 3 shows the embedding capacity of proposed algorithm in compared with DE method. As shown in Table 3, our algorithm is 2.03 times in embedding capacity over DE method. Therefore, the prediction residual has more changeable elements for embedding of watermark. Intuitively, this is because that the prediction residual has more non-zero coefficients than transform coefficients or quantized coefficients.

We also calculate the objective video coding quality variation (PSNR) difference in experiments. Peak Signal-to-Noise Ratio (PSNR) measure has been used to analyze the quality of watermarked video with respect to original video, which is given as Equation (2) shown.

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right) \tag{2}$$

The difference of the value of PSNR before embedding watermark and after embedding watermark $\Delta PSNR$ is defined by:

$$\Delta PSNR = PSNR' - PSNR.$$

Where $PSNR'$ and PSNR are the video coding quality before and after embedded watermark, respectively, and MSE (mean square error) is a measure which is used to quantify the difference between the initial video frame I and the stego video frame $I'$.

If the video frame has a size of $M \times N$ then:

$$MSE = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(I\left(i.j\right) - I'\left(i,j\right)\right)$$

Figure 9 shows the performance evaluation curve between the embedding capacity and PSNR degradation. The proposed algorithm is only 1.2dB in PSNR degradation when embedding capacity is 100000 bit. This means that proposed algorithm efficiently performs the watermarking.

Finally, we compare compression performance of proposed algorithm with that of the standard HEVC lossless mode and residual sample-based prediction mode. Table 4 lists the comparison result for compression performance in case of embedding of 100000 bit watermarks.

Table 4 lists coding performance comparison results among the proposed algorithm, the standard HEVC lossless mode and the residual sample-based prediction.

Where kbps % $^1$ compare with the standard HEVC loss-less mode, and kbps % $^2$ compare with the residual sample-based prediction.

Table 1: Parameters of test sequences

| Class | Sequence name | Resolution | Frame count | Frame rate |
|---|---|---|---|---|
| F | BaskeballDrillText | 480 × 480 | 500 | 50fps |
| F | ChinaSpeed | 1024 × 768 | 500 | 30fps |
| F | SlideEditing | 1280 × 720 | 300 | 30fps |
| F | SlideShow | 1280 × 720 | 500 | 20fps |

Table 2: Distribution of $4 \times 4$ residual in all size of residual for HEVC lossless mode

| Sequence name | Distribution of $4 \times 4$ residual (%) |
|---|---|
| BaskeballDrillText | 93.09 |
| ChinaSpeed | 92.15 |
| SlideEditing | 84.13 |
| SlideShow | 78.34 |

Table 3: Average embedding capacity per frame

| Sequence name | DE | Proposed | Increased capacity (%) |
|---|---|---|---|
| BaskeballDrillText | 39856 | 86 248 | 2.16 |
| ChinaSpeed | 62700 | 163 002 | 2.60 |
| SlideEditing | 45328 | 90 624 | 2.00 |
| SlideShow | 32082 | 43 828 | 1.37 |

Table 4: The coding performance comparison results

| Sequence name | kbps % [1] | kbps % [2] |
|---|---|---|
| BaskeballDrillText | -0.3 | 1.1 |
| ChinaSpeed | -7.2 | 0.6 |
| SlideEditing | -1.2 | 0.9 |
| SlideShow | -12.0 | 1.4 |

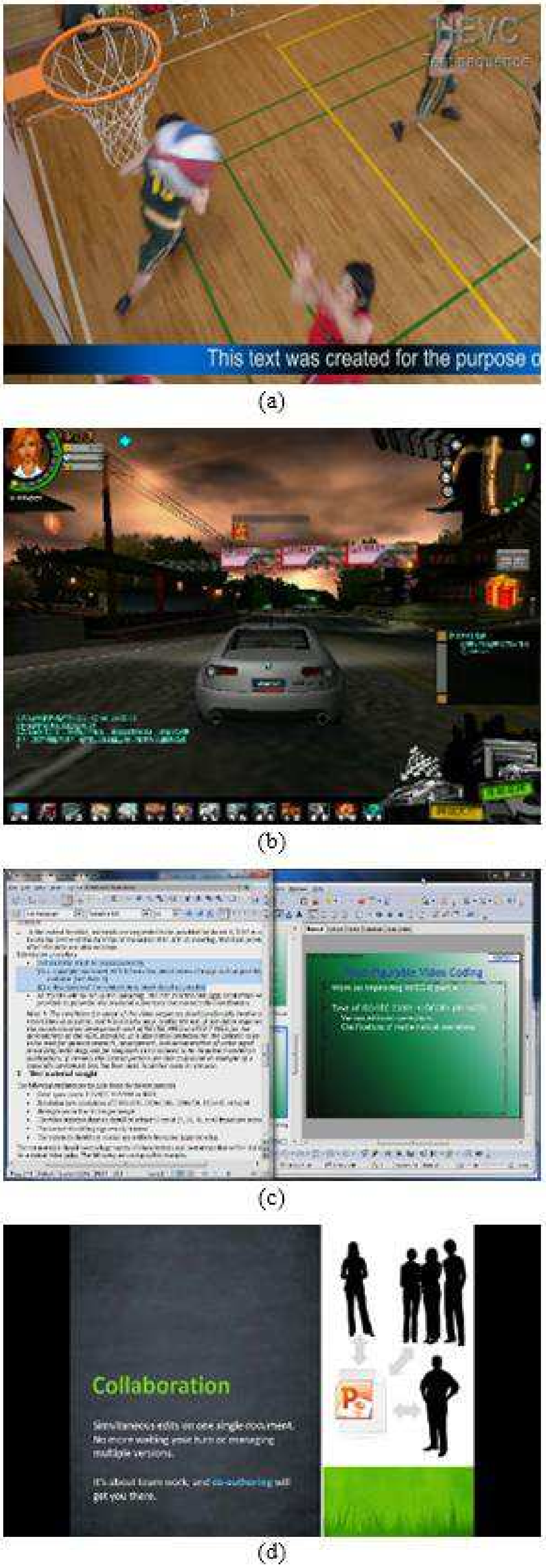
It can be observed that our proposed algorithm outperforms the standard HEVC lossless mode by 5.2% on average in terms of the compression efficiency. Such observation indicates that, compared with the HEVC lossless mode, the proposed algorithm does not degrade the performance of compression efficiency. It is also observed that compared with the residual sample based prediction, our proposed algorithm has a performance loss of 1% with respect to the compression efficiency.

## 6 Conclusions

In this paper, an efficient reversible watermarking algorithm in HEVC lossless mode is proposed. HEVC lossless mode is suitable for the implement of reversible watermarking since it bypasses all nonreversible operations such as transform, quantization and in-loop filtering operations. In this paper, we have analyzed the distribution characteristics of prediction residual of HEVC and reversible watermarking algorithms in spatial domain, and selected the reversible algorithm suitable for prediction residual. The proposed algorithm not only provides enough space for embedding of desired payload but also has the capability of controlling embedding capacity. In addition, compared with the HEVC lossless mode, the proposed algorithm does not degrade the performance of compression efficiency. Our experiment results indicate that the averaged embedding capacity of this algorithm is at least two times more than that of DE technique. It is also shown that the compression efficiency of the proposed algorithm outperforms the standard HEVC lossless mode by 5.2%.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 61401081), the Fundamental Research Funds for the Central Universities (No. N150404005), and the Ministry of Education - China Mobile Research Funds (No. MCM20150103).


HEVC
Test sequence
This text was created for the purpose o
Collaboration
Simultaneous edits on one single document. No more waiting your turn or managing multiple versions.
It's about team work, and co-authoring will get you there.

(a)

(b)

(c)

(d)

Figure 8: Test sequence: (a) BaskeballDrillText, (b) ChinaSpeed, (c) SlideEditing, (d) SlideShow

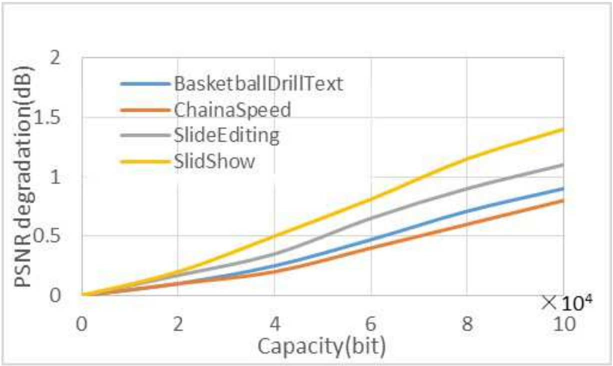


Figure 9: Performance evaluation curve between the embedding capacity and PSNR difference

## References

[1] Samira Bouchama, Hassina Aliane, and Latifa Hamami, "Reversible data hiding scheme for the h.264/avc codec," in *International Conference on Information Science and Applications*, pp. 1–4, Suwon, South Korea, June 2013.

[2] Fangdong Chen and Houqiang Li, "Improved rate-distortion optimization algorithms for hevc lossless coding," in *International Conference on Multimedia Modeling*, pp. 454–465, Sydney, NSW, Australia, January 2015.

[3] Jung Ah Choi and Yo Sung Ho, "Efficient residual data coding in cabac for hevc lossless video compression," *Signal Image & Video Processing*, vol. 9, no. 5, pp. 1055–1066, 2015.

[4] Reuben A. Farrugia, "Reversible visible watermarking for h.264/avc encoded video," in *Eurocon - International Conference on Computer As A Tool*, pp. 1–4, Lisbon, Portugal, April 2011.

[5] Rosewarne C. Flynn D, "Common test conditions and software reference configurations for hevc range extensions," in *Proceedings of the 14th Meeting of Joint Collab-orative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11*, 2013.

[6] Wen Gao, Minqiang Jiang, and Haoping Yu, "On lossless coding for hevc," in *IS&T / SPIE Electronic Imaging*, pp. 8666 – 8666 – 13, Burlingame, California, U.S.A., February 2013.

[7] Wien Hong, Tung Shou Chen, and Chih Wei Shiu, *Reversible data hiding for high quality images using modification of prediction errors.* Elsevier Science Inc., 2009.

[8] Byung Gyu Kim, Kostas Psannis, and Dong San Jun, *Special issue on architectures and algorithms of high-efficiency video coding (HEVC) standard for real-time video applications.* U.S.A.: Springer-Verlag New York, Inc., 2016.

[9] Songbin Li, Jiangyun Fu, Peng Liu, and Yuxin Jiang, "An information hiding approach based on integer transform coefficient and virtual space encoding for

h.264/avc," *Circuits Systems & Signal Processing*, vol. 34, no. 11, pp. 1–22, 2015.

[10] Yunxia Liu, Leiming Ju, Mingsheng Hu, Xiaojing Ma, and Hongguo Zhao, "A robust reversible data hiding scheme for h.264 without distortion drift," *Neurocomputing*, vol. 151, no. 1, pp. 1053–1062, 2015.

[11] Nirmal S. Nair, Tojo Mathew, A. S. Neethu, Viswajith P. Viswanath, Madhu S. Nair, and M. Wilscy, "A proactive approach to reversible data hiding in encrypted images," *Procedia Computer Science*, vol. 46, pp. 1510–1517, 2015.

[12] Fei Peng, Xiaolong Li, and Bin Yang, "Adaptive reversible data hiding scheme based on integer transform," *Signal Processing*, vol. 92, no. 1, pp. 54–62, 2012.

[13] Christine I Podilchuk and Edward J Delp, "Digital watermarking: algorithms and applications," *Signal Processing Magazine IEEE*, vol. 18, no. 4, pp. 33–46, 2001.

[14] G. J. Sullivan, J. Ohm, Woo Jin Han, and T. Wiegand, "Overview of the high efficiency video coding (hevc) standard," *IEEE Transactions on Circuits & Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, 2012.

[15] Wei Liang Tai, Chia Ming Yeh, and Chin Chen Chang, "Reversible data hiding based on histogram modification of pixel differences," *IEEE Transactions on Circuits & Systems for Video Technology*, vol. 19, no. 6, pp. 906–910, 2009.

[16] D. M. Thodi and J. J. Rodrguez, "Expansion embedding techniques for reversible watermarking," *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 16, no. 3, p. 721, 2007.

[17] Jun Tian, "Reversible data embedding using a difference expansion," *IEEE Transactions on Circuits & Systems for Video Technology*, vol. 13, no. 8, pp. 890–896, 2003.

[18] Cabir Vural and Burhan Barakl, "Adaptive reversible video watermarking based on motion-compensated prediction error expansion with pixel selection," *Signal Image Video Processing*, vol. 10, no. 7, pp. 1225–1232, 2016.

# Biography

**KwangHyok Jo** is a Ph.D. candidate with School of Computer Science and Engineering, Northeastern University, China. His current research interests include DRM technology, information hiding, network security.

**Weimin Lei** is a professor and Ph.D supervisor with School of Computer Science and Engineering, Northeastern University, China. His current research interests include protocols and services in IP multimedia system, multipath transport, overlay network, SDN and adhoc network.

**Zhaozheng Li** is a Ph.D. candidate with School of Computer Science and Engineering, Northeastern University, China. His current research interests include IP communication, trusted technology, network security.

**Xiaoshi Song** is a lecturer and doctor with School of Computer Science and Engineering, Northeastern University, China. His current research interests include Computer networks, optical networks.

# Feature Extraction Optimization for Network Intrusion Detection in Control System Networks

Travis Atkison[1], Stanislav Ponomarev[2], Randy Smith[1] and Bernard Chen[3]
(Corresponding author: Travis Atkison)

Computer Science Department, University of Alabama[1]
PO Box 870290, Tuscaloosa, AL 35487, USA
(Email: atkison@cs.ua.edu)
BBN Technologies, Cambridge, MA, USA[2]
Computer Science Department, University of Central Arkansas[3]
201 Donaghey Ave, Conway, AR 72035, USA

(Received May 24, 2017; revised and accepted Aug. 27, 2017)

## Abstract

Security measures for Industrial Control Systems (ICSs), until recently, have come mainly in the form of a physical disconnect by implementing an "air-gap". This disconnect isolated the nodes of an ICS network from other networks, including the Internet. While connecting an ICS network to the Internet is beneficial to both the engineers and companies that operate them, it places these ICSs in a situation where they are vulnerable to attacks as the protocols that are used by several of the ICSs have very little, if any, security mechanisms. This paper focuses on optimization of the feature extraction algorithms used in a continuing effort to develop a Network Telemetry based Intrusion Detection System (IDS). After development and testing of the optimizations described in this paper, the developed IDS was able to achieve 99.99% accuracy when differentiating between machines of an attacker and engineer on the same network.

*Keyword: Control System Security; Network Intrusion Detection; Network Telemetry*

## 1 Introduction

The use of Industrial Control Systems (ICS) has become common place in many businesses. While typically seen in the utilities industries (water, gas, oil, electricity), they are not limited to such. For instance, some factories use ICSs to control their assembly lines. Theme parks can use them to control their rides. Even the International Space Station uses control systems for many different operations. ICSs are designed to take complex data, process it, and complete designated tasks.

Since ICSs are used to control many sensitive operations, security is of extreme importance. Early ICSs used a security scheme known as "air-gap." Every node in an air-gapped ICS network was physically disconnected from other networks [1] and the Internet, as it was not widely used then. This physical disconnection meant that an attacker would need physical access to the system in order to perform an attack. Therefore, cyber security for an ICS was not considered a priority. Instead, the designers focused on the safety of the physical system operations and the availability of the ICS network.

"Air-gap" security measures worked well enough as long as the physical disconnect was maintained. However, as the Internet became widely used and networking equipment became more affordable, companies began to realize that they could save time and money by attaching their ICSs to the Internet. Engineers would no longer need physical access to the individual ICS networks. They would be able to monitor the ICS and address critical issues from anywhere.

Although there were benefits, attaching ICSs to the Internet exposed the ICS networks to new security threats. The main issue was the accessibility of the ICS network to would-be attackers. No longer did an attacker have to gain physical access to the network to launch an attack. Instead, they could use the Internet to access and launch a wide array of attacks against the ICS, such as non-patched system attacks, DDoS attacks, zero-day attacks, *etc.* ICS protocols were originally developed assuming "air-gapped" accessibility; therefore, cyber security measures were not designed into the protocols. By attaching these devices to the Internet, they became much more vulnerable to attacks.

Another issue inherent with an Internet connected ICS network is the difficulty of verifying the authenticity of the host sending the data and the data itself. Information, such as the sender's location, connection properties, configuration and architecture, is kept hidden from the control applications. This makes it much more difficult to authenticate the host transmitting the data [16]. It also makes it easier for attackers to change the data within the data packet but make the packet itself appear authentic [26].

As vulnerabilities were exposed in Internet connected ICSs, it became apparent to the engineers that industry wide standards were lacking, often not used, and some companies were even creating their own. Although a few standards were available, companies were not required to use them. Having a wide array of ISC protocols made it hard to determine what the vulnerabilities were for all ICSs [12]. One protocol might have vulnerabilities that another protocol does not, and vice versa. Therefore, engineers were forced to use more time, energy and resources to find and fix the vulnerabilities.

One approach to addressing cyber-security in an ICS network is a network telemetry-based Intrusion Detection System (IDS). This approach looks at data that is sent through the network that is not used by the transmission protocol. This data is known as network telemetry. It includes information such as packet size, number of dropped packets, when the packet arrives, session sizes and times. The IDS, then, seeks to measure the telemetry data and verify it. In order to verify the data, the IDS must look at all the packets that pass through the network. From these packets, the IDS can calculate the telemetry data's running average and look for anomalies. When anomalies are found, the system triggers an alert for the engineers to investigate.

## 2 Background

Many cyber-attacks have been launched against ICSs [14]. Stuxnet is the most well-known example [10]. It was used to infiltrate the Iranian nuclear facilities' ICSs and change the spinning frequencies of the centrifuge motors. This caused damage to hundreds of Iranian centrifuges. Stuxnet was a well thought out, well written piece of malicious code that was able to duplicate and spread itself across a network. While it was able to spread itself over the Internet, it did not necessarily need an Internet connection. It was able to infect any USB drive that was inserted into an infected computer and use that USB drive to infect other computers [10, 22, 29]. Stuxnet was extremely difficult to find and made the centrifuge breakdowns look like unfortunate accidents.

Stuxnet opened the eyes of many ICS engineers to the need for cyber security measures to be added to ICS protocols. The consequences of a successful ICS cyber attack were too great for engineers to sit back and do nothing. Therefore, significant research is now being done on how to secure cyber-physical systems. The research has covered a wide array of approaches. Some focus on preventing a specific type of attack, i.e. denial of service or man-in-the-middle. Others focus on topics such as state estimation, traffic analysis, and hardware analysis.

One proposed method is to use a hardware fingerprint and duration [24–26]. By using patterns found in the communications of ICS network nodes, the researchers are able to create a digital fingerprint for each node that can be used to authenticate the data. In [33], Wallace *et al.* propose an IDS that can distinguish between the state of an ICS under attack and the state of an ICS operating normally. They suggest that the objects being controlled by the ICS have distinguishable underlying physical properties that can be used for state estimation.

Another approach has been to focus on network traffic. Carcano *et al.* [6] suggest an IDS that looks for state anomalies using ICS network traffic. It creates an image of the ICS from the network traffic that flows through it and uses that image to find state anomalies within the system. Temporal packet data, as seen in [28], can also be used to find anomalies in network traffic. The IDS developed in [28], which focused on the BACnet protocol, used probability functions to examine packet signatures and determine if the packet was authentic. The drawback to their method was the number of false positives resulting from anomalies found in the ICS's physical domain for which the engineers attempted to fix by reprogramming the programmable logic controllers (PLCs).

Cheung *et al.* [8] propose a model-based intrusion detection system for SCADA networks. In their system, an analysis of the communication model using Modbus is done to look for signatures of packet fields. When specific signatures are detected, an alarm is triggered.

Long *et al.* [17] chose to focus on denial of service attacks. They proposed a packet filtering system to help mitigate these types of attacks. Oh *et al.* [23] focused on man-in-the-middle attacks and suggested some modifications be made to the ICS protocols. The National Institute of Standards and Technology recommends using firewalls, MAC address locking and encryption, among other things, to help prevent man-in-the-middle attacks [31].

The latest research by Fovino *et al.* [11] developed a critical state based intrusion detection system that is able to prevent damage to physical plants. For an IDS to follow the state changes of an ICS, protocols used to transmit the data must be parsed, and state changes recorded. Fovino's latest IDS can parse Modbus and DNP3 protocols. Several filtering and monitoring techniques were developed that can describe unwanted states of the ICS. However, there is still no single solution that can guarantee the security of an ICS. While a state-based IDS will monitor the state of the system being secured, some measures to prevent the attackers from modifying the code run on PLCs have to be implemented and some of the control packets have to be blocked. To achieve this, an IDS provides a packet language that can describe signatures of unwanted Modbus and DNP3 packets.

### 2.1 Data Mining/Machine Learning in Network Security

By utilizing principal component analysis in [34], high accuracy state classification of a power grid was achieved. The reduced feature set was compared to new power grid states using Hotelling's $T^2$ value. If the value was too high, then the new feature set could not exist and was determined to be malicious.

Mantere *et al.* [20] states that machine learning IDS can be very useful in a deterministic network such as an ICS network. Unlike typical business networks, an ICS network has a specific periodic packet flow that contains little to no noise during its normal operation. Mantere *et al.* proposes the use of throughput, IP address, average packet size, timing, flow direction, as well as payload data as features used for machine learning.

In their further work in [18], Mantere *et al.* analyzed many features listed in their original research, including network timing features - data that is critical to the research that will be presented in this paper. Mantere *et al.* was not able to detect useful behavior in timing features to detect anomalies. However, the presented research did not target any attacks on the network. The research concluded that ICS networks overall have many anomalies present in the traffic due to misconfiguration of the hardware.

The most recent research by Mantere *et al.* [19] focuses on creating a complementary network security monitoring using Self-Organizing Maps as a means of machine learning. Their approach targets restricted IP networks and does not use any network timing features, but uses packet data instead. The research concludes that deterministic properties of ICS networks make machine learning a viable tool for network anomaly detection.

Gao *et al.* [13] used a Neural Network to classify the ICS network traffic as normal and abnormal based on the operation of the MSU SCADA Security Laboratory water tank control system. The experiment resulted in a 100% accuracy classification of negative false data injection, 95% for positive false data injection, and 84.9% for a random data response injection. Unfortunately, the Neural Network developed achieved only 12.1% accuracy for a replay attack.

Network Telemetry based IDSs, such as the one presented in this paper, are useful in preventing unauthorized access to ICS PLCs. They are even able to differentiate between benign and malicious single-packet attack patterns. In order to demonstrate this ability, a specific type of denial-of-service attack was performed. It is known as a CPU shutdown attack. For this type of attack, a single packet is transmitted to the PLC, which commands its CPU to shutdown. For the system to function properly again, a physical reset of the PLC must be performed. While most systems would ban CPU shutdown commands, the telemetry based IDS did not need to do this. It was able to determine when the single packet shutdown command originated from an attacker and when it originated from the SCADA system.

## 2.2 Network Telemetry

The Internet Protocol (IP) forwards packets on a network between multiple nodes using Ethernet frames until the packet's final destination is reached. For this to happen, the Ethernet frame headers must include the media access control (MAC) address for the destination of the packet as well as the source. MAC addresses are different from normal IP addresses. IP addresses are assigned by self-configuration protocols or network administrators to different machine interfaces; however, MAC addresses are set when the controller circuit for the network interface is manufactured [21].

Because software can be used to spoof both the IP and the MAC addresses, it is often difficult for the server to determine where the packet is actually coming from [30, 35, 36]. To protect data integrity and prevent illegal access through spoofing, encryption is often used for security-critical algorithms [4]. While encryption is an effective method in the deterrence of most attackers, it can be broken. Encryption can, at the very least, be attacked through brute force, no matter the level of security of the cryptographic function [3]. Passwords that are weak can be easily cracked, phishing sites can trick users into giving out their passwords [5], and, like in Stuxnet, reverse engineering can be used on hashes [22]. Network telemetry data is very useful when trying to detect network intrusions. It can help to detect anomalies caused by an attacker using software that is not normally used or even using different machines.

Chandola *et al.* [7] provide a survey of multiple studies focusing on anomaly detection in which machine learning and data mining methods and techniques are used to detect anomalies. One particular data mining application that concentrates on traffic data interactions and correlations is NetMine [2]. Though the methodologies implemented are similar to those that are used in the research effort presented here, their studies are more interested in improving network stability and traffic quality than network security.

By analyzing derived transport layer statistics, Erman *et al.* [9] successfully clustered together packets that were similar. Without extracting packet data, they were able to identify protocols successfully by using DBSCAN and K-Means algorithms [9]. Using the received signal strength (RSS) of packets that are transmitted over wireless networks and analyzing the statistical fluctuations, Shen *et al.* were able to show the ability to detect host spoofing [30].

Wireshark is a software-based network analyzer that captures and displays network traffic in real time. One of Wireshark's many features is that it gives a system administrator the ability to keep all network packets that come in to the network so that they can be analyzed later to find any useful information. For this research effort, Wireshark was used to extract network packet arrival times. This data was placed in comma separated values format to interpret and generate graphs of the data [27].

Although the detection approach that is presented in this effort may not endure the dynamics of a traditional enterprise-wide Local Area Network (LAN), it is beneficial in control system LANs for the detection of spoofed hosts. Control system LANs are unique in their construction in that the hosts will commonly communicate in set intervals that are defined by the particular polling protocol that is used [15]. The detection method that was developed for this research effort has the ability to determine when communication is occurring outside of these set intervals and will trigger a potential intrusion alert when detected. In addition, with the continued use of open source tools that can automate the attack process targeting control systems [32], this effort proves to be successful in differentiating between benign and malicious network packets.

# 3 Telemetry Based Intrusion Detection System

One of the critical components of the developed Intrusion Detection System algorithm is the session capture method. When an attacker is spoofing MAC and IP addresses, there are two methods for them to inject data into the control system network - either by using TCP session spoofing, or instantiating new TCP connections. This section covers the basics of TCP sessions, and discusses experiments performed to evaluate different session instantiation techniques.

The possibility of spoofed addressing, as well as TCP session injections, makes it impossible to differentiate between communication sessions as defined by the TCP flow model. Therefore, to logically group incoming traffic into sessions, a new session definition must be created. Several session aggregation techniques were selected and tested for this research.

## 3.1 TCP Session Flows

An example TCP session flow is shown in Figure 1. TCP is a connection oriented protocol that utilizes internal state variables in order to maintain the connection. TCP session begins with a TCP handshake. The TCP handshake includes a 3-way packet exchange that is commonly explained as SYN, SYN/ACK, ACK. SYN is a synchronization packet sent from the client to the server. This packet starts a TCP session and asks the server to establish a connection. The server then responds with a SYN/ACK packet, acknowledging the connection request. Afterwards, the client acknowledges the server's acknowledgment with an ACK packet. This completes the TCP handshake and allows the data exchange to be started. The TCP session is closed with a FIN, ACK, FIN, ACK sequence. When the client decides to terminate the connection, it sends a FIN packet to the server. The server then sends an ACK packet, acknowledging the request to terminate the connection. After sending the ACK packet, the server has time to release all of the resources used by the connection, and sends a FIN packet of its own to notify the client that all of the resources have been cleared. The client then sends its final packet - ACK - to acknowledge the connection termination. TCP sessions are sets of all of the packets between two nodes that start with the TCP handshake and end with the FIN, ACK, FIN, ACK packet sequence.

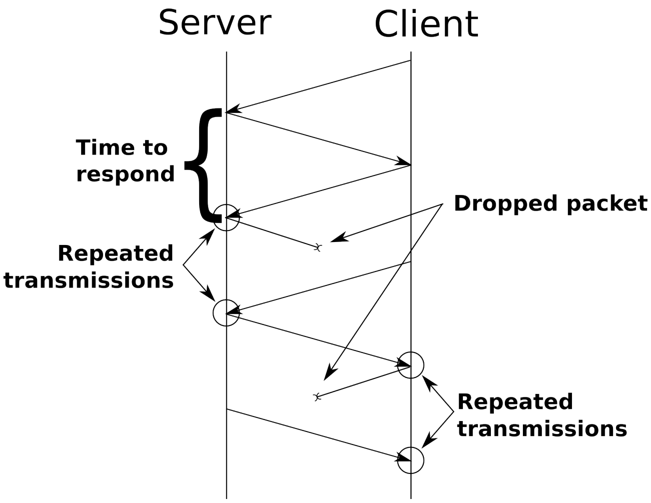


Figure 1: Client-server session graph

Whenever an attacker wants to spoof and inject information into the TCP session, he has to inject packets in the middle of an already established TCP connection; therefore, proper network telemetry can only be obtained if the data is extracted from a session flow model, rather than a by-packet basis because there is no performance information present in the analysis of a single packet. A single packet will only have the time stamp of its arrival to the server. In order to capture relevant performance information of the TCP communication, two methods have been created and tested for this research.

## 3.2 Session Duration Based Instantiation

Session duration based instantiation seeks to identify periods of communication silence when silence is larger than a time threshold. New sessions are defined as a set of packets between the detected silence intervals when the next packet is sent towards the server (PLCs). This method was selected to match the periodical silence of the polling nature of Modbus/TCP SCADA architecture. Silence time threshold was determined experimentally by using 0.1 second interval increments to maximize the C4.5 based classifier accuracy. Table 1 and Figure 2 show session counts for a given time interval, while Table 2 and Figure 3 show accuracies achieved by the C4.5 algorithm, as well as Model Build Times (MBT) in seconds.

A malicious session defines a session that has at least one malicious packet. As the length of the inter-session silence increased, the count of benign sessions decreased, while the count of malicious sessions increased. This trade off happened due to the definition of the malicious session. This is shown in Table 1 where the number of benign sessions drop from 89943 for a silence interval of 0.1 to only

Table 1: Generated session counts for varying silence intervals

| Interval (s) | Malicious Sessions | Benign Sessions | Total Sessions |
|---|---|---|---|
| 0.1 | 51156 | 89943 | 141099 |
| 0.2 | 51021 | 42125 | 93146 |
| 0.3 | 50893 | 29972 | 80865 |
| 0.4 | 50893 | 24458 | 75351 |
| 0.5 | 50891 | 19430 | 70321 |
| 0.6 | 50891 | 14454 | 65345 |
| 0.7 | 50888 | 9648 | 60536 |
| 0.8 | 50886 | 4869 | 55755 |
| 0.9 | 50508 | 403 | 50911 |

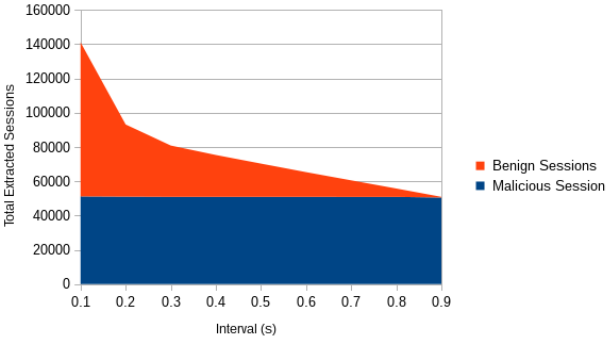


Figure 2: Silence interval detection session counts

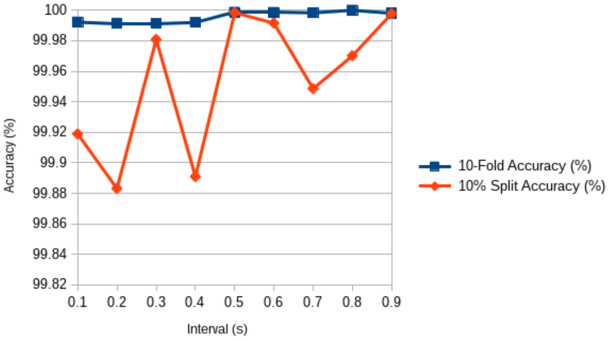


Figure 3: Silence interval detection accuracy

403 as the interval increases to 0.9. A similar drop is also seen in the total number of sessions. Figure 2 pictorially shows this dramatic drop in the number of sessions. A benign session could not have any malicious packets at all, while a malicious session could have any number of benign packets, as long as it had at least one malicious packet.

Table 2 provides accuracies for two different classification experiments: 10-Fold validation, and 10% data split, as well as the time in seconds the C4.5 classifier took to build its model (MBT). The 10-Fold validation experiment takes the dataset and splits it into 10 chunks. The classifier uses the first 9 chunks for training, and the last chunk for accuracy verification. The chunks are then rotated and the experiment is repeated ten times. The accuracy value in Table 2 is the average accuracy of classification for all ten experiments. While this accuracy is not a good estimate of the overall classifier accuracy, it can be used to verify the information distribution of the dataset. If some parts of the dataset contained lower information about the subject, the 10-Fold accuracy would be a lower number. Unlike 10-Fold validation, the 10% Split experiment uses the first 10% of the dataset for training, and the other 90% for validation. The 10% Split accuracy recorded in Table 2 refers to the classifier accuracy when classifying 90% of the dataset. 10% of the dataset translates to approximately 2.5 hours of captured traffic.

Table 2: Classification results of varying silence intervals

| Interval (s) | 10-Fold 1Accuracy (%) | 10% Split Accuracy(%) | MBT (s) |
|---|---|---|---|
| 0.1 | 99.9922 | 99.9189 | 3.33 |
| 0.2 | 99.9914 | 99.8831 | 0.87 |
| 0.3 | 99.9913 | 99.9808 | 0.42 |
| 0.4 | 99.9920 | 99.8909 | 0.33 |
| 0.5 | 99.9986 | 99.9984 | 0.29 |
| 0.6 | 99.9985 | 99.9915 | 0.25 |
| 0.7 | 99.9983 | 99.9486 | 0.17 |
| 0.8 | 100.000 | 99.9701 | 0.49 |
| 0.9 | 99.9980 | 99.9978 | 0.11 |

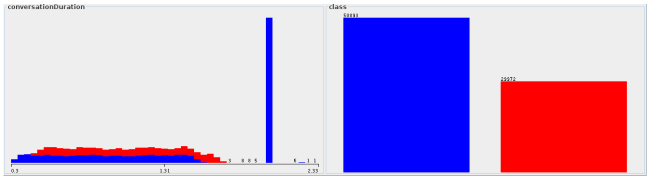


Figure 4: Silence interval session duration and class distributions for 0.3s silence interval

The interval of 0.9 was the last usable interval. Intervals above 0.9 are not listed because only malicious sessions were created while using such a large interval. While Classification accuracy seems to increase as the interval increases (Figure 3), the amount of malicious sessions in the dataset starts to outweigh the amount of benign sessions drastically. For example, if we take the amount of sessions for the 0.9 second interval, just choosing a malicious class over all of the data will result in 99.208% accuracy, as there are 50508 malicious features and only 403 benign features. However, selecting a malicious class for the 0.3 second interval would result in an accuracy of 63.936%, while the C4.5 classifier was able to achieve 99.9808% accuracy for a 10% Split experiment.

Examining the results presented in Table 2 and Figure 3 across all intervals shows an accuracy achievement of between 99.8909% and 99.9984% for the 10% Split experiment and an accuracy between 99.9913% and 100% for the 10-Fold experiment. As shown in Table 2, the MBT are low and are relative to the total session size as noted in Table 1.

Table 3: Generated session counts for varying conversation lengths

| Length (packets) | Malicious Sessions | Benign Sessions | Total Sessions |
|---|---|---|---|
| 1 | 343202 | 255209 | 598411 |
| 2 | 283280 | 157384 | 440664 |
| 3 | 183664 | 99144 | 282808 |
| 4 | 167121 | 53786 | 220907 |
| 5 | 132017 | 49889 | 181906 |
| 6 | 133396 | 21106 | 154502 |
| 7 | 120452 | 16223 | 136675 |
| 8 | 112428 | 3897 | 116325 |

For the use of the silence interval as a session separation method, the best value is determined to be 0.3 seconds, as it provides high classification accuracy (99.9913% and 99.9808% for the 10-Fold and 10% Split experiments respectively) as well as approximately equal amounts of malicious and benign features. Figure 4 shows the distribution of features based on session duration and features class for the value of the 0.3 second interval. In addition, Figure 4 shows a duration based distribution of sessions based on the silence interval of 0.3 seconds.

## 3.3 Session Length Based Instantiation

Session length can be used as an alternative method of extracting sessions from a non-interruptible TCP traffic flow. In these experiments, sessions were separated by counting the amount of packets already in the session. If that number is larger than a given length and the next packet was sent to the server, a new session was instantiated. Table 3, as well as Figure 5, present the data from varying session length separation values.

The same trend found in Table 1 can be noticed in Table 3 for the silence based extraction - the larger the extracted sessions are, the less benign features are extracted due to the benign session being defined as having no malicious packets. However, the amount of malicious sessions extracted is not as stationary as with the silence based extraction. As the session length increases, the amount of extracted malicious sessions decreases from 343,202 malicious sessions for a packet length of 1 down to 112,428 malicious sessions for a packet length of 8, as seen in Figure 5.

At no point of using session length as a session detection parameter was there a comparable amount of malicious and benign sessions extracted as seen in the silence intervals. Moreover, when the amount of extracted sessions was compared to each other, the classification accuracy was significantly lower than that of a silence based extraction.

Similar to the results presented above for silence intervals, Table 4 and Figure 6 present classification and accuracy results for varying conversation lengths respectively. Examining the results presented across all intervals

Table 4: Classification results of varying conversation lengths

| Length (packets) | 10-Fold (%) | 10% Split (%) | MBT (s) |
|---|---|---|---|
| 1 | 97.0796 | 97.1679 | 41.6 |
| 2 | 97.3172 | 97.2468 | 20.9 |
| 3 | 99.9346 | 99.7929 | 13.5 |
| 4 | 99.9570 | 99.8521 | 8.31 |
| 5 | 99.9082 | 99.7306 | 6.47 |
| 6 | 99.9663 | 99.8698 | 4.28 |
| 7 | 99.9188 | 99.5699 | 4.34 |
| 8 | 99.9003 | 99.8147 | 1.62 |

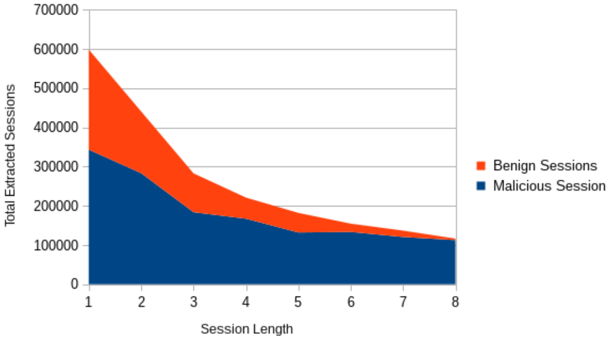


Figure 5: Session length based extraction counts

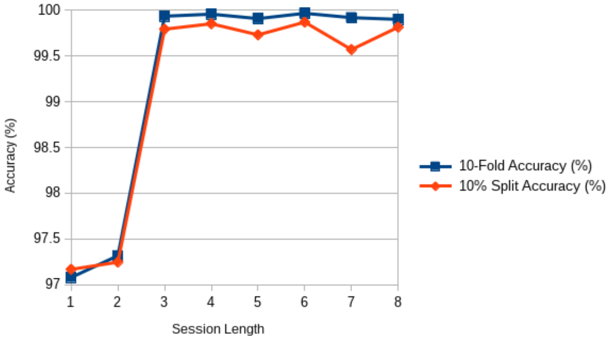


Figure 6: Session length extraction accuracy

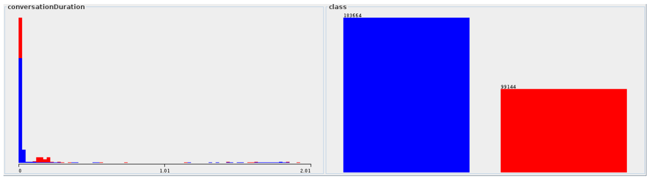


Figure 7: Session length of 3 session duration and class distributions

show an accuracy achievement of between 97.1679% and 99.8698% for the 10% Split experiment and an accuracy between 97.0796% and 99.9663% for the 10-Fold experiment. As shown in Table 4, the MBT are low and are relative to the total session size as noted in Table 3.

Comparable accuracy to that of the silence interval is achieved when the session has at least 3 packets. 10% Split Accuracy of 99.7929 and 10-Fold Accuracy of 99.9346 are achieved, however, it takes 13.5 seconds to build the model. When the distribution of extracted sessions is graphed (Figure 7), the majority of the sessions span very little time in contrast to silence based session extraction.

## 4 Discussion and Future Work

Through the use of different methods there may be improvement in the accuracy of the session classifier. One way to increase the classifier accuracy is by being able to make a distinction between delays that are introduced by a client machine's software and/or hardware and the delays that are introduced by networking hardware, such as routers. Because of idiosyncrasies of a given network's routing algorithm, the number of hops calculated and the delay measurement may not always be directly proportional. As an example, a router may select a connection path with a higher number of hops because of latency due to line congestion. As the separation decreases between the client and server, this problem diminishes, but the problem still needs addressing at the increased separation distances. The detection accuracy for the IDS that is developed and presented in this paper achieved high results. Furthermore, by using a specific set of features for which the attacker has little control over, the IDS provides robustness in network intrusion detections.

There are several threads of future work for this effort. These include looking at how changes in software will affect the detection accuracy, assessing how different hop counts from the target effect the accuracy curve, and expanding the variety of attacks to test against the IDS. Due to the fact that different variations of software and hardware can create unique delays in the propagation of the traffic, a possibility exists to create delay fingerprints that would identify nodes that communicate with the IDS. Work will be done to see if this fingerprinting technique can be accomplished.

## 5 Conclusions

When an attacker is spoofing their machine's addresses and injecting traffic in the middle of an already-established TCP session, it becomes impossible to isolate the attacker's TCP sessions from TCP sessions from a benign machine. In order to collect the information that can be used for an intrusion detection, two session aggregation methods were created and tested. Overall, silence based feature extraction presents better accuracy and efficiency of detecting network intrusions. Both classification accuracy and model build time were better for this method in comparison to session length based extractions. Results presented provide accuracies at 99% when differentiating between the machines of an attacker and an engineer on the same network.

## References

[1] "Recommended practice: Improving industrial control systems cybersecurity with defense-in-depth strategies,". tech. rep., Department of Homeland Security, 2016.

[2] Daniele Apiletti, Elena Baralis, Tania Cerquitelli, and Vincenzo DElia, "Characterizing network traffic by means of the netmine framework," *Computer Networks*, vol. 53, no. 6, pp. 774–789, 2009.

[3] Klaas Apostol, *Brute-force Attack*. SaluPress, 2012.

[4] David Basin, Cas Cremers, Kunihiko Miyazaki, Sasa Radomirovic, and Dai Watanabe, "Improving the security of cryptographic protocol standards," *IEEE Security & Privacy*, vol. 13, no. 3, pp. 24–31, 2015.

[5] Joseph Bonneau, Cormac Herley, Paul C Van Oorschot, and Frank Stajano, "Passwords and the evolution of imperfect authentication," *Communications of the ACM*, vol. 58, no. 7, pp. 78–87, 2015.

[6] Andrea Carcano, Alessio Coletta, Michele Guglielmi, Marcelo Masera, Igor Nai Fovino, and Alberto Trombetta, "A multidimensional critical state analysis for detecting intrusions in scada systems," *Industrial Informatics, IEEE Transactions on*, vol. 7, no. 2, pp. 179–186, 2011.

[7] Varun Chandola, Arindam Banerjee, and Vipin Kumar, "Anomaly detection: A survey," *ACM Computing Surveys (CSUR)*, vol. 41, no. 3, p. 15, 2009.

[8] Steven Cheung, Bruno Dutertre, Martin Fong, Ulf Lindqvist, Keith Skinner, and Alfonso Valdes, "Using model-based intrusion detection for scada networks," in *Proceedings of the SCADA security scientific symposium*, pp. 1–12, 2007.

[9] Jeffrey Erman, Martin Arlitt, and Anirban Mahanti, "Traffic classification using clustering algorithms," in *Proceedings of the 2006 SIGCOMM workshop on Mining network data*, pp. 281–286. ACM, 2006.

[10] Nicolas Falliere, Liam O Murchu, and Eric Chien, "W32. stuxnet dossier," *White paper, Symantec Corp., Security Response*, 2011.

[11] Igor Nai Fovino, Alessio Coletta, Andrea Carcano, and Marcelo Masera, "Critical state-based filtering system for securing scada network protocols," *Industrial Electronics, IEEE Transactions on*, vol. 59, no. 10, pp. 3943–3950, 2012.

[12] Jingcheng Gao, Jing Liu, Bharat Rajan, Rahul Nori, Bo Fu, Yang Xiao, Wei Liang, and CL Philip Chen, "Scada communication and security issues," *Security and Communication Networks*, vol. 7, no. 1, pp. 175–194, 2014.

[13] Wei Gao, Thomas Morris, Bradley Reaves, and Drew Richey, "On scada control system command and response injection and intrusion detection," in *eCrime Researchers Summit (eCrime), 2010*, pp. 1–9. IEEE, 2010.

[14] Béla Genge, István Kiss, and Piroska Haller, "A system dynamics approach for assessing the impact of cyber attacks on critical infrastructures," *International Journal of Critical Infrastructure Protection*, vol. 10, pp. 3–17, 2015.

[15] Niv Goldenberg and Avishai Wool, "Accurate modeling of modbus/tcp for intrusion detection in scada systems," *International Journal of Critical Infrastructure Protection*, vol. 6, no. 2, pp. 63–75, 2013.

[16] L Todd Heberlein and Matt Bishop, "Attack class: Address spoofing," in *Proceedings of the 19th National Information Systems Security Conference*, pp. 371–377, 1996.

[17] Men Long, Chwan-Hwa John Wu, and John Y Hung, "Denial of service attacks on network-based control systems: impact and mitigation," *Industrial Informatics, IEEE Transactions on*, vol. 1, no. 2, pp. 85–96, 2005.

[18] Matti Mantere, Mirko Sailio, and Sami Noponen, "Network traffic features for anomaly detection in specific industrial control system network," *Future Internet*, vol. 5, no. 4, pp. 460–473, 2013.

[19] Matti Mantere, Mirko Sailio, and Sami Noponen, "A module for anomaly detection in ics networks," in *Proceedings of the 3rd international conference on High confidence networked systems*, pp. 49–56. ACM, 2014.

[20] Matti Mantere, Ilkka Uusitalo, Mirko Sailio, and Sami Noponen, "Challenges of machine learning based monitoring for industrial control system networks," in *Advanced Information Networking and Applications Workshops (WAINA), 2012 26th International Conference on*, pp. 968–972. IEEE, 2012.

[21] Jeremy Martin, Erik Rye, and Robert Beverly, "Decomposition of mac address structure for granular device inference," in *Proceedings of the 32nd Annual Conference on Computer Security Applications*, pp. 78–88. ACM, 2016.

[22] Aleksandr Matrosov, Eugene Rodionov, David Harley, and Juraj Malcho, "Stuxnet under the microscope," *ESET LLC (September 2010)*, 2010.

[23] Sangkyo Oh, Hyunji Chung, Sangjin Lee, and Kyungho Lee, "Advanced protocol to prevent man-in-the-middle attack in scada system," *International Journal of Security and its Applications*, vol. 8, no. 2, pp. 1–8, 2014.

[24] Stanislav Ponomarev and Travis Atkison, "Industrial control system network intrusion detection by telemetry analysis," *IEEE Transactions on Dependable and Secure Computing*, vol. 13, no. 2, pp. 252–260, 2016.

[25] Stanislav Ponomarev and Travis Atkison, "Session duration based feature extraction for network intrusion detection in control system networks," in *Computational Science and Computational Intelligence (CSCI), 2016 International Conference on*, pp. 892–896. IEEE, 2016.

[26] Stanislav Ponomarev, Nathan Wallace, and Travis Atkison, "Detection of ssh host spoofing in control systems through network telemetry analysis," in *Proceedings of the CIRSC conference*, pp. 1–12, 2014.

[27] Chris Sanders, *Practical packet analysis: Using Wireshark to solve real-world network problems*. No Starch Press, 2017.

[28] Naoum Sayegh, Imad H Elhajj, Ayman Kayssi, and Ali Chehab, "Scada intrusion detection system based on temporal behavior of frequent patterns," in *Mediterranean Electrotechnical Conference (MELECON), 2014 17th IEEE*, pp. 432–438. IEEE, 2014.

[29] Bruce Schneier, "The story behind the stuxnet virus," *Forbes. com*, 2010.

[30] Yong Sheng, Keren Tan, Guanling Chen, David Kotz, and Andrew Campbell, "Detecting 802.11 mac layer spoofing using received signal strength," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pp. 1768–1776. IEEE, 2008.

[31] Keith Stouffer, Victoria Pillitteri, Suzanne Lightman, Marshall Abrams, and Adam Hahn, "Guide to industrial control systems (ics) security," *NIST Special Publication*, pp. 800–82, 2015.

[32] Nathan Wallace and Travis Atkison, "Observing industrial control system attacks launched via metasploit framework," in *Proceedings of the 51st ACM Southeast Conference*, ACMSE '13, pp. 22:1–22:4, New York, NY, USA, 2013. ACM.

[33] Nathan Wallace, Stanislav Ponomarev, and Travis Atkison, "A dimensional transformation scheme for power grid cyber event detection," in *Proceedings of the CIRSC conference*, pp. 1–12, 2014.

[34] Nathan Wallace, Sean Semple, and Travis Atkison, "Identification of state parameters for stealthy cyber-events in the power grid using pca," in *PES General Meeting— Conference & Exposition, 2014 IEEE*, pp. 1–5. IEEE, 2014.

[35] Guang Yao, Jun Bi, and Athanasios V Vasilakos, "Passive ip traceback: Disclosing the locations of ip spoofers from path backscatter," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 3, pp. 471–484, 2015.

[36] Jaegwan Yu, Eunsoo Kim, Hyoungshick Kim, and Junho Huh, "A framework for detecting mac and ip spoofing attacks with network characteristics," in *Software Security and Assurance (ICSSA), 2016 International Conference on*, pp. 49–53. IEEE, 2016.

# Biography

**Travis Atkison** is an Assistant Professor of Computer Science at the University of Alabama. He is the director of the Digital Forensics and Control System Security Lab (DCSL). His current research efforts focus on the topic of cyber security. These efforts include malicious software detection, threat avoidance, digital forensics, and security in control system environments including transportation.

**Stanislav Ponomarev** is currently a Cyber Security Scientist at BBN Technologies. Dr. Ponomarev received his Master and PhD degrees from Louisiana Tech University in 2015.

**Randy Smith** is an Associate Professor in the Department of Computer Science at The University of Alabama. Dr. Smith is an active researcher in transportation safety with over 15 years experience in crash data analysis, linear referencing systems and safety data integration. His research has been supported by various federal agencies including NASA, NSF, and the USDOT. In addition, Dr. Smith has worked with multiple states on transportation safety matters.

**Bernard Chen** received the Master and PhD degrees in Computer Science from Georgia State University in 2008. Currently, he is an Associate Professor in the Department of Computer Science at The University of Central Arkansas. His research interests include Data Science, Data Mining, Bioinformatics, and Wineinformatics.

# A Secure Routing Protocol Based on Reputation Mechanism

Yanhui Lv, Kexin Liu, Deyu Zhang, and Zhuo Miao
*(Corresponding author: Yanhui Lv)*

College of Information Science and Engineering, Shenyang Ligong University
6 Nanping Middle Rd, Hunnan Qu, Shenyang 110168, China
(Email: yanhuilv@126.com)

*(Received May 4, 2017; revised and accepted Aug. 28, 2017)*

## Abstract

The wireless sensor network is often deployed in harsh and unattended environment and is easily attacked and interfered due to its vulnerability. Most of routing protocols for wireless sensor network were initially designed for saving energy and had less consideration on security. In view of this, a secure routing protocol based on reputation mechanism is proposed. Firstly, considering at the problems of trust mechanism including complex computation of trust value and excessive energy consumption, a binomial distribution reputation mechanism (BDRM) is presented. On this basis, aiming at the deficiency of current routing protocol in security defense, a secure routing protocol ST-GEAR is designed. This protocol has a better ability to defend against attacks by introducing the secure bootstrap model and the BDRM in the routing. The simulation results show that the ST-GEAR routing can improve the transmission rate of packet and reduce the loss rate of packet and energy consumption.

*Keywords: Reputation Mechanism; Secure Bootstrap Model; Secure Routing; Wireless Sensor Network*

## 1 Introduction

The wireless sensor network (WSN) is a distributed network formed by a lot of randomly distributed sensor nodes with perception, computing and processing ability of data by self- organization [4]. The nodes can send the received information after being gathered and integrated to specific nodes in order to play a role of the real- time perception and monitoring to the target region.

WSN was mainly applied in military field at first, while its application has expanded to traffic management, environmental monitoring, smart home, medical care, manufacturing industry and other fields with the continuous development of network technology. WSN can solve some issues that cannot be achieved by traditional network, and this is also an important reason why WSN is highly valued and studied deeply.

Comparing with traditional network, WSN has a larger amount of nodes, and the energy of each node is limited. By the self- organization, the nodes form a network whose topological structure has dynamic changes. These features of WSN itself result in limited defense ability of sensor nodes and vulnerable network which is easily attacked. The nodes in the network adopt the wireless communication transmission, which will easily influence the normal network communication once the transmission is interfered. Besides, WSN is often deployed in harsh and unattended environment, which make it face serious security issues. Only when the security of WSN is guaranteed, can it be applied in corresponding fields to exert its advantages. In order to effectively solve the security issues of WSN, researches are mainly carried out from the perspective of key management, intrusion detection, secure routing and secure data fusion.

In WSN, each node has a possibility to become a routing node. The network layer routing protocol is responsible for processing the received data by nodes and transmitting the data to the target nodes along the set routes. If there are attackers in the node communication process in network, attackers can use some methods to break the normal packet transmission between nodes or intercept some useful information in network, which will have an impact to the network and even lead to network paralysis. To guarantee the data transmission security in WSN, it is necessary to study the secure routing protocol for wireless sensor network [19].

Currently, there are many routing protocols of wireless sensor network being proposed, and most of them are initially designed for saving node energy and extending network lifetime, such as LEACH [5], GEAR [17], and other routing protocols [14]. However, these routing protocols are rarely considerate the security of routing protocol and it is easy to make them not work anymore if they are attacked by attackers in certain ways.

Aiming at defense methods to attacks by analyzing the attack type that routing protocol easily suffers, a secure routing protocol based on reputation mechanism is proposed in this paper. Firstly, a binomial distribution reputation mechanism (BDRM) based on Beta distribution is proposed. The BDRM calculates the node trust value by the node energy and communication process, and simplifies the computation complexity. Then, a secure bootstrap model is given, which can ensure that the information transmitted between nodes is encrypted so that the security of data transmitted in the network can be improved. On this basis, a security trust - geographical and energy aware routing protocol (ST-GEAR) is presented combining the BDRM, secure bootstrap model and GEAR to improve the security of GEAR routing protocol. When ST-GEAR protocol selects the next hop forwarding node, it will consider the following aspects including node energy, distance and node trust value, and select the node with minimum comprehensive cost.

The rest of the paper is organized as follows. Section 2 introduces the research status. Section 3 describes the proposed binomial distribution reputation mechanism (BDRM) in detail. On this basis, Section 4 presents a security trust geographical and energy aware routing (ST-GEAR). Section 5 reports the simulation results. Concluding remarks and future directions are given in Section 6.

## 2 Research Status

In 2003, Karlof and Wagner initiated to study the secure routing issues of WSN, analyzed the situation that existing routing protocols are easily attacked in details and also studied the security measures for defending attacks [6].

Reference [1] explains the ways that the WSN routing protocols are easily attacked in details and also concludes the defense mechanism of each attack. Reference [18] summarizes the data security standard. The data security index mainly includes data acknowledgement, data authorization, data integrity and data update. The reference also proposes SNEP (Secure Network Encryption Protocol) and TESLA (Micro Timed Efficient Streaming Loss-tolerant Authentication Protocol). SNEP can realize data confidentiality, usability, integrity and timeliness; TESLA applies delay to send keys one by one to achieve the digital signature and further realizes the data acknowledgement. Reference [2] introduces 4 security mechanisms based on symmetric key technology in details. The introduction of security mechanism into routing protocol can effectively protect routing protocol from attacks to improve network security.

Besides, there are also some representative secure routing protocols as follows.

The trust routing TRANS [15] based on location establishes the trust routing according to geographical location of nodes and isolates the malicious nodes in the network. The protocol judges whether a node is safe or not by the node trust value. The nodes with large trust value are safer and therefore such nodes should be chosen when routing and the nodes with small trust value should be isolated. The TRANS protocol uses the loose time synchronization mechanism in the process of authentication requests between nodes and most of operations are completed by base station.

The secure routing LKHW [12] based on logical key hierarchy is proposed for solving security issues of DD routing protocol. The logical key hierarchy LKH is applied in order to improve the security of DD protocol in multicast. The key that each node obtains in the network is allocated by base station one by one. Logically, the keys obtained by all nodes can be considered as a tree with the base state as the root. Only two nodes that are the child nodes under the same root node have the same key to communicate with each other.

The SEEM [10] called secure and energy-efficient multipath routing protocol selects the routing by the base station that will choose the path with most optimal energy as the next hop forwarding path. In this protocol, each node only needs to know its own routing information to the base station, which can reduce the energy consumption and avoid being attacked.

A two-layer authentication protocol with anonymous routing TAPAR for wireless ad hoc networks is proposed in [7]. A novel solution is introduced without resorting to PKI operations to achieve anonymity between two communication entities over insecure networks.

Aiming at the security issues that LEACH routing protocol is easily attacked by Hello Flood and so on, reference [16] introduces security mechanism and frame of SPINS into the LEACH, proposes the secure low- consumption self- adaptive SLEACH routing protocol and applies the encryption method to guarantee the security of routing protocol. Moreover, other typical algorithms can reference [8, 11, 13].

However, the existing routing protocols at present still have some disadvantages in security defense. For this reason, aiming at the ways that network layer routing protocol are easily attacked, this paper designs a secure routine protocol to improve the network security.

## 3 BDRM

Considering the problem that routing protocols easily suffer insider attack, a binomial distribution reputation mechanism (BDRM) is presented. The node trust value is calculated by the energy and the communication process of the node. The comprehensive node trust value can be used as a basis for identifying a malicious node and will be applied to the routing in the next section.

### 3.1 Energy Trust Value

In the trust mechanism, if the trust value of a node is large, the number that the nodes participate the routing will increase and then will increase the energy consumption of the nodes. This can cause these nodes to fail prematurely. In order to avoid this problem, the node residual energy as well as other factors is considered when computing the node trust value in BDRM. For the sending node $i$ and receiving node $j$, the calculation formula of node energy trust value $ER_{ij}$ of node $i$ against node $j$ is shown in Equations (1) and (2).

$$ER_{ij} = \begin{cases} \frac{E_j}{E} & ET_{ij} \geq \theta \\ 0 & ET_{ij} < \theta \end{cases} \tag{1}$$

$$ET_{\mathrm{ij}} = \frac{E_j}{(E_{tx} + E_{rx})} \tag{2}$$

Where, $E$ represents the initial energy of node; $E_j$ represents the current residual energy of node $j$; $E_{tx}$ represents the energy consumption of node $i$ by transmitting the message; $E_{rx}$ represents the energy consumption of node $j$ by receiving the message; $ET_{ij}$ represents the energy trust evaluation parameter of node $i$ against node $j$; represents the setting threshold of node energy.

When $ET_{ij}$ is greater than or equal to the threshold, it represents that node $j$ can be trusted. At this time, the energy trust value of node $j$ can be calculated. Otherwise, it represents the energy of node $j$ cannot be trusted.

## 3.2 Calculation of Direct Trust Value

A node has two basic communication processes for forwarding the packet. One is that the node forwards the packet normally, and the other is that the node forwards the packet abnormally. This packet forwarding interaction can be simulated by the binomial distribution, and the direct trust value of node can be calculated.

Suppose the times of packet forwarding between two nodes is $m$, where, the times of normal packet forwarding is denoted by $a$, and the times of abnormal packet forwarding is denoted by $b$. At the same time, the probability of normal packet forwarding is assumed as $p$. Then, the probability distribution of $p$ can be obtained by the binomial distribution as shown in Equation (3).

$$f(p) = \frac{(a+b)!}{a!b!} p^a (1-p)^b \tag{3}$$

Because $f(p)$ is the probability distribution function, the maximum value of the function can be used to present the value of $p$. Set

$$f'(p) = [\frac{(a+b)!}{a!b!} p^a (1-p)^b]' = 0 \tag{4}$$

By solving Equations (4), the value of $p$ can be obtained as follows.

$$p = \frac{a}{a+b} \tag{5}$$

Assume that the trust value of node $i$ against node $j$ is represented by $TD_{ij}$, then the value of $p$ is the value of $TD_{ij}$ that obeys the binomial distribution B $(a+b,\, a)$, i.e. $TD_{ij} \sim$ B $(a+b,\, a)$. Thus, Equation (6) is gotten.

$$TD_{ij} = \frac{a}{a+b} \tag{6}$$

Within the communication range, a node judges whether its neighbor node normally receives the packet by a way of single- hop acknowledgement. The specific description is shown as follows.

The node $i$ sends packet to node $j$ and require node $j$ to return to it an Ack after receiving the packet. Node $i$ compares the copy it saves with the received Ack. If they are the same, it proves that node $j$ normally receives the packet from node $i$; otherwise, it means the node $j$ does not receive the packet.

In addition, a node uses the two- hop acknowledgement to judge whether its neighbor node normally forwards the received packet. The specific description is shown as follows.

After the node $i$ has confirmed the node $j$ normally receives the packet, node $j$ should further forward the received packet. Suppose the forwarding node is node $k$, after node $k$ receives the packet, it needs to send two Acks. One is for node $j$ to let the node $j$ know that node $k$ normally receives the packet. The other is for node $i$, and node $i$ will also compare the Ack with the copy after receiving the Ack form node $k$. If they are the same, it proves that node $j$ is normally forwarding the packet of node $i$; otherwise, it indicates that node $j$ does not normally forward the packet of node $i$. There may be something wrong with node $j$; perhaps, the node $j$ itself is a problematical node.

## 3.3 Calculation of Indirect Trust Value

For the node $i$ and node $j$, they may have some common neighbor nodes. If node $i$ is able to ensure the credibility of these neighbors, it evaluates the node $j$ by these neighbors, which is the indirect trust value of node $i$ against node $j$. Let node $N_1$ is one of the common neighbor nodes of node $i$ and node $j$, next, we use an example to illustrate the computing process of the indirect trust value.

Firstly, node $i$ sends a request packet to node $N_1$ for calculating the indirect trust value of node $j$. After $N_1$ receives the request packet, it will send a response packet to node $i$, which includes the latest trust parameter value $(a_{N_{1j}}, b_{N_{1j}})$ on the evaluated node $j$.

According to the trust parameter value of node $j$ sent by N1, node $i$ calculates the indirect trust value $TD_{N_{1j}}$ against node $j$, and the calculation formula is shown in Equation (7).

$$TD_{N_{1j}} = \frac{a_{N_{1j}}}{a_{N_{1j}} + b_{N_{1j}}} \tag{7}$$

If node $i$ receives $n$ valid indirect trust values from neighbors, and they are $TD_{1j}$, $TD_{2j}$, $TD_{3j}$, $\cdots$, $TD_{nj}$, respectively. Take the average value of these values as the indirect trust value $TI_{ij}$ of node $i$ against the evaluated node $j$, and the calculation formula is shown in Equation (8).

$$TI_{ij} = \overline{TD_{Nkj}} = \frac{\sum\limits_{k=1}^{n} TD_{Nkj}}{n} \tag{8}$$

### 3.4 Calculation of Comprehensive Trust Value of Nodes

From Equations (1), (6) and (8), we can obtain the node energy trust value $ER_{ij}$, direct trust value $TD_{ij}$ and indirect trust value $TI_{ij}$ respectively, and the direct and indirect trust value are actually the trust value of node in the communication process. So, the trust value in the communication process is the integration of the value $TD_{ij}$ and $TI_{ij}$ and is represented by $TC_{ij}$. The calculation formula is shown in Equation (9).

$$TC_{ij} = \alpha TD_{ij} + \beta TI_{ij} \tag{9}$$

Where, $\alpha$ and $\beta$ represent the weights of $TD_{ij}$ and $TI_{ij}$, respectively, and they satisfy the equation $\alpha$+$\beta$=1, $\alpha$>0, $\beta$>0.

The choice of weights of and will influence the trust value of node in the communication process. When $\alpha$>0.5, it signifies the node has more trust in $TD_{ij}$ and less consideration to $TI_{ij}$. While, the result will be opposite if $\beta$ selects a larger value.

Oh this basis, the comprehensive trust value $T_{ij}$ of node is calculated based on the energy and the communication process of node. It is the integration of the trust values of $TC_{ij}$ and $ER_{ij}$ and the calculation formula is shown in Equation (10).

$$T_{ij} = \omega TC_{ij} + \gamma ER_{ij} \tag{10}$$

Where, $\omega$ and $\gamma$ denote the weights of $TC_{ij}$ and $ER_{ij}$, and they satisfy the equation $\omega$+$\gamma$=1, $\omega$ >0, $\gamma$ >0.

### 3.5 Update of Node Trust Value

Each transmission of the trust value between nodes will consume some energy of the nodes. Considering the energy consumption of node itself, the BDRM uses periodic updates to update the node trust value. Each node is to update its trust value according to its own trust update period (TUP), and the calculation method of TUP is shown in Equation (11).

$$TUP = 3 \times (bint + drate \times bint) \tag{11}$$

Where, *bint* denotes the time interval of communication between nodes, and *drate* denotes the number of the packets sent per second.

Within the TUP of node, the communication process and energy of node will have some changes and therefore, the comprehensive trust value of node will also change. The following example will have a specific explanation.

For the node $i$ and node $j$, suppose that the two nodes have $m$ ($m = a + b$) communication process, after that, they interacts $w$ times again. Among the $w$ times, the times of normal and abnormal packet forwarding are $r$ and $s$ respectively, i.e. w=r+s. The latest direct trust value $TD_{ij}^{new}$ of node i against node j within the TUP also obeys the binomial distribution B(m+w, a+r), i.e. $TD_{ij}^{new} \sim$ B $(m+w, a+r)$ that is shown in Equation (12).

$$TD_{if}^{new} = \frac{a+r}{m+w} = \frac{a+r}{a+b+s+r} \tag{12}$$

Correspondingly, the energy trust value and indirect trust value of node $i$ against node $j$ will change, and the updated value $T_{new}$ will be obtained.

## 4 Design of ST-GEAR Secure Routing Protocol

Among the geographical location based routing protocols, comparing with other routing protocols, the GEAR protocol has more advantages, and that is why this paper takes GEAR protocol as a research object. Aiming at the deficiencies in the defense of GEAR protocol, a security trust - geographical and energy aware routing (ST-GEAR) protocol is proposed to improve the security of GEAR. The BDRM and the secure bootstrap model are introduced in the ST-GEAR, which makes it have a better ability to resist attacks.

### 4.1 Design of Secure Bootstrap Model

The secure bootstrap model is the protection mechanism of wireless sensor network. The nodes in the network form a secure communication network based on identity authentication, encryption key and authentication key, and among them the core is to establish the key. After confirming the key between nodes, the packet transmitted between nodes is the encrypted packet. Only after decrypting the packet with decryption key, can the nodes obtain the original data, which improves the security of information transmission between nodes.

In WSN, there are many key management schemes. The simplest one is the pre- shared key scheme which means that only one symmetric key is shared among all nodes in the network, but this way is too dependent on the sink node. In fact, the random key pre- distribution scheme is more widely used. There is a large key pool in the random key pre- distribution scheme. Each node has a portion of keys in the key pool, and only the nodes with the same pair of key can establish a connection to form a secure transmission path. In the random key predistribution scheme, a node just needs to store a portion of keys of the key pool, which can reduce the key storage space. For each sensor node In WSN, the storage capacity and computing power is limited, so it is feasible to choose the random key pre- distribution scheme. Considering that the performance of the random key pre- distribution scheme can be improved to some extent by combining it with the geographic location of node, a random key predistribution scheme based on grid deployment model is proposed in this paper.

Suppose there are $n$ sensor nodes being randomly deployed in WSN, and each node can store $m$ keys. The location information of each node can be known in advance, so we will consider it as the ID of node in order to distinguish each node, showing as $ID_i(i = 1, 2, 3, \cdots, n)$.

1) Initialization Phase of Key
In the network initialization, each node should be allocated with key. Each node ID$i$ randomly selects $m$ nodes from other nodes around it and generates $m$ node pairs. Each node pair (ID$i$, ID$j$) is allocated with corresponding key. Thus, an n×m common matrix $P$ can be created according to the node number $n$ and node number m being randomly selected.

   Meanwhile, the sink nodes in the network can build a random n×n symmetric matrix $S$ which is only known by the sink nodes themselves and confidential to others. Thus, the matrix $A = (SP)^T$ can be calculated. Next, store the element of $i^{th}$ row of A and the $i^{th}$ column of $P$ in node $i$; store the element of $j^{th}$ row of A and the $j^{th}$ column of $P$ in node $j$. When the shared key for communication between node $i$ and node $j$ is needed to be established, they should exchange their own information, and take the product of the $i^{th}$ row of $A$ and the $j^{th}$ column of $P$ as the key $k^{ij}$ of node $i$ and consider the product of the $j^{th}$ row of $A$ and the $i^{th}$ column of $P$ as the key $k^{ji}$ of node $j$. because $S$ is symmetric, it is easy to get the Equation (13).

$$\begin{aligned} K &= (SP)^T P = P^T S^T P \\ &= P^T SP = (AP)^T = K^T \end{aligned} \quad (13)$$

   Therefore, the node pair (ID$i$, ID$j$) takes $k_{ij}$=$k_{ji}$ as the shared key.

2) Grid Deployment Model
In WSN, the communication range of sensor node is round, so there must be certain overlapping parts within the communication range between nodes. To reduce the overlapping parts, the grid deployment model is used. The grid deployment model is to deploy several virtual regular polygons in WSN, making them cover the whole network as much as possible. By comparing regular polygons including square, equilateral triangle and regular hexagon, we can know that the overlapping area of regular hexagon is smaller than that of both. Therefore, we choose the regular hexagon to be as the deployment model [9].

3) Key Establishment
The next phase is to establish the key. Firstly, each node should broadcast its own location information to the nodes around itself. According to the location information, it is easy to know whether there is a shared key between the two nodes.

   In addition, each regular hexagon unit uniquely shares a common matrix with its adjacent unit. Thus in the key establishment phase, each node in a unit and the node in adjacent unit can exchange their ID in the network to confirm the key pair. The key pair is unique to each node pair.

   After finishing the key establishment, each node pair with shared key has a connection line. According to the connection lines, the whole network forms a connection graph. Further, based on each connection line in the connection graph, each node in the network can find the node with the shared key one by one, and the connection path composed by these nodes is safe.

## 4.2 Design of ST-GEAR Secure Routing

When a node selects a neighbor node as the next hop forwarding node, it needs to consider the following factors including the location, residual energy and trust value of the node, and chooses the node with minimum comprehensive cost value from the neighbor node to the target node. For a forwarding node $j$, the calculation formula of comprehensive cost $Cc$ ($j$, $D$) from the node $j$ to target node $D$ is shown in Equation (14).

$$Cc = \alpha \times (\beta d(j, D) + (1 - \beta)E_j) + (1 - \alpha)(1 - T_{ij}). \quad (14)$$

Where, $d$ $(j,$ $D)$ denotes the distance from the node $j$ to node $D$; $\alpha$ and $\beta$ are the adjustable ratio parameters between 0 and 1.

For a node $i$, when it needs to select a neighbor node as the next hop forwarding node, the specific description of routing algorithm is shown as follows.

1) The node $i$ needs to confirm whether it has neighbor nodes or not.

2) If the node $i$ has neighbor nodes, it needs to choose a node such as node $j$. Next, to judge whether the energy of node $j$ satisfies the energy threshold set in the network. If it does, go to step (3); if not, the node $i$ needs to broadcast the message that the node $j$ is with less energy.

3) In the case that the energy of node $j$ satisfies the requirement, the node $i$ needs to further determine whether the node $j$ is reliable by its trust value. If the trust value of node $j$ is smaller than the trust threshold set in the network, the node $i$ needs to broadcast the message that the node $j$ is with small trust value.

4) Only when the energy and trust value of node $j$ are greater than corresponding threshold, can the comprehensive cost of node $j$ be calculated. For node $i$, there may be multiple neighbor nodes satisfy above requirements. The node $i$ will select the node with minimum comprehensive cost as the next hop forwarding node.

In addition, the data transmitted between nodes is the encrypted packet. Only after decrypting the packet with decryption key, can the nodes obtain the original data. In

ST-GEAR protocol, the sink node will query data in target region to obtain the corresponding information. According to the location of target region, the sink node will send the query command to the target region. This process will be divided into two parts: the first is that the sink node sends query commands to target region, and the second is that the query command is transmitted in target region.

When the sink node sends a query command to target region, it selects the next hop node to forward the message. The message sent by sink node is encrypted, and only the selected next hop node can decrypt the encrypted message, can the node send the message continuously. For node $j$, when it receives the message sent by node $j$, it needs to decrypt the message with the shared key $k_{ij}$ to get the location information of the target node. Repeat this process until the target node is found. After a query command sent by the sink node arrives in the target region, it can be transmitted using different strategies according to the distribution of nodes in target region. Firstly, set a threshold for the number of nodes according to the node distribution in target region. If the number of nodes is greater than the threshold, the recursion method to transmit the message can be used. Otherwise, the flooding method can be used to broadcast the message directly.

## 5 Simulation Experiment and Analysis

In this section, the simulation environment and performance index will be given firstly. Then, the simulation results of ST-GEAR protocol are presented.

### 5.1 Simulation Environment and Performance Index

The simulation scenario is set to a square monitoring region that the side length is 100 m. There are 100 wireless sensor nodes are deployed randomly. The specific simulation parameters are shown in Table 1.

### 5.2 Analysis on Simulation Results

There are many ways to attack the routing protocol. In this simulation experiment, the selective forwarding attack is adopted that means a node will not forward the packet after receiving. The specific simulation results and analysis are shown as below.

- Simulation analysis of the BDRM

Figure 1 shows the trust value simulation results of normal and abnormal nodes. The initial trust value of node is set to 0.5. After a number of packet transmission, the trust value of a normal node tends to 0.9, while the trust value of an abnormal node tends to 0.1. If the node trust value is small, this node can be considered as a malicious node that will not be chosen in the routing.

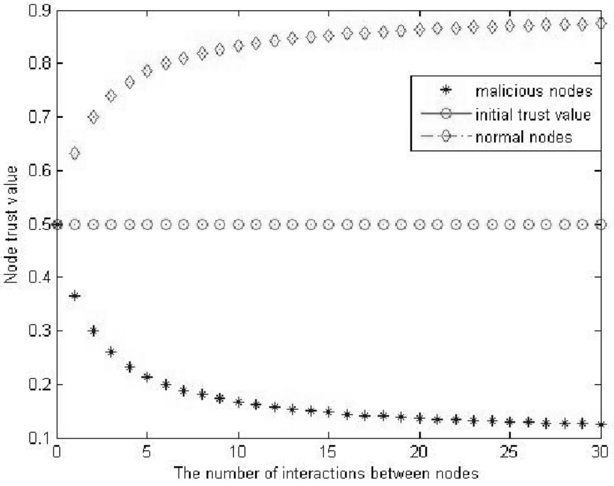


Figure 1: Node trust value

Next, make a simulation comparison between BDRM proposed in this paper with classic BRSN [3], and the results are shown in Figure 2 and Figure 3. The node trust value is an important basis for routing algorithm and the node can be judged whether it can be trusted by the node trust value. In BDRM, the node trust value is calculated based on the node energy and communication process; while in BRSN, it does not take into account the energy factor. From Figure 2, we can see that there are some difference between node trust value calculated by BDRM and BRSN. The trust value of BRSN obeys Beta distribution, while the BDRM obeys binomial distribution. The uptrend of normal node trust value calculated by BDRM is more stable than that of node trust value calculated by BRSN.

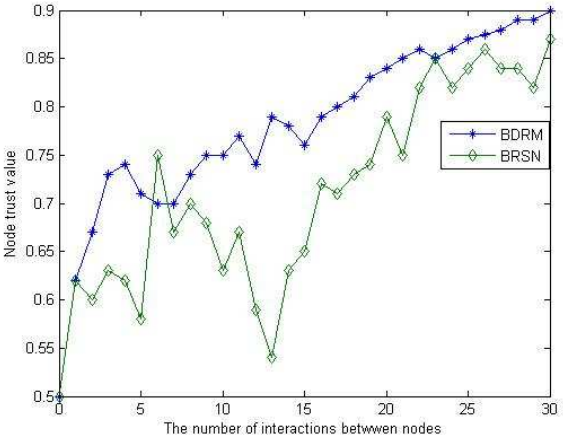


Figure 2: Trust value of normal node

On the contrary, when there are malicious nodes in the network, the BDRM and BRSN can recognize the

Table 1: Specific parameters of simulation

| **Simulation Parameters** | **Values** |
|---|---|
| *number of nodes* | 100 |
| *communication radius of node* | 50m |
| *initial energy of node* | 2J |
| *energy consumed by sending or receiving a packet* | 0.001J |
| *packet length* | 36bit |

malicious node by the node trust value. As can be seen from Figure 3, the trust value of malicious node calculated by BDRM is on a downward trend from the initial trust value of 0.5. Comparing with BRSN, the trust value of malicious node in BDRM changes significantly, this can improve the probability of identifying the malicious nodes in the network.

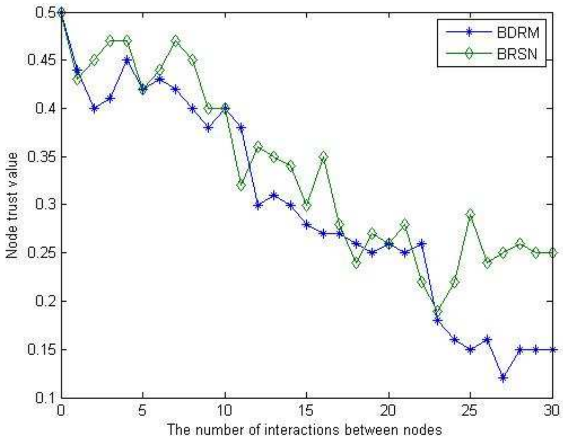


Figure 3: Trust value of malicious node

- Simulation analysis of ST-GEAR protocol.

The simulation results of routing, transmission rate of packet, loss rate of packet and energy consumption rate are given as follows.

1) Routing Simulation from Source Node to Target Node
   Set a source node and a destination node randomly. Define the source node as the current node and calculate the distance from the current node to destination node.

   Normally, when the network is not attacked, the selected path from current node to destination node is shown in Figure 4. The node will choose the path with the minimum cost according to the distance to target node and the residual energy of node.

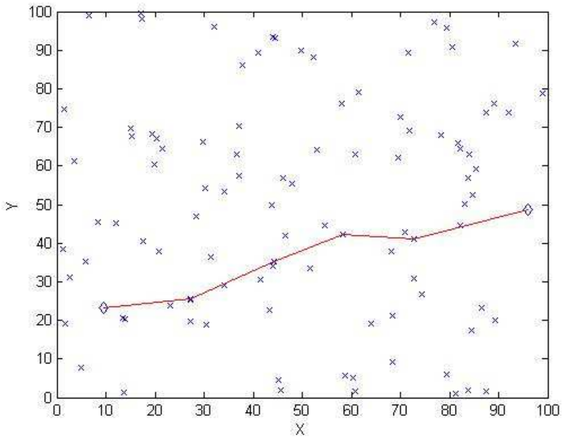


Figure 4: Normal routing

When the network is attacked which means there are some malicious nodes in the network, the nodes with small trust value can be found based on BDRM. In the routing process, the node will choose the path with the minimum cost according to the node trust value as well as the distance and the residual energy of nodes. The node with small trust value can not participate in the routing process. The simulation result is shown in Figure 5.

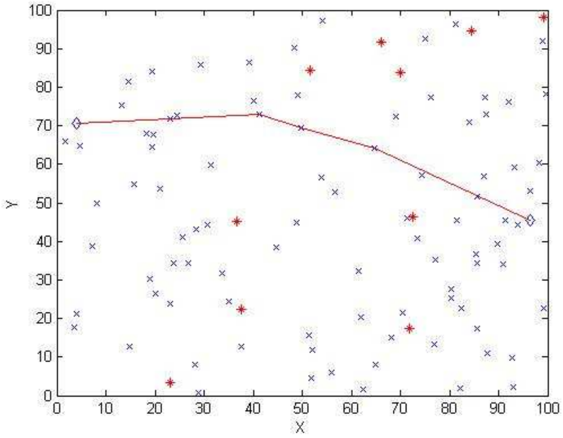


Figure 5: Routing with malicious nodes

2) Transmission Rate of Packet

When there are attacking nodes in the network, the transmission rate of packet in ST-GEAR, E-GEAR and GEAR are shown in Figure 6.

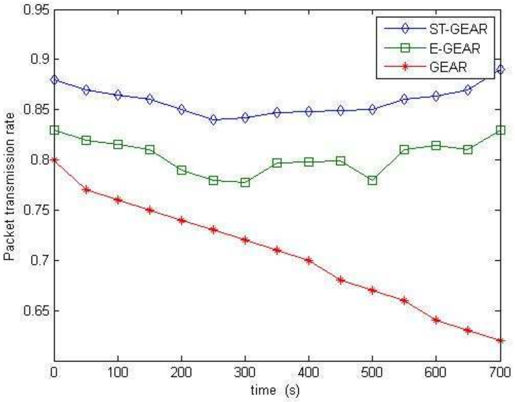


Figure 6: Transmission rate of packet

As shown in Figure 6, the transmission rate of packet in GEAR decreases gradually with time. While, the transmission rates in ST-GEAR and E-GEAR are relatively stable. This is because the node trust mechanism is introduced into the ST-GEAR and E-GEAR, and the node trust value is used as an important factor when routing. The trust value of attacking node will be reduced over time, and when a node selects the next hop, the node with small trust value will not be chosen. In addition, the transmission rate of packet in ST-GEAR is greater than that of E-GEAR, because the BDRM is used in ST-GEAR.

3) Loss Rate of Packet
The simulation result of the loss rate of packet is shown in Figure 7 when there are attacking nodes in the network. It can be seen that the loss rate of packet in ST-GEAR is small compared with other two protocols. In addition, the loss rate of packet in ST-GEAR is smaller than that of E-GEAR. This is because the node trust mechanism used in BDRM takes into account the node energy as well as the communication process of node.

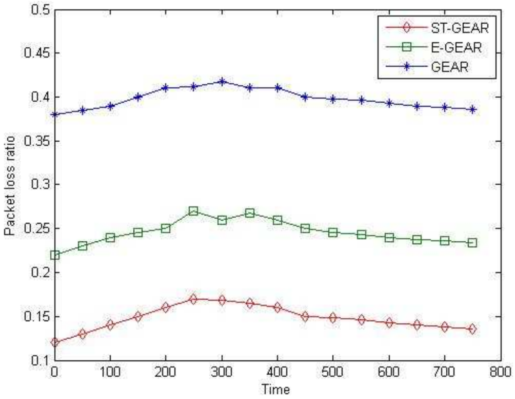


Figure 7: Loss rate of packet

4) Energy Consumption Rate
Figure 8 presents the simulation results of energy consumption rate in ST-GEAR, E-GEAR and GEAR. The energy consumption rate of GEAR is the minimal compared with two protocols. This is because the introduction of the reputation mechanism into the two other protocols consumes some energy. In addition, the energy consumption of ST-GEAR is less than that of E-GEAR. This is because the computational complexity of calculating the node trust value in BDRM is reduced, and the energy consumption is correspondingly reduced.

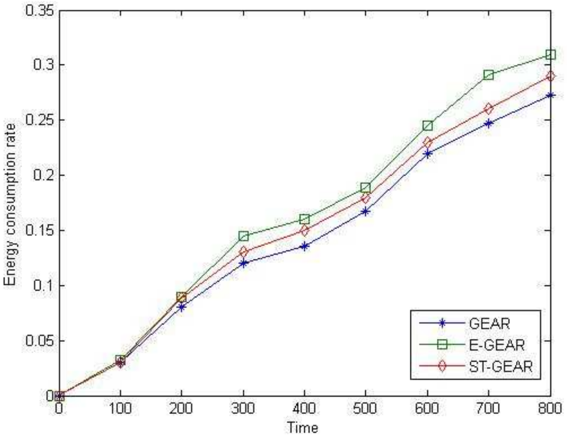


Figure 8: Energy consumption rates

## 6 Conclusions

This paper takes the security of wireless sensor network as the research background and proposes a ST-GEAR secure routing protocol. Firstly, aiming at the problems of complex computation of trust value and excessive energy consumption of node in current trust mechanism, a reputation mechanism BDRM based on binomial distribution is presented. The node trust value is calculated by the residual energy and the communication process, which can be taken a basis of identifying the malicious nodes in the network. Secondly, a secure bootstrap model based on regular hexagon grid deployment is given. On this basis, the ST-GEAR protocol is proposed. In the routing, the protocol selects the node with minimum cost as the next hop forwarding node considering the node energy, distance and the trust value. This protocol can ensure that the information transmitted between nodes is encrypted, which can prevent the malicious nodes in the network to attack the network and improve the security and robustness of the network. The simulation results show that the ST-GEAR protocol can improve the transmission rate of packet and reduce the loss rate of packet and energy consumption compared with GEAR and C-GEAR protocols.

The proposed ST-GEAR protocol can solve the problems including the selective forwarding, Sybil attack and false routing attack that are vulnerable to the GEAR, but there still are some disadvantages. Firstly, the introduction of reputation mechanism BDRM in the protocol has some influence on the energy consumption of the network. Secondly, the consideration of the type of malicious nodes is not very comprehensive. There are other types of malicious nodes. These will be improved in the future researches.

## Acknowledgments

This work is supported by the Natural Science Foundation of Liaoning Province (CN) (20170540777). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

## References

[1] Wu Bo and Li La-Yuan, "Secure routing algorithm based on power-efficient for wireless sensor networks," in *Pacific-Asia Conference on Circuits,communications and System*, pp. 35–38, 2009.

[2] Benjie Chen, Kyle Jamieson, Hari Balakrishnan, and Robert Morris, "Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks," *Wireless Networks*, vol. 8, no. 5, pp. 481–494, 2002.

[3] Saurabh Ganeriwal, Laura K. Balzano, and Mani B. Srivastava, "Reputation-based framework for high integrity sensor networks," *Acm Transactions on Sensor Networks*, vol. 4, no. 3, pp. 1–37, 2008.

[4] Vehbi C. Gungor and Gerhard P. Hancke, "Industrial wireless sensor networks: Challenges, design principles, and technical approaches," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 10, pp. 4258–4265, 2009.

[5] W. B Heinzelman, A. P Chandrakasan, and H Balakrishnan, "An application specific protocol architecture for wireless microsensor networks," in *IEEE Transactions on Wireless Communication*, pp. 660–670, 2002.

[6] C. Karlof and D. Wagner, "Secure routing in wireless sensor networks: attacks and countermeasures," in *IEEE International Workshop on Sensor Network Protocols and Applications, 2003. Proceedings of the First IEEE*, pp. 113–127, 2003.

[7] Chun Ta Li and Min Shiang Hwang, "A lightweight anonymous routing protocol without public key en/decryptions for wireless ad hoc networks," *Information Sciences*, vol. 181, no. 23, pp. 5333–5347, 2011.

[8] W. T. Li, T. H. Feng, and M. S. Hwang, "Distributed detecting node replication attacks in wireless sensor networks: A survey," *International Journal of Network Security*, vol. 16, no. 5, pp. 323–330, 2014.

[9] Yanhui Lv, Zhuo Miao, and Xiurong Wei, "An energy gradient based hexagon clustering protocol for wireless sensor networks," *Icic Express Letters Part B*, vol. 6, 2015.

[10] Nidal Nasser and Yunfeng Chen, "Seem: Secure and energy-efficient multipath routing protocol for wireless sensor networks," *Computer Communications*, vol. 30, no. 11C12, pp. 2401–2412, 2007.

[11] Ganesh R. Pathak and Suhas H. Patil, "Mathematical model of security framework for routing layer protocol in wireless sensor networks ," *Procedia Computer Science*, vol. 78, pp. 579–586, 2016.

[12] R Di Pietro, L. V Mancini, Yee Wei Law, and S Etalle, "Lkhw: a directed diffusion-based secure multicast scheme for wireless sensor networks," in *International Conference on Parallel Processing Workshops, 2003. Proceedings*, pp. 397–406, 2003.

[13] Sohini Roy and Ayan Kumar Das, *Secure Hierarchical Routing Protocol (SHRP) for Wireless Sensor Network*. Springer Berlin Heidelberg, 2014.

[14] Eliana Stavrou and Andreas Pitsillides, *A survey on secure multipath routing protocols in WSNs*. Elsevier North-Holland, Inc., 2010.

[15] S. Tanachaiwiwat, P. Dave, R. Bhindwale, and A. Helmy, "Location-centric isolation of misbehavior and trust routing in energy-constrained sensor networks," in *IEEE International Conference on Performance, Computing, and Communications*, pp. 463–469, 2004.

[16] WANGXiao-yun, YANGLi-zhen, and CHENKe-fei, "Sleach: Secure low.energy adaptive clustering hierarchy protocol for wireless sensor networks," *Wuhan University Journal of Natural Sciences*, vol. 10, no. 1, pp. 127–131, 2005.

[17] Yan Yu, Ramesh Govindan, and Deborah Estrin, "Geographical and energy aware routing: a recursive data dissemination protocol for wireless sensor networks," *Marine Pollution Bulletin*, vol. 20, no. 1, p. 48, 2011.

[18] Yihui Zhang, Li Xu, and Xiaoding Wang, "A cooperative secure routing protocol based on reputation system for ad hoc networks," *Journal of Communications*, vol. 3, no. 6, pp. 283–285, 2008.

[19] L. I. Zhi-Yuan and Ru Chuan Wang, "A survey of secure routing in wireless sensor networks," *Journal of Nanjing University of Posts and Telecommunications)*, vol. 30, no. 1, pp. 77–87, 2010.

## Biography

**Yanhui Lv** biography. She received the master's degree in computer application technology from Shenyang Ligong University in 2005, and the Ph.D. degree in computer application technology from Northeastern University in 2010. Now, she is a professor working in the College of Information Science and Engineering, Shenyang Ligong University. Her current research interests include wireless sensor network and system simulation.

**Kexin Liu** biography. She received the B.E. degree in computer science and technology from Jiangxi University of Traditional Chinese Medicine in 2016. Currently, she is a master degree candidate for computer application technology at Shenyang Ligong University. Her research interests include wireless network and software engineering.

**Deyu Zhang** biography. He received the Ph.D. degree in computer science from Nanjing University of Science and Technology in 2005. Currently, he is a professor working in the College of Information Science and Engineering, Shenyang Ligong University. His research interests include wireless network, artificial intelligence and embedded system.

**Zhuo Miao** biography. She received the B.E. degree in computer science and technology from Shenyang Ligong University in 2014, and the master's degree in computer application technology from Shenyang Ligong University in 2017. Her research interests include wireless sensor network and embedded system.

# An Efficient and Secure Cipher-Text Retrieval Scheme Based on Mixed Homomorphic Encryption and Multi-Attribute Sorting Method Under Cloud Environment

Lin Teng, Hang Li, Jie Liu, and Shoulin Yin
(Corresponding author: Hang Li)

Software College, Shenyang Normal University
No.253, HuangHe Bei Street, HuangGu District, Shenyang, P.C 110034, China
(Email: 1451541@qq.com)

*(Received May 11, 2017; revised and accepted Aug. 26, 2017)*

## Abstract

Cipher-text retrieval plays an important role in data encryption storage service under cloud environment. The present sorting search algorithms have low precision due to corresponding score computing with single local attribute, which cannot accurately sort according to the similarity. To solve this problem, we propose an efficient and secure cipher-text retrieval scheme based on mixed homomorphic encryption and multi-attribute sorting method under cloud environment. This new scheme is divided into four steps: 1) Constructing multiple attribute characteristic vector safety index for documents; 2) Constructing reverse index for the uploaded documents and generating vector set of document, then computing module of each document vector; 3) Encrypting document vector set with homomorphic encryption and uploading them into cloud; 4) Adopting multiple attribute score formula to calculate document relevance score, according to the scores ranking to return the interesting retrieval results for users. Experiments show that our method has the higher retrieval speed and the better retrieval efficiency under cloud environment.

*Keywords: Cipher-Text Retrieval; Cloud Environment; Homomorphic Encryption; Multi-Attribute Sorting Method; Relevance Score*

## 1 Introduction

Cloud computing [24, 25] is the comprehensive development of parallel computing, distributed computing and grid computing. Cloud has attracted widespread attention and recognition as it transfers the traditional computing and storage functions into the cloud environment, which saves lots of hardware cost for users. Currently, the typical cloud platforms are EC2 [3], Google App Engine [11] and Microsoft Live Mesh. With the development of cloud, more sensitive information (such as medical records, financial information and important documents of company) are stored in cloud [1, 2]. Once the data are received by cloud provider, users lose the directly control for their data, which can cause the leak of privacy data. Encryption is an effective method to protect privacy of users' data. However, this way loses many features and can lead to difficult encryption [5, 6, 9, 21]. Especially, how to conduct encrypted data query in untrusted cloud environment has aroused people's attention.

Currently, most ciphertext retrieval schemes do not support sort search. Especially, under cloud computing environment with a large scale of data, there may be lots of documents including one keyword [15, 18]. How to find the closest document in some documents is difficulty. Many researchers had proposed a lot of schemes to improve it. Tan [26] proposed a non-circuit based Ciphertext Policy-Attribute Based Homomorphic Encryption scheme to support outsourced cloud data computations with a fine-grained access control under the multiuser scenario. First, he incorporated Attribute Based Encryption scheme into homomorphic encryption scheme in order to provide a fine-grained access control on encrypted data computation and storage. Then, the proposed scheme was further extended into non-circuit based approach in order to increase the practical efficiency between enterprise and cloud service providers. The results had greatly reduced the computation time and ciphertext size. But the new scheme had a low retrieval efficiency. Gong [12] presented an encryption scheme, based on the composite degree residuosity classes, which could block chosen ciphertext attack while maintaining homomorphism. Elhoseny [10] proposed a novel encryption schema based on Elliptic Curve Cryptography and homomorphic encryption to secure data transmission in

WSN. To reduce energy consumption of cluster head, homomorphic encryption was used to allow cluster head to aggregate the encrypted data without having to decrypt them. He demonstrated that the proposed method was capable to work with different sensing environments that needed to capture text data as well as images. Lu [22] proposed encryption of all genotype and phenotype data to maintain the privacy of subjects. Chen [8] analyzed the secrecy capacity between the source station and the destination station based on the Shannon third theorem on channel capacity, which introduced some secure cooperative communication system. Hou [14] proposed a robust remote authentication scheme with privacy protection, which achieved the efficiency. Arup [7] proposed scheme by using simple Boolean based encryption and decryption of the data files, which was low in computational cost. Zhang [29] presented a pairing-based multi-user homomorphic encryption scheme to privately outsource computation of different users. And there are many other methods, such as [19, 20, 28]. But retrieval efficiency is still low in their methods.

Therefore, this paper proposes an efficient and secure cipher-text retrieval scheme based on mixed homomorphic encryption and multi-attribute sorting method. This new scheme can execute distinction sorting according to the different sorting ways selected by the user, and return the most suitable retrieval results. From the keywords local property and global property, multi-attribute sorting comprehensively reflects the document characteristics, which not only fully considers the difference under the same keywords in different documents, but ensures the documents' quality and authority. Meanwhile, it adopts fully homomorphic encryption method that not only ensures the safety of the user's data, but can directly execute addition and multiplication operation for cipher text, this will greatly improve the retrieval efficiency. Compared with the existing schemes, it improves the precision with ensuring data security.

This paper is organized as follows. Section 2 detailed introduces ciphertext retrieval model based on multi-attribute sorting and homomorphic encryption. We give the experiment analysis in Section 3. There is a conclusion in Section 4.

# 2 Ciphertext Retrieval Model Based on Multi-Attribute Sorting and Homomorphic Encryption

Data has characteristics with frequent using and a large scale in a cloud computing environment. How to quickly and accurately find the data you need in this environment is a very important problem. Obviously, it is inappropriate to use a linear search algorithm and public key encryption algorithm for retrieval process. In that it needs much logarithmic operation. Though the existing ciphertext sort search algorithms have small calculation and high speed, the precision is lower. This paper presents a data retrieval model based on multi-attribute sorting to make up for this shortcoming as shown in Figure 1.

From Figure 1, we can know that this new model includes two main processes: document preprocessing and cryptograph retrieval correlation sorting. Document preprocessing is used to extract multi-attribute eigenvector and construct reverse index. Function of cryptograph retrieval correlation sorting is that it adopts multi-attribute scoring function to calculate the document correlation score according to user's retrieval request and return sorting result.

## 2.1 Determining Keywords

In this step, initial document is as the input. Words will be separated by segmentation method. It will form result set $K = (k_1, \cdots, k_n)$ which will be filtered. In this results, it chooses the keywords that can reflect document meaning clearly.

## 2.2 Multi-Attribute Eigenvector Extraction

Multi-attribute eigenvector extraction is a key step for document preprocessing. It can extract the multi-attribute eigenvector of keywords. The followings are some definitions.

**Definition 1.** Keyword local property. *The contribution of keyword for document can be affected by local factors. These influence factors are defined as keyword local property, such as TFIDF value, word property, word length and word position* etc.

**Definition 2.** Keyword global property. *It also can be called document property. The contribution of keyword for document can be affected by global factors. These influence factors are defined as keyword local property, such as document quoted frequency and download number* etc.

**Definition 3.** Local property eigenvector. *Eigenvector using keyword local property can be called local property eigenvector.*

**Definition 4.** Global property eigenvector. *Eigenvector using keyword global property can be called global property eigenvector.*

**Definition 5.** Single property eigenvector. *If document eigenvector only uses one keyword property that it can be called single property eigenvector.*

**Definition 6.** Multi-property eigenvector. *It consists of keyword local property and global property as Figure 2.*

The *keywords* are extracted from document. *Attributes* indicate local and global properties.

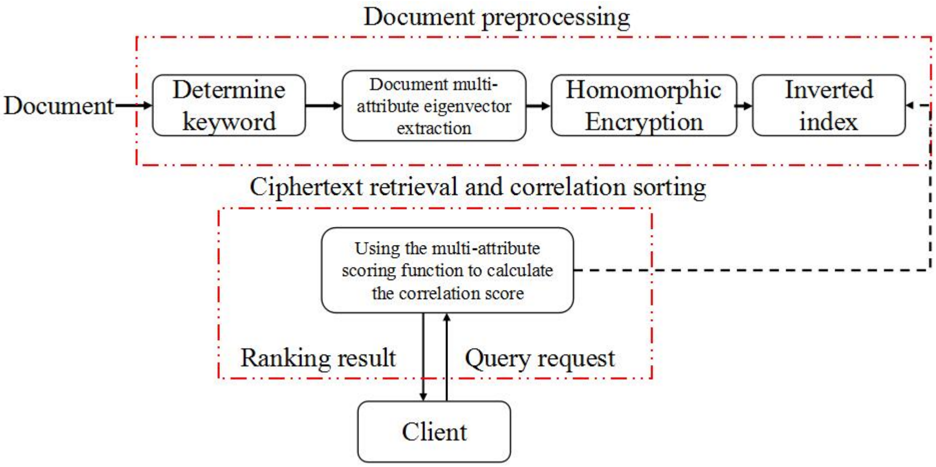

Document preprocessing
Document
Determine keyword
Document multi-attribute eigenvector extraction
Homomorphic Encryption
Inverted index
Ciphertext retrieval and correlation sorting
Using the multi-attribute scoring function to calculate the correlation score
Ranking result
Query request
Client

Figure 1: Ciphertext retrieval model based on multi-attribute sorting

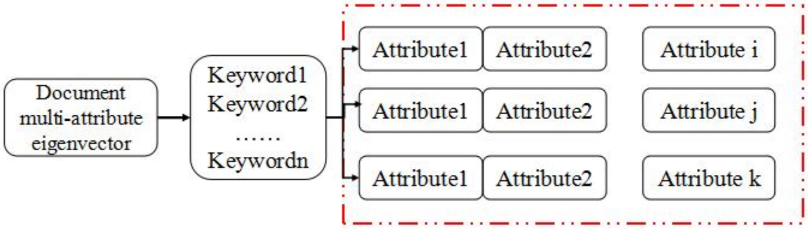


Figure 2: Document multi-attribute eigenvector

## 2.3 Homomorphic Encryption

Homomorphism is the concept of modern algebra [23]. Supposing $< G, \cdot >$ and $< H, * >$ are two algebra systems. If $\forall a, b \in G$, then $f(a \cdot b) = f(a) * f(b)$. The $f$ can be called homomorphic mapping from $G$ to $H$. Cryptography promotes the concept of homomorphic mapping in modern algebra. Encryption is a mapping from the plaintext to cryptographic [4]. And if the encryption mapping is a homomorphic mapping, we can say that it is a homomorphic encryption scheme. In this paper, homomorphic encryption is based on integer modulo arithmetic, its processes are described as follows:

- *Kengen*. Randomly select a $P-digit$ big prime number as key $p$.
- *Encryption*. Randomly select a $Q - digit$ big prime number $q$, and $P > Q > plaintextlength$. Two random numbers $r_1$ and $r_2$, $N = pq$. Ciphertext $c = (m + pr_1 + pqr_2) \bmod N$.
- *Decryption*. Plaintext $m = c \bmod p$.

Homogeneity analysis. Assuming that two plaintexts $m_1$ and $m_2$, and corresponding ciphertexts $c_1$ and $c_2$. So

$$\begin{aligned} c_1 &= (m_1 + pr_{11} + pqr_{12}) \bmod N. \\ c_2 &= (m_2 + pr_{21} + pqr_{22}) \bmod N. \end{aligned}$$

Additive homogeneity analysis.

$$\begin{aligned} c_1 + c_2 &= (m_1 + m_2 + p(r_{11} + r_{21}) \\ &\quad + pq(r_{12} + r_{22}) \bmod N \\ &= c(m_1 + m_2). \end{aligned}$$

Multiplication homogeneity analysis.

$$\begin{aligned} c_1 \cdot c_2 &= (m_1 \cdot m_2 + pm_1 r_{21} + pqm_1 r_{22} \\ &\quad + pm_2 r_{11} + p^2 r_{11} r_{21} \\ &\quad + p^2 q r_{11} r_{22}) \bmod N. \\ c_1 \cdot c_2 \bmod p &= m_1 \cdot m_2. \end{aligned}$$

In summary, the homomorphic encryption is based on integer modulo arithmetic, which not only meets the additive homogeneity, but meets multiplication homogeneity.

## 2.4 Reverse Index

Reverse index [16,30] is a data structure used to describe the relation between keywords set and documents set. It stores the storage location mapping of keyword in one document under the full-text retrieval.

The traditional reverse index structure consists of an index file and an reverse file. The index list is a collection of all the keywords in the document. It is composed of all records, each record contains the keywords and keywords' corresponding pointer. Pointer points to the corresponding logical address in the reverse file. The reverse

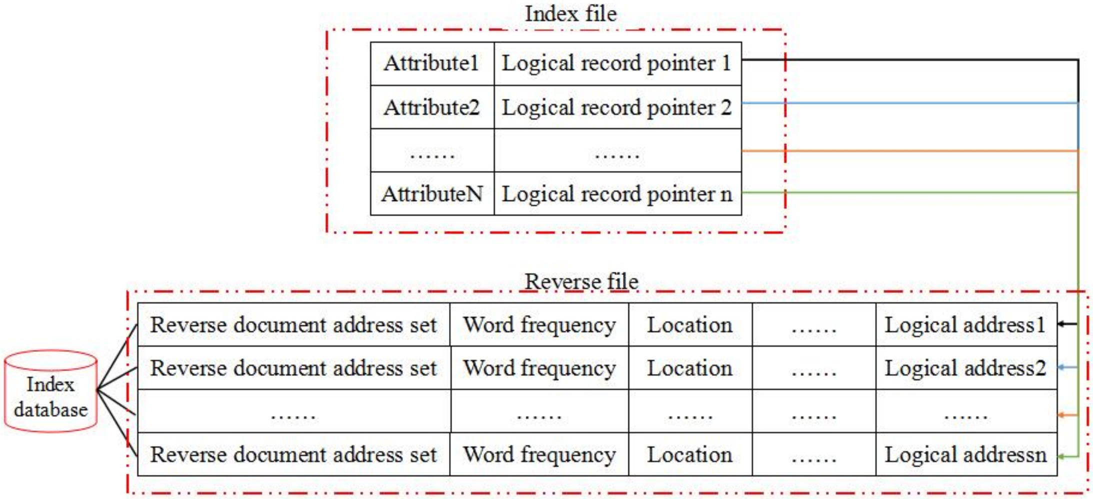


Figure 3: Multi-attribute reverse index

file indicates that which documents contain this keyword, the frequency information of keywords and addresses set of these documents. Our new reverse index model is as Figure 3. According to multi-attribute eigenvectors, it will change the single attribute keyword frequency in traditional reverse file as multi-attribute including keyword frequency, location, indexed frequency *etc.*

## 2.5 Score Function of Document

Multi-attribute score function is adopted to execute relevance sorting. Score function can calculate the correlation score between retrieval word and a document.

In our new model, the relevance score computing is according to local property and global property. Due to the different importance degree of attribute, we introduce attribute weight. For example, for attribute $(\rho_1, \rho_2, \cdots, \rho_n)$ of keyword $K$, its corresponding weight value $(o_1, o_1, \cdots, o_n)$, and $\sum_1^n o_i = 1$. Therefore, attribute score function of our new scheme can be defined as:

$$score = \sum_{i=1}^{n} \rho_i \times o_i. \tag{1}$$

Where $o_i$ is the weight of attribute $\rho_i$, which can be dynamic adaptively adjusted according to different sort methods.

# 3 Experience Analysis

In order to verify the effectiveness of new scheme, we make comparison experiments under MATLAB environment.

First, we select evaluation criteria: Recall rate ($RE$), Precision rate($PE$) and $MAP$(Mean of Average Precision).

$$RE = \frac{SRE}{TRE}. \tag{2}$$

$$PE = \frac{SRE}{TNE}. \tag{3}$$

$$MAP = \sum_{i}^{r} \frac{1}{r}. \tag{4}$$

Where $SRE$ and $TRE$ denote system retrieved relevant files and total number of relevant files respectively. $TNE$ is the total number of retrieved files. If the retrieved file is closer to the top, the $MAP$ is likely to rank higher.

Second, we select 50 papers from Google Scholar and set global attribute for all papers. When constructing retrieve model, each paper will be preprocessed. We set *title*, *content*, *keywords* and *summary* as retrieve information. CP-ABHER-LWE [27], HACC [17] and ECCH [13] are compared with our method (abbreviated to HEMAS) to show the better results of our scheme under the same experiment environment.

It conducts same query in single attribute cryptograph retrieval system and multi-attribute cryptograph retrieval system and returns to the results of top 20, 30, 40. Table 1 presents the results. It can be seen that our method is better than other three schemes.

Table 1: Experimental evaluation results

| Index | CP-ABHER-LWE | HACC | ECCH | HEMAS |
|---|---|---|---|---|
| RE | 93.4% | 95.8% | 94.5% | 96.8% |
| PE | 94.6% | 96.2% | 95.5% | 98.4% |
| MAP | 94.9% | 96.1% | 94.9% | 97.5% |

Similarly, it conducts same query in single attribute cryptograph retrieval system and multi-attribute cryptograph retrieval system and returns to the results of top 10, 20, 30, 40. Table 2 is the retrieval time comparison.

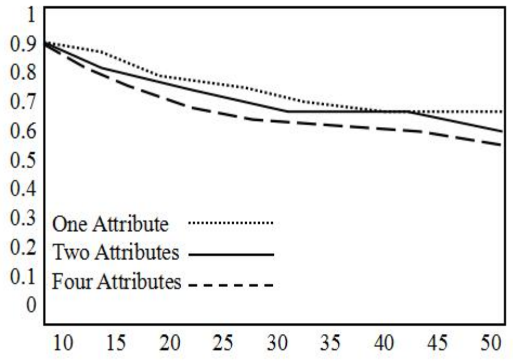


Figure 4: Precision comparison with new scheme

From the table, the retrieval time is greatly reduced with new scheme.

Table 2: Retrieval time

| Index | CP-ABHER<br>-LWE | HACC | ECCH | HEMAS |
|---|---|---|---|---|
| Top 10 | 188ms | 190ms | 205ms | 163ms |
| Top 20 | 191ms | 192ms | 208ms | 164ms |
| Top 30 | 193ms | 193ms | 212ms | 166ms |
| Top 40 | 195ms | 195ms | 214ms | 168ms |

We only make experiment with our new method using different attributes: one attribute, two attributes, four attributes as shown in Figure 4.

We can see that precision will become more higher with the increasing of attribute. In summary, this paper proposes the efficient and secure cipher-text retrieval scheme based on mixed homomorphic encryption and multi-attribute sorting method under cloud environment, which is effectiveness for improving retrieval accuracy rate.

### 3.1 Performance Analysis

We also make a comparison to CP-ABHER-LWE [27], HACC [17] and ECCH [13] with our HEMAS method. The following is the explanation of symbols in this section: $p$: bilinear operation. $e$: exponent operation. $s$: point multiplication operation in $G_1$. $|G_1|$: the element length of corresponding group. $|m|$: the length of message. $|U|$: the length of user identity.

Table 3 shows the calculation about the four algorithms. And we can know that users with HEMAS only need one pairing operation and one point multiplication operation in Homomorphic encryption stage less than CP-ABHER-LWE, HACC and ECCH. In Reverse index stage, HEMAS needs $2n - 1$ pairing operations obviously superior to ECCH. The pairing operation number is more than CP-ABHER-LWE, HACC and ECCH. In that our new scheme dose not need exponent operation, the total calculation is superior to CP-ABHER-LWE, HACC and ECCH when $n$ is big.

In order to specifically analyze running time, we use the A type elliptic curve to test in jpbc database. Then we record the running time with the above schemes as table4 from MATLAB platform.

Table 4 shows that the running time with our new scheme is less than other schemes. It is the optimal scheme.

## 4 Conclusions

We present a mixed retrieval scheme in this paper, which shows the following merits.

1) **Fast retrieval speed**. The new scheme only needs to search the safety index list that is very suitable for the data with a large scale in the cloud environment.

2) **Less calculation**. Due to homomorphic encryption, it greatly decreases the calculation time.

3) **Precision improved**. This scheme introduces multi-attribute eigenvectors including keywords global properties and local properties, which makes comprehensive evaluation for documents from several aspects, so as to return the most relevant documents.

4) **Shorten retrieval time**. In that the precision of score computing greatly reduces the sorting time, so retrieval time is reduced too.

## References

[1] D. S. AbdElminaam, "Improving the security of cloud computing by building new hybrid cryptography algorithms," *International Journal of Electronics and Information Engineering*, vol. 8, no. 1, pp. 40-48, 2018.

[2] S. Almulla, Y. Y. Chan, "New secure storage architecture for cloud computing," *Future Information Technology*, pp. 75-84, 2011.

[3] B. Amazon, "Amazon elastic compute cloud (Amazon EC2)," *Virtual Grid Computing*, 2010.

[4] Z. Cao, L. Liu, Y. Li, "Ruminations on Fully Homomorphic Encryption in Client-server Computing Scenario," *International Journal of Electronics and Information Engineering*, vol. 8, no. 1, pp. 32–39, 2018.

[5] T. Y. Chang and M. S. Hwang, W. P. Yang, "An improved multi-stage secret sharing scheme based on the factorization problem," *Information Technology and Control*, vol. 40, no. 3, pp. 246-251, 2011.

[6] T. Y. Chang, M. S. Hwang, W. P. Yang, K. C. Tsou, "A modified ohta-okamoto digital signature for batch verification and its multi-signature version," *International Journal of Engineering and Industries (IJEI'12)*, vol. 3, no. 3, pp. 75-83, Sep. 2012.

Table 3: Calculation comparison with different schemes

| Index | CP-ABHER-LWE | HACC | ECCH | HEMAS |
|---|---|---|---|---|
| Homomorphic encryption | $3ne + s$ | $n(2p + s + 3e)$ | $3n(p + 2e)$ | $n(s + p)$ |
| Reverse index | $4ns - 2s + 2np + p$ | $n(2p + s)$ | $n(2e + 3p)$ | $(2n - 1)p$ |

Table 4: Calculation time comparison with different schemes

| Index | CP-ABHER-LWE | HACC | ECCH | HEMAS |
|---|---|---|---|---|
| Homomorphic encryption | 258.647ms | 241.323ms | 109.752ms | 50.219ms |
| Reverse index | 352.458ms | 288.943ms | 209.546ms | 68.252ms |

[7] A. K. Chattopadhyay, A. Nag, K. Majumder, “Secure data outsourcing on cloud using secret sharing scheme,” *International Journal of Network Security*, vol. 19, no. 6, pp. 912, 2017.

[8] J. S. Chen, C. Y. Yang, M. S. Hwang, “The capacity analysis in the secure cooperative communication system,” *International Journal of Network Security*, vol. 19, no. 6, pp. 863-869, 2017.

[9] S. F. Chiou, M. S. Hwang, S. K. Chong, “A simple and secure key agreement protocol to integrate a key distribution procedure into the DSS,” *International Journal of Advancements in Computing Technology (IJACT'12)*, vol. 4, no. 19, pp. 529-535, Oct. 2012.

[10] M. Elhoseny, H. Elminir, A. Riad, *et al.*, “A secure data routing schema for WSN using elliptic curve cryptography and homomorphic encryption,” *Journal of King Saud University - Computer and Information Sciences*, vol 28, no. 3, pp. 262-275, 2015.

[11] B. Ferriman, T. Hamed, Q. H. Mahmoud, “Storming the cloud: A look at denial of service in the google app engine,” in *IEEE International Conference on Computing, Networking and Communications*, pp. 363-368, 2015.

[12] L. Gong, S. Li, Q. Mao, *et al.* ”A homomorphic encryption scheme with adaptive chosen ciphertext security but without random oracle,” *Theoretical Computer Science*, vol. 609, pp. 253-261, 2016.

[13] M. Q. Hong, P. Y. Wang, W. B. Zhao, “Homomorphic encryption scheme based on elliptic curve cryptography for privacy protection of cloud computing,” in *IEEE International Conference on Big Data Security on Cloud*, pp. 152-157, 2016.

[14] G. Hou, Z. Wang, “A robust and efficient remote authentication scheme from elliptic curve cryptosystem,” *International Journal of Network Security*, vol. 19, no. 6, pp. 904-911, 2017.

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

[16] M. S. Hwang and C. C. Lee, “Research issues and challenges for multiple digital signatures,” *International Journal of Network Security*, vol. 1, no. 1, pp. 1-7, 2005.

[17] C. Lan, H. Li, C. Wang, “A new key-aggregate encryption scheme with chosen ciphertext security,” in *IEEE Sixth International Conference on Information Science and Technology*, pp. 229-233, 2016.

[18] C. C. Lee, Shih-Ting Hsu, M. S. Hwang, “A study of conjunctive keyword searchable schemes,” *International Journal of Network Security*, vol. 15, no. 5, pp. 311–320, 2013.

[19] H. Li, S. L. Yin, Chu Zhao and Lin Teng, “A proxy re-encryption scheme based on elliptic curve group,” *Journal of Information Hiding and Multimedia Signal Processing*, vol. 8, no. 1, pp. 218-227, Jan. 2017.

[20] J. Liu, S. L. Yin, H. Li and L. Teng, “A density-based clustering method for K-anonymity privacy protection,” *Journal of Information Hiding and Multimedia Signal Processing*, vol. 8, no. 1, pp. 12-18, Jan. 2017.

[21] L. Liu, W. Kong, Z. Cao, J. Wang, “Analysis of one certificateless encryption for secure data sharing in public clouds,” *International Journal of Electronics and Information Engineering*, vol. 6, no. 2, pp. 110-115, 2017.

[22] W. J. Lu , Y. Yamada, J. Sakuma ”Privacy-preserving genome-wide association studies on cloud environment using fully homomorphic encryption,” *BMC Medical Informatics and Decision Making*, vol. 15, 2015.

[23] M. Ogburn, C. Turner, P. Dahal, “Homomorphic encryption,” *Procedia Computer Science*, vol. 20, pp. 502-509, 2013.

[24] Z. Qingchen, L. T. Yang, C. Zhikui, “Privacy preserving deep computation model on cloud for big data feature learning,” *IEEE Transactions on Computers*, vol. 65, no. 5, 2016.

[25] Z. Qingchen, Z. Hua, L. T. Yang, C. Zhikui, B. Fanyu, “PPHOCFS: Privacy preserving high-order CFS algorithm on cloud for clustering multimedia data,” *ACM Transactions on Multimedia Computing, Communications and Applications*, 2016. DOI: 10.1145/2886779.

[26] S. F. Tan, A. Samsudin, “Ciphertext policy-attribute based homomorphic encryption (CP-ABHER-LWE) scheme: A fine-grained access control on outsourced cloud data computation,” *Journal of Information*

*Science and Engineering*, vol. 33, no. 3, pp. 675-694, 2017.

[27] S. F. Tan, A. Samsudin, "Ciphertext policy-attribute based homomorphic encryption (CP-ABHER-LWE) scheme: A fine-grained access control on outsourced cloud data computation," *Journal of Information Science & Engineering*, vol. 33, no. 3, pp. 675-694, 2017.

[28] S. L. Yin and J. Liu, "A K-means approach for map-reduce model and social network privacy protection," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 7, no. 6, pp. 1215-1221, Nov. 2016.

[29] W. Zhang, "A pairing-based homomorphic encryption scheme for multi-user settings," *International Journal of Technology & Human Interaction*, pp. 11, 2016.

[30] H. Zhu and R. Wang, "Multi-party password-authenticated key exchange scheme with privacy preserving using chaotic maps in random oracle model," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 8, no. 1, pp. 42-53, Jan. 2017.

# Biography

**Lin Teng** received the B.Eng. degree from Shenyang Normal University, Shenyang , Liaoning province, China in 2016 . Now, she is a laboratory assistant in Shenyang Normal University. Her research interests include Multimedia Security, Network Security, Filter Algorithm and Data Mining. She had published more than 10 international journal papers on the above research fields. Email:1532554069@qq.com.

**Hang Li** is a full professor in Software College, Shenyang Normal University. He received his B.S. and M.S. degrees from Northeastern University. He is also a master's supervisor. He has research interests in wireless networks, mobile computing, cloud computing,social networks, network security and quantum cryptography. Professor Li had published more than 30 international journal papers (SCI or EI journals) on the above research fields. Email: 1451541@qq.com.

**Jie Liu** is a full professor in Software College, Shenyang Normal University. He received his B.S. and M.S. degrees from Harbin Institute of Technology. He is also a master's supervisor. He has research interests in wireless networks, mobile computing, cloud computing,social networks, network security and quantum cryptography. Professor Liu had published more than 30 international journal papers (SCI or EI journals) on the above research fields. Email: nan127@sohu.com.

**Shoulin Yin** received the B.Eng. And M.Eng. degree from Shenyang Normal University, Shenyang , Liaoning province, China in 2013 and 2015 respectively. His research interests include Multimedia Security, Network Security, Filter Algorithm and Data Mining. He received School Class Scholarship in 2015. Yin had published more than 50 international journal papers (SCI or EI journals) on the above research fields. Email:352720214@qq.com.

# A New Scheme for Source-location Privacy in Wireless Sensor Networks

Shaoquan Jiang, Min Li, and Zailiang Tang
*(Corresponding author: Min Li)*

Institute of Information Security, Mianyang Normal University
166 Mianxing Rd. West, High-Tech District, Mianyang 621000, China
(Email: shaoquan.jiang@gmail.com)

*(Received May 16, 2017; accepted Aug. 26, 2017)*

## Abstract

In this work, we study the source location privacy in wireless sensor networks (SLP-WSN), where a source wishes to send a message to a base station while preserving the privacy of its location. This problem has a lot of applications including habitat monitoring and military surveillance. A well-known strategy is to divide the message transmission into two stages. The first stage routes the source message to a random position in the network while the second stage routes the source message to the base station through the shortest path. To preserve the source location privacy, the system must securely mask the trace of the routing. Thus, the first stage routing is essential. The literature proposed many approaches. However, they either assume a sensor's awareness of its location or have high energy consumption or have an obvious privacy weakness. In this work, we propose a new method for SLP-WSN without these defects. We first help a sensor node to obtain distances to three reference nodes by flooding. Then, we design the first stage routing probabilistically, where the next sensor is chosen according to a carefully designed distribution such that a sensor close to the random point has a higher probability. Our experiment shows that the energy consumption and the message delay only have a small factor expansion, relative to directly routing via the shortest path to the base station. It is also shown to have a good privacy.

*Keywords: Information Security; Location Privacy; Security Protocol; Sensor Network Security*

## 1 Introduction

With the advances of the network technology, wireless sensor networks (WSNs) are widely believed to be a good solution for monitoring unattended or dangerous environments. The network is formed with many small and cheap sensors [?]. Sample applications include environmental monitoring and military surveillance.

Due to its unattended nature, it is vulnerable to various attacks. The standard security problems such as confidentiality and authentication can be solved using well-known cryptographic techniques. However, some issues can not be solved cryptographically [?, ?, ?]. The location privacy is one of them. It can be well interpreted by the panda monitoring problem, where sensors are deployed in the panda's habitat to monitor its activity and report to the base station. However, since the transmission is wireless in nature, it can be eavesdropped by hunters who attempts to trace back to the panda through the message path. This tracing procedure does not need to know the meaning of a message (if encrypted) and hence it is not concerned with the secret keys. Thus, a careful design (especially the routing protocol) is needed to combat such an attack. In this sense, a regular routing scheme for WSN is certainly not enough.

### 1.1 Related Works

The representative technique for location privacy in WSN is phantom routing proposed by Ozturk *et al.* [?] (enhanced by Kamat *et al.* [?]), where the message routing consists of two stages: in stage one, it follows a directed random walk of $h$ steps and in stage two, the message is routed to the base station via flooding or a single path. Li *et al.* [?, ?] extended the phantom routing by first sending the message to a random intermediate rode and then routing it to a network mixing ring (NMR) before sending it to the base station. Li *et al.* [?, ?] considered multi random intermediate nodes by choosing their polar coordinates as $(d_1, \alpha), (d_2, 2\alpha), \cdots, (d_n, n\alpha)$, where $\alpha$ is a random angle. A single path routing to a random intermediate point from a restricted area was studied in [?]. Yao and Wen [?] considered the random shortest path from the source to the base station. This approach does not have a good location-privacy as with high probability the physical routing path will stay around the line between the source and the base station and hence the attacker will be relatively easy to capture many packets and trace back to the source in a hop-by-hop manner. The phantom routing technique was further studied in [?, ?, ?].

Location privacy for a receiver was studied in the literature. Although this is different from our problem, it can be regarded as the dual problem to the source location privacy. In [?, ?, ?], the protocol is based on a multi-path routing and a fake packet injection. However, their fake packet injection has a constant rate and hence is energy consuming. Jian *et al.* [?] considered the location privacy of a moving receiver in a wireless sensor network. Their method involves a fake packet injection by each intermediate sensor and the fake packet needs to transmit sufficiently far. Hence, their method is still energy-consuming. Luo *et al.* [?] proposed a variant of [?]. But they still need to forward the fake packet and hence is energy consuming again. Phantom routing and fake source techniques are combined in [?].

Yao *et al.* [?] proposed a new method for source-location privacy in WSN, where they considered the notion of *ring* that is the set of nodes with the same distance to the base station. They first routes the message to a random ring $c$ and then routes it on the ring toward a fixed direction for some steps. Next, they routes to a random ring $b$ to do the similar thing and finally routes the message directly to the base station. It is assumed that a node on a ring knows which neighbors are in a given direction and which are not. This can not be implemented only by a landmark (as done by the east-west separation in [?]) because nodes are placed circularly. So it requires the awareness of its own and neighbors' positions.

Yang *et al.* [?] considered the clustered wireless sensor network with a cluster head more powerful than a common sensor and the location-privacy is achieved through the faking message simulation by cluster heads. This method is not suitable for a homogenous network such as our setting since the message simulation will consume the power quickly. Location-privacy in wireless sensor network against a global eavesdropper was studied in [?, ?, ?, ?, ?, ?], where the adversary can eavesdrop the whole network and the privacy is achieved by fake packets. Since a global attacker is more powerful than our local eavesdropper, this method is relevant to our setting. However, it always results in a large energy consumption and we also feel that a global eavesdropper model is too strong. Hence, we will not consider such an adversary.

### 1.2 Contribution

In this work, we propose a new method for SLP-WSN. Our scheme lies in the well-known two-stage strategy: the first stage routes the message to a random intermediate position and the second stage routes the message to the base station. However, we do not assume any location information for each sensor (beyond the knowledge of three reference nodes' coordinates). Instead, we help a sensor to obtain distances (termed *hop distance* which is the number of hops between two nodes) to reference nodes and use the sensor's hop distance tuple as its location information. We also find out a method to estimate the hop distance between two positions while each position is represented by its hop distance tuple. Our scheme starts the first stage routing using a probabilistic method. Specifically, starting from the source, the next sensor is chosen probabilistically among the current sensor's neighbors, where the probability distribution is carefully designed such that a sensor close to the random intermediate position has a better chance of being selected. Our choice of the probability distribution allows the next-hop sensor is chosen probabilistically so that it is hard to trace back, while the message can still steadily move toward the random intermediate position. The second stage is a shortest path routing. Under our design, we evaluate its privacy and efficiency. We consider the message delay measure (called *PathRatio*), defined as the ratio of our routing path length to the shortest path from the source to the base station. Our *PathRatio* only has a small factor. We also consider the energy consumption ratio (termed *EnergyRatio*), defined as the length of total messages sent in the whole network to the length of total messages sent by nodes when only using the shortest path routing from the source to the base statio. Our scheme, *EnergyRatio* is equal to *PathRatio* and is small. We also consider the *SafetyPeriod*, defined as the number of source messages the source can send before its location privacy is broken. We build a model to quantify this and find our protocol has a good privacy. Finally, assuming the network is almost fully connected, our scheme can deliver the source message reliably. In comparison with existing works [?, ?, ?, ?, ?, ?, ?], our protocol either has a much smaller *EnergyRatio* or a better privacy or removes a node's awareness of its location; see Section 4.3 for details.

## 2 Model

We now formalize the source-location privacy in wireless sensor networks (SLP-WSN). This consists of the system model, assumptions, location privacy and efficiency measures.

### 2.1 System Model

SLP-WSN is a system that enables a source $S$ to transmit a message $m$ to a base station under the help of some sensors such that no adversary, who eavesdrops the traffic at some points, can determine the location of $S$. In our model, $S$ is an entity that can communicate, where the motivation example is the soldier in a battle field or a sensor carried by a monitored animal. In some works, $S$ is a sensor that monitors the environment such as a wild animal. Obviously, these two presentations have no essential differences in the location privacy technology. We use $v_1, \cdots, v_{n-1}$ to denote *sensors* and use $v_0$ to denote *base station*. For simplicity, a sensor or base station is called a *node*.

We suggest a three-stage model for a SLP-WSN system, which consists of a deployment stage, a preprocessing stage and a message transmission stage. Details are as follows.

**Deployment.** In this stage, a system manager will deploy nodes in a desired area. The position of $S$ is undetermined and it can move arbitrarily. For simplicity, we assume that the deployed area is planar with radius $R$. However, our work can be easily generalized to the three-dimensional case.

**Preprocessing.** In this stage, $v_0, \cdots, v_{n-1}$ jointly execute a protocol. At the end, each $v_i$ obtains an internal state, which will be crucial for the next-stage protocol.

**Message transmission.** This is the main part of the system. It helps source $S$ transmit a source message $m$ to base station $v_0$. Toward this, $S$ will find a node $v_{i_0}$ nearby and send $m$ to it. Then, $v_{i_0}$ will find an adjacent node $v_{i_1}$ and send $m$ to it. This process continues until $m$ reaches $v_0$.

## 2.2 Assumptions

Our system will make the following assumptions.

- We assume that the location of a sensor is fixed throughout the system. This assumption will be used to keep the internal state of a node obtained in the preprocessing stage unchanged. However, as long as the network topology does not change frequently, it can be relaxed by executing the preprocessing stage periodically. This will be further discussed in Section 5.2. This slow changing topology is suitable for applications such as the habitat monitoring.
- The sensor network as an undirected graph is almost fully connected, where an edge $(v_i, v_j)$ means that $v_i$ is in the hearing range of $v_j$ and vice versa. This in fact is the necessary assumption for any useful sensor network. We use the term "almost" as a few unconnected nodes do not affect the system validity and it is easy to satisfy through a uniformly random deployment.
- We assume that the message between sensor nodes are authenticated. This is the assumption that has been made in the literature. It can be waived through a key management. This assumption essentially means that we only consider the eavesdropping attack. Strictly, this belongs to the formulation of the adversarial model. But we put here to remind the readers this restriction on the adversarial power.

## 2.3 Location Privacy

In this subsection, we consider the privacy of a SLP-WSN system. Toward this, we first specify feasible adversarial behaviors and then define the location privacy.

**Adversarial behaviors.** The adversary has $\nu$ devices to perform eavesdropping attacks in the network, each of which has a hearing range $h$. For any selected position, he can place an eavesdropping device. Any signal transmitted in its hearing range will be captured. For each captured signal, the attacker can localize its immediate sender node.

**Remark 1.**

1) *In this work, we only consider the eavesdropping attack above. An active attack such as a man-in-the-middle attack is not considered as a convention of known SLP-WSN systems. This can be amended through a key distribution protocol to allow any two neighboring nodes to share a key with which the authentication and confidentiality can be achieved. This is obviously out of the scope of this work and we will not explore this.*

2) *Some works in the literature (e.g., [?, ?, ?]) considered the global eavesdropping attack, where an adversary can eavesdrop any message sent in the network. We feel this attack is too strong and unnecessary. To secure against such an attack, the system must sacrifice the efficiency. For example, the systems in [?, ?, ?] emit many faking packets in order to mislead the attacker. This results in a large energy consumption and is not desired.*

**Location privacy.** The location privacy is to require that an attacker can not find the physical location of $S$ by performing an eavesdropping attack above. The location privacy is quantified by the number of source messages that $S$ can send before it is localized by the attacker and we call it *SafetyPeriod*. Certainly, *SafetyPeriod* is expected to be as large as possible for better privacy.

## 2.4 Efficiency Measures

In this subsection, we define three measures to evaluate the efficiency of a SLP-WSN system.

**PathRatio.** We use *PathRatio* to denote the ratio of the number of edges on the real path that a source message will travel from $S$ to $v_0$, relative to the number of edges on the shortest path from $S$ to $v_0$. Note that a small *PathRatio* indicates a small transmission delay and hence is desired.

**EnergyRatio.** We use *EnergyRatio* to denote the ratio of the total length of messages that nodes in the whole network have sent in order to transmit one source message from $S$ to $v_0$, relative to the length of messages sent by the nodes when $S$ routes the message only through the shortest path to $v_0$. This measure is concerned with the energy consumption of the network and hence is better to be as small as possible. If $S$ sends $m$ to $v_0$ by simply flooding it into the network, then *EnergyRatio* will be $n/d_{Sv_0}$, where $d_{Sv_0}$ is the shortest path length from $S$ to $v_0$ and $n$ is the network size.

**DeliveryRate.** The *DeliveryRate* is the percentage of the source messages from $S$ that have been successfully delivered to $v_0$. Usually, for a useful protocol, *DeliveryRate* should be almost 100%.

# 3 Construction

In this section, we present a new SLP-WSN protocol. In our protocol, $v_3, \cdots, v_{n-1}$ are *ordinary sensors* while $v_1, v_2$ are *reference sensors*. Reference sensors are functionally identical to ordinary sensors, except that they will, together with the *base station* $v_0$, play as anchors to help all sensors determine their locations. We present our protocol in stages.

## 3.1 Deployment

In this stage, a system manager deploys the nodes in a region of radius $R$ as follows.

- Randomly deploy $\{v_3, \cdots, v_{n-1}\}$ in the desired area (of radius $R$) and place $v_0, v_1, v_2$ on the perimeter such that any two of them are separated by an angle $2\pi/3$ (as in Figure 1).

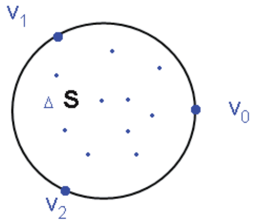


Figure 1: The placement of $v_0, v_1, v_2$

## 3.2 Preprocessing

In this stage, each node will obtain some state information for its future execution. This includes its neighboring information, the shortest path to $v_0$ and the minimum distances to $v_0, v_1, v_2$, where the minimum distance between $v$ and $v'$ is the minimal number of nodes to traverse from $v$ to $v'$.

Let $\mathcal{N}(v)$ be the set of nodes within the hearing range of node $v$ and the hearing range of a node is $r_0$. So, $v_i \in \mathcal{N}(v_j)$ if and only if $v_j \in \mathcal{N}(v_i)$. Clearly, $\mathcal{N}(v)$ is defined only after the deployment.

The preprocessing stage has two sub-protocols. The first one helps each $v$ to learn $\mathcal{N}(v)$ while the second one helps $v$ to learn its location information.

**Compute $\mathcal{N}(v)$.** In this protocol, each $v_i$ broadcasts hello. When a node $v_j$ receives hello from $v_i$, it means that $v_i$ lies in its hearing range $r_0$ and so it adds $v_i \in \mathcal{N}(v_j)$. Since every node $v$ broadcasts hello to its neighbors, any $v_j$ can compute $\mathcal{N}(v_j)$. The formal description is as follows.

1) For $i = 0, \cdots, n-1$, $v_i$ broadcasts $v_i|$hello.

**loop** Upon $v_i|$hello, each $v_j$ adds $v_i$ into $\mathcal{N}(v_j)$.

2) If no more hello is heard, $v_j$ stores $\mathcal{N}(v_j)$.

We remind again that, as other related works, the privacy model in this paper only considers eavesdropping attacks and hence $\mathcal{N}(v)$ above can be correctly computed. However, as remarked in the privacy model, the privacy against active attacks can be achieved easily through a key management scheme.

**Compute the location information.** Now we show how to help $v_i$ to compute its minimum distances to $v_0, v_1, v_2$ respectively. If the distance were Euclidean, then they are sufficient to uniquely locate $v_i$ (by elementary mathematics). However, we do not assume a sensor to be equipped with any positioning tool such as GPS, the Euclidean distance is hard to obtain. Instead, we use the length of the shortest path (i.e., the number of edges on the shortest path) between two nodes as a representation for the distance between them and call it *hop distance*. Under this, a node's distance vector (to $v_0, v_1, v_2$) should approximately represent its relative location in the network[1].

To compute the hop distance from each $v_i$ to $v \in \{v_0, v_1, v_2\}$, we run the one-to-all shortest path algorithm for three times. Let $d_{vu}$ be the hop distance from node $v$ to node $u$. Then, the three executions of the algorithm will respectively compute $\{d_{v_0v_i}\}_{i=0}^{n}$, $\{d_{v_1v_i}\}_{i=0}^{n}$ and $\{d_{v_2v_i}\}_{i=0}^{n}$. The protocol is to continuously update $d_{vv_j}$ for $v = v_0, v_1, v_2$ and finally obtain the correct $d_{vv_j}$. Since the symbols in $\{d_{v_0v_i}\}_i$, $\{d_{v_1v_i}\}_i$ and $\{d_{v_2v_i}\}_i$ do not overlap, we can present the three executions in parallel. We use $\texttt{pre}(u)$ to record the next node on the shortest path from $u$ to $v_0$ in computing $\{d_{v_0v_i}\}_i$. Clearly, starting from any node $u$ and iteratively following $\texttt{pre}(\cdot)$, the shortest path to witness $d_{v_0u}$ is given. Note that the shortest path to witness $d_{v_1u}$ or $d_{v_2u}$ (other than hop distances $d_{v_1u}, d_{v_2u}$) is not interesting to us and hence is not considered. Details are in Figure 2.

[1] It is possible that two neighboring nodes have the same distance vector. However, it is unlikely that two nodes of several hops away still share the same distance vector. Thus, although the distance vector can not accurately localize a node, it certainly approximates its *relative* location very well.

1. Each $v_i$ lets $d_{vv_i} = \infty$ for $v \in \{v_0, v_1, v_2\}$ and $\texttt{pre}(v_i) = \perp$ . Then, $v \in \{v_0, v_1, v_2\}$ updates $d_{vv} = 0$ and sends $v|v|0$ to $\mathcal{N}(v)$.
2. [**loop**] When $v_i$ receives $v_j|v|\ell$ with $v \in \{v_0, v_1, v_2\}$ from $v_j \in \mathcal{N}(v_i)$, it proceeds only if $d_{vv_i} > \ell{+}1$. In this case, it updates $d_{vv_i} = \ell{+}1$, sends $v_i|v|d_{vv_i}$ to $\mathcal{N}(v_i)$, and (if $v = v_0$) also updates $\texttt{pre}(v_i) = v_j$.
3. If no more update is heard, $v_i$ defines $\mathbf{d}_i = (d_{v_0v_i}, d_{v_1v_i}, d_{v_2v_i})$ and broadcasts $v_i|\mathbf{d}_i$ to $\mathcal{N}(v_i)$.
4. $v_j$ keeps $(v_i, \mathbf{d}_i)$ for $v_i \in \mathcal{N}(v_j) \cup \{v_j\}$, and $\texttt{pre}(v_j)$.

Figure 2. Compute $\texttt{pre}(v_i), d_{v_0v_i}, d_{v_1v_i}$ and $d_{v_2v_i}$ for each $v_i$.

At the end of this stage, $v_i$ obtains $\mathbf{d}_i, \mathcal{N}(v_i)$, $\{\mathbf{d}_j\}_{v_j \in \mathcal{N}(v_i)}$ and $\texttt{pre}(v_i)$ and it keeps them for the use in the next stage. Again, we remind that as the privacy model considers eavesdropping attack only, all the information can be correctly computed.

## 3.3 Message Transmission

In this stage, we show how $S$ sends message $m$ to $v_0$ under the help of some sensors. As our concern is to hide the location of $S$, we can not route $m$ through the shortest path to $v_0$. Otherwise, an attacker can stay around $v_0$ to eavesdrop signals and trace back hop-by-hop. This is feasible, as the shortest path from $v$ to $v_0$ is fixed (recall that $\texttt{pre}(v_i)$ is fixed after the preprocessing stage) and one signal allows him to trace one step back (hence $d_{Sv_0}$ signals will lead to $S$).

The idea of our protocol is as follows. Source $S$ first sends $m$ to an adjacent node $v_{i_0}$. Then, $v_{i_0}$ chooses a random point in the deployed area through sampling a vector $\mathbf{d}$ that represents a random position's hop distances to $v_0, v_1, v_2$. It then probabilistically chooses a node $v_{i_1} \in \mathcal{N}(v_{i_0})$ according to a certain distribution and requests $v_{i_1}$ to route $m$ to location $\mathbf{d}$. Here the choice of $v_{i_1}$ has a property that a node closer to $\mathbf{d}$ has a better chance to be selected. Upon $m$, $v_{i_1}$ chooses a node $v_{i_2}$ and requests it to route $m$ to $\mathbf{d}$. This process continues until $m$ reaches a node $v_{i_N}$ close to $\mathbf{d}$. In this case, $v_{i_N}$ routes $m$ to $v_0$ via the shortest path.

Before proceeding, we introduce or recall some notations. Some of them will not be used until Section 4. But we put them together for an easy reference later.

- $||\mathbf{d} - \mathbf{d}'|| = \sqrt{\sum_{i=0}^{2} |d_i - d_i'|^2}$.
- $\mathbf{d}_j$ is the vector of the hop distance from $v_j$ to $v_0, v_1, v_2$.
- $\mathsf{ord}_i(v_j)$ is the order of $||\mathbf{d} - \mathbf{d}_j||$ in the decreasingly sorted list of $\{||\mathbf{d} - \mathbf{d}_t||\}_{v_t \in \mathcal{N}(v_i)}$ for a given $\mathbf{d}$, where $v_j \in \mathcal{N}(v_i)$.
- Given a constant $\omega > 0$ and for $v \in \mathcal{N}(v_i)$, let

$$Q_{v_i,\mathbf{d}}(v) = \frac{|\mathcal{N}(v_i)|\omega + \mathsf{ord}_i(v)}{|\mathcal{N}(v_i)|^2(\omega + .5) + .5|\mathcal{N}(v_i)|}. \quad (1)$$

- *elementary triangle formulae*: if a triangle has two sides of lengths $\ell_1, \ell_2$ with angle $\theta$ between them, then the third side has a length

$$L(\ell_1, \ell_2, \theta) = (\ell_1^2 + \ell_2^2 - 2\ell_1\ell_2 \cos\theta)^{1/2}. \quad (2)$$

- $R$ is the radius of the deployed area with origin $O$; $r_0$ is the hearing of a sensor node; $h$ is the hearing range of an adversarial device.
- A point in the deployed area has polar coordinates $(r, \theta)$, with origin $O$ so that $v_0$ has the coordinates $(R, 0)$. Thus, $v_1$ is at $(R, \frac{2\pi}{3})$ and $v_2$ is at $(R, \frac{4\pi}{3})$.
- $\delta$ is a small constant (e.g., about 3) and its concrete value is not important.
- $\zeta$ is a constant scaling factor (see details at the end of this section).
- The area within radius $r_1$ of $S$ is called *threat area* and $r_1$ is called *threat radius*.

Recall that $v_i$ has the following state from the previous stage.

- $\mathcal{N}(v_i)$: the set of neighbors of node $v_i$.
- $\mathbf{d}_i = (d_{v_0v_i}, d_{v_1v_i}, d_{v_2v_i})$: $d_{vu}$ is the hop distance (i.e., the minimum number of hops) from $v$ to $u$.
- $\texttt{pre}(v_i)$: the first node on the shortest path from $v_i$ to $v_0$.

Notice that $Q_{v_i,\mathbf{d}}(\cdot)$ is a probability distribution over $\mathcal{N}(v_i)$. We will use this distribution to select the next node before approaching $\mathbf{d}$. This selection has the property that a node $v_j$ with $\mathbf{d}_j$ closer to $\mathbf{d}$ will have a larger probability $Q_{v_i,\mathbf{d}}(v_j)$. One might wonder why we can not simply replace $|\mathcal{N}(v_i)|\omega + \mathsf{ord}_i(v_j)$ with $||\mathbf{d} - \mathbf{d}_j||$. In theory, this is possible. But it has a drawback that when $v_i$ is far from $\mathbf{d}$, $||\mathbf{d} - \mathbf{d}_j||$ will remain almost constant when $v_j$ goes over $\mathcal{N}(v_j)$. Thus, $Q_{v_i,\mathbf{d}}(\cdot)$ will be almost uniformly random. Under this, statistics tells us that the routing will remain about $S$ even after a long time. The constant factor $\omega$ is used to adjust the "gap" between $Q_{v,\mathbf{d}}$ and a uniformly random distribution. The gap will decrease when $\omega$ increases. Given $\omega$, a smaller $|\mathcal{N}(v)|$ implies a larger gap and so we can increase $\omega$ to reduce the gap. A more considerate design is to allow a sensor $v$ to choose its own $\omega$ (w.r.t. the value $|\mathcal{N}(v)|$). In this work, we just use a global $\omega$ for the convenience of analysis.

In our experiment, we found that choosing $v_{i_{j+1}}$ solely according to $Q_{v_i,\mathbf{d}}(\cdot)$ is problematic. For some networks, the routing will be stuck in a small area for a long time. To avoid this, we make a special rule such that if $v_{i_j}$

1. If $S$ wishes to send $m$ to $v_0$, it sends $m$ to an adjacent node $v_{i_0}$. Upon $m$, the latter chooses $\theta \in [0, 2\pi)$ uniformly randomly and $r \in [0, R]$ with density $P(r) = \frac{2r}{R^2}$. Next, it computes $\mathbf{d}$ with $d_u = \zeta \cdot L(r, R, \frac{2u\pi}{3} - \theta)/r_0$ for $u = 0, 1, 2$, and samples $v_{i_1}$ from $\mathcal{N}(v_{i_0})$ w.r.t. $Q_{v_{i_0},\mathbf{d}}(\cdot)$ and *special rule* below. Finally, it sends $0|v_{i_0}|m|\mathbf{d}$ to $v_{i_1}$.
2. [**loop**] (*randomized routing*) If $v_{i_j}$ receives $0|v_{i_{j-1}}|m|\mathbf{d}$, it does the following. If $||\mathbf{d} - \mathbf{d}_{i_j}|| \leq \delta$ (small constant), then it moves to step 3; otherwise, it samples $v_{i_{j+1}}$ w.r.t. $Q_{v_{i_j},\mathbf{d}}(\cdot)$ and *special rule* below, and sends $0|v_{i_j}|m|\mathbf{d}$ to $v_{i_{j+1}}$.
3. [**loop**] (*deterministic routing*) Upon $1|v_{i_{j-1}}|m|$ or when transferred from step 2, if $v_{i_j} = v_0$, then $m$ arrives at $v_0$ successfully; otherwise, $v_{i_j}$ sends $1|v_{i_j}|m$ to $v_{i_{j+1}} = \texttt{pre}(v_{i_j})$.

Figure 3. Message transmission. *Special rule:* The next sensor policy is amended such that if the current node $v_{i_j}$ has been visited three times, then the next node $v_{i_{j+1}}$ will be the node among $\mathcal{N}(v_{i_j})$ with the minimal distance to $\mathbf{d}$ and that if the current node $v_{i_j}$ has been visited seven times, then $v_{i_j}$ moves to step 3 to route $m$ via the shortest path to $v_0$.

was visited three times, then the next node $v_{i_{j+1}}$ will be chosen as the node in $\mathcal{N}(v_{i_j})$ closest to $\mathbf{d}$. Under this, the message will not linger around $v_{i_j}$. Besides this, we also found that there exists a certain bad $\mathbf{d}$ so that no $v_i$ lies in the distance $\delta$ to $\mathbf{d}$. In this case, the message will move around $\mathbf{d}$ forever. To avoid this problem, we also make the special rule such that if $v_{i_j}$ has been visited seven times, then it directly routes $m$ to $v_0$ via the shortest path. Of course, here the threshold three and seven can be modified to other values but we found they work well in our experiments.

For a point at $(r, \theta)$, we can see that their *Euclidean distances* to $v_0, v_1, v_2$ are respectively $L(R, r, \frac{2u\pi}{3} - \theta), u = 0, 1, 2$. However, what we need is a measure that is compatible with the hop distance vectors $\mathbf{d}_i$'s. Fortunately, we find in the experiment that, for a randomly chosen $(r, \theta)$, if we scale the Euclidean distance by a factor $\zeta/r_0$, then this scaled distance vector will well approximate a hop distance vector. Further, this $\zeta$ only depends on the average node degree $\mathbf{E}(|\mathcal{N}(v)|)$ (e.g., $\zeta = 1.478$ if the average node degree is 7). Intuitively, $\zeta$ is affected by two factors: (1) the average hop length in a shortest path is smaller than the full hop length $r_0$; (2) the shortest path between two nodes is not in a straight line. Both factors will result in $\zeta$ larger than 1.

With the above discussion, we are now ready to present our protocol; see Figure 3.

# 4 Analysis

In this section, we analyze the location privacy of our protocol and discuss its efficiency. When necessary, please see Section 3.3 to recall some notations.

## 4.1 Location Privacy

Now we consider the location privacy of source $S$. Before our analysis, we give a sample path for the randomized routing to get a picture of what it looks like; see Figure 4.

We will analyze the following attack framework. The attacker has $\nu$ eavesdropping devices. He can place his devices at any location for eavesdropping. We assume that if the attacker captures a message within the radius $r_1$ of $S$, then he can localize $S$. We call this area *threat area* and $r_1$ the *threat radius*. To properly capture the capability of a real adversary, $r_1$ can not be large (a large $r_1$ also implies a strong model and we can not result in an interesting result). We suggest $r_1 = 8r_0$ although our analysis does not depend this. Toward discovering the threat area, an attacker can try any strategy to place his eavesdropping device. Our effort in this section is to find the number of source messages that $S$ can send before the attacker can discover the threat area. This is, we derive the *SafetyPeriod*. This is done in several steps.

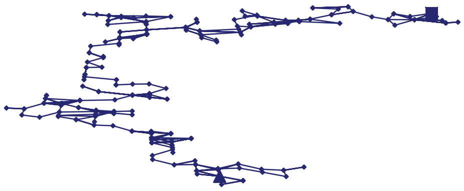


Figure 4: A sample path from source ▲ to random intermediate point ■, with $R = 63r_0$ and $\mathbf{E}[|\mathcal{N}(v)|] = 7$

For convenience, we call the path from $v_{i_0}$ to $v_{i_N}$ *Path I* and the path from $v_{i_N}$ to $v_0$ *Path II*.

**An eavesdropping attack on Path II is useless.** Since Path II starts at $v_{i_N}$, an eavesdropping attack on Path II can at most trace back to $v_{i_N}$. This is useless as $\mathbf{d}$ is the uniformly random in the deployed area (independent of the source $S$).

In the following, we only consider the eavesdropping attack on Path I. We first show that catching the message from $v_{i_j}$ to $v_{i_{j+1}}$ can not imply any information about the location of $v_{i_{j-1}}$ (although it indeed indicates the location of $v_{i_j}$ as assumed).

Under this, one eavesdropped message only exposes one edge on the transmission path. Then, we analyze the edges exposed by eavesdropped messages and derive the probability that an attacker can discover the threat area. From this, *SafetyPeriod* will be computed.

**Tracing back more than one edge from one eavesdropped message on Path I is impossible.** In this part, we argue that an attacker can not trace back more than one edge from one eavesdropped message on the routing Path I. We will not directly prove this rigorously. Instead, we compare our next sensor policy with the uniformly random policy, in which the next sensor node is chosen uniformly randomly from the neighbors of the current node[2]. In contrast, our policy is to choose the next sensor $v_{i_{j+1}}$ according to $Q_{v_{i_j},\mathbf{d}}(\cdot)$ and *special rule* (see Figure 3). Since the next node of the uniformly random policy is independent of the current and previous nodes, tracing back more than one edge is obviously impossible. On the other hand, the next sensor policy through the shortest path to $\mathbf{d}$ is deterministic and hence is the worst. The strategy for proving our policy is good is to find a measurement under which, our policy is close to the uniformly random policy while the shortest path policy is the worst.

Toward this, we define $\mu_j$ to be the angle between $\overrightarrow{v_{i_0}v_{i_j}}$ and the average vector $\mathbf{E}(\overrightarrow{v_{i_j}v_{i_{j+1}}})$, where $\mathbf{E}(\cdot)$ is taken over the distribution of $v_{i_{j+1}}$ (for fixed $\mathbf{d}$ and $v_{i_j}$). We also define $\theta_j$ to the angle between $\overrightarrow{v_{i_0}v_{i_j}}$ and $\overrightarrow{v_{i_j}v_{i_{j+1}}}$. If the next sensor policy is good, then $\theta_j$ should vary a lot around $\mu_j$. For the uniformly random policy, no matter what $\mu_j$ is, $\theta_j$ is uniformly random over $[0, \pi]$. In contrast, for the shortest path policy, $\theta_j \equiv \mu_j$. Thus, we consider $\Delta = \mathbf{E}(|\theta_j - \mu_j|)$ as the measurement for the performance of the next sensor policy, where $\mathbf{E}(\cdot)$ is over the distribution of $v_{i_j}, v_{i_{j+1}}, \mathbf{d}$ and $\mu_j$ (note that $\mu_j$ depends on $v_{i_j}$ and $\mathbf{d}$). Note that $\Delta$ varies with $j$. We run simulations to see how $\Delta$ in our scheme performs. We take $R = 63r_0, \omega = 1$, and the average node degree $\mathbf{E}[|\mathcal{N}(v)|] = 7$. The result for $(\Delta, j)$ is shown in Figure 5, where we notice that the shortest path policy has $\Delta = 0$ while the uniformly random policy has[3] $\Delta = \pi/3$. From the experimental result, we can see that our $\Delta$ is very close to that of the uniformly random policy and hence demonstrates its excellent performance against the tracing-back attack for more than one edge.

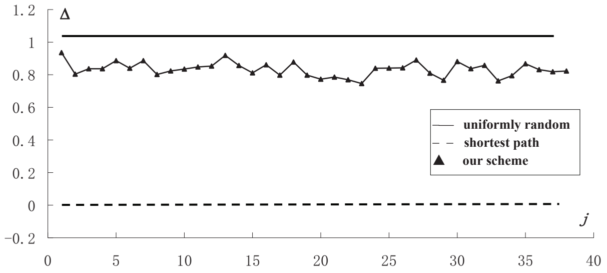


Figure 5: $\Delta = \mathbf{E}(|\theta_j - \mu_j|)$ varies with $j$: our scheme vs uniformly random policy (constant $\pi/3$) vs the shortest path policy (constant 0).

**The probability to discover the threat area.** In the above, we have demonstrated that one eavesdropped message can only expose the underlying edge (in a routing path). Now if an attacker can capture many messages, then their underlying edges in the routing paths are exposed. If one of these messages lies in the threat area, then the attacker may realize it and compromise the location privacy of $S$. In the following, we will calculate the probability that an attacker can catch one message in the threat area when $S$ sends one message. Then, we will use it to compute the *SafetyPeriod*.

If a device is placed on the radius $r$ of $S$, the probability that an outgoing message from $S$ will be eavesdropped, can be calculated as follows. If $r \le h$, then certainly the message will be catched as the hearing range of the device covers $S$. If $r > h$, it is easy to see that the device covers the angle (centered at $S$) of at most $2\beta = 2\arcsin(h/r)$; see Figure 6. Hence, the message is catched with probability at most $\frac{\arcsin(h/r)}{\pi}$.

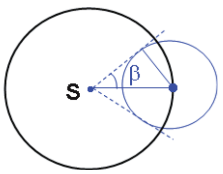


Figure 6: Hearing coverage of an eavesdropping device with respect to $S$

Now since $S$ is uniformly random in the deployed area, a particular eavesdropping device is on the radius $r$ to $S$ with the probability density $\frac{2r}{R^2}$ (here for simplicity, we assume the maximum distance is still $R$). Thus, the probability that a single device is placed within the threat radius $r_1 (> h)$ to $S$ and also catches a message, is

$$\begin{aligned}
&\int_0^h \frac{2r}{R^2}dr + \int_h^{r_1} \frac{2r}{R^2}\frac{\arcsin(h/r)}{\pi}dr \\
=&\frac{r^2 * \arcsin(h/r) + h\sqrt{r^2 - h^2}}{R^2\pi}\Big|_h^{r_1} + \frac{h^2}{R^2} \text{ (use Maple)} \\
=&\frac{r_1^2 * \arcsin(h/r_1) + h\sqrt{r_1^2 - h^2}}{R^2\pi} + \frac{h^2}{2R^2} \qquad (3)
\end{aligned}$$

We can assume $h \le 0.3r_1$, under which, it is well approximated that $\arcsin(h/r_1) \approx h/r_1$ and

[2] Note since a uniformly random policy does not allow a message to go far, this strategy actually is not recommended in the literature. However, since it obviously prevents an attacker from tracing back more than one edge, it is an ideal standard to evaluate the performance of our next sensor policy.

[3] Note that under the uniformly random policy, $\theta_j$ is uniformly random over $[0, \pi]$, while $\mu_j$ is uniformly random over $[0, \pi]$ (over the distribution of $v_{i_j}$ and $\mathbf{d}$). Thus, $\Delta = \pi/3$, which can be easily calculated from $\mathbf{E}(|\theta_j - \mu_j|)$.

$\sqrt{r_1^2 - h^2} \approx r_1$. Then, Eq. (3) is approximately $\frac{2r_1h}{R^2\pi} + \frac{h^2}{2R^2}$. This is the probability for one device. Since the attacker has $\nu$ devices, the probability he catches a message within radius $r_1$ of $S$ (when one source message is sent) is

$$\frac{2\nu r_1 h}{R^2\pi} + \frac{\nu h^2}{2R^2}. \tag{4}$$

**Calculating SafetyPeriod.**

Now we calculate the *SafetyPeriod*. If $S$ sends $W$ messages in the *SafetyPeriod*, then an attacker can identify the *threat area* of $S$ only if he can catch at least one message within the radius $r_1$ of $S$. Thus, to preserve the location privacy of $S$, we can use Equation (4) to demand the average number of catched messages in the *threat area* to satisfy $\frac{2Wr_1h\nu}{R^2\pi} + \frac{Wh^2\nu}{2R^2} < 1$. This implies $W < \frac{R^2}{0.5\nu h^2 + 0.636\nu r_1 h}$. Note any $W$ satisfies this restriction will be good. Thus,

$$SafetyPeriod = \lceil \frac{R^2}{0.5\nu h^2 + 0.636\nu r_1 h} \rceil - 1. \tag{5}$$

Take $r_1 = 8r_0, h = 3r_0$. We have *SafetyPeriod*$= \lceil 0.0506R^2/(r_0^2\nu) \rceil - 1$. If $R = 63r_0$ and $\nu = 1$, then *SafetyPeriod*=200. Note we take $\nu = 1$ as the literature does not consider the case $\nu > 1$ and this will be convenient for us to compare. The curve $W = 0.0506R^2/r_0^2$ is shown in Figure 7 with *SafetyPeriod* being the integer value just below the curve.

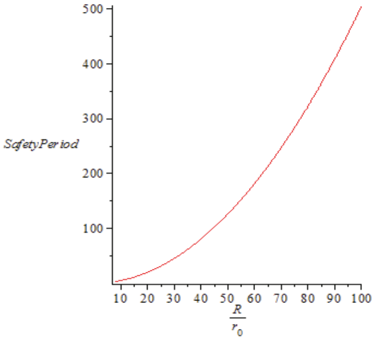


Figure 7: SafetyPeriod vs $R/r_0$, with $\nu = 1, r_1 = 8r_0, h = 3r_0$. *SafetyPeriod* for $R/r_0$ is the integer just below the curve.

**Remark 2.** *The hearing range of adversary $h$ would not be significantly larger than the sensor hearing range $r_0$, as the signal by the sensor will die out quickly beyond the range of $r_0$ (or the noise will be more powerful than the signal), under which the adversarial device can not decode correctly (even if it is good). Thus, our sample choice $h = 3r_0$ is reasonable.*

## 4.2 Efficiency

In this subsection, we discuss the *PathRatio*, *EnergyRatio* and *DeliveryRatio* of our scheme. For the definitions of these measures, see Section 2.4.

In our assumption at Section 2.2, we assume that the network is almost fully connected and thus the message *DeliveryRate* is almost 100%. Our *PathRatio* = *EnergyRatio* also has a good performance. The experimental result is shown in Figure 8, where we depicted the *PathRatio* vs $\omega$ for $R = 63r_0$ and average node degree 7. We also depicted the experimental result *PathRatio* vs the average node degree *deg* for $R = 63r_0$ in Figure 9 (where a node degree vs $\Delta_{18}$ is also shown). We can see that our *PathRatio* is typically very small while the privacy is preserved well (in Figure 7). Although our experiment is done on average node degree 7, we prefer a smaller degree, because the smaller degree gives a smaller *PathRatio* and *EnergyRatio*, as seen in Figure 9(a). The only problem is that we need to satisfy the almost full connectivity. If there is a strategy to satisfy this for a smaller degree, it is certainly good for our application[4]. Since it is obviously out of the scope of this work, we will not explore this.

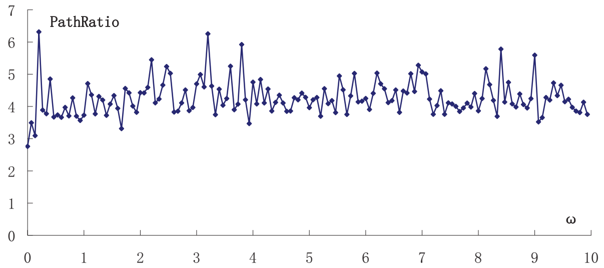
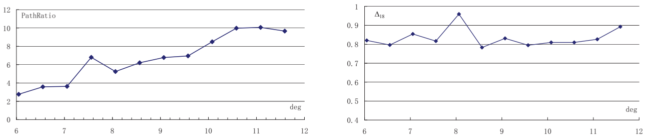


Figure 8: *PathRatio* vs $\omega$ ($R = 63r_0$ and average degree $\mathbf{E}[|\mathcal{N}(v)|] = 7$).



(a) node degree vs PathRatio (b) node degree vs $\Delta_{18}$

Figure 9: The effect of node degree to PathRatio and $\Delta$, where $R = 63r_0$ is fixed. We can see that PathRatio increases significantly with node degree while it does not impact $\Delta$ significantly.

## 4.3 Comparison

In this subsection, we compare our scheme with related works and the summary of the comparison appears in Table 4.3. We believe that our criteria and result are satisfactory.

All the schemes [?, ?, ?, ?, ?, ?, ?] use the two-stage routing strategy as we do. We assume the network is

[4] Indeed, *SafetyPeriod* is not significantly affected by a reduced degree, as our *SafetyPeriod* analysis in Section 4.1 depends on the average node degree only through $\Delta$ while $\Delta$ does not depend significantly on this average degree: $\Delta_{18}$ does not change significantly with the average node degree (see Figure 9(b)) and $\Delta_j$ does not change significantly with $j$ (see Figure 5).

almost fully connected and so all these works (including ours) have almost 100% *DeliveryRate*.

Flooding-based methods were proposed in [?, ?]. They all have a large energy ratio (which is 125 in their sample experiment) although their *PathRatio* is 1. Since our *PathRatio* can be made small (Figure 8 and Figure 9(a)) as long as the average node degree is not large, our protocol is certainly advantageous to theirs.

Phantom single path method in [?] starts the first phase routing with $\ell$ steps of directed random walk. However, the directional information is visible and as observed in [?], this reduces the attack complexity by a factor of $2^{\ell}$. This places a concern on their *SafetyPeriod*. The directed random walk was designed as a weak version of a purely random walk, where the latter has the problem that the message will only move nearby the source node. Our method is to approximate the random walk while it moves toward the random intermediate position. This is achieved by carefully choosing the distribution function for the next-hop sensor.

In comparison with [?, ?, ?, ?, ?], our obvious advantage is to remove a node's awareness of its personal location. This is important as it might need an extra hardware such as GPS to realize. Since our *EnergyRatio* and *PathRatio* can be made small as mentioned above, our gain is interesting. In addition, [?, ?] used a mixing ring to hide the routing and thus the ring nodes have a fast power drainage (although authors suggested a leverage strategy, the effect is limited). Lightfoot *et al.* [?] directly routes the message to a random intermediate point deterministically in the first stage and results in a smaller *PathRatio*. As seen before, a deterministic routing is certainly not advantageous as our probabilistic routing. Li and Ren [?, ?] proposed several schemes. The most interesting one (in our view) is the multi-intermediate method. We do not have an obvious advantage over their method other than removing a node's awareness of its own position. For *EnergyRatio* and *PathRatio* in [?, ?], the comparison with us depends on their choice of the number of intermediate nodes.

As a summary, we can safely conclude that our protocol in comparison with [?, ?] has either a much smaller *EnergyRatio* or a better privacy and in comparison with Li *et al.* [?, ?, ?, ?, ?] has the advantage of removing a node's awareness of its own position. As our *EnergyRatio* and *PathRatio* can be made small (with a moderate average node degree), our advantages are interesting.

# 5 Other Issues

In this section, we discuss some other issues that are important for a useful SLP-WSN system.

## 5.1 Localizing $S$ from a Base Station

Our SLP-WSN system is to prevent an attacker from localizing the source $S$. However, in some situations, the base station (operated by a personnel) might wish to localize $S$. In this case, $S$ can send $\mathbf{d}_{i_0}$ of node $v_{i_0}$ to $v_0$ through our routing system (whenever he moves to a new location). Of course, to be secure, this should be encrypted and authenticated using a secret key shared between $S$ and the base station. When the base station receives $\mathbf{d}_{i_0}$, he can find $S$ by moving toward $\mathbf{d}_{i_0}$. For this to work, he might need to query a node $v_j$ (on his way to $S$) for its distance vector $\mathbf{d}_j$.

Table 1: Performance comparison (undesired result marked double black).

| | Path Ratio | Energy Ratio | own loction Awareness | Safety Period |
|---|---|---|---|---|
| SinglePath Phantom [?] | small | small | partial | **small** |
| Flood [?, ?] | 1 | **large** | not required | large |
| [?, ?]<br>[?, ?, ?] | vary | vary | **required** | large |
| ours | small | small | not required | large |

## 5.2 Service Availability

**Network robustness.** Network is robust if the network is connected when a small fraction of nodes are out of order. Thus, it is significant to design a robust network. However, since this issue is common in a general wireless sensor network. We will not discuss it here and assume that the network remains almost fully connected even if a small fraction of nodes are out of order.

**Routing information update.** In our system, we rely on each node and their neighbors' information about the shortest path to $v_0$ and distance vectors $\mathbf{d}_i$'s. If all nodes are alive and located at the fixed locations, then this information will remain unchanged. If a few nodes die out, then the system can be maintained to continue functioning. Toward this, the system can use a standard network routing update strategy to maintain a node's internal state. Since the number of broken sensors is small, this will be a small workload only. We may also try to keep it working without an update. The only concern is $\mathsf{Pred}(\cdot)$, where when $\mathsf{Pred}(v)$ dies out, our system does not specify what is the next sensor for $v$. To patch this, $v$ can send the message to the node $v_j$ with the second smallest $d_{j0}$ among its neighbors $v_j \in \mathcal{N}(v)$. If this $v_j$ still dies out, then it can try $v_t$ with the third smallest $d_{t0}$. It continues until the message is sent out. If the message can not send out at a node $v$, then the path to $v_0$ is broken and the delivery fails. If only a few nodes are broken in the network, this bad event should not occur with a noticeable probability. Now if the system runs for a longer period, then many nodes might move or die out. In this case, the system needs a significant work to update. It might be better to run the preprocessing stage once again. If this update does not occur frequently, we believe that the re-execution is a feasible solution.

## 6 Conclusion

In this work, we studied the source location privacy in wireless sensor networks, where the source $S$ wants to route a message $m$ to the base station $v_0$ while keeping its own location private. We proposed a new scheme for this problem. Our scheme routes $m$ through a single path to $v_0$. Our idea is to first route $m$ to an intermediate position $\mathbf{d}$ while the choice of the next sensor is randomized according to a well-designed probabilistic strategy. Our strategy has the property that a node close to $\mathbf{d}$ has a better chance to be selected. When $m$ approaches $\mathbf{d}$, the underlying node then routes it directly to $v_0$ through a shortest path. Our protocol performs well at the energy consumption, delivery rate, time delay and the safety period. Importantly, we do not assume that a node is equipped with a localization tool such as GPS. Our scheme has significant advantages over existing protocols.

## References

## Biography

**Shaoquan Jiang** received the B.S. and M.S. degrees in mathematics from the University of Science and Technology of China, Hefei, China, in 1996 and 1999, respectively. He received the Ph.D degree in Electrical and Computer Engineering from the University of Waterloo, Waterloo, ON, Canada, in 2005. From 1999 to 2000, he was a research assistant at the Institute of Software, Chinese Academy of Sciences, Beijing; from 2005 to 2013, he was a faculty member at the University of Electronic Science and Technology of China, Chengdu, China; from 2013 to now, he is a faculty member at Mianyang Normal University, Mianyang, China. He was a postdoc at the University of Calgary from 2006 to 2008 and a visiting research fellow at Nanyang Technological University from Oct. 2008 to Feb. 2009. His research interests are key stream generators, public-key based secure systems and secure protocols.

# Application of FSM Machine and S-Box in KASUMI Block Cipher to Improve Its Resistance Against Attack

Raja Muthalagu and Subeen Jain
*(Corresponding author: Raja Muthalagu)*

Department of Electrical and Electronic Engineering, Birla Institute of Technology and Science, Pilani
345055 Dubai International Academic City, Dubai, United Arab Emirates
(Email: raja.m@dubai.bits-pilani.ac.in)

*(Received July 2, 2017; revised and accepted Oct. 22 & Nov. 5, 2017)*

## Abstract

In this paper, modifications in the original KASUMI block cipher is proposed by introducing a finite-state machine (FSM) and substitution box (S- box) to provide better confidentiality and integrity function in global system for mobile communications (GSM) and 3G networks. The FSM constitute the nonlinear combiner of SNOW 3G block cipher and it uses two S-box to provide strong diffusion. The addition of FSM in KASUMI is introducing the non-linearity in output bits and it will increase the complexity for an attacker to make an attack. Also, few changes are made in a KI and KL keys that are used in a different rounds of KASUMI to prevent various attacks such as a rectangle attack, sandwich attack, single key attack, etc. The simulation results show the performance improvement of the proposed modified KASUMI design is compared with the conventional KASUMI in terms both the encryption speed and encryption time taken.

*Keywords: Finite-state Machine (FSM); KASUMI; S-box; SNOW-3G*

## 1 Introduction

The rapid growth of mobile communications has increased the requirement of having secure network/communication between the users. Multiple ways of attacking or hacking a network are used by attackers. As the wireless mode of communications provides feasibility and ease to the users, the security of information being exchanged within two users or group of users is always at threat. Many algorithms have been proposed which are used for different encryption purposes [5, 6, 16]. Some have proved resistant towards attacks while some have high security issues with them and hence, proved as weaker one by attackers. For having secure network, encryption services and algorithms involved need to be robust and secure enough to provide end-to-end secure transmission of data among various users. It poses a challenge for designers to design highly secure and attack-resistant algorithm for encryption of data.

The KASUMI block cipher is used for providing encryption services in mobile networks like GSM, universal mobile telecommunications system (UMTS) and general packet radio service (GPRS). In UMTS, KASUMI is used in the confidentiality (f8) and integrity algorithms (f9) with names UEA1 and UIA1, respectively. In GSM, KASUMI is used in the A5/3 key stream generator and in GPRS in the GEA3 key stream generator. The KASUMI is evolved from MISTY1 algorithm to provide users safe and secure way for exchange of data. The KASUMI is a slightly modified version of MISTY1. And it provides easy hardware implementation that meets security requirement of 3G mobile communications. In [3, 4, 11, 12, 14] attacks related to KASUMI were discussed which indicate that KASUMI is weak algorithm. Though these attacks were very powerful and posed a threat on it, they were not considered due to impractical assumptions made as suggested in [11] by 3GPP society and thus, inapplicable to real-life attack on full KASUMI. But as there are chances of these or other attacks related to them being carried out practically, new algorithms are designed and worked upon to provide high confidentiality and security of data. The experimental study of the obtained encryptor from various researchers are demonstrated its effectiveness in protecting from many existing types of attacks aimed at block cipher algorithms [13]. Also Reference [10] present a new concept called a certificateless key insulated encryption scheme (CL-KIE).

In this paper, it proposed the modified KASUMI block cipher to improve its resistance against attack. The SNOW-3G block cipher is a another encryption algorithm for mobile networks and the concept of SNOW-3G is widely used in our proposed method. The SNOW-3G is used as UEA2 and UIA2 algorithm for providing confidentiality and integrity in 3rd Generation Partnership Project (3GPP) [7, 8]. It is seen as strong enough for carrying any attack as it has Rijndael's SR box and SQ box and LFSR. It also uses three 32-bit registers R1, R2, R3 and 16 s-boxes and each having capacity to hold 32-bits. As suggested in [9] inclusion of R3 had increased resistance of SNOW-3G towards algebraic attacks along with use of two S-box, and it can be used strengthen the proposed modified KASUMI. The FSM is known to provide resistance towards differential and linear attack and, algebraic attack as it uses $\mathbf{S}_1$ and $\mathbf{S}_2$ 32-bit boxes along with three registers R1, R2, and R3. The proposed KASUMI is using a part of SNOW-3G security module which are FSM machine and two S-boxes. We have taken into use three 32-bit Shift Registers two of them providing input to FSM. Besides introducing SNOW-3G, small change in KL keys of 1st and 8th round as well as change in KI keys is also made which is discussed in further section of paper.

This paper is organized in following way: Section 2 gives briefly an overview of KASUMI. Different functions and keys used in each rounds for their respective functions are described. Section 3 contains brief description of SNOW-3G and describes about its two modules LFSR and, FSM. Functioning of initialization and keystream mode for generation of key-stream is also discussed. Section 4 discusses about our proposed work related to changes in KASUMI. Section 5 provides results of our work done. Finally conclusions are given in Section 6.

Throughout this paper, $\oplus$ stands for EX-OR operation, $\|$ stands for Concatenation operation and $\boxplus$ is addition modulo $2^{23}$.

## 2 Overview of KASUMI

KASUMI is modified form of cipher algorithm MISTY1. It is Fiestel network of 8-rounds taking input of 64-bit and giving output of 64-bit by using 128-bit key for each round. Functions of KASUMI are FI, FO and FL functions performed by them are completely different from each other and they use key values for doing their operations. The key values are KI for FI function, KO for FO function and, KL for FL function for all 8 rounds. Figure 1 provides Fiestel structure of 8-round KASUMI algorithm [2]. For odd numbered rounds, function FL comes before FO, FI functions while in even numbered rounds, FO, FI comes before FL function. Brief description of three functions of algorithm is given below along with key-value operations done in them.

### 2.1 FL Function

This function has two rounds of operation as shown in Figure 1 It takes 32-bit input and performs operation on it using 32-bit key, KL, to produce 32-bit output. Key, KL, is divided in two 16-bit values for each of the two rounds as shown below:

$$\mathbf{KL}_{\rm i} = \mathbf{KL}_{\rm i,1} \| \mathbf{KL}_{\rm i,2}$$

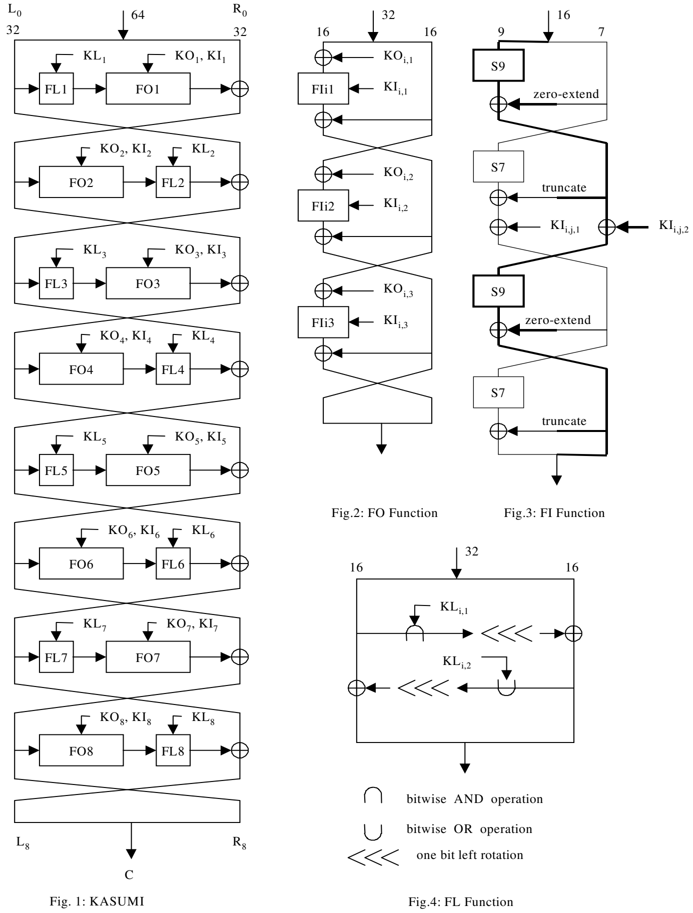


Fig. 1: KASUMI

Fig.2: FO Function

Fig.3: FI Function

Fig.4: FL Function

Figure 1: The original KASUMI algorithm, FO, FI and FL functions.

Input is divided in two 16-bit values, **L** (left) and **R** (right) as given below:

$$\begin{aligned}\mathbf{R}^{'} &= \mathbf{R} \oplus ROL(\mathbf{L} \cap \mathbf{KL}_{\rm i,1})\\ \mathbf{L}^{'} &= \mathbf{L} \oplus ROL(\mathbf{R} \cap \mathbf{KL}_{\rm i,2})\end{aligned}$$

$\mathbf{R}^{'}$ and $\mathbf{L}^{'}$ are each 16-bit values obtained after performing operations on **R** and **L**. These values obtained are then concatenated to make 32-bit output.

### 2.2 FO Function

This function has three rounds of operation as shown in Figure 1. It first takes 32-bit input, divides it in two equal 16-bit values and, performs operation on it using 48-bit key KO and, another 48-bit sub-key KI used in FI function. Figure for FO function is shown in Figure 1. The 48-bit sub-keys are subdivided into three 16-bit subkeys where:

$$\begin{aligned}\mathbf{KO}_{\rm i} &= \mathbf{KO}_{\rm i,1} \| \mathbf{KO}_{\rm i,2} \| \mathbf{KO}_{\rm i,3}\\ \mathbf{KI}_{\rm i} &= \mathbf{KI}_{\rm i,1} \| \mathbf{KI}_{\rm i,2} \| \mathbf{KI}_{\rm i,3}\end{aligned}$$

For each integer j (number of rounds within FO) with $1 \leq \mathrm{j} \leq 3$, $\mathbf{R}_\mathrm{j}$ and $\mathbf{L}_\mathrm{j}$ are given as:

$$\begin{aligned} \mathbf{R}_\mathrm{j} &= FI(\mathbf{L}_{\mathrm{j}-1} \oplus \mathbf{KO}_{\mathrm{i,j}}, \mathbf{KI}_{\mathrm{i,j}}) \oplus \mathbf{R}_{\mathrm{j}-1} \\ \mathbf{L}_\mathrm{j} &= \mathbf{R}_{\mathrm{j}-1} \end{aligned}$$

We then, return the 32-bit concatenated value obtained $(\mathbf{L}_3 \| \mathbf{R}_3)$ after completion of FO function.

### 2.3 FI Function

This function involves four rounds of operation. FI function takes 16-bits of input data and gives 16-bits of output. Data is split into 9-bit, say $\mathbf{L}_0$, which goes to left side of FI function and, 7-bit value, say $\mathbf{L}_0$, which goes to right side of FI function. Two **S**-boxes used are **S9** box used when 9-bit data is operated and, **S7** box used when 7-bit data is operated. FI-box is given in Figure 1. Truncation of 9-bit data done by removing 2 most significant bits when it is operated with 7-bit data and, 7-bit data is zero padded by adding two zeros as most significant bits when it is operated with 9-bit data. We define the following series of operations:

$$\begin{aligned} \mathbf{L}_1 &= \mathbf{R}_0, \\ \mathbf{R}_1 &= \mathbf{S9}[\mathbf{L}_0] \oplus \mathbf{ZE}(\mathbf{R}_0), \\ \mathbf{L}_2 &= \mathbf{R}_1 \oplus \mathbf{KI}_{\mathrm{i,j},2}, \\ \mathbf{R}_2 &= \mathbf{S7}[\mathbf{L}_1] \oplus \mathbf{TR}(\mathbf{R}_1) \oplus \mathbf{KI}_{\mathrm{i,j},1}, \\ \mathbf{L}_3 &= \mathbf{R}_2, \\ \mathbf{R}_3 &= \mathbf{S9}[\mathbf{L}_2] \oplus \mathbf{ZE}(\mathbf{R}_2), \\ \mathbf{L}_4 &= \mathbf{S7}[\mathbf{L}_3] \oplus \mathbf{TR}(\mathbf{R}_3), \\ \mathbf{R}_4 &= \mathbf{R}_3 \end{aligned}$$

The function returns the 16-bit value $(\mathbf{L}_4 \| \mathbf{R}_4)$. The Figure 1 shows the structure of KASUMI [2] along with three functions used in it.

## 3 Overview of SNOW-3G

SNOW-3G is word-oriented cipher algorithm which performs operations by using 128-bit key and, 128-bit Initialization Vector (IV). It has two security modules, LFSR (Linear Feedback Shift Register) and FSM [9]. The LFSR is made up of 16 $s$-blocks each having the capacity to hold 32-bit data and, the feedback is defined by a primitive polynomial over the finite field GF $(2^{32})$. The second module is FSM which consists of three registers **R1**, **R2** and **R3** and it perform functions by using two substitution boxes **S1** and **S2**. The algebraic operations used in the FSM are EX-OR and addition modulo $2^{32}$.

The SNOW-3G algorithm is working under two different modes, one is initialization mode and other is key-stream generation mode. In initialization mode, 32-bit output **F** is generated and it is discarded at the beginning. After that, the algorithm goes into key-stream mode generation and produces 32-bit output called as **F**. The **F** value generated from key-stream is used in the feedback part of LFSR as shown in Figure 2 Various 32-bit values of the key-stream are generated which are then used for encryption of different input data. Brief description about working of LFSR and, FSM is given below.

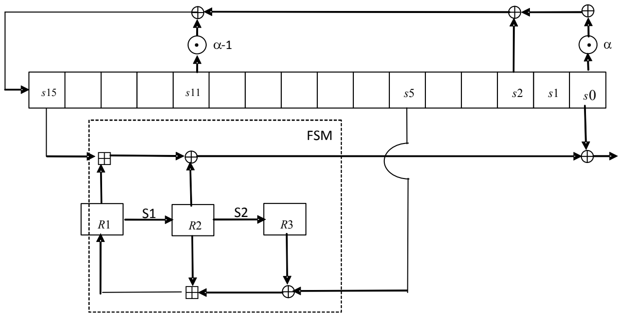


Figure 2: The SNOW-3G algorithm [9]

### 3.1 LFSR

It has 16 **s**-boxes $\mathbf{s}_0, \mathbf{s}_1, ..., \mathbf{s}_{15}$ and each box having the holding capacity of 32-bit. As seen from Figure 2, the feedback values are taken only from $\mathbf{s}_0, \mathbf{s}_2$ and $\mathbf{s}_{11}$ boxes and given as feedback to the $\mathbf{s}_{15}$ after some mathematical feedback computations that are $MUL_\alpha$ and $DIV_\alpha$.

### 3.2 FSM

The FSM takes two input values from LFSR which are $\mathbf{s}_5$ and $\mathbf{s}_{15}$ and it produces 32-bit output word **F** defined as follows:

$$\mathbf{F} = (\mathbf{s}_{15} \boxplus \mathbf{R1}) \oplus \mathbf{R2}$$

The registers (**R1**, **R2** and **R3**) are updated with the inflow of new values. Intermediate value **r** is calculated before getting in operation with register **R1** and so, it is given as:

$$\mathbf{r} = \mathbf{R2} \boxplus (\mathbf{R3} \oplus \mathbf{s}_5)$$

The values corresponding to the Registers **R3**, **R2**, **R1** are computed in the following way:

$$\begin{aligned} \mathbf{R3} &= \mathbf{S2}(\mathbf{R2}) \\ \mathbf{R2} &= \mathbf{S1}(\mathbf{R1}) \\ \mathbf{R1} &= \mathbf{r} \end{aligned}$$

## 4 Proposed Work on KASUMI

In this proposed design, the FI function is completely removed from all 8 rounds of KASUMI. But the keys that are used in FI function are retained. The three 16-bit keys of FI function used in each round of KASUMI are now used differently. For a particular round, three KI

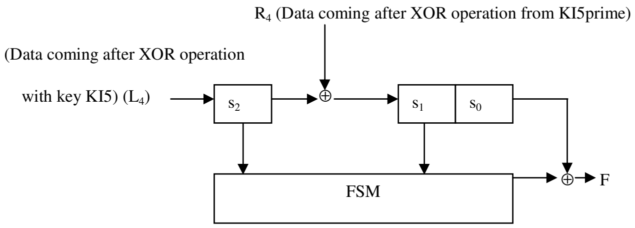

R$_4$ (Data coming after XOR operation from KI5prime)
(Data coming after XOR operation
with key KI5) (L$_4$)
S$_2$
S$_1$
S$_0$
FSM
F

Figure 3: Structure used at end of the 4th round of KASUMI

keys, $\mathbf{KI}_{i1}$, $\mathbf{KI}_{i2}$ and $\mathbf{KI}_{i3}$ are used. The first two 16-bit KI sub-keys, $\mathbf{KI}_{i1}$ and $\mathbf{KI}_{i2}$ are concatenated to form a 32-bit key. Similarly, $\mathbf{KI}_{i2}$ and $\mathbf{KI}_{i3}$ are concatenated to form another 32-bit key, i.e.,

$$\begin{aligned} \mathbf{KI}_1 &= \mathbf{KI}_{i1} \| \mathbf{KI}_{i2} \\ \mathbf{KI}_2 &= \mathbf{KI}_{i2} \| \mathbf{KI}_{i3} \end{aligned}$$

The same operation is performed with KI keys of the other rounds of KASUMI to form two separate 32-bit key. Final key values are given below and their implementation is shown in Figure 4 where 'n' denotes $n^{\text{th}}$ round value,

$$\begin{aligned} \mathbf{KI3} &= (\mathbf{KI}_{i1}[n] \| \mathbf{KI}_{i2}[n]) \oplus \\ &\qquad (\mathbf{KI}_{i1}[n+2] \| \mathbf{KI}_{i2}[n+2]) \\ \mathbf{KI3}_{\text{prime}} &= (\mathbf{KI}_{i2}[n] \| \mathbf{KI}_{i3}[n]) \oplus \\ &\qquad (\mathbf{KI}_{i2}[n+2] \| \mathbf{KI}_{i3}[n+2]) \end{aligned}$$

Similarly, other sets of 32-bit key values are defined as follows:

$$\begin{aligned} \mathbf{KI5} &= (\mathbf{KI}_{i1}[n+3] \| \mathbf{KI}_{i2}[n+3]) \oplus \\ &\qquad (\mathbf{KI}_{i1}[n+6] \| \mathbf{KI}_{i2}[n+6]) \\ \mathbf{KI5}_{\text{prime}} &= (\mathbf{KI}_{i2}[n+3] \| \mathbf{KI}_{i3}[n+3]) \oplus \\ &\qquad (\mathbf{KI}_{i2}[n+6] \| \mathbf{KI}_{i3}[n+6]). \\ \mathbf{KI7} &= (\mathbf{KI}_{i1}[n+1] \| \mathbf{KI}_{i2}[n+1]) \oplus \\ &\qquad (\mathbf{KI}_{i1}[n+4] \| \mathbf{KI}_{i2}[n+4]) \\ \mathbf{KI7}_{\text{prime}} &= (\mathbf{KI}_{i2}[n+1] \| \mathbf{KI}_{i3}[n+1]) \oplus \\ &\qquad (\mathbf{KI}_{i2}[n+4] \| \mathbf{KI}_{i3}[n+4]). \\ \mathbf{KI8} &= (\mathbf{KI}_{i1}[n+5] \| \mathbf{KI}_{i2}[n+5]) \oplus \\ &\qquad (\mathbf{KI}_{i1}[n+7] \| \mathbf{KI}_{i2}[n+7]) \\ \mathbf{KI8}_{\text{prime}} &= (\mathbf{KI}_{i2}[n+5] \| \mathbf{KI}_{i3}[n+5]) \oplus \\ &\qquad (\mathbf{KI}_{i2}[n+7] \| \mathbf{KI}_{i3}[n+7]). \end{aligned}$$

It is well known that 8 rounds are involved in KASUMI. Let's say 0th round is the first round of KASUMI. If $[n+(integer)] > 7$, then MSB will be masked and value corresponding to binary value is obtained. Then the obtained value will be considered for further operations (for example, the binary value 110 gets changed to the binary value 010 after masking MSB). Two $n^{\text{th}}$ round values are required for producing 32-bit KI key and it will be chosen as per the procedure given in the above eqnarray*s. These KI keys are used as corresponding round values.

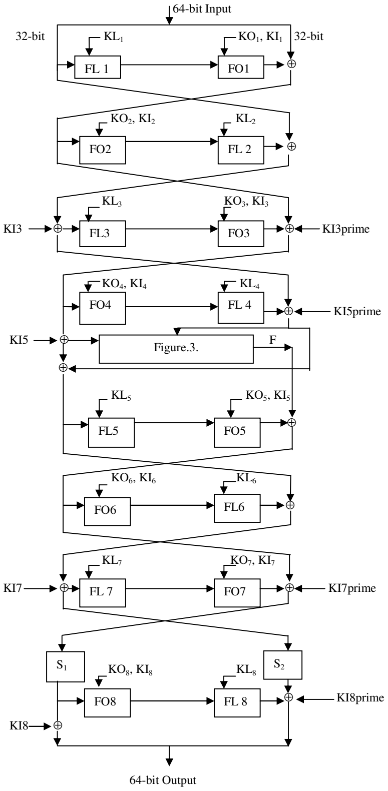


Figure 4: Proposed modified KASUMI algorithm (using combined $\mathbf{S}_1$ and $\mathbf{S}_2$ at the 7th round)

In proposed design, in addition to making changes in KI key, small change in second KL-key produced by KL function is also done in a proposed design. This change is done only in last and first rounds. As suggested in reference [15], KL key value used in the 8th round of the KASUMI is weak and it is having some bits as same as that of the KL key value that used in the second round of the KASUMI. The chances for the attack can be extended to more rounds of the algorithm if the attack has been done already in the 6th or 7th rounds of the KASUMI. The changes in the KL function is introduced only in the first and last rounds of the KASUMI. The **S3** box is used to changing KL key value of 2nd stage of FL function that present in the first and last rounds. This $S$-box is same as that used in SNOW-3G (32-bit Rijndael's $S$-box) but it performs the 16-bit data operation as the length of the key is 16-bit.

Another important change that made in the proposed design is the insertion of FSM machine along with three shift registers and each of it is having the capacity of holding 32-bits data as is shown in Figure 3 The functioning of FSM is same as in SNOW-3G. But the SNOW-3G is basically well known for key-stream generation but in our case it is used different purpose. The FSM is used only after the end of 4th round, as shown in Fig.7 and Fig.8. The output which is generated from FSM for the first time will be discarded. After discarding this first value from

FSM, next output bits generated will be used as an input for next round of the KASUMI. In purposed design the FSM is initialized for only one time rather than 32 times as done in SNOW-3G algorithm. The input to FSM is given through two shift registers that are $\mathbf{s}_1$ and $\mathbf{s}_2$. The left input value from the end of fourth round is entered into $\mathbf{s}_2$. This $\mathbf{s}_2$ is then EX-ORed with right input value that coming from the end of the fourth round. This value produced is then passed into $\mathbf{s}_1$ shift register followed by $\mathbf{s}_0$. Fig.6. illustrates the the above explanations.

The 32-bit value of $\mathbf{F}_1$ is generated from FSM and it is EX-ORed with $\mathbf{s}_0$ to produce the final output $\mathbf{F}$. This $\mathbf{F}$ is referred as a right input data ($\mathbf{R}_5$) for next round (i.e. $5^{\text{th}}$ round) of the KASUMI. The left 32-bit input data ($\mathbf{L}_5$) is r $4^{\text{th}}$ round of the KASUMI.

$$\begin{aligned}\mathbf{L}_5 &= \mathbf{L}_4 \oplus \mathbf{R}_4\\ \mathbf{R}_5 &= \mathbf{F} = (\mathbf{F}_1)(32 - bitoutoutfromFSM) \oplus \mathbf{s}_0\end{aligned}$$

Third modification which have been done in the proposed design is the insertion of $\mathbf{S}_1$ and $\mathbf{S}_2$ boxes which are based on Rijndael's $\mathbf{S}_{\mathrm{R}}$ and $\mathbf{S}_{\mathrm{Q}}$ box, respectively. The functionality of the $\mathbf{S}_1$ and $\mathbf{S}_2$ boxes are same as in SNOW-3G. But these boxes are used only after the end of the $7^{\text{th}}$ round. It considers two different cases, one is the use of both the $\mathbf{S}_1$ and $\mathbf{S}_2$ boxes while other use only the $\mathbf{S}_1$ box at the end of the 7th round which are illustrated in Fig.7 and Fig8, respectively. Since the $\mathbf{S}_2$ box is not strong as $\mathbf{S}_1$ box, it considered to use both the $\mathbf{S}_1$ and $\mathbf{S}_2$ boxes [8]. The 32-bit output data from left side is EX-ORed with '**KI8**' key value and the 32-bit right side data is EX-ORed with **KI8** prime' key to obtain the final 32-bit left and 32-bit right data. This is illustrated in Figure 4 and Figure 5 As given in [15], the last two rounds are observed to be weaker. In order to strong, the nonlinearity output can be produced with the help of **S**-box that can be leads to increase the complexity for the attacker.

The FSM is implemented after the $4^{\text{th}}$ round, but this produced more delay in the encryption process. If the FSM is implement before $4^{\text{th}}$ round, then also no difference in encryption time. Our focus was to merely propose modifications in KASUMI to make it more robust and also it consumes approximately the same time for encryption compared to standard KASUMI.

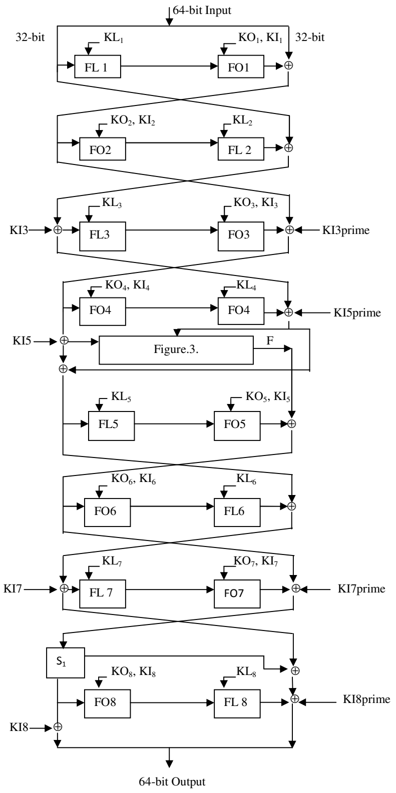


Figure 5: Proposed modified KASUMI algorithm (using only $\mathbf{S}_1$ at the $7^{\text{th}}$ round)

## 5 Simulation Results

The proposed algorithm was tested using National Institute of Standards and Technology (NIST) [14] statistical test suite where all 15 tests were passed. This suite is used for testing of randomness produced in designed/modified algorithm. The encryption time required by the modified algorithm indicates an increase by few seconds compared to original KASUMI. Comparison is done based on the 'No. of iterations' which means number of times algorithm remains in operational mode continuously before producing final output value. There are Federal Information Processing Standard (FIPS) or Standardization Administration of China (SAC) standards also used in previous works as alternate standard for testing an algorithm. In this proposed design, a large array of numbers at a single time is given to find out an approximate time taken when the number of iterations are entered. The testing was based on the software (MobaXtermlinux application) and it done on Intel (R) Core(TM) i7-4770 CPU @ 3.4GHz.

As one can be seen from the simulation results and the tables presented in this paper, the performance improvement of the proposed modified KASUMI design is compared with the conventional KASUMI in terms both the encryption speed and encryption time taken. In addition to that, it also present the simulation result for comparing the two different proposed design based on different usage of the **S**-boxes ($\mathbf{S}_1$ box alone and combined $\mathbf{S}_1$ and $\mathbf{S}_2$). The Table 1 illustrate the values for encryption speed and encryption time taken for various number of numbers entered for the proposed modified KASUMI algorithm (using combined $\mathbf{S}_1$ and $\mathbf{S}_2$ box at the $7^{\text{th}}$ round). And Table 2 illustrate the values for encryption speed and encryption time taken for various number of numbers entered for the proposed modified KASUMI algorithm (using only $\mathbf{S}_1$ box at the $7^{\text{th}}$ round). And Table 3 illustrate the values for encryption speed and encryption time taken for various number of numbers entered for the conventional KASUMI algorithm.

First, Figure 6 shows the performance comparisons of the conventional KASUMI with that of the proposed modified KASUMI (using combined $\mathbf{S}_1$ and $\mathbf{S}_2$ at the $7^{\text{th}}$ round) in terms of encryption time taken Vs number of numbers entered. The main purpose of this simulation to show the performance in terms of encryption time taken of the proposed modified KASUMI is almost identical to conventional KASUMI. Figure 7 shows the performance comparisons of the conventional KASUMI with that of the proposed modified KASUMI (using combined $\mathbf{S}_1$ and $\mathbf{S}_2$ box at the $7^{\text{th}}$ round) in terms of encryption speed Vs Number of numbers entered. As can be seen from the Figure 7, the proposed modified KASUMI speed is reduced for higher values of number of numbers entered compared to the conventional KASUMI. Figure 8 and Figure 9 also shows the similar type of performance comparisons as in Figure 6 and Figure 7 but for the case proposed modified KASUMI (using only $\mathbf{S}_1$ at the $7^{\text{th}}$ round). Again, the performance degradation of our proposed design over the conventional design is clearly observed from Figure 8 and Figure 9.

Figure 10 examines the performance comparison of the proposed modified KASUMI using combined $\mathbf{S}_1$ and $\mathbf{S}_2$ box at the $7^{\text{th}}$ round Vs proposed modified KASUMI using only $\mathbf{S}_1$ box at the $7^{\text{th}}$ round in terms of encryption time taken Vs number of numbers entered. From this result, we can observe that while only the $\mathbf{S}_1$ at the $7^{\text{th}}$ round is used, the time taken for the for encryption is reduced by 1 to 2 seconds for the cases of $5 * 10^7$ to $1 * 10^8$ iterations. Figure 11 shows the proposed modified KASUMI using combined $\mathbf{S}_1$ and $\mathbf{S}_2$ box at the $7^{\text{th}}$ round Vs proposed modified KASUMI using only $\mathbf{S}_1$ box at the $7^{\text{th}}$ round in terms of encryption speed Vs number of numbers entered. As can be seen from the Figure 11, the encryption speed of the proposed modified KASUMI using only $\mathbf{S}_1$ box at the $7^{\text{th}}$ is increased by 20 to 60 Kbps for the cases of $5 * 10^7$ to $1 * 10^8$ iterations. Though, the encryption speed of the proposed modified KASUMI using combined $\mathbf{S}_1$ and $\mathbf{S}_2$ box at the $7^{\text{th}}$ round would get reduced little, but it can be acceptable to have a delay of 1 to 2 by achieving more resistance against attacks.

Table 1: proposed modified KASUMI algorithm (using combined $\mathbf{S}_1$ and $\mathbf{S}_2$ box at the $7^{\text{th}}$ round)

| No. of iterations | Encry. time taken in sec. | Encry. speed in bps |
|---|---|---|
| $1 * 10^6$ to $2 * 10^6$ | 1 | 2000000.00 |
| $4 * 10^6$ | 2 | 2000000.00 |
| $6 * 10^6$ | 3 | 2000000.00 |
| $8 * 10^6$ | 4 | 2000000.00 |
| $1 * 10^7$ | 5 | 2000000.00 |
| $2 * 10^7$ | 11 | 1818181.82 |
| $4 * 10^7$ | 22 | 1769132.49 |
| $5 * 10^7$ | 28 | 1785714.29 |
| $6 * 10^7$ | 34 | 1764705.88 |
| $7 * 10^7$ | 40 | 1750000.00 |
| $8 * 10^7$ | 46 | 1739130.43 |
| $9 * 10^7$ | 51 | 1764705.88 |
| $1 * 10^8$ | 57 | 1754385.96 |

Table 2: proposed modified KASUMI algorithm (using only $\mathbf{S}_1$ box at the $7^{\text{th}}$ round)

| No. of iterations | Encry. time taken in sec. | Encry.n speed in bps |
|---|---|---|
| $1 * 10^6$ to $2 * 10^6$ | 1 | 2000000.00 |
| $4 * 10^6$ | 2 | 2000000.00 |
| $6 * 10^6$ | 3 | 2000000.00 |
| $8 * 10^6$ | 4 | 2000000.00 |
| $1 * 10^7$ | 5 | 2000000.00 |
| $2 * 10^7$ | 11 | 1818181.82 |
| $4 * 10^7$ | 22 | 1818181.82 |
| $5 * 10^7$ | 27 | 1821851.85 |
| $6 * 10^7$ | 33 | 1818181.82 |
| $7 * 10^7$ | 38 | 1832105.26 |
| $8 * 10^7$ | 44 | 1818181.82 |
| $9 * 10^7$ | 49 | 1826734.69 |
| $1 * 10^8$ | 55 | 1818181.82 |

## 6 Conclusion

In this paper,it is proposed the use the FSM machine, Rijndael's SR box and SQ box in the existing KASUMI algorithm to make it as more robust against attacks. Also it is proposed to modify the KL keys by using SR box. And this changes are is done in the KL keys which are corresponding to the 1st and 8th. I addition to that, it is also proposed to use KI keys in different ways by using Ex-OR and concatenation operations in $t^{\text{th}}$ and 8th as explained in this paper. From the simulation results, it observed that the proposed modified KASUMI algorithm will take more encryption time compared to standard KASUMI algorithm, but the proposed modified KASUMI algorithm reduce linear and differential attacks in KASUMI.

Table 3: Original conventional KASUMI algorithm

| No. of iterations | Encry. time taken in sec. | Encry.speed in bps |
|---|---|---|
| $1 * 10^6$ to $2 * 10^6$ | 1 | 2000000.00 |
| $4 * 10^6$ | 2 | 2000000.00 |
| $6 * 10^6$ | 3 | 2000000.00 |
| $8 * 10^6$ | 4 | 2000000.00 |
| $1 * 10^7$ | 5 | 2000000.00 |
| $2 * 10^7$ | 10 | 2000000.00 |
| $4 * 10^7$ | 20 | 2000000.00 |
| $5 * 10^7$ | 26 | 2000000.00 |
| $6 * 10^7$ | 30 | 1973076.92 |
| $7 * 10^7$ | 35 | 2000000.00 |
| $8 * 10^7$ | 40 | 2000000.00 |
| $9 * 10^7$ | 45 | 2000000.00 |
| $1 * 10^8$ | 51 | 1970784.31 |

# References

[1] 3GPP TR 55.919 V6.1.0, *Specification of the A5/3 Encryption Algorithms for GSM and ECSD, and the GEA3 Encryption Algorithm for GPRS*, Technical Report 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; Document 4: Design and evaluation Report (Release 6), 2002-2012.

[2] 3GPP TS 35.202 V14.0.0, *Specifications of the 3GPP Confidentiality and Integrity Algorithms*, Technical

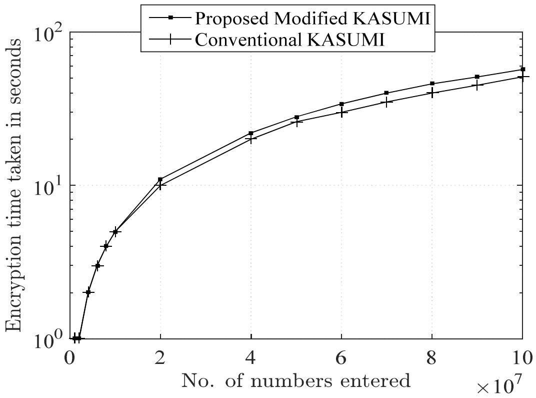


Figure 6: Performance comparisons of the conventional KASUMI Vs the proposed modified KASUMI (using combined $\mathbf{S}_1$ and $\mathbf{S}_2$ at the 7th round) in terms of encryption time taken Vs number of numbers entered.

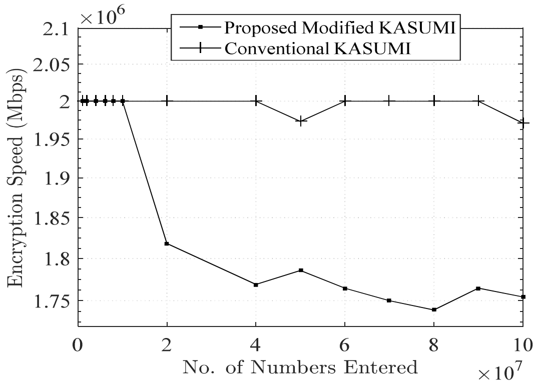


Figure 7: Performance comparisons of the conventional KASUMI Vs the proposed modified KASUMI (using combined $\mathbf{S}_1$ and $\mathbf{S}_2$ at the 7th round) in terms of encryption speed Vs number of numbers entered.

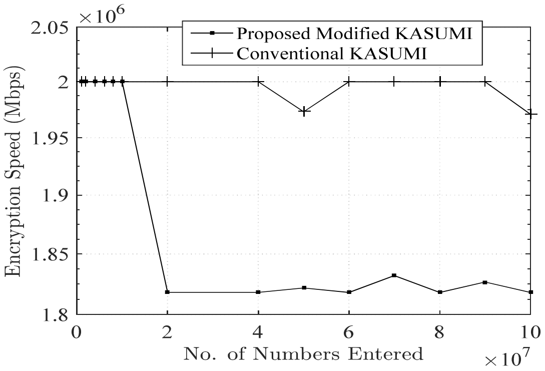


Figure 8: Performance comparisons of the conventional KASUMI Vs the proposed modified KASUMI (using only $\mathbf{S}_1$ at the 7th round) in terms of encryption time taken Vs number of numbers entered.

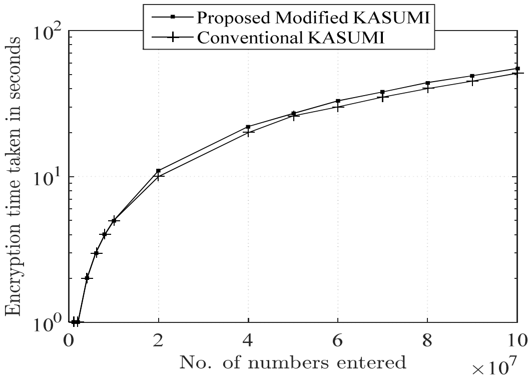


Figure 9: Performance comparisons of the conventional KASUMI Vs the proposed modified KASUMI (using only $\mathbf{S}_1$ at the 7th round) in terms of encryption speed Vs number of numbers entered.

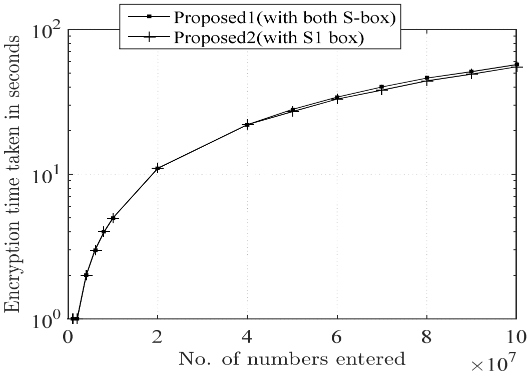


Figure 10: Proposed modified KASUMI using combined $\mathbf{S}_1$ and $\mathbf{S}_2$ box at the 7th round Vs proposed modified KASUMI using only $\mathbf{S}_1$ box at the 7th round in terms of encryption time taken Vs number of numbers entered.

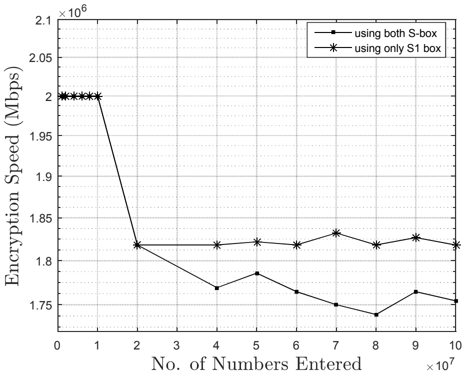


Figure 11: Proposed modified KASUMI using combined $\mathbf{S}_1$ and $\mathbf{S}_2$ box at the 7th round Vs proposed modified KASUMI using only $\mathbf{S}_1$ box at the 7th round in terms of encryption speed Vs number of numbers entered.

Report 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; Document2: KASUMI Specifications V14.0.0, 2017.

[3] E. Biham, O. Dunkelman, N. Keller, "The rectangle attack - Rectangling the serpent," in *Proceedings of the EUROCRYPT*, Lecture Notes in Computer Science, pp. 340–357, Springer, May 2001.

[4] E. Biham, O. Dunkelman, N. Keller, "Related-key rectangle attack on full kasumi," in *Proceedings of the 11th International Conference on Theory and Application of Cryptology and Information Security*, pp. 443–461, Dec. 2005.

[5] C. S. Chen, X. Yu, Y. X. Xiang, X. Li, T. Li, "An improved DPA attack on DES with forth and back random round algorithm," *International Journal of Network Security*, vol. 19, no. 2, pp. 285-294, 2017.

[6] K. Chetioui, G. Orhanou, S. El Hajji, "New protocol e-DNSSEC to enhance DNSSEC security," *International Journal of Network Security*, vol. 20, no. 1, pp. 19-24, 2018.

[7] H. Choudhury, B. Roychoudhury, D. Kr. Saikia, "Security extension for relaxed trust requirement in non-3GPP access to the EPS," *International Journal of Network Security*, vol. 18, no. 6, pp. 1041-1053, 2016.

[8] ETSI/SAGE, *Specification of the 3GPP Confidentiality and Integrity Algorithms UEA2 and UIA2*, Technical Report Document 5: Design and Evaluation Report, Ver. 1.0, 2006.

[9] ETSI/SAGE, *Specification of the 3GPP Confidentiality and Integrity Algorithms UEA2 and UIA2*, Technical Report Document 2: SNOW 3G Specification, Ver. 1.1, 2006.

[10] L. He, C. Yuan, X. Hu, and Z. Qin, "An effcient and provably secure certificateless key insulated encryption with applications to mobile internet," *International Journal of Network Security*, vol. 19, no. 6, pp. 940-949, 2017.

[11] N. Keller, O. Dunkelman, A. Shamir, "A practical-time attack on the A5/3 cryptosystem used in third generation GSM telephony," *IACR Cryptology Eprint archive*, 2010.

[12] T. Saito, "A single-key attack on 6-round KASUMI," *IACR Cryptology Eprint archive*, 2011.

[13] M. Styugin, "Establishing systems secure from research with implementation in encryption algorithms," *International Journal of Network Security*, vol. 20, no. 1, pp. 35-40, 2018.

[14] Z. Wang, X. Dong, K. Jia, J. Zhao, "Differential fault attack on kasumi cipher used in gsm telephony," *Mathematical Problems in Engineering*, vol. Article ID 251853, pp. 7, 2014.

[15] W. Yi, S. Chen, "Multidimensional zero-correlation linear cryptanalysis of the block cipher KASUMI", *IET Information Security*, vol. 10, no. 4, pp. 215-221, 2016.

[16] H. Zhu, Y. Zhang, and Y. Sun, "Provably secure multi-server privacy-protection system based on Chebyshev chaotic maps without using symmetric cryptography," *International Journal of Network Security*, vol. 18, no. 5, pp. 803-815, 2016.

# Biography

**Raja Muthalagu** received his Ph.D. in Wireless Communication from National Institute of Technology (NIT), Tiruchirappalli, India in 2014. He joined the Department of Electrical and Electronics Engineering, BITS, Pilani, Dubai Campus, in 2015, where he is currently a full Assistant Professor. His research interests include orthogonal frequency division multiplexing (OFDM), multiple-input and multiple-output (MIMO) systems, and network security.

**Subeen Jain** received hid B.E. (Honors) in Electronics and Communications engineering from BITS-Pilani Dubai campus. His areas of interest include security algorithms mainly related to mobile security and networking and, areas related to telecommunications.

# An Efficient Approach to Resolve Covert Channels

Muawia A. Elsadig[1] and Yahia A. Fadlalla[2]
(Corresponding author: Muawia A. Elsadig)

College of Computer Science and Technology, Sudan University of Science and Technology[1]
Khartoum, Sudan
Lead Consultant/Researcher, InfoSec Consulting, Hamilton, Ontario, Canada[2]
(Email: muawiasadig66@gmail.com)

(Received May 8, 2017; revised and accepted Oct. 21, 2017)

## Abstract

The competitive edge of many companies and public trust in government institutions can often depend on the security of the information held in their systems. Breaches of that security, whether deliberate or accidental, can be profoundly damaging. Therefore, security is a highly topical issue for both designers and users of computer systems. A system is said to be secure if it supports the policy of a security model in a demonstrable way. Two users, or processes operating on their behalf, are communicating indirectly or covertly in such a system if they are communicating through means that violate the interpretation of the supported security model. Research to eliminate or resolve covert communication channels is limited compared to the real, rapid, and often dangerous threats these channels continuously pose. That is due; at large; to their ingenious, inventive, and numerous scenarios. In order for any two users to establish a covert channel, they both must know one another's identity. This paper proposes a design that is based on the fact that it is impossible inside a system for any process to recognize any user, for whom other processes are invoked, in order to covertly communicate with him or her - identities of all users are hidden. Our design is sought to eliminate covert channels that are known to a system and those that are unknown and waiting to be discovered and potentially utilized illicitly. The design is sought to eliminate covert channels indifferent to the scenario they employ.

*Keywords: Covert Channels; Channel bandwidth; Cryptography; Multilevel Security; Network Protocols; Security Model; Security Policy*

## 1 Introduction

It is well known that a large number of databases contain data that needs to be categorized into different security levels to securely manage the access to this data. In addition, the database user must obtain security clearance that allows them to access a particular data level [16]. In other words, any system must ensure that the users obtain the data which they are authorized for [35]. Therefore, civilian and military government agencies are used to building up their relational database systems based on the hierarchical classification levels such as Top Secret, Secret, Confidential and Unclassified.

The Mandatory Access Control (MAC) is a common security access control model that requires users and resources to be classified and assigned security labels [14], In other words, objects and subject are labelled with different security levels [13]. MAC is an approach to restrict unauthorized users from accessing objects that hold sensitive information. It is a B1 level requirement of the Orange Book [9], and interested readers can see more about the Orange Book in [23]. When a system ororganizes its data into different classification levels and mandates with its access control utilizing the MAC model, then the system will be defined as a Multilevel Secure System (MLS). The MLS is an implementation of MAC. It is mainly developed for databases and computers that belong to highly sensitive government organizations such as the U.S. Department of Defense [10]. In Multilevel Secure Database Management Systems, users are cleared at different clearance levels (*i.e.* Top-secret, Secret, Confidential and Unclassified. While data is given different sensitivity levels (i.e. Top-secret, Secret, Confidential and Unclassified) [38]. A covert channel scenario exists if labelled data is being transferred to an unauthorized user without violating Mandatory Access Controls. The unauthorized user in this case is a user that hasn't the appropriate clearance to access this labelled data. Commonly, a covert channel is a method of exploiting a communication channel by a process in order to pass information without violating the system security policy. It is noteworthy to mention that any proposal to solve a covert channel problem should take into account the system usability and usefulness. In other words, a useful covert channel solution is a solution that doesn't diminish the usability and performance of the overt channel - a legitimate channel that is supposed to allow labelled data to be accessed by a legitimate user. This paper introduces a design that is supposed to fully eliminate any potential covert channel scenario that is intended to leak labelled information to unauthorized parties.

The rest of this paper is organized as follows: The next section gives a general overview of covert channel concepts and their development. Section 3 illustrates the typical covert channel model, which reflects the general concept of the covert channel scenario through addressing so called prisoners problem. Section 4 sheds some light onto some common fundamental concepts to give a concrete idea that facilitates a full understanding of our proposed design which is presented in Section 5. Then Section 6 outlines the evaluation criteria of the proposed design. The paper is concluded in Section 7 and subsequently the future work is presented in Section 8.

## 2 Covert Channel Development

A covert channel allows people to exchange hidden information in an undetectable way - in a way that doesn't diminish legitimate communication procedures, which complicates the detection of such kinds of threats. In addition, a covert channel can also be exploited to pass malicious activities such as Trojans and viruses *etc.* so traditional firewall systems couldn't recognize them.

It is illogical to attain full elimination of covert channels; however, it is possible to reduce them through an efficient and careful system design. Initially, the covert channel was introduced in stand-alone systems and was later extended to exploit computer network environments. Accordingly, there are two scenarios in which covert channels exist: network-based system covert channels and stand-alone system covert channels. In stand-alone covert channels, two processes of different security levels communicate with each other covertly to leak hidden information, (*i.e.* a high security level process leaks secret data to another process with low security level). In, contrast, a network covert channel exploits network protocol to carry covert messages [4].

Recently, different techniques have been developed that increasingly magnify covert channel threats. These rapidly developed techniques present security experts with a real challenge to fight against this ongoing threat. Interested readers are referred to [6], for in depth information on the rapid development of covert channels, in which a lack of covert channel countermeasures is clearly noticed. In the same context, Elsadig *et al.* introduced a valuable concept, the network covert channel triangle (DSM - Development, Switching and Microprotocol), which involves three elements that have the most direct impact in developing network covert channel technology. The DSM triangle reflects the importance of network covert channels and the security challenge that is posed [6]. Another good contribution in network steganography is done by Wendzel *et al.* [40]. They introduced a unified description that assist in categorizing network hiding methods. This valuable effort provides a unified taxonomy of hiding techniques, enables the comparison between them and offers an evaluation framework to assess hiding methods novelty.

It is noteworthy to mention that covert channels are not always used to threaten information security. Some practical uses for covert channels were presented, [11, 12, 34, 36] such as use of covert channels by network administrators to distribute secret information among the network users, use of covert channels to secure authentication processes, *etc.* As an example of using covert channels legitimately, Singh *et al.* proposed an approach that uses covert communication to enhance Vehicular Adhoc Networks (VANETs) security [34]. VANETs have become a hot area of research as they pose many security challenges [8]. Moreover, any suggested security solution for VANETs has to ensure an acceptable overhead that doesn't affect their protocols performance [7]. Singh's model exploits a storage covert channel approach to convey secure data while the transmission of unimportant data can be done through the system overt channel. However, this trend, the trend of using covert channels for useful purposes, doesn't change the fact that covert channels are ongoing, devolved and a dangerous threat, as covert channel techniques are developed according to the rapid development of computer system and network protocols.

## 3 Typical Covert Channel Model

This section introduces the typical concept of covert channels, which is illustrated through the common scenario that is known as the prisoners problem [33].

Alice and Bob are prisoners who wish to communicate to each other, keeping in mind the end goal is to arrange their escape. The possible communication channel that can be used to speak to each other is under monitoring by so called Wendy (Warden). Wendy is dedicated to watch the communication between them. When Wendy catches any suspicious information, Alice and Bob will be moved into solitary confinement and that results in killing any hope for them to exchange any piece of information. Therefore, in order to avoid this situation, Alice and Bob should find a secret channel to exchange their messages in a manner such that Wendy cannot be able to detect them. In this case, this channel is known as a covert channel which allows two communication parties to exchange their secret information covertly without being detected by a monitoring system.

When Alice and Bob are establishing their communication via networked computers, then the scenario is representing another type of covert channel which is know as a network covert channel [5].

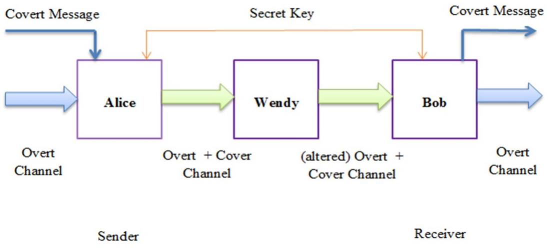


Figure 1: Typical covert channel model

# 4 Fundamental Concepts

This section has introduced fundamental security concepts that are required to understand our proposed design which is illustrated in Section 5. These concepts encompass Authentication, Multilevel Security and Cryptosystems. In addition, this section sheds some light on a storage covert channel that threatens multilevel security systems.

## 4.1 Authentication

Authentication, authorization and accounting (AAA) is a framework to apply policies, control and manage access to resources and examine the usage of these resources [42]. Authentication is a process of verifying the identity of a subject so as to ensure no subject can gain access to an information resource (object) unless their identity is verified. The next step after verification is Authorization which determines what a subject accesses after authentication. Authorization is classed into two types, Course Authorization and Fine Authorization. As an example, having access to a payroll system is a Course Authorization while determining which function is allowed to be accessed after getting into a payroll system is a Fine Authorization.

Accountability is concerned with recording what a subject does, when it does it and where it does it.

These combined processes (authentication, authorization and accounting) are considered vital for effective network management and security. Moreover, authentication is considered the foundation of all security systems [28].

An effective and successful authentication procedure is heavily based on the efficiency of the verification procedure that is being used. The trusted computing base (TCB) is the mechanism that is used to perform the authentication procedures as a part of its whole security mission. It maintains enforcement of the security policy of a given system. The careful design and implementation of a TCB for any system is paramount to its overall security. System users are considered outside the system boundaries, so the TCB doesn't deal directly with the system users. The TCB is dealing with the processes inside the system that act on behalf of the real system users. Normally, inside a system boundary, the system creates processes to represent users, so the processes request and consume the system resources on behalf of the associated users.

In our proposed design, the Extended TCB Theater (ETCB) that is illustrated in Section 5, a user is assumed to have neither stolen nor forged the identity of a legitimate user that is given to provide authorized access to the TCB system.

Commonly, there are three types of authentication mechanisms: password-based [1, 15, 20, 25, 27, 39], token-based and biometrics-based [17, 21, 22, 43]. Password-based is the most popular one and it uses something a user knows (*i.e.* password, personal identity number (PIN) *etc.*). While token-based uses something a user possesses (*i.e.* smart cards, physical keys *etc.*). The last type is biometrics-based which is known as something a user is and does (*i.e.* fingerprint matching, iris scanning, voice recognition *etc.*) [28].

Pham *et al.* discussed the shortcomings of the aforementioned authentication mechanisms and recommended the new mechanism that has emerged recently, which is known as electroencephalography (EEG) [28]. EEG is a type of biometrics and combines the advantages of password-based and biometrics-based authentication mechanisms without their shortcomings. Accordingly, Pham *et al.* proposed an authentication system based on EEG signals to be used in multilevel security systems.

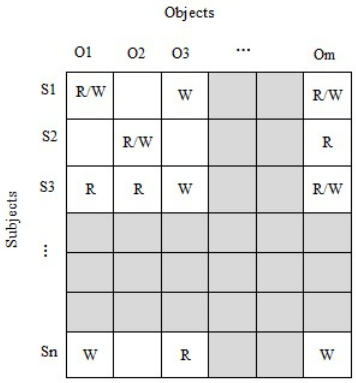
## 4.2 Multilevel Security

The development of database systems and their utilization by various people with various interests makes it crucial to design completely secured database systems [31].

Commonly, any secure Database Management System (DBMS) uses some rules to control access to its data. The setting of these rules is defined as the system security policy, in which any company enforces its security policy to allow only authorized users to access what they are authorized for. The access process encompasses subject and object. The subject is an active process request to access an object, while the object is a passive entity such as information resources. The rules that control the access process can be abstracted by what is called an access control matrix. Each element of this matrix represents an authorized mode of access as illustrated in Figure 2.

Each subject is assigned a clearance, and each object a classification [18]. Clearances and classifications are formed into so called access classes. Each access class involves two parameters (a hierarchical and a group of nonhierarchical categories). Top secret and Secret are examples of hierarchical components while Navy, Military is an example of a group of nonhierarchical categories [9].

MAC, which is mainly based on the Bell-LaPadula model [10, 26] allows a labelled object to be transferred to a subject if and only if the object access class is dominated by that of the subject. The TCB mechanism is closely similar to MAC concepts. It is the set of security components in which to enforce a system security policy.


Objects
Subjects

| | O1 | O2 | O3 | … | | Om |
|---|---|---|---|---|---|---|
| S1 | R/W | | W | | | R/W |
| S2 | | R/W | | | | R |
| S3 | R | R | W | | | R/W |
| | | | | | | |
| ⋮ | | | | | | |
| | | | | | | |
| Sn | W | | R | | | W |

Figure 2: An example of TCB model

It assigns labels to objects, users and processes. The TCB should be carefully designed and protected to enforce access controls effectively.

Fadlalla asserted that covert channels still exist in many systems despite the fact that many effective criteria were presented to disallow any attempt at covert communication between two processes. These criteria include CTCPEC, TCSEC and the Bell-LaPadula model *etc.* [9].

As mentioned, in Multilevel Secure Database Management Systems, users are cleared at different clearance levels (*i.e.* Top-secret, Secret, Confidential and Unclassified. Data is given different sensitivity levels (*i.e.* Top-secret, Secret, Confidential and Unclassified) [38]. To illustrate the concept of a security level classifications approach an example is given below.

**Example 1.** *Imagine a database system with five user classifications as follows: Top Secret, Secret, Confidential and Unclassified. According to these classifications, the database would be classified into the same security level classifications as shown in Figure 3.*

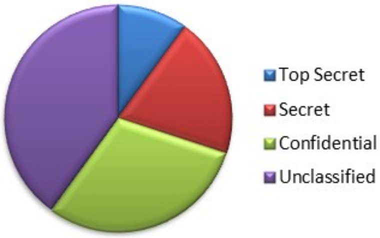


Figure 3: Database security level classification

A user can access the database according to two parameters (their security level and database security level). These classifications are hierarchical in nature. In other words, an unclassified user can only see the unclassified database while the Top-Secret user can see all other database security levels (Top Secret, Secret, Confidential and Unclassified). Confidential can see Confidential and Unclassified security levels and so on as illustrated in Figure 4.

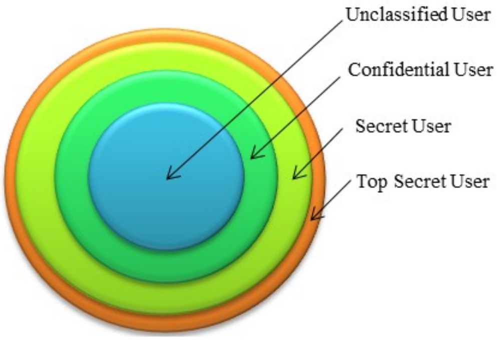


Figure 4: User clearance levels

Generally, it is commonly known that any system enforces the two rules: (1) simple security property and (2) star property providing multilevel security [30, 37]. Multilevel Security is expected to help in the decision to access a domain with security classifications, a domain where its information has security levels (*i.e.* Confidential, Secret *etc.*) [29]. Based on some specific requirements and definitions, the National Security Agency (NSA) has defined various levels of security for computer and network systems. These definitions were stated on the evaluation criteria: Trusted Computer System Evaluation Criteria (TCSEC) and Trusted Network Interpretation of the Trusted Computer System Evaluation Criteria. Accordingly, six hierarchical ratings levels are defined: A1, B3, B2, B1, C2 and C1. A1 is the most secure level while C1 is the least secure. The requirements for MLS are presented in Division B, which includes three sub-levels (B3, B2 and B1). Generally, we summarize that MLS refers to a system in which at least two classification levels of information or more are processed at the same time, and not all users are cleared for all levels of information [41].

## 4.3 Covert Channel

When dealing with a data object, as one subject writes to and another reads from, in this case the data object is considered an overt channel because this entity (data object) is mainly intended to hold data (*i.e.* files, buffers *etc.*). When subjects exchange information through a non-data object entity, then a covert channel exists. That means covert channels use a non-data object (an object that isn't used to hold data) to send data from one subject to another. There are two types of covert channel, storage and timing covert channels [3]. In a storage channel, the unauthorized information is being exchanged through a non-data object (*i.e.* one high process writes information to a non-data object, and then a low process reads that information). This scenario is legally accepted because the two processes can write to or read from a non-data object. However, from a different perspective, the scenario is violating the system security policy because the information is exchanged between a high process and low process, which is prohibited. Therefore, it is noteworthy to say that detection of covert channels is highly difficult since the covert channel doesn't diminish the system legal operations. Timing covert channels occur by means of one process that can modulate signals or secret messages to another process that is not authorized to gain such information. After modulation of the secret message by the sender process, then the intended process (receiver) observes and decodes the secret message [4].

Commonly, most covert channel detection methods rely heavily on identifying illegal flow of information in source code or top-level specifications. As a matter of fact, some excluded resources in the interpretation phase of any security model represent fruitful areas for developing covert channels. These resources include design detail, implementation detail *etc.*

### 4.4 Cryptosystems

A cryptographic system has five components:

- A plaintext message.
- A ciphertext message.
- A key space.
- An enciphering transformation: the transformation of a plaintext into ciphertext.
- A deciphering transformation: the transformation of a ciphertext into plaintext.

Commonly, there are two rules that each cryptographic system should adhere to, these rules are:

1) The two main operations of any cryptographic system, enciphering and deciphering transformations, have to work for all keys of the key space.

2) Any cryptographic system must depend on maintaining the security of the keys not on the secrecy of the encryption/decryption algorithms.

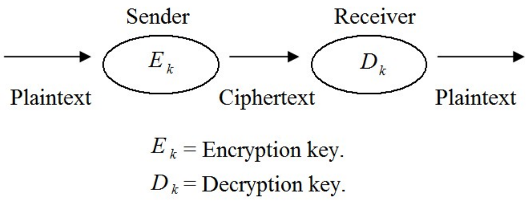


Figure 5: Diagram of a cryptographic system [9]

The Cryptosystems are categorized into two classes: symmetric cryptosystems and asymmetric cryptosystems. For symmetric cryptosystems, a same key is used for both operations, encryption and decryption, and this key should be kept secret. In asymmetric cryptosystems, the encryption key and the decryption key are different. So, the decryption key is kept secret while the encryption key is made public [9].

## 5 Extended TCB Theater

The extended TCB Theater is a security design that aims to eliminate any potential covert channel that is supposed to pass classified information to unauthorized users covertly. The design is an extended version of our design, TCB Theater presented in [9]. The enhanced version, called Extended TCB Theater (ETCBT) is a combination of two approaches, authentication and cryptography. This is to ensure that the identity of a user is kept secret from other users. So, covert processes - that work on behalf of users - would never know each other and this is expected to fully eliminate any potential covert channel that intends to leak confidential information to unauthorized parties. The ETCBT is controlling the request that is made by a subject to access an object; the subject would be any process that wishes to utilize object resources. As a matter of fact, some processes are legitimate processes that work on behalf of legitimate users, while some processes can be malicious (*i.e.* processes infected by Trojan horses or by any other means of malicious activity) that work on behalf of covert users -legitimate users that communicate with each other covertly. Therefore, this design is mainly built to break the connection between covert processes through hiding the true users' identities.

### 5.1 ETCBT Assumptions

There are three assumptions that our proposed system heavily relies on:

1) The design assumes that a user can't be seen on the ETCBT system after assigning a process to work on their behalf.

2) The design assumes that more than two processes are running at the same time. It is commonly expected that most times this assumption is valid, but in case this assumption is not valid, the ETCBT introduces a new approach to ensure the validity of this assumption. The ETCBT ensures the presence of this assumption by introducing a confusing process in the case that only two processes are running at a given time. If two covert processes try to infer that they are the only communicated processes at a given time, the confusing process defeats that.

3) The design assumes that a trusted, secure and direct communication path should be attained between a user and the ETCBT system. To ensure the two communication parties are authenticated by each other and their exchanged data is secured.

## 5.2 ETCBT Components and Description

The ETCBT system consists of four essential components listed below and illustrated in Figure 6.

1) Theater Office.
2) TCB.
3) Confusing Process Generator.
4) The user.

### 5.2.1 Theater Office

This unit is the same as the "Box Office" unit illustrated in the previous model [5]. A user is required to obtain "pass-information" to be able to access the TCB. This pass-information is provided by the Theater Office. When the Theater Office provides the user with the pass-information, it sends encrypted information about the user to the TCB at the same time. The detail of how this unit works is discussed in the next section.

### 5.2.2 TCB

This unit consists of four components: the Database Trusted Guard, the Reference Monitor, the Authenticator, and the Secure Data Block. The Secure Data Block has a direct communication channel with the Theater Office, whereas the rest of the components are logically connected to the Theater Office.

### 5.2.3 Confusing Process Generator

This unit works only under special conditions, when there are only two processes running at the same time. When this condition is met, the Reference Monitor sends a request to the Confusing Process Generator, and then the latter unit generates a confusing process. The Reference Monitor receives the confusing process and prompts the Database Trusted Guard to give the confusing process access to the same object that is being used by the aforementioned two processes. In this case, the two processes will be confused about which process has performed the last operation on the shared object. Therefore, this breaks up any potential covert communication, if we assume that the two processes have an intention to leak classified information covertly.

### 5.2.4 The User

This unit represents a user who wishes to access the TCB. In other words, the user who wishes to access an object in a classified database (Multilevel Security System).

## 5.3 How ETCBT Works

A user is granted an access class, a list of user allowed processes to be executed on the TCB, and a permanent login identifier. These are permanent privileges unless a change is required by the system security policy. The Theater Office is responsible for providing the aforementioned privileges. Then the user uses his log on identifier and encryption key/keys to only communicate with the Theater Office.

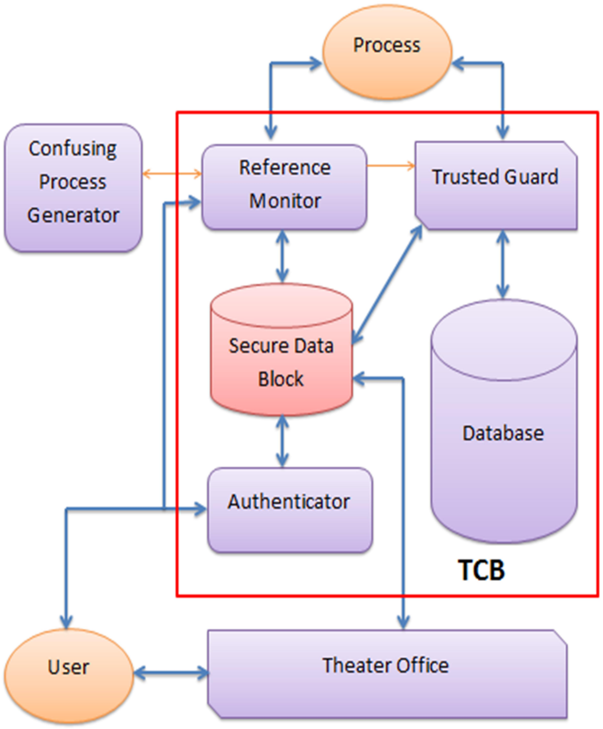


Figure 6: ETCBT design diagram

When a user needs to access an object, they send their plain login identifier to the Theater Office - unencrypted identifier. That means the user must be identified by the Theater Office. Then the Theater Office does the following scenarios simultaneously:

1) Send encrypted "pass information" to the user, which includes a temporary login identifier along with a new user encryption key to be used to communicate with the TCB.

2) Send a "ticket" to the TCB through the Secure Data Block unit, which is connected to all other TCB's components. The ticket is an encrypted packet that consists of the following parts:

   - The "pass-information" which is encrypted with a shared key between the TCB Authenticator and the Theater Office.
   - Users' login identifiers and a list of their allowed processes to be executed on the TCB. This information is encrypted with a key. This key is shared between the Theater Office and the TCB Reference Monitor.
   - Users' login identifiers and their access classes are both encrypted with a key. This key is

shared between the Theater Office and the TCB Trusted Guard.

These scenarios ensure separate communication between the Theater Office and the TCB's components. This design relies on cryptography to secure the communication between its components as described above. So, each component has a logical separate secure path to the other components, which ensures the separation between the users and the processes are working on their behalf. In return, this will successfully prevent any attempts to establish any covert communication.

## 6 Approach Evaluation Criteria

Our evaluation is based on implementations of real scenarios that are expected to reflect realistic results which prove our approach's expectations. Our design is sought to completely stop any two users - usually with different access levels - to benefit from their covert communication. To initiate a covert channel, two users who are presented inside the systems by two different processes, have to identify each other. One of the two users is the sender of the illicit information and the other is the receiver. Our design hides all users' identities from the TCB. Our design does not prevent; nor does it worry about; illicit signaling between processes, rather, it prevents illicit signaled information to reach its intended receiver. Hiding users' identities prevents any potential covert communication between any two users. Our approach establishes a trusted path service which ensures users direct and uncompromised secure communication with TCB; the Trusted Guard is placed between a process and the database. Each TCB component is forced to know the least information about users necessary to serve them; and is certainly not aware of their identities. Therefore, our evaluation is dependent on the cryptosystem used; we assume it to be effective in terms of its efficiency and usability. The evaluation procedure may consider either symmetric or asymmetric cryptosystem. Taking into account the development in cryptosystem techniques; *e.g.*, Elliptic Curve Cryptography (ECC) challenges RSA. ECC attains better security in case of using a smaller key compared to RSA [19]; consequently; it decreases the processing overhead and that is one important criterion when it comes to evaluating a cryptosystem's performance. Any suitable supported programming language may be used to implement our design, *e.g.*, C++. while this section demonstrates our approach's evaluation criteria, the implementation and evaluation results are left as future work, as indicated in Section 8.

## 7 Conclusion and Discussion

Based on hiding all user identities from the TCB and thus from the processes that are operating on their behalf, an efficient design has been proposed by this paper to prevent any two users with different classes (*i.e.* Top Secret level and Confidential level) from initiating a covert communication channel. As a matter of fact, the design is not preventing two processes from being communicated covertly. However, the process that passes a covert message never knows to which user the receiving process belongs. So, at the level of processes communication, the processes will never identify each other on the level of the true users that they are operating on their behalf. This design is expected to totally overcome the covert channel problem in multilevel security systems even if there are only two processes running at the same time. When there are only two active processes, the design introduces a so called "confusing process" to confuse the two active processes and thus prevents them from having a successful covert communication.

With the rapid development of computer networks, many enterprises build up their local network. Defiantly, these networked computers store a great amount of data or information and the users continuously exchange information over the network. This phenomenon has boosted the need for secure communication [24]. In this sense, some recent woks have been presented such as those at [2, 32].

An extension of the Bell-Lapadula model to suit local area networks, called the L-BLP model, has been proposed in [32]. This is where each host is assigned a security level, and then according to these levels the monitoring device controls all communications between these hosts by applying the L-BLP security policy. The L-BLP defines system topologies and builds up new state transition rules to control the flow of information securely. This represents a multilevel security local area network MSL. The L-BLP model allows a low-level host to send information to another with a high level and prevent a high-level host sending information to a low-level host. However, this scenario isn't valid for a TCP/IP network, which requires a Knowledge message (ACK) to ensure packet delivery. When a packet is sent by a low host to a high one, the low host waits to receive the ACK message, but the high host is prevented by L-BLP model policy from sending any information to the low host. Solving this problem (by allowing the high host to send an ACK message to the low one) causes a covert channel scenario as a high host can exploit the ACK message to pass information to a low host which is against security rules [24]. Our design is expected to solve this problem as it is mainly based on the separation between a user and the process that operates on their behalf. If a user (a Host in terms of network) acknowledges another user (another Host), this acknowledgment is within the level of the users. Therefore, the processes are not being involved as per our design policy and thus prevent any attempt to handle any covert communication.

# 8 Future Work

Our future work would be focused in verifying our design through a real or simulated environment to give realistic results. In addition, this design is expected to be extended to suit a network environment as an approach to fix network-based covert channels. The network covert channel techniques are dramatically increased, developed and pose a real challenge.

# References

[1] A. Asimi, Y. Asimi, A. Abdellah, and Y. Sadqi, "Strong zero-knowledge authentication based on virtual passwords," *International Journal Network Security*, vol. 18, no. 4, pp. 601–616, 2016.

[2] M. Casado, M. J. Freedman, J. Pettit, J. Luo, N. McKeown, and S. Shenker, "Ethane: Taking control of the enterprise," in *ACM SIGCOMM Computer Communication Review*, vol. 37, pp. 1–12, 2007.

[3] P. Dong, H. Qian, Z. Lu, and S. Lan, "A network covert channel based on packet classification," *International Journal Network Security*, vol. 14, no. 2, pp. 109–116, 2012.

[4] M. A. Elsadig and Y. A. Fadlalla, "Survey on covert storage channel in computer network protocols: Detection and mitigation techniques," in *The International Conference on Advances in Information Processing and Communication Technology (IPCT'16)*, pp. 79–85, 2016.

[5] M. A. Elsadig and Y. A. Fadlalla, "Survey on covert storage channel in computer network protocols: Detection and mitigation techniques," *International Journal of Advances in Computer Networks and Its Security*, vol. 6, no. 3, pp. 11–17, 2016.

[6] M. A. Elsadig and Y. A. Fadlalla, "Network protocol covert channels: Countermeasures techniques," in *9th IEEE-GCC Conference and Exhibition (GCCCE'17)*, pp. 706–714, 2017.

[7] M. A. Elsadig and Y. A. Fadlalla, "Performance analysis of popular manet protocols," in *9th IEEE-GCC Conference and Exhibition (GCCCE'17)*, pp. 1085–1089, 2017.

[8] M. A. Elsadig and Y. A. Fadlalla, "Vanets security issues and challenges: A survey," *Indian Journal of Science and Technology*, vol. 9, no. 28, 2016.

[9] Y. Fadlalla, *Approaches to Resolving Covert Storage Channels in Multilevel Secure Systems*, 1997. ISBN:0-612-23861-X.

[10] O. S. Faragallah, E. M. El-Rabaie, F. E. A. El-Samie, A. I. Sallam, and H. S. El-Sayed, *Multilevel Security for Relational Databases*, pp.304, 2014.

[11] T. S. A. Fatayer, "Generated un-detectability covert channel algorithm for dynamic secure communication using encryption and authentication," in *Palestinian International Conference on Information and Communication Technology (PICICT'17)*, pp. 6–9, May 2017.

[12] T. S. Fatayer and K. A. A. Timraz, "Mlscpc: Multilevel security using covert channel to achieve privacy through cloud computing," in *World Symposium on Computer Networks and Information Security (WSCNIS'15)*, pp. 1–6, 2015.

[13] D. Gollmann, "Computer security," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 5, pp. 544–554, 2010.

[14] Y. Huang and X. Ma, "A security model based on database system," in *International Conference on Electrical and Control Engineering (ICECE'10)*, pp. 4954–4957, 2010.

[15] M. S. Hwang, S. K. Chong, and T. Y. Chen, "Dos-resistant id-based password authentication scheme using smart cards," *Journal of Systems and Software*, vol. 83, no. 1, pp. 163–172, 2010.

[16] S. Jajodia and R. Sandhu, "Toward a multilevel secure relational data model," *ACM SIGMOD Record*, vol. 20, no. 2, pp. 50–59, 1991.

[17] O. Kaiwartya, M. Prasad, S. Prakash, D. Samadhiya, A. H. Abdullah, and A. O. A. Rahman, "An investigation on biometric internet security.," *International Journal Network Security*, vol. 19, no. 2, pp. 167–176, 2017.

[18] N. Kaur, R. Singh, M. Misra, and A. K. Sarje, "Concurrency control for multilevel secure databases," *International Journal Network Security*, vol. 9, no. 1, pp. 70–81, 2009.

[19] A. V. N. Krishna, A. H. Narayana, and S. K. Murthy, "A hybrid digital signature scheme on dependable and secure data," *International Journal of Electronics and Information Engineering*, vol. 6, no. 2, pp. 87–93, 2017.

[20] C. C. Lee, C. H. Liu, and M. S. Hwang, "Guessing attacks on strong-password authentication protocol," *International Journal Network Security*, vol. 15, no. 1, pp. 64–67, 2013.

[21] C. T. Li and M. S. Hwang, "An efficient biometrics-based remote user authentication scheme using smart cards," *Journal of Network and computer applications*, vol. 33, no. 1, pp. 1–5, 2010.

[22] C. T. Li and M. S. Hwang, "An online biometrics-based secret sharing scheme for multiparty cryptosystem using smart cards," *Network*, vol. 3, no. 4, pp. 5, 2010.

[23] S. B. Lipner, "The birth and death of the orange book," *IEEE Annals of the History of Computing*, vol. 37, no. 2, pp. 19–31, 2015.

[24] X. Liu, H. Xue, X. Feng, and Y. Dai, "Design of the multi-level security network switch system which restricts covert channel," in *IEEE 3rd International Conference on Communication Software and Networks (ICCSN'11)*, pp. 233–237, 2011.

[25] Y. J. Liu, C. C. Chang, and S. C. Chang, "An efficient and secure smart card based password authentication scheme," *International Journal of Network Security*, vol. 19, no. 1, pp. 1-10, 2017.

[26] W. Meanrach and S. Chittayasothorn, "A bitemporal multilevel secure database system," in *Africon*, pp. 1–7, 2007.

[27] J. Moon, D. Lee, J. Jung, and D. Won, "Improvement of efficient and secure smart card based password authentication scheme," *International Journal Network Security*, vol. 19, no. 6, pp. 1053–1061, 2017.

[28] T. Pham, W. Ma, D. Tran, P. Nguyen, and D. Phung, "Eeg-based user authentication in multilevel security systems," in *International Conference on Advanced Data Mining and Applications*, pp. 513–523, 2013.

[29] F. S. Prass, L. M. Fontoura, and O. M. D. Santos, *A Framework Based on Security Patterns for Transformations*, pp. 319-331.

[30] P. Sapra and S. Kumar, "Development of a concurrency control technique for multilevel secure databases," in *International Conference on Optimization, Reliabilty, and Information Technology (ICROIT'14)*, pp. 111–115, 2014.

[31] P. Sapra, S. Kumar, and R. K. Rathy, "Performance analysis of decomposition techniques in multilevel secure relational database systems," in *Proceedings of the Second International Conference on Computational Science, Engineering and Information Technology*, pp. 544–549, 2012.

[32] T. G. Si, Y. X. Zhang, and Y. Q. Dai, "L-blp security model in local area network," *Dianzi Xuebao (Acta Electronica Sinica)*, vol. 35, no. 5, pp. 1005–1008, 2007.

[33] G. J. Simmons, "The prisoners' problem and the subliminal channel," in *Advances in Cryptology*, pp. 51–67, 1984.

[34] A. Singh and K. Manchanda, "Establishment of bit selective mode storage covert channel in vanets," in *IEEE International Conference on Computational Intelligence and Computing Research (ICCIC'15)*, pp. 1–4, 2015.

[35] P. D. Stachour and B. Thuraisingham, "Design of ldv: A multilevel secure relational database management system," *IEEE Transactions on Knowledge and Data Engineering*, vol. 2, no. 2, pp. 190–209, 1990.

[36] Y. Sun, X. Guan, and T. Liu, "A new method for authentication based on covert channel," in *Proceedings of the 8th IFIP International Conference on Network and Parallel Computing (NPC'11)*, pp. 160–165, 2011.

[37] M. Thiyagarajan, C. Raveendra, and V. Thiagarasu, "Web service authentication and multilevel security," *Indian Journal of Science and Technology*, vol. 8, no. 15, 2015.

[38] B. Thuraisingham, "Multilevel secure database management system," in *Encyclopedia of Database Systems*, pp. 1789–1792, 2009.

[39] C. S. Tsai, C. C. Lee, and M. S. Hwang, "Password authentication schemes: Current status and key issues," *International Journal Network Security*, vol. 3, no. 2, pp. 101–115, 2006.

[40] S. Wendzel, W. Mazurczyk, and S. Zander, "Unified description for network information hiding methods," *Computer Science*, vol. 22, no. 11, pp. 1456–1486, 2016.

[41] T. C. Williams, *Multi-Level Security Network System*, June 27, 2006. US Patent 7,069,437.

[42] C. C. Yang, C. W. Lee, T. Y. Chang, M. S. Hwang, "A solution to mobile IP registration for AAA", *Lecture Notes in Computer Science*, vol. 2524, pp. 329–337, 2003.

[43] M. Zhang, B. Yang, W. Zhang, and T. Takagi, "Multibiometric based secure encryption, authentication scheme with fuzzy extractor," *International Journal Network Security*, vol. 12, no. 1, pp. 50–57, 2011.

# Biography

**Muawia A. Elsadig** obtained his MSc degree in Computer Network and Bachelor degree in Computer Engineering. His research interests lie in the area of Information Security, Network Security, Cybersecurity, Wireless Sensor Networks, Network Protocols, and Information Extraction; ranging from theory to design to implementation. Elsadig worked at different accredited international universities and has many publications at recognized international journals and conferences.

**Dr. Fadlalla** is the lead researcher and scientist at InfoSec Consulting in Ontario, Canada. He earned B.Sc. in Computer Science from the State University of New York at Utica, N.Y., USA; M.Sc. and Ph.D. in Computer Science from the University of New Brunswick at Fredericton, New Brunswick, Canada in 1985, 1992, and 1996, respectively. Professor Fadlalla taught at the University of New Brunswick at Fredericton and at the State University of New York at Albany. He is frequently invited nationally and globally to speak on contemporary information security issues. He appears and was featured in Canadian television and newspapers as an information and national security expert; likewise, he was featured in newspapers in France, Morocco, Sudan, and England. Professor Fadlalla is and was an information security consultant to different private companies and government agencies in Canada. He has extensive publications record in the areas of Computer Security, Information Assurance, Cryptography, and Cyber Security. He is a member of numerous professional associations and societies worldwise.

# A Secure and Efficient Ciphertext Encryption Scheme Based on Attribute and Support Strategy Dynamic Update Via Hybrid Encryption Method

Yanhua Wang[1], Yaqiu Liu[1] and Kun Wang[2]
(Corresponding author: Kun Wang)

College of Information and Computer Engineering, Northeast Forestry University[1]
Harbin 150040, China
College of Information and Electronic Technology, Jiamusi University[2]
148 Xuefu Road, Xiangyang District, Jiamusi City, Heilongjiang Province 154007, China
(Email: wk_116@126.com)

*(Received Sept. 21, 2017; revised and accepted Dec. 8, 2017)*

## Abstract

Traditional ciphertext encryption scheme easily leaks individual data privacy information. Therefore, this paper proposes a secure and efficient ciphertext encryption scheme based on attribute and support strategy dynamic update via hybrid encryption method. The asymmetric encryption algorithm (ASE) is adopted to encrypt the data and the attribute-based encryption algorithm (ABEA) to encrypt the sharing key. The results are stored in cloud with the form of ciphertext, which ensures the security of stored data. Meanwhile, it reduces the number of ciphertext variable parameters, this way saves the energy consumption of physiological sensors. Access strategy of dynamic update technology is based on attribute encryption algorithm which provides an extension features, it is the effective way to achieve fine-grained access control. Dynamic update technology allows data owner to dynamically update ciphertext access information in the clouds. Through comparing the proposed scheme with the state-of-the-art schemes in terms of performance, the experimental results show that the proposed scheme can ensure the security of the stored data. In addition, it has a low communication and computation overhead, which proves that our new scheme has higher performance than other encryption schemes.

*Keywords: Asymmetric Encryption; Ciphertext Encryption Scheme; Data Privacy; Energy Consumption; Support Strategy Dynamic Update*

## 1 Introduction

In recent years, how to design an efficient and secure public key encryption scheme is a hot issue. Ciphertext encryption system is a special public key encryption system [1,16]. In a ciphertext-based encryption scheme, any string (such as an E-mail address), address, ID card number *etc.*, can be used as a legitimate public key. The private key of user needs to be generated by the private key generation center. Due to the widely use and flexible application of ciphertext encryption, researchers propose a variety of secure ciphertext encryption schemes, but the common drawback is that the safety certificate is not compact, or difficult to solve the Diffie-Hellman (BDH) problem [6,9,14]. For data storage security privacy protection, Jivanyan [4] proposed to send the user key of encrypted the data through the security channel to ensure the stored data privacy security. To reduce the transmission overhead,

Li [5] proposed a new encrypted storage scheme that used the user's private key to encrypt the data and store it on the server side. The user access server could decrypt the data directly. But the above two programs need to maintain a large number of keys, and it is not easy to control. Zhou [18] proposed a new construction of Ciphertext Policy Attribute-Based Encryption (CP-ABE), named Privacy Preserving Constant CP-ABE (denoted as PP-CP-ABE) that significantly reduced the ciphertext to a constant size with any given number of attributes. Furthermore, PP-CP-ABE leveraged a hidden policy construction such that the recipients' privacy was preserved efficiently. As far as we know, PP-CP-ABE was the first construction with such properties [2]. Furthermore, he developed a Privacy Preserving Attribute-Based Broadcast Encryption (PP-AB-BE) scheme. Compared to existing Broadcast Encryption (BE) schemes, PP-AB-BE was more flexible because a broadcasted message could be encrypted by an expressive hidden access policy, either with or without explicit specifying the receivers. Hu [?] proposed a secure and efficient data communication protocol. After establishing a communication channel by a two-factor authentication, it used CP-ABE technology and signature methods to make protection and privacy certification for the data, but the weak point was that the energy consumption was large. But they has a common problem-low computational efficiency.

In view of the above problems, this paper presents an attribute-based security and efficient ciphertext encryption scheme via hybrid encryption method. The asymmetric encryption algorithm encrypts the data and attribute-based encryption algorithm encrypts the sharing key respectively to ensure the security of the stored data and realize the user's attribute-based access control. The traditional encryption algorithm is improved by reducing the number of ciphertext variable parameters, it has improved communication efficiency and reduced computational overhead.

The rest of the paper is organized as follows. Section 2 introduces public key encryption definition and security model. Section 3 outlines the proposed scheme to analyze detailed processes. Experience and security analysis are given in Section 4. Section 5 finally concludes this paper.

## 2 Public Key Encryption Definition and Security Model

A public key encryption scheme consists of three algorithms.

1) Key generation algorithm ($Setup(k)$). Given secure parameter $k$, this algorithm outputs public key $pk$ and corresponding private key $sk$.

2) Encryption algorithm ($Enc(pk, m)$). Given public key $pk$ and plaintext $m$, this algorithm outputs $C$.

3) Decryption algorithm ($Dec(sk, C)$). Given private key $sk$ and ciphertext $C$, this algorithm outputs plaintext $M$ or $N$. $N$ indicates that it is illegal ciphertext.

IND-CCA2 secure model of public key encryption can be defined by the game between challenger $Ch$ and adversary $A$.

**Stage 1.** System building. Challenger $Ch$ runs $Setup$ to generate public-private key pair $(pk, sk)$ and reserve $sk$. $pk$ is sent to adversary.

**Stage 2.** Query1. Adversary does a series of decryption queries. When adversary decrypts ciphertext $C$, challenger runs $Dec(sk, C)$ and sends results to adversary.

**Stage 3.** Challenge. When adversary decides to finish query1 stage, it will output two equal plaintexts $m_1$ and $m_2$. Challenger will randomly select $\alpha \in (0, 1)$ and set challenge ciphertext as $C' = Enc(pk, m_\alpha)$. Then $C'$ will be returned to adversary.

**Stage 4.** Query2. Adversary does a series of decryption queries. Challenger uses the same method like stage1 to respond. Here, adversary cannot decrypt $C'$.

**Stage 5.** Guessing. Finally, adversary outputs the guess $\alpha'$ of $\alpha$.

The above adversary $A$ is defined as IND-CCA2 adversary. And the advantage of public key encryption is as:

$$Adv_A^{IND-CCA2} = |Pr[\alpha' = \alpha] - \frac{1}{2}|.$$

For any adversary $A$ with probabilistic polynomial time, the advantage $Adv_A^{IND-CCA2}$ of public encryption scheme can be ignored. Therefore, this public encryption scheme is IND-CCA2 security.

## 3 Secure and Efficient Ciphertext Encryption Scheme

The system mainly involves four basic technologies: access tree [3], bilinear map [8], Shamir secret Sharing [17] and CABE [12].

### 3.1 Access Tree

Supposing $T$ is an access tree, and each non-leaf node $x$ in $T$ represents a threshold structure described by its child node $num_x$ and a threshold $k_x$, where $num_x$ represents the number of child nodes of node $x$. $k_x$ represents the threshold value of $x$, and $k_x \in [1, num_x]$. When $k_x = 1$, node $x$ denotes a logic OR relation. When $k_x = num_x$, it represents a logic AND relationship. Each leaf node $x$ is described by an attribute and a threshold $k_x = 1$.

Access tree involves several related functions in the practical application. Here is a brief introduction. The function $parent(x)$ represents the parent of node $x$. The function $att(x)$ represents the associated attribute of return leaf node $x$. In order to conveniently invoke nodes in the access tree, access tree sorts each node, and the function $index(x)$ represents the index value of the return node $x$.

### 3.2 Bilinear Map

Supposing $G_0$ and $G_1$ are two $p - order$ multiplicative cyclic groups. $g$ is a generator of $G_0$ and $e$ is a bilinear map, namely $e : G_0 \times G_0 \rightarrow G_1$, then for any $i, j, k \in G_0$ and $a, b \in Z_p$, the map $e$ has the following properties:

1) Bilinear: $e(i^a, j^b) = e(i, j)^{ab}$.

2) Non-degenerative: $e(g, g) \neq 1$.

3) Polymerizability: $e(i \cdot j, k) = e(i, k) \times e(j, k)$.

If the group operation is highly computable in $G_0$ and the map $e : G_0 \times G_0 \rightarrow G_1$, then the group is called bilinear. So map $e$ is commutative: $e(g^a, g^b) = e(g, g)^{ab} = e(g^b, g^a)$.

## 3.3 Shamir Secret Sharing

Shamir secret sharing technology can be explained that distributors distribute the sharing key $s$ to $n$ participants $P_1, P_2, \cdots, P_n$, any $t$ participants can restore the key, and less than $t$ participants can not restore the key. The main principle of Shamir secret sharing is to divide the key $s$ into $n$ parts and distribute them to $n$ participants. Randomly selecting the $t$ integers $a_0, a_1, \cdots, a_{t-1}$ from finite field $GF(q)$. Supposing $a_0 = s$, we can form a polynomial $f(x) = s + a_1x + a_2x^2 + \cdots + a_{t-1}x^{t-1}$. Calculating $f(x_i)(i = 1, 2, \cdots, n)$, and then sharing $x_i$ and $f(x_i)$ with the participant $P_i$. When the collected sub-secrets are greater than or equal to that provided by $t$ participants, we can rebuild the key $S$ by the following formula.

$$s = f(0) = \sum_{i=1}^{t}(f(x_i) \prod_{j\in[1,t],j\neq i\frac{0-x_j}{x_i-x_j}}).$$

## 3.4 Ciphertext Attribute-based Encryption

Ciphertext attribute-based encryption (CABE) is a method that associates a user's private key with a set of user attributes and expresses the attribute discrimination condition as an access tree and is deployed in a ciphertext. The user can decrypt the ciphertext only if the attribute set of the user's private key satisfies the attribute judgment condition in the ciphertext.

In general, CABE includes the following four steps:

1) System initialization *Setup*: input a random parameter, output the main key $MK$ and public parameter $PK$.

2) Message Encryption $CT = Encrypt(PK, M, T)$: input encrypted message $M$, public parameter $PK$ and access tree $T$, output ciphertext $CT$.

3) Key generation $SK = KeyGen(MK, S)$: input attribute set $S$ and the main key $MK$, output the user's public key $PK$ and private key $SK$.

4) Ciphertext decryption $M = Decrypt(CT, SK)$: input ciphertext $CT$ and user private key $SK$. And decrypt ciphertext $CT$ to get the plaintext $M$.

## 3.5 Proposed Security and Efficient Ciphertext Encryption Scheme

Because of the excellent characteristics of CABE, it can be used for our ciphertext access control scenario, but CP-ABE belongs to asymmetric encryption algorithm, it is difficult to encrypt a lot of data. So in our scheme, we first use an AES sharing secret key $K$ to encrypt the data, a RSA public key is used for signing and verifying $K_{pub}$, $K_{priv}$ for $K$. Then the $K$ is encrypted using the improved CP-ABE to reduce the number of ciphertext variables and ensure that $K$ can only be accessed by legitimate users.

The following is the detailed realizing process for our new scheme.

### 3.5.1 System Initialization

1) Randomly select a $p - order$ cyclic group $G_0$, a generator $g$ and a bilinear pair $e : G_0 \times G_0 \rightarrow G_1$. Select a hash function $H : G_0^2 \rightarrow Z_p$ and $(x_1, x_2, x_3 \in G)$. Compute $g_1 = g^x$;

2) Define a Lagrangian coefficient $\Delta_{i,x} : \Delta_{i,x}(x) = \prod_{j\in[1,t],j\neq i\frac{0-x_j}{x_i-x_j}})$, $i$ represents the element in user attribute set $S$;

3) Randomly select a parameter $\alpha$ and a hash function $H : 0, 1^* \rightarrow 0, 1^l$;

4) Generate the master key $MK = (\alpha, g^\alpha)$ and the parameter $PK = G_0, g, e(g, g)^a$;

5) Use $PK$ and user attribute set $S$ to generate the user private key $SK = (\forall_j \in S : D_j = g^{\frac{a}{H(j)}}$.

### 3.5.2 Data Encryption

Given a plaintext $M \in G_T$, it first randomly selects $\tau \in Z_p^*$.

1) Physiological sensor $SN_i$ perceives information $M_1$, generates a hash value $hash(M_1)$ for $M_1$;

2) $SN_i$ uses $K$ to encrypt information $M_1$ and get $E_K(M_1, hash(M_1))$, uses $K_{priv}$ to sign $K$ and get $SIG_{K_{priv}}(K)$;

3) Use Algorithm 1 to encrypt $K$ and get $CK = (T, C = Ke(g, g)^{as}$, $y \in Y$, $C_y = g^{H(att(y))qy(0)}$;

4) $SN_i$ sends $E_K(M_1,\ hash(M_1))$ and $\{SIG_{K_{priv}}(K)$, $CK\}$ to $PDA$, $SN_i \rightarrow PDA$: $\{ID_{SN_i}$, $ID_{PDA}$, $E_K(M_1,\ hash(M_1))$, $SIG_{K_{priv}}(K)$, $CK\}$;

5) When the $SN_i$ uploads the $i{-}th$ data $M_1$, performing step1 and step2, sends $E_K(M_i, hash(M_i))$ to PDA: $SN_i \rightarrow PDA : ID_{SN_i}, ID_{PDA}, E_K(M_1, hash(M_1))$;

   In each time period, $SN_i$ only sends the ciphertext data $CK$ and the signature data $SIG_{K_{priv}}(K)$ to PDA, the next transmission is happened when the sharing key $K$ is updated or invalidated.

6) Then it calculates $C_0 = e(g_1, x_3)^r M$, $C_1 = g^r$ and $C_2 = (x_1^w, x_2^\tau, x_3)^r$. Here, $w = H(C_0, C_1)$. Finally, it outputs ciphertext $CK$.

Algorithm 1. $Encrypt(PK, K, T)$

- Input: public parameter $PK$, sharing key $K$ and access tree $T$ with root node $R$.

- Output: ciphertext $CK$.

**Step 1.** Set a polynomial $q_x$ with $d_x = k_x - 1$ power for each node in the access tree.

**Step 2.** Select a random number $s \in Z_p$ and set $s = q_R(0)$, meanwhile, select $d_R$ random numbers $a_i \in Z_p$. Define the polynomial $q_R$ for root node $R$.

**Step 3.** Set $q_x(0) = q_{parent}(x)(index(x))$ for any other nodes $x$ in the access tree, and select $d_x$ random numbers from $Z_p$ to define $q_x$.

**Step 4.** Let $Y$ be the set of leaf nodes, and the ciphertext constructed by access tree $T$ is $CK = (T, C = Ke(g,g)^{as}, y \in Y : C_y = g^{H(att(y))q_y(0)})$.

**Step 5.** Output ciphertext $CK$.

#### 3.5.3 Support Strategy Dynamic Update

Support strategy dynamic update (SSDU) contains attribute authority, cloud server, data owner and data user.

- Attribute authority. In the system, it has multiple attribute authorities, each attribute authority does not depend on other attributes authorities. It is only responsible for managing their domain attribute of users. Attribute authority is responsible for generating the public and private key pair. And according to the domain users, they generate their own private key respectively.

- Cloud server. It is responsible for storing data and providing file access service for the data users. Cloud server is responsible for updating old ciphertext access strategy tasks at the same time. After updating, the completion of the new ciphertext is equivalent to the ciphertext directly generated by the new access strategy.

- Data owner. It is responsible for access strategy and using these strategies to encrypt data. Data owner is responsible for generating policy to update the key at the same time, and requesting a cloud server to update the old access strategy.

- Data user. The system generates an unique identity for each data user.

In SSDU, there are seven polynomial time algorithms: *GlobalSetup*, *AuthoritySetup*, *SKeyGen*, *Encrypt*, *Decrypt*, *UKeyGen* and *CTUpdata*.

- $GlobalSetup(\lambda) \rightarrow GP$. Select a random security parameter $\lambda$ as input and output the public parameter $GP$.

- $AuthoritySetup(GP, AID) \rightarrow (SK_{AID}, PK_{AID})$. In the system, each attribute authority performs this algorithm to complete initialization operation. Use public parameter $GP$ and unique identity $AID$ to generate private key and public key $(SK_{AID}, PK_{AID})$.

- $SKeyGen(GID, GP, S_{GID,AID}, SK_{AID})$. $AID$ produces attribute private key $SK_{AID}$ for $GID$. $SK_{AID}$, $GID$, $S_{GID,AID}$ and $GP$ is as input. $SK_{GID,AID}$ is as output.

- $Encrypt(PK, GP, m, A) \rightarrow CT$. Public parameter $GP$, access structure $A$ and public key set $PK$ are used to encrypt message $m$. Output ciphertext $CT$.

- $Decrypt(CT, GP, SK_{GID,AID} \rightarrow m$. $GP$ and $SK_{GID,AID}$ are used to decrypt ciphertext $CT$. If the attribute of user private key can satisfy the access structure in ciphertext $CT$, then the encryption algorithm can correctly decrypt ciphertext and output message $m$.

- $UKeyGen(PK, EnInfo(m), A, A') \rightarrow UK_m$. Data owner executes the cryptographic strategy update key generation algorithm. Public key set $PK$, enciphered message $EnInfo(m)$ of message $m$, old access structure $A$ and new access structure $A'$ are as input. Then it generates ciphertext update key $UK_m$.

- $CTUpdata(CT, UK_m) \rightarrow CT'$. Ciphertext update algorithm is performed by the cloud server. It uses $UK_m$ to update $CT$ and outputs the ciphertext $CT'$ related to new access structure $A'$.

#### 3.5.4 Decryption Stage

Given ciphertext $C = (\tau, C_0, C_1, C_2)$, computing $w = H(C_0, C_1)$ and verifying $e(C_1, x_1^w, x_2^\tau, x_3) = e(g, C_2)$. Plaintext will be recovered with private key $M \leftarrow \frac{C_0}{e(C_1, x_3)^x}$.

1) The user obtains the ciphertext $\{ID_{SN_i}$, $E_K(M_1$, $hash(M_1)$, $\cdots$, $E_K(M_n$, $hash(M_n)))\}$ and $SIG_{K_{priv}}(K), CK$ of $SN_i$ from the PDA;

2) The user uses Algorithm 2 to decrypt $CK$ and get $K$;

3) The user uses $K_{pub}$ to verify the correctness of $SIG_{K_{priv}}(K)$, if it is correct, then it performs the next step, otherwise returns to the first step;

4) User uses $K$ to decrypt $E_K(M_1$, $hash(M_1)$, $E_K(M_2$, $hash(M_2)$, $\cdots$, $E_K(M_n$, $hash(M_n)))$ and get $M_1'$, $M_2'$, $\cdots$, $M_{n1}'$, $hash(M_1)$, $hash(M_2)$, $\cdots$, $hash(M_n)$;

5) User verifies batched the correct of information $M_1', M_2', \cdots, M_{n1}'$, if $hash(M_i') = hash(M_i), \cdots, hash(M_n), (1 \leq i \leq n)$, then the information acquired by users is valid, otherwise give up it;

Algorithm 1 is an improved algorithm for CP-ABE encryption. And Algorithm 2 is the corresponding decryption algorithm. In Algorithm 1, calculating one Tate pair needs time $O(|p|)$, outputting ciphertext $CK$ needs to calculate two Tate pairs, so the time complexity of the whole algorithm is $O(2|p|)$. While the CP-ABE encryption algorithm needs to calculate four Tate pairs with a time complexity $O(4|p|)$.

Algorithm 2. $Decrypt(CK, SK)$.

- Input: ciphertext $CT = T, C, C_y$, user privacy key $SK = \forall j \in S : D_j$ and user attribute set $S$.
- Output: Sharing key $K$.

**Step 1.** If node $x$ is a leaf node, executing $DecryptNode(CT, SK, x)$ to get $e(D_i, C_x) = e(g^{\frac{a}{H(i)}}, g^{H(att(x)q_x(0))}) = e(g,g)^{aq_x(0)}$;

**Step 2.** If node $x$ is a non-leaf node, it calls $DecryptNode(CT, SK, y)$ for each child node $y$ to obtain $e(g,g)^{aq_x(0)}$;

**Step 3.** Executing the formula $F_X = \prod_{y \in s_x} e(g,g)^{aq_x(i)\cdot\Delta_{i,s'_x(0)}}$ for all child nodes $y$ and get $e(g,g)^{aq_x(0)}$;

**Step 4.** Repeat step 1, 2, 3, until $e(g,g)^{aq_R(0)}$ of the root node $R$ of access tree $T$ is obtained;

**Step 5.** Calculating $\frac{C}{e(g,g)^{aq_R(0)}}$ to get $K$;

**Step 6.** Output the sharing key $K$.

# 4 Experiment and Analysis

In this section, we compare our new scheme (abbreviated to AHEM) with state-of-the-art schemes PPCPA [13], FEAC [11], EABE [10]. Our experimental equipment is Intel i5-4200U 2.30GHz processor, 8G memory, 64-bit Windows 8.1 operating system.

Because the PDA has a wealth of energy resources, but physiological sensor energy resources are limited. The experiment mainly analyzes calculation overhead of data encryption operation and transmission data communication costs for the physiological sensor. Since the size of information is directly related to the communication overhead, we begin to analyze the information size.

## 4.1 Information Size

Any physiological sensors $SN_i$ needs to upload two types of data for PDA: one is the encryption and signature data $E_K(M, hash(M), SIG_{K_{priv}}(K)$, and the other is ciphertext data $CK = T, C, C_y$. Therefore, the information size of our new scheme is:

$$T^1 = |ID_{SN_i}| + |ID_{PDA}| + |T| + |C| + |C_y| + |SIG_{K_{priv}}(k)| + |E_K(M, hash(M))|.$$

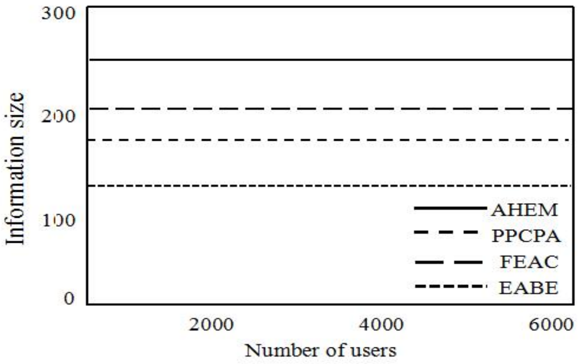


Figure 1: Relation between information size and users' number

At the same time, we also give the information size of PPCPA, FEAC, EABE as shown in following equations respectively.

$$\begin{aligned} T^2 &= |ID_{SN_i}| + |ID_{PDA}| + |T| + |\tilde{C}| + |C_y| \\ &\quad + |C| + |AES(K,M)| + |\sigma| + |C'|. \\ T^3 &= |ID_{SN_i}| + |ID_{PDA}| + |2T| + |\tilde{C}| + |C_y| \\ &\quad + |2C| + |AES(K,M)| + |3C'|. \\ T^4 &= |ID_{SN_i}| + |ID_{PDA}| + |3T| + |2C_y| \\ &\quad + |5C| + |AES(K,M)| + |4\sigma| + |C'|. \end{aligned}$$

We set the $ID_{SN_i}$ of physiological sensor $SN_i$ and the $ID_{PDA}$ of the PDA is 1byte, the size of $T$ is 4bytes, size of $SIG_{K_{priv}}(K)$ is 40bytes, the size of $E_K(M, hash(M)$ is 16bytes. $C$ and $C_y$ are variable. In our scheme, the bilinear map $e$ uses Tate pairs. The elliptic curve $E$ is defined as $q - order$ 160-bit prime in $F_p$, groups $G_1$ and $G_2$. According to the literature [7,15], in order to provide a security level equivalent to 1024-bit RSA, if the group $G_2$ is defined as a $q - order$ subgroup in the multiplicative group of finite fields $F^*_{p^2}$, then $p$ is a 64bytes prime. In addition, in the finite field $F^*_{p^3}$, $p$ is 42.5bytes and $p$ is 20bytes in the finite field $F^*_{p^6}$. For uniform comparison, we take $|p|$=64bytes. Therefore, the information size of PPCPA is 175bytes, the information size of FEAC is 201 bytes, the information size of the EABE is 131bytes (m=64). Figure 1 shows the relationship between the size of the information and the number $N$ of users. From the curve we can see that the new scheme's information size and the number of users are independent of each other.

## 4.2 Communication Overhead

Assuming that during time period, the physiological sensor $SN_i$ transforms $n$ data to PDA. Total communication costs AHEM, PPCPA, FEAC, EABE are shown as following equations respectively.

$$\begin{aligned} T^1_n &= |ID_{SN_i}| + |ID_{PDA}| + |T| + |C| + |C_y| \\ &\quad + |SIG_{K_{priv}}(k)| + |E_K(M, hash(M))| \times n. \end{aligned}$$

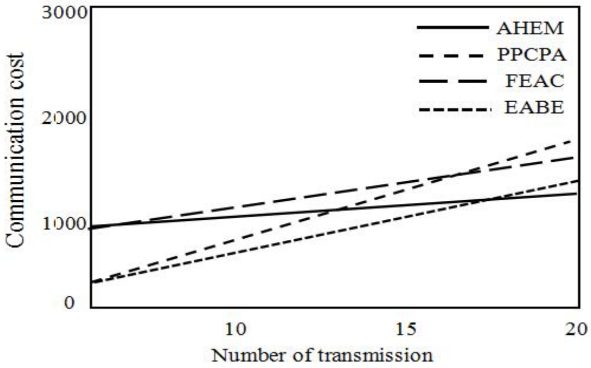


Figure 2: Relation between communication cost and data transmission number

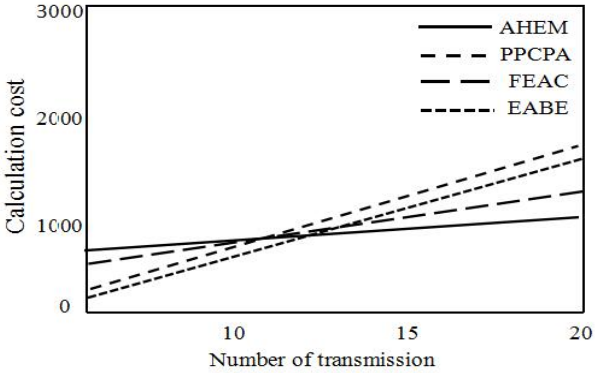


Figure 3: Relation between calculation cost and data transmission number with high cost equipment

$$\begin{aligned} T_n^2 &= |ID_{SN_i}| + |ID_{PDA}| + |T| + |\tilde{C}| + |C_y| \\ &\quad +|C| + |AES(K,M)| \times n + |\sigma| + |C'|. \\ T_n^3 &= |ID_{SN_i}| + |ID_{PDA}| + |2T| + |C| \\ &\quad +|SIG_{K_{priv}}(k)| \times n + |E_K(M, hash(M))|. \\ T_n^4 &= |ID_{SN_i}| + |ID_{PDA}| + |\tilde{C}| + |C_y| \times n \\ &\quad +|1.5C| + |AES(K,M)| + |\sigma| + |2C'|. \end{aligned}$$

Figure 2 shows the relationship between the communication overhead and the number of data transmissions $n$. It can be seen from the curve that when the number of $n$ increases, the communication overhead of all schemes increases. The growth rate of the EABE is the largest, and the growth rate of AHEM is the smallest.

### 4.3 Calculation Overhead

At the 64-bit Intel i5-4200U processor with running speed 2.30GHz, calculating one Tate needs about 61.03ms. In addition, the certification of a 160-bit ECDSA signature takes about 18.48ms. Note that we omit the computational overhead of hash operation and symmetric encryption operation. In that they have a significantly lower computational cost. Assuming that the physiological sensor $SN_i$ transmits $n$ data to PDA during the time period. In the EABE scheme, $SN_i$ encrypts the data, and the computational overhead is three Tate pairs. So the total computational cost is $3 \times 61.03 = 181.09ms$. In the FEAC scheme, the calculation is five Tate pairs with a total computational overhead $5 \times 61.03 = 305.15ms$. The computational overhead of PPCPA scheme is mainly generated by two ECDSA signatures, with a total computational cost 36.96nms. The computational overhead of the AHEM scheme is generated primarily by an ECDSA signature with a total computational effort of 18.48nms.

Figure 3 shows the relationship between the computational cost and the number of data transmissions $n$. From the curve, our scheme has a low computational overhead and is not affected by the number of data transfers, since it only needs to perform a Tate operation on $n$ data transmission. But to illustrate the performance of proposed algorithm, we also perform an experiment with low cost equipment. Results are as Figure 4.

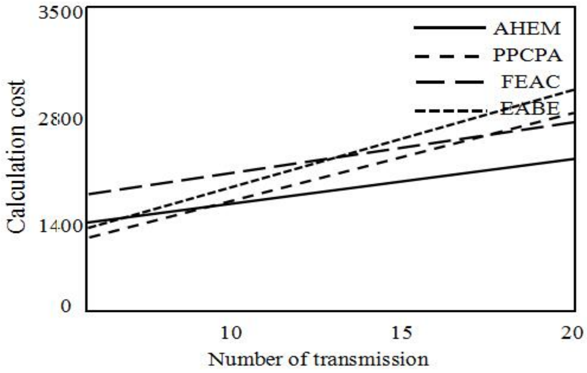


Figure 4: Relation between calculation cost and data transmission number with low cost equipment

## 5 Conclusion

In this paper, a secure and efficient ciphertext encryption scheme based on attribute and support strategy dynamic update via hybrid encryption method is proposed. The new scheme encrypts the data through asymmetric encryption technology, which guarantees the security of the stored data, and associates the user key with a set of attributes. Associating the sharing key with a set of attribute discrimination criteria, the user can decrypt the ciphertext only if the attribute discrimination condition is satisfied avoiding the cost of distributing the sharing key for each user. Finally, experiments for the proposed scheme, the results show that our new scheme has very low computational and communication overhead. In the future work, we will carry out the proposed program, so as to further improve the effectiveness of privacy protection.

## 6 Acknowledgments

This work was supported by the Independent Innovation Fund for Young Teachers (2572015BB24); This work was also supported by the National Natural Science Foundation of China (31370565) and School-level Project (JMSUJCMS2016-005).

## References

[1] J. S. Chen, C. Y. Yang, M. S. Hwang, "The capacity analysis in the secure cooperative communication system," *International Journal of Network Security*, vol. 19, no. 6, pp. 863-869, Nov. 2017.

[2] P. S. Chung, C. W. Liu, and M. S. Hwang, "A study of attribute-based proxy re-encryption scheme in cloud environments", *International Journal of Network Security*, vol. 16, no. 1, pp. 1-13, 2014.

[3] P. Hu, H. Y. Gao, "A key-policy attribute-based encryption scheme for general circuit from bilinear maps," *International Journal of Network Security*, vol. 19, no. 5, pp. 704-710, 2017.

[4] M. S. Hwang, S. M. Chen and C. Y. Liu, "Digital signature with message recovery based on factoring and discrete logarithm," *IETE Journal of Research*, vol. 62, no. 3, pp. 415-423, Sep. 2016.

[5] H. Li, S. L. Yin, C. Zhao and L. Teng, "A proxy re-encryption scheme based on elliptic curve group," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 8, no. 1, pp. 218-227, Jan. 2017.

[6] L. H. Li, S. F. Tzeng, M. S. Hwang, "Generalization of proxy signature based on discrete logarithms", *Computers & Security*, vol. 22, no. 3, pp. 245–255, 2003.

[7] H. J. Liu, Y. H. Chen, H. Tian, T. Wang and Y. Q. Cai, " Attribute-based secure and efficient ciphertext encryption scheme," *Journal of Chinese Computer Systems*, vol. 38, no. 8, pp. 1708-1711, 2017.

[8] J. Liu, S. L. Yin, H. Li and L. Teng, "A density-based clustering method for k-anonymity privacy protection," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 8, no. 1, pp. 12-18, Jan. 2017.

[9] E. J. L. Lu, M. S. Hwang, and C. J. Huang, "A new proxy signature scheme with revocation", *Applied Mathematics and Computation*, vol. 161, no. 3, PP. 799-806, Feb. 2005.

[10] Y. Mao, Y. Zhang, M. R. Chen, Y. Li and Y. Zhan, "Efficient attribute-based encryption schemes for secure communications in cyber defense," *Intelligent Automation & Soft Computing*, vol. 22, no. 3, pp. 1-7, 2016.

[11] R. Y. Sreenivasa, R. Dutta, "Fully secure bandwidth-efficient anonymous ciphertext-policy attribute-based encryption," *Security & Communication Networks*, vol. 8, no. 18, pp. 4157-4176, 2016.

[12] W. Susil, Y. Jiang, Y. Mu and F. Guo, "Ciphertext-policy attribute-based encryption supporting access policy update and its extension with preserved attributes," *International Journal of Information Security*, pp. 1-16, 2017.

[13] G. S. Tamizharasi, B. Balamurugan, H. A. Gaffar, "Privacy preserving ciphertext policy attribute based encryption scheme with efficient and constant ciphertextsize," in *International Conference on Inventive Computation Technologies*, pp. 1-5, 2017.

[14] L. Teng, H. Li, "A high-efficiency discrete logarithm-based multi-proxy blind signature scheme," *International Journal of Network Security*, v0. 20, no. 5, 2018.

[15] C. H. Wei, M. S. Hwang, A. Y. H. Chin, "A secure privacy and authentication protocol for passive RFID tags," *International Journal of Mobile Communications*, vol. 15, no. 3, pp. 266-277, 2017.

[16] N. I. Wu, M. S. Hwang, "Development of a data hiding scheme based on combination theory for lowering the visual noise in binary images," *Displays*, vol. 49, pp. 116-123, Sep. 2017.

[17] S. L. Yin, L. Teng and J. Liu, "Distributed searchable asymmetric encryption," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 4, no. 3, pp. 684-694, 2016.

[18] Z. Zhou, D. Huang and Z. Wang, "Efficient privacy-preserving ciphertext-policy attribute based encryption and broadcast encryption," *IEEE Transactions on Computers*, vol. 64, no. 1, pp. 126-138, 2016.

## Biography

**Yanhua Wang** received the B.Eng. degree from Northeast Forestry University, Harbin, 150040 China in 2013. Now, she is working for a doctor degree in College of Information and Computer Engineering, Northeast Forestry University. Her research interests include Multimedia Security, Network Security, artificial intelligence and Data Mining.

**Yaqiu Liu** obtained his Ph.D. degree in Information Science and Engineering from Northeast Forestry University. Yaqiu Liu is a full professor of the College of Information and Computer Engineering at Northeast Forestry University. He is also a doctor's supervisor. He has research interests in wireless networks, cloud computing, social networks and quantum cryptography. Prof. Liu had published more than 50 papers on the above research fields.

**Kun Wang** received the B.Eng. degree from Jiamusi University, Jiamusi 154007 China in 2010. Now, she is working for a doctor degree in College of Information and Computer Engineering, Northeast Forestry University. Her research interests include Multimedia Security, Network Security and Data Mining.

# A Grudger Based AIS Approach to Coerce Selfish Node Cooperation in MANET

Lincy Elizebeth Jim and Mark A. Gregory
(Corresponding author: Mark A. Gregory)

Department of Electrical and Electronics Engineering, RMIT University
124 La Trobe St, Melbourne VIC 3000, Australia
(Email: mark.gregory@rmit.edu.au)

(Received Aug. 14, 2017; revised and accepted Nov. 11, 2017)

## Abstract

Mobile Ad hoc Networks (MANET) utilize multi-hop communications to forward packets across the network consuming power, processing, and memory resources. In an ideal MANET the nodes are unselfish and forward packets on demand. The real-time, ad hoc and open characteristics of MANET make it susceptible to selfish and malicious nodes affecting performance. In a MANET, some of the nodes may decide to selfishly cooperate or to not cooperate, with other nodes. The selfish nodes reduce the overall effectiveness of network communications, decrease packet delivery rates and increase packet delivery time. This paper investigates an approach to harness selfish node energy and transmission capacity to share network load. This paper utilizes a Grudger Artificial Immune Systems based trust model to study the impact of selfish nodes in the network. The proposed algorithm demonstrates an increase in the packet delivery ratio.

*Keywords: Artificial Immune System; Grudger; Mobile Ad hoc Networks*

## 1 Introduction

In Mobile Ad hoc Networks (MANET) information is sent across the self-organizing network utilizing node to node communications. MANET is an ideal candidate for mobile communications in regions with limited access to fixed infrastructure and for emergency and disaster relief operations. The nodes that form a part of the network have limited battery power and utilize the help of other nodes in the network for packet forwarding. The traditional MANET routing protocols like Dynamic Source Routing (DSR) [1] and Ad hoc on Demand Distance Vector (AODV) [2] function on the assumption that all nodes are highly cooperative and truthful. The dynamic MANET topology and communication demands, particularly as a relay, can take a toll on the limited node battery power, and it is possible that nodes will adopt a selfish stance to prevent further power drain by relay requests. The selfish nodes continue to consume the resources of other nodes while preserving their resources.

Selfish nodes limit MANET performance, and it is reasonable in certain situations to adopt an approach that isolates selfish nodes upon identification or to encourage selfish nodes to change their behavior before isolation is imposed. MANET is a networking approach that can be utilized for disaster management, military and rescue operations. In each of these scenarios, for MANET to be effective there is a need to limit the number of selfish nodes. MANET effectiveness is increased when the nodes within the network are active participants thereby reducing the amount of traffic that is resent due to nodes failing to relay traffic as requested.

This paper is divided into four sections. Section 2 provides a description of earlier work relating to selfish nodes found in the literature. Section 3 provides an overview of the Artificial Immune System (AIS) and selected AIS algorithms. Section 4 describes the proposed Grudger Artificial Immune System framework (GrAIS). In Section 5, simulation results for scenarios with different mission-critical workloads are presented. Finally, Section 6 concludes the paper and outlines future work.

## 2 Related Work

Routing protocols developed for MANET can be classified as proactive, reactive and hybrid. The effect of node selfishness on routing and node resource utilization efficiency has not been a focus for earlier research. In [3] misbehavior in MANET was first identified, defined and the focus of this work was to alleviate node misbehavior. Research found in the literature appears to focus on how to detect and isolate selfish nodes. This approach does not penalize the selfish nodes nor to coerce the selfish nodes to forward packets. The malicious nodes are rewarded if they're identified and removed from routing paths. In [4] a review of node selfishness in MANET is provided. This paper summarizes existing approaches to dealing with the selfishness problem and the authors provide a proposed solution to mitigate the selfishness problem. The operation of DSR [5] is explored and as energy depletes node selfishness occurs. Various types of selfishness are defined and the problems arising because of selfish nodes co-existing in the network is investigated. In [6] the data flows between MANET nodes are observed and when a selfish node does not forward a packet, the neighbor node waits for a pre-defined threshold number of packet transmission failures to be exceeded before triggering an alarm.

In [7] the impact of selfish nodes on the quality of service in MANET is explored. This work analyses parameters including throughput, average hop count and packets dropped. The hop count increases as the selfish node concentration increases. The authors found that there is an increase in the number of packets dropped along with a significant decrease in throughput as the selfish node concentration increases. In [8] the MANET nodes are encouraged to be altruistic and the nodes are given positive or negative scores depending on their behavior. The altruistic nodes utilize their energy to relay for other nodes but they relay for selfish nodes only once. This approach does not call for the participation of selfish nodes for any communication.

The Combined Immune Theories Algorithm (CITA) [15, 16] utilizes the basic principles of well-known immune theories including the Dendritic Cell Algorithm (DCA), Clonal Selection (CS), and the Negative Selection Algorithm (NSA). Cerri and Ghioni compare this algorithm with the Secure Ad hoc on Demand Distance Vector algorithm (SAODV) and demonstrate its improved performance [14]. DCA is used to obtain contextual information. Dendritic Cells (DC) are associated with a subset of neighboring nodes called elements, which are responsible for DC maturation. Element subsets are monitored using adjacent Immature Detectors (ID), adjacent Mature Detectors (MTD) and Memory Detectors (MMD). During the learning phase the network is configured with trusted nodes, parameters, alarms and the nodes have a set of detectors. The CITA algorithm shows an improved performance regarding the packet delivery ratio in the presence of malicious nodes performing a denial of service attack.

# 3 Artificial Immune Systems

AIS are adaptive systems inspired by theoretical immunology and observed immune functions, principles, and models, which are applied to complex problem domains [17]. Research into the immune system is gaining in significance due to its unique ability to solve complex issues. AIS research, as a branch of computational intelligence, has attained importance since its genesis in the 1990's. There are, to date, four major AIS algorithms upon which AIS research is based. They are 1) Artificial Immune Networks (AIN); 2) NSA; 3) CS; and 4) Danger Theory and the DCA. The AIS research field combines the Immunology, Computer Science, and Engineering disciplines to solve complex problems. The prominent AIS features include learning, memory and pattern recognition. Forrest *et al.* initially proposed the NSA [18] to differentiate between self and non-self cells based on the generation of T-cells. This approach was originally applied to computer virus detection. Based on the work of Forrest *et al.* variations of the NSA have been formulated keeping in mind the fundamental properties of the original algorithm.

An immune system is the defender of the human body against pathogens. There has been a significant amount of work in recent research about how to use Human Immune System (HIS) [16] concepts to solve complex problems. The HIS is capable of processing information, learning and memorizing salient information. The AIS borrows principles from the HIS and attempts to apply the fundamental concepts to other applications. The primary task of the HIS is to keep the body healthy and protect it from pathogens. The HIS consists of organs, cells, and tissue that work together to identify and attack dangerous invasive threats like bacteria and viruses. In the event of an attack by a pathogen, a series of steps called an immune response is launched, thereby distinguishing, and protecting cells and tissue from harmful pathogens.

The key to a healthy immune system is the ability to make a distinction between self and non-self cells. The immune defenders launch an attack on anything they identify as foreign. Antigens are foreign objects that trigger an immune response. Transplants from another person may also relate to non-self, can lead to an attack by the HIS and, as a result, to limit the probability that this unwanted outcome can occur, masking drugs have been developed.

The AIN model was redefined by Timmis *et al.* [19]. The AIS coterie has produced many versatile sets of immune inspired algorithms to solve real world as well as computational problems. An insight into the mathematical immuno-computing strategies was provided by Tarakanov *et al.* [20].

## 3.1 Dendritic Cell Algorithm

Steinman and Cohn [21] identified the DC characteristic as an antigen presenting cell. A DC is mainly composed of leukocytes and is present in all tissue. DCs inside the tissues segregate and mature during an appropriate trigger; once they mature they move to secondary lymphoid tissues and present antigen to T-cells to launch an immune response. Immature DCs are found on the body surfaces including the skin and is also found in blood. When the pathogens are identified, captured and processed by the immature DC the DCs migrate to the thymus and spleen where they mature and induce an immune response.

The change of state of the DC [22, 23] is facilitated by the identification of signals such as the pathogen-associated molecular pattern (PAMP), danger signals and apoptotic signals (safe signals) as seen in Figure 1. The signals are described as follows: (1) PAMPs activate the immune response thereby protecting the host from infection; (2) danger signals are released during tissue cell damage, their strength is lower than PAMPs; (3) safe signals are given out when programmed/normal cell death occurs; and (4) inflammatory cytokines are given out when general tissue distress occurs and amplify the effect of the other three signals. The immune response of the T-cell is determined by the corresponding weights of the four signal types. The DCA was proposed by Greensmith *et al.* [24] and combines various signals to investigate the current circumstance of the environment and non-parallel sampling of another data stream (antigen). A fuzzy margin derived corresponding to the concentration of co-stimulatory molecules is an indicator for a DC to stop antigen collection and migrate to a virtual lymph node. The DCA works on the input signals with presumed weights to produce output signals.

The algorithm proposed in [25] consists of the following stages: initialization, update, and aggregation. During the initialization phase, training and initial values are set. The update stage consists of two stages namely tissue update and cell cycle analysis. The DC is designed as a Libitissue server [26]. The cell cycle is a well-defined process that occurs at a user-defined rate. As soon as the antigen data is processed, the process of the cell cycle and tissue updating stops. During the concluding stage, aggregation of the collected antigens occurs together with analysis, and the Mature Context Antigen Value (MCAV) per antigen is derived.

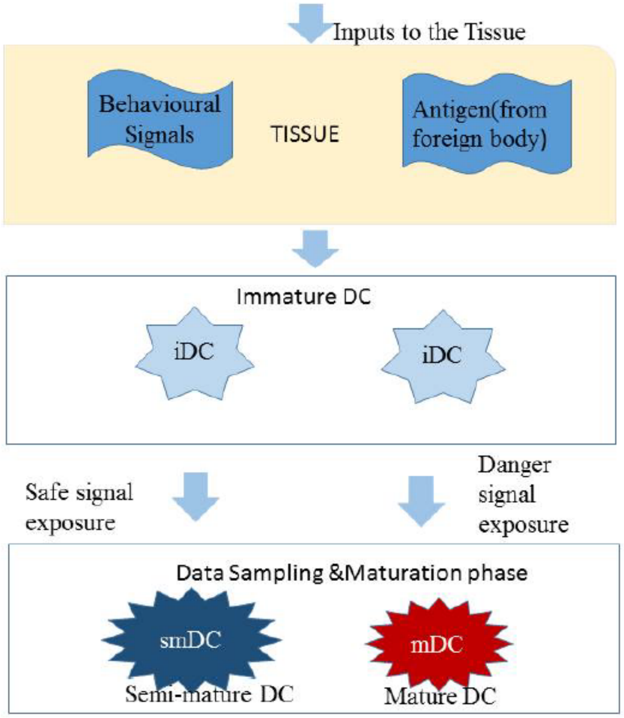


Figure 1: Dendritic cell algorithm schematic

### 3.2 Danger Theory

The Danger Theory, proposed by Matzinger, emphasized that the "foreignness" of a microbe is not the primary factor that ignites a response [27]. Danger Theory states that antigen-presenting cells are activated by danger/alarm signals from exerted cells. Danger signals will not be sent by robust cells or by cells experiencing normal cell death. Any cell that dies unnaturally sends out a danger signal, and antigens near the dying cell are captured by antigen presenting cells like macrophages and are then dispensed to the lymphocytes. B-cells also secrete antibodies. The antibodies that identify a match with the antigens present in the danger zone will be triggered. Those antibodies that do not identify a match with the antigens will not be in the danger zone and therefore will not be triggered.

The Danger Theory has its drawbacks and Aickelin *et al.* proposed applications of Danger Theory that highlight: (1) the presence of an Antigen Presenting Cell (APC) is crucial for a danger signal; (2) a danger signal does not need to be related to threatening events; (3) danger signals can be positive or negative (presence or absence of signal); and (4) conceptual ideas were also proposed on how the Danger Theory can be used for anomaly detection [29]. Based on the Danger Theory, danger signals always spark an immune response. In a computing application of Danger Theory, low or high memory usage, fraudulent disk activity, and other events could be indicated by danger signals. The immune system reacts to the antigens in the danger zone once a danger signal is created. After the critical components are recognized, they are then sent to a part of the system for further verification. The Two-Signal Model extended by Bretscher *et al.* [28] explains Danger Theory in a different way where two signals are needed to activate the lymphocytes: 1) antigen recognition; and 2) co-stimulation. The co-stimulation signal indicates that the antigen is threatening.

## 4 Trade-Off Between Selfishness and Altruism

The motivation for this paper stems from the observation that it is not beneficial to the operation of a MANET to ignore or isolate selfish nodes. The approaches presented in Section **??** isolated the selfish nodes with a bad reputation. Initially, all nodes in the network have the same classification and as time changes some nodes tend to become selfish. One of the reasons for nodes to become selfish is due to the relay load that the node may have experienced. Traffic workload has a direct effect on energy consumption and as energy reduces the nodes can become selfish for various reasons including observation of the number and state of neighboring nodes. The good nodes tend to overlook selfish nodes and continue to render service to the selfish nodes irrespective of any service in return. This paper provides a trust model framework that incorporates the good and selfish nodes. The proposal is that for high traffic volumes the routing task should be carried out by all nodes, including selfish nodes. In [11] the author uses a model that considers how birds clean each other of parasites in hard to reach places, therefore helping with individual and group survival. The author defines three different model behaviors:

1) Sucker-Birds that blindly help other birds without expecting anything in return.
2) Cheat-Birds that take advantage of all the help they can get but do not offer anything in return.
3) Grudger-Birds that help others and recall who they have helped. In case the same bird does not reciprocate, they will not help that bird again.

The routing model proposed in this paper categorizes good nodes into a sucker group, cheat nodes into selfish and Dendritic Cell (DC) nodes into a grudger group. In the proposed GrAIS algorithm, as seen in Figure 2, each node is modeled as a Grudger Dendritic Cell (gDC). This DC node is analogous to the HIS DCs as they are the first line of defense in HIS. The initiator gDC node sends a Route request to the nodes in the network. The nodes that already have a path to the destination will send back a Route Reply. Upon receipt of the Route Reply, the source gDC node calculates the Probability of communication nearness ($P_{com}$) [9] of those nodes from which a Route Reply was obtained. The packet is sent to the node that responded with the highest $P_{com}$ value.

During this phase, the source node expects nodes to Acknowledge (ACK) packet receipt. In the case of a selfish node that does not send an ACK, a high priority PAMP signal is raised by the gDC node and the initiator node is also notified. The selfish node is forced to acknowledge receipt of high priority PAMP signal as this high priority packet overwrites the selfish node's buffer and there upon the packet signal will be transmitted as high priority by the previous gDC node.

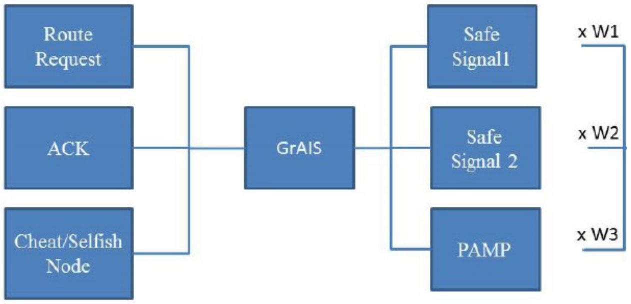


Figure 2: GrAIS model. The interaction types between nodes are shown along with the incorporation of the trust model in order to launch an immune response

Similarly the gDC nodes, when they do not receive a response from the intermediate selfish node, inform the sender node and raises the high priority PAMP signal to validate the presence or absence of a selfish node. In our AIS based trust model, three trust signals are proposed:

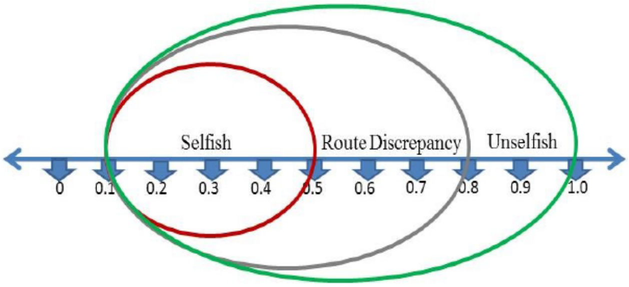


Figure 3: Trust number-line model

- Safe Signal 1 (SS1) - This is generated upon receipt of Route Reply;
- Safe Signal 2 (SS2) - This is generated upon receipt of ACK;
- PAMP - This signal helps validate selfish node behavior. PAMP activates the immune response thereby protecting the host from infections in HIS. Similarly PAMP, being a high priority signal, overwrites the node buffer, and the selfish node will acknowledge receipt of the PAMP.

The trust value $T^{TP}_{a,,b}(t)$ is evaluated by *Node a* towards *Node b* at time $t$, $TP$ is the trust purpose. $T^{TP}_{a,b}(t)$ is represented as a real number in the range [0, 1] as seen in Figure 3 where 1 indicates unselfish nodes, [0.5-0.8] indicates route error discrepancies and indicates a selfish behavior.

$$T^{TP}_{a,b}(t) = w_1 T^{SS1}_{a,b} + w_2 T^{SS2}_{a,b} + w_3 T^{PAMP}_{a,b} \qquad (1)$$

Where $w_1, w_2, w_3$, are the weights related to the trust components, with $w_1 + w_2 + w_3 = 1$.

Instead of assigning individual weights to each of the trust elements a priority signal, PAMP, is used and a signal, SAFE, to indicate the nodes are behaving correctly. The weight of the PAMP priority signal is shown by $w_{PAMP}$. The weight of the safe signal is shown by $w_{SAFE}$. Equation (1) can be rewritten as:

$$T_{a,b}(t) = w_{PAMP}[T^{PAMP}_{a,b}] + w_{SAFE}[T^{SS1}_{a,b} + T^{SS2}_{a,b}]. \qquad (2)$$

Where $w_{PAMP} + w_{SAFE} = 1$. A sliding window transmission approach is used to decrease the effect of conditions arising out of a network that could affect the trust calculation. We use a timing window ($\Delta t$) to determine the number of successful and unsuccessful packets sent between nodes. Let us consider a scenario where *Node a* evaluates *Node b* based on its behavior; thereby making *Node a* trustor and *Node b* the trustee. *Node c* sits beyond *Node b*. The trust relationship between nodes $a, b$ and $c$as shown in Figure 4 is given by $(a, b) = (a, b) : (b, c)$.

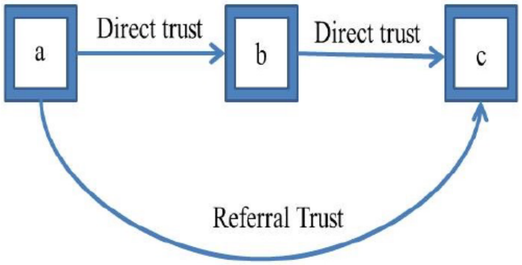


Figure 4: Trust relationship

Let the Trust Purpose be defined as "the node should be *good*." The trust between *Node a* and *Node b*will be direct therefore it's a functional (direct) level of trust whereas the trust between *Node a* and *Node c* will be indirect (referral) as well as an exponential decay factor of trust $\rho$ is also considered therefore it's a referral level [13] of trust.

$$T^{TP}_{a,b,c}(t) = T^{TP}_{a,b}(t)P_{com} + T^{TP}_{b,c}(t)P_{com} + e^{-\rho\Delta t}T^{TP}_{a,c}P_{com}.$$

To compute$TRIALRESTRICTION$, we consider the number of interactions between nodes $a, b$and $c$over the maximum possible number of interactions that could occur with any neighbor node during the interval [*0, t*]. The hop count measure calculated by P$_{com}$ [9] and the Effective Energy of each node ($EE_{node}$) [10] is detrimental during the interaction between nodes in the GrAIS model. The flow chart of the GrAIS algorithm is as seen in Figure 5. In this approach, the following interaction categories, with regards to an unselfish node are considered, given that *Node a*:

- Sending Request;
- Receiving Reply;
- Selection of node based on highest value of P$_{com}$;
- Send Packet;
- If no ACK, send PAMP;
- If PAMP received, classify node as selfish node;
- gDC node will resend packet to selfish node.

$P_{com}$ is an important factor while evaluating the trust purpose ($TRIALRESTRICTION$) between any two nodes as the packet will be sent to the node that responds with a route reply and highest $P_{com}$ value. In this approach $TRIALRESTRICTION$ between any two neighboring nodes is computed by taking into account the number of communications between nodes $a$and $b$over the maximum possible number of interactions that could occur with any neighbor node during the interval [*0,t*]. The trust purpose for *Safe Signal 1* is computed by taking the ratio of the total number of route replies (N$_{RREP}$) received with the total number of route requests sent (N$_{RREQ}$). The trust purpose for *Safe Signal 2* is computed by taking the ratio of the total number of acknowledgment packets$TRIALRESTRICTION$ received by the sender with the route reply packets sent by the destination/intermediate ($TRIALRESTRICTION$ node. The trust purpose for the PAMP signal is computed by taking the ratio of the total number of PAMP sent for every route reply received by the sender and no acknowledgment sent by the destination/neighbor node.

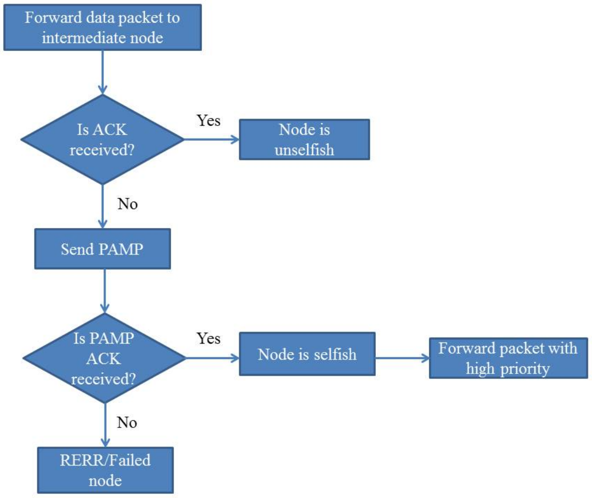


Figure 5: Flow of events in the proposed GrAIS model. The importance of PAMP signal is depicted in the flow chart.

$$\begin{aligned} T^{SS1}_{a,b}(t) &= [\frac{N_{RREP}}{N_{RQ}}]P_{com} \\ T^{SS2}_{a,b}(t) &= [\frac{N_{ACK}}{N_{RREP}}]P_{com} \\ T^{PAMP}_{a,b}(t) &= [\frac{N_{PAMP}}{N_{RREP}}]P_{com} \end{aligned}$$

The intermediate node informs the source node of a neighboring node that appears to be selfish. The source node sends a PAMP signal to overwrite the selfish node buffer and this selfish node is added to a blacklist to prevent it being used in future communications if it does not respond to the PAMP signal. The high priority PAMP signal plays a vital role in this process. The "Activate DC" mode that is switched on due to a selfish node being identified sets in motion the response process. The effect of the PAMP signal in the presence of selfish nodes and its impact on packet loss ratio can be seen in Figure 6. As the PAMP effect to deal with the selfish nodes, the packet loss ratio is reduced.

The source node sends a PAMP signal, $PAMP_{send}$ and each node have to acknowledge receipt by sending back a PAMP receive signal, $PAMP_{recv}$. The selfish node that did not formerly acknowledge receipt of the packet will be forced to respond with a PAMP receive signal, $PAMP_{recv}$, as the PAMP signal is a high priority message and it overwrites the node buffer.

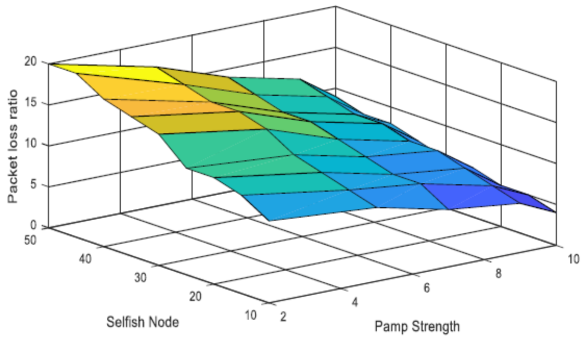


Figure 6: Weight of PAMP signal strength

Once *Node* $a$ obtains $T^{TP}_{a,b}$ for $TP = Safe$ Signal1, Safe Signal 2 and PAMP then $T^{TP}_{a.b}$ is calculated based on Equation (2).

- $T^{SS1}_{a,b}(t)$: Measures the number of times any intermediate (trustee) node generated a route reply. Here a settlor node evaluates the unselfish and honest behavior of the trustee node. This trust component is computed based on the number of interactions between the trustor and trustee node.
- $T^{SS2}_{a,b}(t)$: The trust element is evaluated when the trustee node sends back an acknowledgment of receipt of a packet.
- $T^{PAMP}_{a,b}(t)$: Analysed by observing if the intermediate node received no acknowledgment to the data packet but it did send a route reply earlier, and then the PAMP signal is sent to validate selfish behavior in a node.

The GrAIS utilizes the concept of a Price of Anarchy [9] for load calculation. Consider there are $N$ nodes in the network. In the GrAIS model, nodes that perform routing task employ a trust purpose $T^{TP}_{a,b}$ between any two nodes $a, b$. The Effective Energy of the *Node* $b$ and Trust value of *Node* $b$ as observed by *Node* $a$is taken into consideration. Therefore, the Workload *(WL)* in a routing task undertaken by any *Node* $b$ is

$$WL_b = \frac{1}{EE_{node(b)}T^{TP}_{a,b}(t)} \tag{3}$$

The workload is dependent on the energy of a node or inversely proportional to node energy and trust. Equation (3) shows that as workload increases the nodes expend more energy to carry out networking tasks. As the node energy consumption increases the trust value could reduce.

## 5 Simulation Results and Analysis

The simulations were carried out using NS-3 and MATLAB. Energy-aware workload [12] distribution is the most efficient approach to reduce energy consumption and stimulate cooperation of selfish nodes. In the traditional applications of MANETs, the workloads are very simple and wireless communication is usually the most energy-intensive process. However, as the applications of MANETs become more complex, it becomes necessary to efficiently distribute the workloads by considering both the trust, hop count and communication energy consumption. In this paper, we consider workload in terms of the trust metric and energy consumption during packet transmission. The workload in terms of packet transmission is considered to reveal the tradeoff between sucker (good) nodes and selfish (cheat) nodes. In our simulations, there are three workload scenarios explored with the packet delivery workload increasing from Workload-1 to Workload-3.

Table 1: Simulated parameters

| | |
|---:|---|
| Simulator | Ns-3.23 |
| Mobility Model | Random waypoint |
| Simulation Time | 1000s |
| Number of selfish nodes | 10-50 |
| Number of nodes | 150 |
| Traffic Type UDP Network Area | 300m*1500m |
| Packet size | 130 bytes |
| Mobility | 20 m/s |
| Transmission Range | 50m |

In Figure 7 and Figure 8, it can be observed that initially good nodes maintain trust while the selfish nodes choose to conserve their energy. As the workload is initially light and all nodes have more energy, the unselfish characteristic amongst participating nodes becomes a crucial factor when determining trust. The prominent drop in trust amongst the good nodes was observed at $t = 300$ min when the good nodes had depleted their energy to a point where they began to look for alternative pathways that would conserve the energy of known good nodes. The GrAis model performs better as time progresses due to selfish nodes being forced to co-operate. The selfish node maintains trust for a longer $t$ as it would have conserved energy to this point.

In Figure 8, the trend is similar to Figure 7, except that the time during which good nodes start to show a dip in trust occurs earlier than when it occurred in the GrAIS model. This is due to the increase from Workload-1 to Workload-2. As workload increases the energy consumption would also increase and good nodes would diminish their energy store at a corresponding rate whilst cheat nodes act to retain energy. The GrAIS model facilitates traffic flows using selfish nodes and as a result the GrAIS model is able to function more effectively as time progresses when compared to a model that relies upon good nodes to transfer traffic flows. In Figure 9, a new trend is seen with the cheat nodes acting to conserve energy

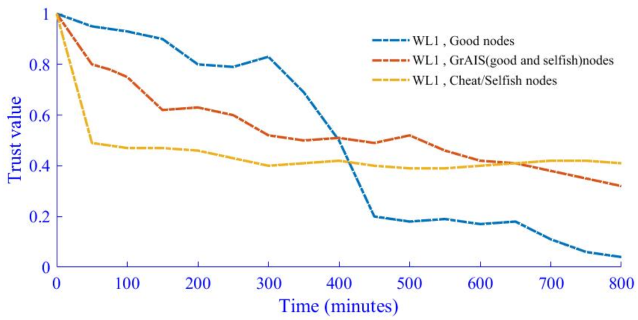


Figure 7: Trust values plotted against time under Workload-1

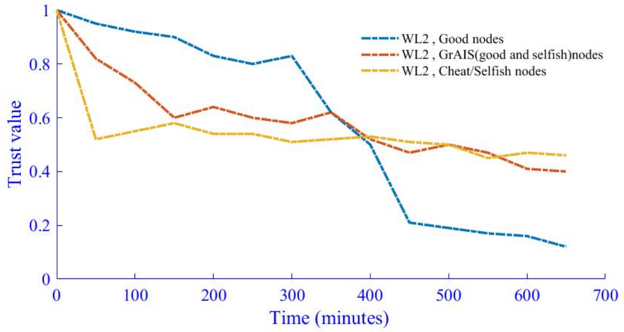


Figure 8: Trust values plotted against time under Workload-2

earlier due to the higher workload and this results in a lack of cooperation from the point where trust dips. The GrAIS model approaches the good node model by forcing the selfish nodes to cooperate with the help of the high priority PAMP signal.

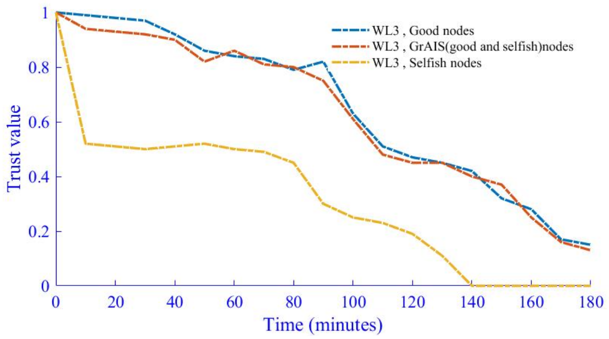


Figure 9: Trust values plotted against time under Workload-3

Using a Packet Delivery Ratio (PDR) metric, as shown in Figure 10, we can evaluate the performance of the proposed GrAIS algorithm. The PDR is a performance metric used in MANET to evaluate the performance of a routing protocol. It is the ratio of the number of data packets delivered to the number of packets sent. GrAIS shows an improved packet delivery ratio while SAODV [14] exhibits a decrease in packet delivery ratio as the number of selfish nodes increases while CITA-AODV [15] follows closely behind GrAIS.

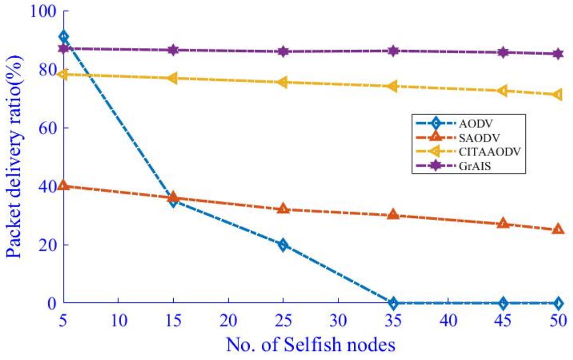


Figure 10: Packet delivery ratio v/s number of selfish nodes

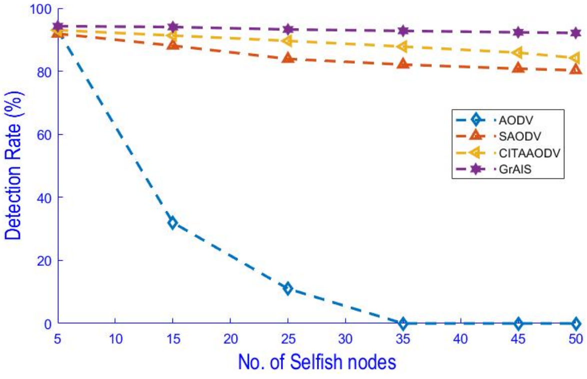


Figure 11: Detection ratio v/s number of selfish nodes

The detection rate has been compared against the number of selfish nodes in the network. A detection rate of 93.41% was achieved while for SAODV the detection rate achieved was 85.34%. This shows that as the number of selfish nodes increases, GrAIS is able to detect the selfish nodes due to its better trust framework.

# 6 Conclusion and Future Work

The GrAIS model approach presented in this paper shows that selfish nodes need not be identified and isolated as there should be an opportunity to force the selfish nodes to participate using high priority signals thereby spreading the load and resource utilization. It is important for network survivability and successful task completion that all MANET nodes cooperate and participate. In some scenarios, it is deemed necessary to include selfish nodes by forcing them to cooperate instead of overlooking their selfish behavior or isolating them from the MANET, as this would reduce the opportunity to leverage this resource. The GrAIS model utilizes the principles of AIS and probability to create a model incorporating good and selfish nodes to combat selfishness in MANET. The results obtained from the simulations have shown that the GrAIS model outcomes are an improvement over models that ignore or isolate selfish nodes as time progresses in spite of increasing workload. A balance between energy utilization, due to good nodes transferring traffic, and energy conservation, due to selfish nodes refusing to transfer traffic, has been achieved by forcing selfish nodes to participate at an appropriate point in the MANET life cycle. A MANET that combines selfishness and unselfishness can be shown to be beneficial when resources, particularly energy, become limited. As future work, a more complex model could be developed exclusively for higher workloads by considering the stability of the GrAIS model over a longer time interval.

## References

[1] C. Perkins, E. Belding-Royer, and S. Das, *Ad hoc On-Demand Distance Vector (AODV) Routing*, RFC, 2003.

[2] S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehavior in mobile ad hoc networks," in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, pp. 255-265, 2000.

[3] Y. Yoo and D. P. Agrawal, "Why does it pay to be selfish in a MANET?," *IEEE Wireless Communications*, vol. 13, pp. 87-97, 2006.

[4] D. G. Kampitaki, E. D. Karapistoli, and A. A. Economides, "Evaluating selfishness impact on MANETs," in *International Conference on Telecommunications and Multimedia (TEMU'14)*, pp. 64-68, 2014.

[5] S. Buchegger and J. Y. Le Boudec, "Nodes bearing grudges: Towards routing security, fairness, and robustness in mobile ad hoc networks," in *10th Euromicro Workshop on Parallel, Distributed and Network-based Processing*, pp. 403-410, 2002.

[6] S. Gupta, C. Nagpal, and C. Singla, "Impact of selfish node concentration in MANETs," *International Journal of Wireless & Mobile Networks*, vol. 3, pp. 29-37, 2011.

[7] M. T. Tran and V. Simon, "Can altruism spare energy in ad hoc networking?" in *Proceedings of the 9th International Conference on Advances in Mobile Computing and Multimedia*, pp. 214-217, 2011.

[8] L. E. Jim and M. A. Gregory, "State analysis of Mobile Ad Hoc Network nodes," in *International Conference on Telecommunication Networks and Applications Conference*, pp. 314-319, 2015.

[9] L. E. Jim and M. A. Gregory, "Packet Storage Time attack-a novel routing attack in Mobile Ad hoc Networks," in *26th International Conference on Telecommunication Networks and Applications Conference*, pp. 127-132, 2016.

[10] R. Dawkins, *The Selfish Gene*, New York: Oxford, 2006.

[11] W. Yu, Y. Huang, and A. Garcia-Ortiz, "Modelling optimal dynamic scheduling for energy-aware workload distribution in wireless sensor networks," in *International Conference on Distributed Computing in Sensor Systems*, pp. 116-118, 2016.

[12] Jøsang, A. and S. Pope, "Semantic constraints for trust transitivity," *Proceedings of the 2nd Asia-Pacific Conference on Conceptual Modelling,* vol. 43, pp. 59-68, 2005.

[13] D. Cerri and A. Ghioni, "Securing AODV: the A-SAODV secure routing prototype," *IEEE Communications,* vol. 46, no. 2, 2008.

[14] A. Khannous, A. Rghioui, F. Elouaai, and M. Bouhorma, "Securing MANETS using the integration of concepts from diverse immune theories," *Journal of Theoretical and Applied Information Technology*, vol. 88, pp. 35, 2016.

[15] K. D. Elgert, *Immunology: Understanding the Immune System*, John Wiley & Sons, 2009.

[16] M. Abdelhaq, R. Hassan, M. Ismail, R. Alsaqour, D. Israf, "Detecting sleep deprivation attack over MANET using a danger theory-based algorithm," *International Journal on New Computer Architectures and Their Applications*, vol. 1, pp. 534-541, 2011.

[17] S. Forrest, A. S. Perelson, L. Allen, R. Cherukuri, "Self-non self discrimination in a computer," *IEEE Symposium on Research in Security and Privacy,* pp. 202-212, 1994.

[18] J. Timmis, M. Neal, J. Hunt, "An artificial immune system for data analysis," *Biosystems,* vol. 55, pp. 143-150, 2000.

[19] A. O. Tarakanov, V. A. Skormin, S. P. Sokolova, *Immunocomputing: Principles and Applications*, Springer, New York, 2003.

[20] R. Steinman and Z. Cohn, "Identification of a novel cell type in peripheral lymphoid organs mice," *The Journal of Experimental Medicine*, vol. 137, no. 5, pp. 1142-1162, 1973.

[21] M. L. Kapsenberg, "Dendritic-cell control of pathogen-driven T-cell polarization," *Nature Reviews Immunology*, vol. 3, pp. 984-993, 2003.

[22] T. Jamie and U. Aickelin, "Towards a conceptual framework for innate immunity," in *3rd International Conference on Artificial Immune Systems*, pp. 112-125, 2004.

[23] J. Greensmith and U. Aickelin, "The deterministic dendritic cell algorithm," in *7th International Conference on Artificial Immune Systems*, pp. 291-302, 2008.

[24] J. Greensmith, U. Aickelin, J. Twycross, "Articulation and clarification of the dendritic cell algorithm," in *5th International Conference on Artificial Immune Systems*, pp. 404-417, 2006.

[25] J. Twycross and U. Aickelin, "Libtissue-implementing innate immunity," in *IEEE Congress on Evolutionary Computation*, pp. 499-506, 2006.

[26] P. Matzinger, "The danger model: A renewed sense of self," *Science*, vol. 296, no. 5566, pp. 301-305, 2002.

[27] P. Bretscher and M. Cohn, "A theory of self-non self discrimination," *Science*, vol. 169, pp. 1042-1049, 1970.

[28] U. Aickelin and S. Cayzer, "The danger theory and its application to artificial immune systems," in *1st International Conference on Artificial Immune Systems (ICARIS'02)*, pp. 141-148, 2002.

[29] P. K. Suri and K. Taneja, "Exploring selfish trends of malicious mobile devices in MANET," arXiv preprint arXiv: 1005.5130, 2010.

## Biography

**Lincy Elizebeth Jim** received a PhD from School of Engineering RMIT University, Melbourne, Australia in 2017; she is a recipient of the Student Travel Award-ITNAC 2016. She has also received the Juniper Networks Certified Associate (JNCIA-Junos) certification. She has more than 5 years of working experience as an Oracle Technical Analyst. She received her Master degree in Electronics and Communication Engineering from National Institute of Technology, Calicut, India in 2007 and her Bachelor of Electronics and Communication Engineering from Cochin University of Science and Technology in 2003.

**Mark A Gregory** (SM'99) is a Fellow of the Institute of Engineers Australia and a Senior Member of the IEEE. Mark A Gregory received a PhD from RMIT University, Melbourne, Australia in 2008, where he is an Associate Professor in the School of Engineering. In 2009, he received an Australian Learning and Teaching Council Citation for an outstanding contribution to teaching and learning. He is the Managing Editor of two international journals (AJTDE and IJICTA) and the General Co-Chair of ITNAC. Research interests include telecommunications, network design and technical risk.

# Efficient Implementation of Password-based Authenticated Key Exchange from RLWE and Post-Quantum TLS

Xinwei Gao[1], Jintai Ding[2], Lin Li[1], Saraswathy RV[2], and Jiqiang Liu[1]
*(Corresponding author: Jintai Ding and Lin Li)*

Beijing Key Laboratory of Security and Privacy in Intelligent Transportation, Beijing Jiaotong University[1]
No.3 ShangYuanCun, Haidian District, Beijing, 100044, P.R.China
Department of Mathematical Sciences, University of Cincinnati[2]
French Hall, 2815 Commons Way, Cincinnati, Ohio 45219, United States
(Email: jintai.ding@gmail.com; lilin@bjtu.edu.cn)

*(Received Aug. 6, 2017; revised and accepted Nov. 21, 2017)*

## Abstract

Two post-quantum password-based authenticated key exchange (PAKE) protocols were proposed at CT-RSA 2017. Following this work, we give much more efficient and portable C++ implementation of these two protocols. We also choose more compact parameters providing 200-bit security. Compared with original implementation, we achieve 21.5x and 18.5x speedup for RLWE-PAK and RLWE-PPK respectively. Compare with quantum-vulnerable J-PAKE protocol, we achieve nearly 8x speedup. We also integrate RLWE-PPK into TLS to construct a post-quantum TLS ciphersuite. This allows simpler key management, mutual authentication and resistant to phishing attack. Benchmark shows that our ciphersuite is indeed practical.

*Keywords: Authenticated Key Exchange; Implementation; Post-quantum; RLWE; TLS*

## 1 Introduction

### 1.1 Key Exchange and Post-Quantum World

With the groundbreaking work "New Directions in Cryptography" from Diffie and Hellman in 1976 [9], the idea of key exchange (KE) and public key cryptography come into reality. Key exchange is a very important cryptographic primitive. With properly designed protocols, two or more parties can agree on same session key for message or data encryption using symmetric encryption algorithms over adversary-controlled network. Well known protocols including Diffie-Hellman key exchange (DH), elliptic curve DH (ECDH) *etc.* However, these protocols cannot authenticate user's identity, *i.e.*, man-in-the-middle (MITM) attack can compromise the security of communication. Authenticated key exchange (AKE) is a solution to this problem. AKE protocols can negotiate key and authenticate identity of communicating parties simultaneously. Well studied protocols including HMQV [20], authenticated DH *etc.* As long as each party has certified public key, two parties can use various explicit or implicit authentication mechanisms to verify the identity of communicating party. The widely-deployed approach is combining unauthenticated key exchange with digital signatures and trusted certificates [21]. This is also known as Public Key Infrastructure (PKI)-based authentication.

Another important line of AKE is password-based AKE (PAKE). PAKE utilizes human-memorable password (or passphrase) which is cryptographically insecure to authenticate and negotiate symmetric session key. PAKE is very strong in the sense that user does not reveal passwords to others [7]. In PAKE, password (or its variant) is pre-shared by both parties. Since only communicating parties know this password, it can be intuitively used as authentication mechanism. Advantages of PAKE include simpler key management since PAKE does not rely on certificates and signatures to authenticate, secure against webpage spoofing, phishing and man-in-the-middle (MITM) attacks since attacker does not know the password, user-friendly authentication (using human-memorable password), prevents offline dictionary attack *etc.* Plenty of PAKE protocols have been standardized and deployed in various applications. Examples and real-world applications of PAKE include PAK & PPK [5], Password Authenticated Key Exchange by Juggling (J-PAKE) [16], secure Pre-Shared Key (PSK) authentication for Internet Key Exchange (IKE) protocol in RFC 6617 [17], elliptic curve J-PAKE ciphersuites in TLS [8], Secure Remote Password protocol (SRP) in RFC 2945 [29] and patented protocols like Encrypted Key Exchange (EKE) [3], Simple Password Exponential Key Exchange (SPEKE) [18] *etc.* Also OpenSSL supports J-PAKE and Firefox Sync service adopted J-PAKE for authentication and key exchange. Some other works related to key exchange include [12, 13, 24] *etc.*

With the advent of quantum computers during past decades, people have realized the untapped potential from quantum computers and huge threats to current cryptographic constructions, especially public key algorithms. Two best-known attacks from quantum computers are Shor's algorithm [28] and Grover's algorithm [14]. Shor's quantum algorithm is widely conjectured to be effective against all mainstream public-key algorithms that designed based on integer factorization problem, discrete logarithm problem *etc.*, including RSA, Diffie-Hellman and elliptic curve-based ones. If large quantum computer exists, it is believed that most public key algorithms can be broken very efficiently. For Grover's quantum algorithm, it attacks symmetric encryption algorithms. Result shows that $n$-bit key provides $n/2$-bit security on quantum computers. Quantum brute force can be defeated by doubling key size without switching to new algorithms. Larger key size works for symmetric and hash algorithms but current public key algorithms will be broken regardless of key size. In 2017, D-Wave 2000Q quantum computer breaks the 2000-qubit barrier. Despite there are controversies around D-Wave on whether they are building truly quantum computer or not, they show the potential of building practical and very powerful quantum computers within coming years. In 2015, National Security Agency (NSA) announced plan to switch to quantum-resistant cryptography in near future. At PQCrypto 2016 conference, NIST announced their call for quantum-resistant cryptographic algorithms for future and plans for post-quantum cryptography standards. Therefore, it is imperative to build quantum-resistant and efficient algorithms, implementations and gain real-world deployment.

### 1.2 Contributions

In this paper, we first present a very efficient implementation of two RLWE-based post-quantum password-based authenticated key exchange protocols proposed at CT-RSA 2017 (denoted as PAKE17) [10]. We also choose more compact parameters providing at least 200-bit security. Our implementation achieve 21.5x and 18.5x performance improvement over original implementation of RLWE-PAK and RLWE-PPK protocol respectively. We also compare performance with J-PAKE, which is deployed in real-world applications but vulnerable to quantum computers. Our implementation is more efficient and achieves 8.5x and 7.4x speedup for RLWE-PAK and RLWE-PPK respectively. Benchmark proves that our work is indeed efficient, which is even faster than current widely deployed and quantum-vulnerable PAKE protocol. Our implementation is a portable C++ implementation and does not rely on new instruction set (*eg*, AVX2) to achieve high performance.

Second, we introduce a post-quantum TLS ciphersuite and present our proof-of-concept implementation. We integrate our efficient RLWE-PPK implementation into TLS ciphersuite in a similar way as pre-shared key ciphersuites. Pre-shared password in RLWE-PPK in this context is a pre-shared key. Advantages of our ciphersuite more convenient key management compared with PKI-based authentication, resistant to phishing attacks and mutual authentication. Benchmark of our implementation shows that our post-quantum TLS ciphersuite is truly practical.

### 1.3 Organization

We recall necessary background knowledge in Section 2. In Section 3, we revisit PAKE17 protocol and introduce new parameter choice with security level estimation, much more efficient implementation, comparison with original work and analysis. In Section 4, we introduce our post-quantum TLS ciphersuite based on RLWE-PPK, proof-of-concept implementation, performance and discussions. We conclude the paper in Section 5.

## 2 Preliminary

### 2.1 Post-Quantum Cryptography

Due to Shor's algorithm, major public key cryptosystems nowadays (RSA, Diffie-Hellman, ECDH *etc.*) are no longer secure when large quantum computer is available. Constructions built on these hard problems: integer factorization problem, discrete logarithm problem or elliptic-curve discrete logarithm problem can be broken on sufficient large quantum computer. Although it is hard to predict the exact time that efficient quantum computers can be built, most scientists believe that they will be built within decades.

Post-quantum cryptography refers to designing and building cryptosystems that can resist attacks from quantum computers. Generally, quantum-resistant cryptosystems can be achieved by these approaches: lattice-based, multivariate-based, hash-based, code-based and symmetric ciphers with larger key size. In this paper, we focus on lattice-based ones since they have strong provable security, high efficiency, simple structure and much smaller key sizes compared with other approaches. A lattice is a set of points in an $n$-dimensional space with periodic structure. The lattice $L(b_1, \cdots, b_n) = \sum_{i=1}^{n} x_i b_i : x_i \in Z$ is formed by linear combinations of $n$ linearly independent vectors $b_1, \cdots, b_n \in R^n$. These vectors are called "lattice basis". With the groundbreaking work of Ajtai [1], cryptographic constructions based on lattice come to existence. The security of lattice-based constructions can be reduced to hard lattice problems, including Shortest Vector Problem (SVP), Closest Vector Problem (CVP) *etc.* Till now, no known efficient classical or quantum algorithm can solve these lattice problems. Lattice-based cryptosystem is one of the most promising constructions for post-quantum world.

### 2.2 Learning with Errors and Ring Variant

In 2005, Oded Regev showed a problem called Learning with Errors (LWE) [26]. LWE is proven to be as hard as solving several worst-case lattice problems when LWE instances are properly initiated. Regev and Peikert [25] showed both quantum and classical reductions from LWE to standard lattice problems. Properly chosen error term keeps LWE problem very hard to solve. Decision version of LWE is to distinguish truly uniform generated samples from $b_i = a_i \cdot s + e_i$, where $s$ is secret key, $e_i$ is the error term. Typically, $s$ and $e_i$ are sampled from discrete Gaussian distribution, $a_i$ is uniformly random generated. Search version is to recover secret $s$ given multiple LWE samples. If one can solve LWE problem, then underlying lattice problem can be solved as well due to reduction theorem. Till now, there are no public known efficient algorithms (both on classical and quantum computers) that can solve lattice problems.

Ring Learning with Errors (RLWE) problem [22] is the analogue of LWE in ring setting. The hardness of RLWE can be reduced to hard problems in ideal lattice. Compared with LWE, main attraction of RLWE problem mainly lies on its high efficiency since schemes based on RLWE reduces quadratic overhead in LWE. Solving RLWE problem can be reduced to finding shortest or closest vectors in ideal lattice. Until now, there are no efficient attacks that especially takes advantage of the ring structure, therefore attacks work for LWE may also work for RLWE.

LWE and RLWE are extremely versatile with available cryptography constructions including encryption, signature, key exchange, identity-based encryption, attribute-based encryption, function encryption, homomorphic encryption *etc.*

## 3 Efficient PAKE17 Implementation

In this section, we first briefly revisit the password-based authenticated key exchange proposed at CT-RSA 2017 in [10]. Two protocols in PAKE17 can be regarded as RLWE analogues of classical PAK and PPK protocol. Second, we introduce our efficient and portable C++ implementation for both three-pass explicit authenticated key exchange protocol (RLWE-PAK) and two-pass implicit authenticated key exchange protocol (RLWE-PPK). We also give more compact parameter choice, analysis of security level, efficient implementation, benchmark and comparisons with original work and J-PAKE protocol.

### 3.1 Revisit PAKE17

PAKE17 can be categorized to an important line of AKE protocols that takes advantage of human-memorable password to achieve authentication. Since only communicating parties share same password (or equivalents), it can be utilized to construct password-based AKE protocols. To the best of our knowledge, only [19] and PAKE17 provide post-quantum PAKE solutions. [19] can be viewed as a general lattice-based construction but it is based on Common Reference String (CRS). CRS-based protocols are more complex and weaker in security proofs. PAKE17 is based on Random Oracle Model (ROM) and this work gives two RLWE-based PAKE protocols: RLWE-PAK and RLWE-PPK. This work follows the idea and structure of PAK & PPK [5] but they are constructed based on RLWE key exchange of [11] (denoted as DING12). They also prove the security of both protocols following [3] with new techniques to adapt to RLWE setting. Proof-of-concept implementation shows that two protocols in PAKE17 are efficient. This is the first work that gives practical and provably secure RLWE-based PAKE constructions.

Figure 1 recalls three-pass explicit authenticated protocol RLWE-PAK. Figure 2 recalls two-pass implicit authenticated protocol RLWE-PPK.

### 3.2 New Parameter Choice and Security Estimation

We first choose new and more compact parameters for PAKE17 considering correctness, security and implementation efficiency. Parameters are given as follows: We choose prime $q = 1073479681$ (approximately 30 bits). $q$ can ensure the correctness of PAKE17 by following theorem 3 in Section 3.2 of PAKE17 and it holds $q \equiv 1 \mod 2n$ that can realize NTT efficiently. We also remark that our prime p choice is more compact than original work (modulus $q = 2^{32}-1$ and it is not a prime). $n = 1024$ and standard deviation $\sigma = 8/\sqrt{2\pi} \approx 3.192$ are identical to original work.

We also analyze security level of our parameter choice. We use LWE estimator proposed in [2] to get security level of our parameter choice. LWE estimator gives a thorough security estimation for both LWE and RLWE-based cryptosystems. It evaluates security level of cryptosystems by computing attack complexity of exhaustive search, coded Blum-Kalai-Wassermann algorithm (BKW), lattice reduction, decoding, reducing BDD to unique-SVP, standard and dual embedding attacks *etc.* This estimation approach is considered as one of the most standardized and complete security estimation tool for LWE and RLWE-based cryptosystems. Authors of [2] also suggested it can be used to evaluate RLWE instances since no known attacks take advantage of ring structure.

Commands for executing security estimation script are given as follows:

- load(“https://bitbucket.org/malb/lwe-estimator/

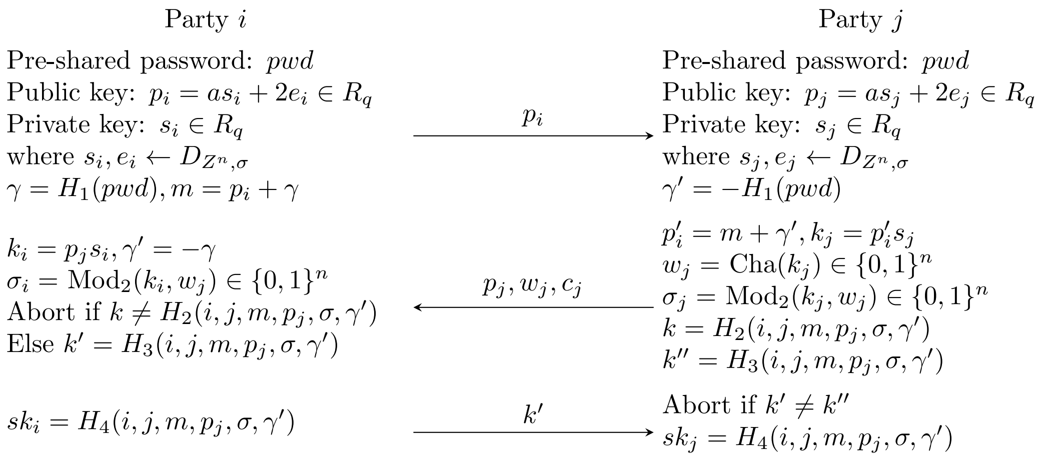

Party $i$
Pre-shared password: $pwd$
Public key: $p_i = as_i + 2e_i \in R_q$
Private key: $s_i \in R_q$
where $s_i, e_i \leftarrow D_{Z^n,\sigma}$
$\gamma = H_1(pwd), m = p_i + \gamma$
$k_i = p_j s_i, \gamma' = -\gamma$
$\sigma_i = \text{Mod}_2(k_i, w_j) \in \{0,1\}^n$
Abort if $k \neq H_2(i, j, m, p_j, \sigma, \gamma')$
Else $k' = H_3(i, j, m, p_j, \sigma, \gamma')$
$sk_i = H_4(i, j, m, p_j, \sigma, \gamma')$
$p_i$
$p_j, w_j, c_j$
$k'$
Party $j$
Pre-shared password: $pwd$
Public key: $p_j = as_j + 2e_j \in R_q$
Private key: $s_j \in R_q$
where $s_j, e_j \leftarrow D_{Z^n,\sigma}$
$\gamma' = -H_1(pwd)$
$p_i' = m + \gamma', k_j = p_i' s_j$
$w_j = \text{Cha}(k_j) \in \{0,1\}^n$
$\sigma_j = \text{Mod}_2(k_j, w_j) \in \{0,1\}^n$
$k = H_2(i, j, m, p_j, \sigma, \gamma')$
$k'' = H_3(i, j, m, p_j, \sigma, \gamma')$
Abort if $k' \neq k''$
$sk_j = H_4(i, j, m, p_j, \sigma, \gamma')$

Figure 1: Explicitly authenticated PAKE17 protocol: RLWE-PAK

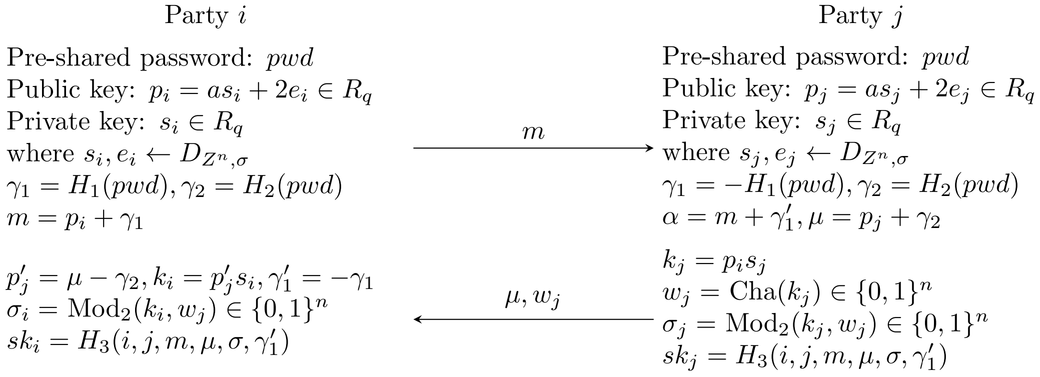

Party $i$
Pre-shared password: $pwd$
Public key: $p_i = as_i + 2e_i \in R_q$
Private key: $s_i \in R_q$
where $s_i, e_i \leftarrow D_{Z^n,\sigma}$
$\gamma_1 = H_1(pwd), \gamma_2 = H_2(pwd)$
$m = p_i + \gamma_1$
$p_j' = \mu - \gamma_2, k_i = p_j' s_i, \gamma_1' = -\gamma_1$
$\sigma_i = \text{Mod}_2(k_i, w_j) \in \{0,1\}^n$
$sk_i = H_3(i, j, m, \mu, \sigma, \gamma_1')$
$m$
$\mu, w_j$
Party $j$
Pre-shared password: $pwd$
Public key: $p_j = as_j + 2e_j \in R_q$
Private key: $s_j \in R_q$
where $s_j, e_j \leftarrow D_{Z^n,\sigma}$
$\gamma_1 = -H_1(pwd), \gamma_2 = H_2(pwd)$
$\alpha = m + \gamma_1', \mu = p_j + \gamma_2$
$k_j = p_i s_j$
$w_j = \text{Cha}(k_j) \in \{0,1\}^n$
$\sigma_j = \text{Mod}_2(k_j, w_j) \in \{0,1\}^n$
$sk_j = H_3(i, j, m, \mu, \sigma, \gamma_1')$

Figure 2: Implicitly authenticated PAKE17 protocol: RLWE-PPK

raw/HEAD/estimator.py”);

- $n$, $\alpha$, $q = 1024$, $f(8, 1073479681)$;
- set_verbose(1);
- _ = estimate_lwe($n, \alpha, q$, skip=[“arora-gb”]).

We remark that our approach is the same as [4]. With above estimation approach, our parameter choice offers at least 200-bit security. Since LWE estimator is constantly updated, our result might slightly different from estimation using latest version of LWE estimator script.

## 3.3 Implementation and Performance

We use NFLlib library [23] to implement RLWE-PAK and RLWE-PPK protocols. NFLlib is a very efficient Number Theoretic Transform (NTT)-based C++ library dedicated to RLWE-based cryptography implementation. It contains various algorithms and programming optimizations for lattice cryptography. NTT, inverse-NTT and sampling from discrete Gaussian operations are very efficient. For fast polynomial multiplication, we adopt NTT and inverse NTT since our parameter $q$ is chosen to instantiate NTT efficiently. Note that NFLlib provides non-constant time implementation and takes advantage of SSE/AVX2 instruction set to optimize NTT and inverse NTT computation.

Our choice for hash function is SHA3-256. We also utilize an extendable-output function (XOF) in our implementation-SHAKE-128. In RLWE-PAK, we have $H_1$ = SHAKE-128($pwd$, 4096), $H_2$ = SHA3-256($C$, $S$, $m$, $\mu$, $\sigma$, $\gamma'$). In RLWE-PPK, we have $H_1$ = SHAKE-128($pwd$, 4096), $H_2$ = SHAKE-128(SHA3-256($pwd$), 4096), $H_i$ = SHA3-256($H_{i-1}$) ($i = 3, 4$). We safely set statistical distance from sampled distribution to discrete Gaussian distribution as $2^{-128}$ to preserve high statistical quality and security. More implementation techniques are introduced in Section 3.4. For Cha() and Mod$_2$(), since input is an polynomial in $R_q$, we take each coefficient in the polynomial and compute its corresponding value.

We test our performance of new implementation, original work and J-PAKE on same server equipped with a 3.4GHz Intel Xeon E5-2687W v2 processor and 64GB memory running 64-bit Ubuntu 14.04. Both PAKE17 implementations are compiled by g++ version 4.8.4 with ’-O3 -fomit-frame-pointer -march=native -m64’ options to achieve maximum performance. Note that this CPU does not support AVX2 instruction set, therefore the performance of our implementation only benefits from SSE instruction set. Recall that the only difference between our parameter choice and PAKE17 is modulus $q$ since we choose same dimension $n$ and standard deviation $\sigma$ for sampling.

We also compare performance of our implementation with J-PAKE, which is known to be vulnerable to quantum computers. J-PAKE implementation we use is later integrated in OpenSSL and is compiled by gcc version 4.8.4 with '-O3 -fomit-frame-pointer -march=native -m64' options. All implementation codes only runs on single core and does not utilizes parallel computing mechanisms. We report average runtime over 10,000 executions of our implementation, original work and J-PAKE in Table 1.

Table 1: Performance of this work, original PAKE17 and J-PAKE implementation

| | | **Client (ms)** | **Server (ms)** |
|---|---|---|---|
| *This work* | RLWE-PAK | 0.176 | 0.175 |
| | RLWE-PPK | 0.203 | 0.203 |
| *Original implementation* | RLWE-PAK | 3.472 (19.7x) | 4.053 (23.2x) |
| | RLWE-PPK | 3.488 (17.2x) | 4.041 (19.9) |
| *J-PAKE* | - | 1.499 | 1.495 |

Number in parentheses is number of times of corresponding runtime for original implementation compared with this work as baseline. Our implementation achieves 21.5x and 18.5x speedup over original implementation for both RLWE-PAK and RLWE-PPK. Compared with J-PAKE, our implementation is 8.5x and 7.4x faster for RLWE-PAKE and RLWE-PPK respectively. This result highlights that our post-quantum PAKE implementation is much more efficient and even much faster than legacy PAKE protocol implementation that is vulnerable to quantum computers.

## 3.4 Discussion

Compared with original PAKE17 implementation, we believe following improvements and optimizations make this work much more efficient:

1) Efficiency from NFLlib library. We use this library to implement major functions related to PAKE17 and this contributes to good performance. NTT, inverse-NTT and sampling computation costs 0.008ms, 0.010ms, and 0.011ms respectively and they are 48x, 45.6x and 13.9x faster than original implementation. There are 2 sampling, 2 NTT and 1 inverse NTT computation operations for RLWE-PAK and RLWE-PPK for both party $i$ and $j$. NTT and inverse NTT functions take advantage the power of SSE instruction set, therefore it is much more efficient. In original work, they use NTL library which is much slower than NFLlib (performance comparison of NFLlib and NTL can be found in [23]). In original work, they adopted FFT for polynomial multiplication and implementation of discrete Gaussian sampler is less efficient. NTT, inverse-NTT and sampling take up around 25% of total running time in our implementation.

2) We adopt a much smarter hashing strategy combined with efficient hashing implementation. We choose SHA3-256 as hash function and SHAKE-128 as XOF. These efforts deliver much faster hashing computations. In original implementation, when adding hashed value $\gamma = \mathrm{H}(pwd)$ to public key $p_i$ $(m = p_i + \gamma)$, they hash a long and randomly generated number using SHA2-256 for $\frac{1024}{256/32} = 128$ rounds. Since SHA2-256 has 256-bit output and each coefficient of polynomial takes up 32-bit, therefore in each SHA2-256 hashing computation round, $\frac{256}{32} = 8$ polynomial coefficients are generated. By repeating 128 times, all 1024 coefficients are derived. We use an alternative and more efficient approach to achieve this. Since output of SHAKE family functions can be extended to any desired length, we adopt SHAKE-128 and extend the output to 4096 bytes to initiate 1024 polynomial coefficients. In our efficient implementation, generate $\gamma = \mathrm{H}(pwd)$=SHAKE-128($pwd$,4096) only costs 0.027ms. There is 1 $\gamma$ computation and 2 $\gamma$ computations in RLWE-PAK and RLWE-PPK respectively. In original work, it costs total of 1.578ms and 3.125ms to generate $\gamma$ for RLWE-PAK and RLWE-PPK respectively, which take up more than 50% of total running time. Our approach is much faster than original implementation by 58.4x (0.027ms) and 57.9x (0.054ms) for RLWE-PAK and RLWE-PPK respectively. Since $\gamma$ computation is very costly, our approach has a significant improvement over original implementation. $\gamma$ computation takes up around 20% of total running time in our implementation.

3) We avoid slower multiplication, division and modular operations by make full use of bitwise operations to boost performance. We also make full use of macros and various programming techniques to reduce unnecessary overhead, expensive memory copy operations *etc.* to improve efficiency. All these approaches also contribute to the efficiency of our implementation.

Moreover, we remark that our implementation is portable. Our code does not take advantage the power of new instruction sets (*eg*, AVX2) to speedup performance, therefore our implementation is portable and can run on more outdated devices.

We note that most time-consuming computation of our PAKE17 implementation is hashing six-tuple $(i, j, m, \mu, \sigma, \gamma')$ using SHA3-256 since they take up more than 8KB of storage. Hashing this tuple costs more time compared with other operations. It costs 0.077ms which takes up 43.75% and 36.95% of total running time for RLWE-PAK and RLWE-PPK respectively.

We believe that PAKE17 and our implementation are indeed practical for post-quantum world.

# 4 Post-quantum TLS Ciphersuite and Implementation

We construct our post-quantum TLS ciphersuite by integrating RLWE-PPK into TLS. It can be regarded as a post-quantum variant of DHE_PSK and ECDHE_PSK. We build our ciphersuite based on DHE_PSK in TLS v1.2 and adapt the notion of password in RLWE-PPK to a key for symmetric encryption which is also pre-shared by two parties (same as definition of PSK in DHE_PSK). Advantages of TLS and PAKE are well-inherited and integrated in our ciphersuite, including easier key management (without PKE-based certificates), mutual authentication, better performance (efficient implementation and no signatures), anti-phishing attack, simplified message flow (inherit from TLS and RLWE-PPK) *etc.* We introduce cryptographic primitive combination of our ciphersuite, implementation based on Botan open source C++ library, benchmarks and discussions.

## 4.1 Introduction

TLS is designed to ensure secure communications over adversary controlled network, providing secrecy and data integrity between two communicating parties. TLS is consisted with two major components: handshake protocol and record protocol. In handshake protocol, two parties negotiate and establish secure connection. In record protocol, two parties transmit encrypted and authenticated data securely. TLS is widely deployed in real world with applications like HTTPS (HTTP over TLS), IMAPS (IMAP over TLS), SMTPS (SMTP over TLS) *etc.* and it has already comprised more than half of total web traffic. It supports various methods for key exchange (Diffie-Hellman and elliptic curve variant, RSA *etc.*), authentication (pre-shared key, RSA, ECDSA *etc.*), encryption (AES, stream ciphers *etc.*) and message authentication code (MAC) algorithms. Two parties can agree on a premaster key using various key exchange algorithms. Other keys are generated through premaster key. For authentication, certificates and signatures are more preferred and widely deployed.

As TLS is so important and we are moving into a post-quantum world, we consider TLS should also adopt post-quantum cryptographic primitives. However, ciphersuites in the latest version of TLS fail to meet the demands since available key exchange and signature algorithms can be broken by quantum computers. Recently, a project called "Open Quantum Safety" (OQS) was launched [27]. It provides prototype open-source software implementation which integrate several unauthenticated post-quantum key exchange protocols into TLS. By combining post-quantum key exchange and classical signatures (RSA, ECDSA *etc.*), OQS project suggests several post-quantum TLS ciphersuites. Google did experiments on integrating a post-quantum key exchange into Chrome browser in canary channel and few Google domain TLS servers with ciphersuite called "CECPQ1" [6]. These works show that post-quantum cryptographic primitives are very promising for real-world applications.

## 4.2 Our Post-Quantum TLS Ciphersuite

Project OQS and Google's effort are based on adopting unauthenticated key exchange protocol with RSA/ECDSA signature to achieve authentication. We want to achieve authentication in a way that discards quantum-insecure signatures. Authenticate two parties using pre-shared key is a very practical and efficient approach as pointed out in [15]. In TLS v1.2, there are various ciphersuites that supports authentication using pre-shared key (PSK), where PSK is a pre-shared symmetric key for encryption and authentication. PSK ciphersuites including standalone PSK, DHE_PSK (Diffie-Hellman ephemeral for key exchange+PSK for authentication, ECDHE_PSK (elliptic curve DHE+PSK), RSA_PSK (RSA for key exchange+PSK) *etc.* Advantages for PSK-based ciphersuites including avoid expensive public key computations for authentication (signing and verifying), mutual authentication, avoid phishing attacks, simpler key management *etc.* It can be established with various methods and it is not the focus of this work.

As suggested in [15], integrating PAKE protocols in TLS does not require too much changes in TLS standards, therefore we can safely take this approach. To the best of our knowledge, there are no existing works that integrate post-quantum PAKE protocols into TLS. Since RLWE-PPK is a two-pass implicit authenticated key exchange protocol and it shares very similar structure with Diffie-Hellman, it can be regarded as post-quantum alternative for pre-shared key ciphersuite in TLS. Analogously, we can replace DHE_PSK in standard TLS with RLWE-PPK without modify structure of TLS and message flow significantly. As mentioned before, PSK is a pre-shared symmetric encryption key for authentication, therefore we pre-share key for both parties in our implementation.

Our post-quantum TLS ciphersuite "RLWE _PPK _WITH _AES _256 _GCM _SHA384" is designed and implemented based on TLS v1.2. We use our efficient implementation of RLWE-PPK to perform key exchange and authentication, thus our ciphersuite can achieve mutual authentication, forward security and resistant to quantum computing attacks. We give detailed cryptographic primitive combination of our post-quantum TLS ciphersuite:

- Key exchange and authentication: We integrate RLWE-PPK protocol revisited in Section 3.1 into TLS to achieve post-quantum key exchange and authentication. This ciphersuite can realize mutual authentication in a more convenient way than PKI-based approach, where client may not have user certificate. Parameter choice for RLWE-PPK follow Section 3.2.

- Authenticated encryption: We choose AES-256-GCM. It provides confidentiality, integrity and authenticity assurances on data.
- Hash function: We choose SHA-384. Our choice followed the principle proposed by NIST of deprecating SHA-1.

### 4.3 Implementation and Performance

Our proof-of-concept implementation is based on Botan library. Botan is a C++ library that provides implementations of a variety of cryptographic algorithms and protocols. We integrate our RLWE-PPK implementation into Botan library according to TLS v1.2 handshake and RLWE-PPK message flow. We also implement test programs that simulate TLS session between client and server using our ciphersuite. Both client and server program are run on localhost. Server listens on port 443 and client communicates with server. We measure runtime from the very beginning of session initiation and stop after handshake completes. Test programs run on a PC equipped with a 2.7GHz Intel Core i7-6820HQ processor and 4GB RAM running Ubuntu 14.04 64-bit version. Test programs are compiled by g++ 4.8.4 with optimization flags "-O3 -m64 -fstack-protector" and execute 1,000 times using single core. Average runtime over 1,000 executions of our ciphersuite for client and server handshake is 4.83ms and 4.94ms respectively.

Although our PAKE-PPK implementation is very efficient, our ciphersuite has larger communication cost. Size of PAKE-PPK key exchange messages from client to server and server to client is 3.75KB and 3.875KB respectively and this is larger than DHE/ECDHE/RSA key exchange messages (around 1-2KB). This might be a disadvantage of our ciphersuite. Also setting up pre-shared materials securely require more works.

## 5 Conclusion

In this paper, we present a portable and truly efficient post-quantum PAKE implementation for RLWE-PAK and RLWE-PPK protocols. We implement these two PAKE protocols in portable C++ style so that our code can run on a variety of devices. We achieve 21.5x and 18.5x performance improvement on both RLWE-PAK and RLWE-PPK over original implementation. Performance of RLWE-PAK and RLWE-PPK in this work is also 8.5x and 7.4x faster than J-PAKE, which is known to be widely deployed but vulnerable to quantum computers. We also integrate RLWE-PPK into TLS as post-quantum TLS ciphersuite. Proof-of-concept implementation based on Botan library shows that our ciphersuite is very practical. Our work shows that post-quantum cryptographic primitives like PAKE17 and real-world post-quantum applications like our post-quantum TLS ciphersuite can be truly efficient.

## Acknowledgement

This work is supported by China Scholarship Council, National Natural Science Foundation of China (Grant No. 61672092) and Fundamental Research Funds for the Central Universities (Grant No. 2017YJS038). Jintai Ding is partially supported by NSF grant DMS-1565748 and US Air Force grant FA2386-17-1-4067.

## References

[1] M. Ajtai, "Generating hard instances of lattice problems," in *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, pp. 99–108, 1996.

[2] M. R. Albrecht, R. Player and S. Scott, "On the concrete hardness of learning with errors," *Journal of Mathematical Cryptology*, vol. 9, no. 3, pp. 169–203, 2015.

[3] S. M. Bellovin and M. J. Merritt, *Cryptographic Protocol for Remote Authentication*, US Patent 5,440,635, Aug. 8, 1995.

[4] J. W. Bos, C. Costello, M. Naehrig and D. Stebila, "Post-quantum key exchange for the TLS protocol from the ring learning with errors problem," in *2015 IEEE Symposium on Security and Privacy*, pp. 553–570, 2015.

[5] V. Boyko, P. MacKenzie and S. Patel, "Provably secure password-authenticated key exchange using diffie-hellman," in *Advances in CryptologyEurocrypt 2000*, pp. 156–171, 2000.

[6] M. Braithwaite, S. Engineer, "Experimenting with post-quantum cryptography" July 7, 2016.

[7] T. Y. Chang, W. P. Yang, M. S. Hwang, "Simple authenticated key agreement and protected password change protocol", *Computers & Mathematics with Applications*, vol. 49, pp. 703–714, 2005.

[8] R. Cragie, F. Hao, "Elliptic curve j-pake cipher suites for transport layer security (tls)_draft-cragie-tls-ecjpake-00," pp. 24, 2016.

[9] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, vol. 22, no. 6, pp. 644–654, 1976.

[10] J. Ding, S. Alsayigh, J. Lancrenon, R.V. Saraswathy and M. Snook, "Provably secure password authenticated key exchange based on rlwe for the post-quantum world," in *Cryptographers Track at the RSA Conference*, pp. 183–204, 2017.

[11] J. Ding, X. Xie and X. Lin, "A simple provably secure key exchange scheme based on the learning with errors problem," *IACR Cryptology ePrint Archive*, vol. 2012, pp. 688, 2012.

[12] M. S. Dousti and R. Jalili, "Forsakes: A forward-secure authenticated key exchange protocol based on symmetric key-evolving schemes," *Advances in Mathematics of Communications*, vol. 9, no. 4, pp. 471–514, 2015.

[13] S. Gonzlez, L. Huguet, C. Martnez and H. Villafae, "Discrete logarithm like problems and linear recurring sequences," *Advances in Mathematics of Communications*, vol. 7, no. 2, pp. 187–195, 2013.

[14] L. K. Grover, "A fast quantum mechanical algorithm for database search," in *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, pp. 212–219, 1996.

[15] P. H. Griffin, "Transport layer secured password-authenticated key exchange," *Information Systems Security Association Journal*, vol. 13, no. 6, 2015.

[16] F. Hao, "J-pake: Password-authenticated key exchange by juggling," in *International Workshop on Security Protocols*, pp. 159-171, 2017.

[17] D. Harkins, "Secure pre-shared key (PSK) authentication for the internet key exchange protocol (IKE)," *Internet Engineering Task Force (IETF'12)*, RFC 6617, 2012.

[18] D. P. Jablon, "Cryptographic methods for remote authentication," May 1, 2001.

[19] J. Katz and V. Vaikuntanathan, "Round-optimal password-based authenticated key exchange," *Journal of Cryptology*, vol. 6597, no. 4, pp. 293–310, 2011.

[20] H. Krawczyk, "Hmqv: A high-performance secure diffie-hellman protocol (extended abstract)," in *Annual International Cryptology Conference*, pp. 546, 2005.

[21] C. C. Lee, M. S. Hwang, L. H. Li, "A new key authentication scheme based on discrete logarithms", *Applied Mathematics and Computation*, vol. 139, no. 2, pp. 343-349, July 2003.

[22] V. Lyubashevsky, C. Peikert and O. Regev, "On ideal lattices and learning with errors over rings," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 1–23, 2010.

[23] C. A. Melchor, J. Barrier, S. Guelton, A. Guinet, M. O. Killijian and T. Lepoint, "Nfllib: NTT-based fast lattice library," in *Cryptographer Track at the RSA Conference*, pp. 341–356, 2016.

[24] G. Micheli, "Cryptanalysis of a noncommutative key exchange protocol," *Advances in Mathematics of Communications*, vol. 9, no. 2, pp. 247–253, 2015.

[25] C. Peikert, "Public-key cryptosystems from the worst-case shortest vector problem," in *Proceedings of the Forty-first Annual ACM symposium on Theory of Computing*, pp. 333–342, 2009.

[26] O. Regev, "On lattices, learning with errors, random linear codes, and cryptography," *Journal of the ACM*, vol. 56, no. 6, pp. 34, 2009.

[27] D. Stebila and M. Mosca, "Post-quantum key exchange for the internet and the open quantum safe project," *IACR Cryptology ePrint Archive*, vol. 2016, pp. 1017, 2016.

[28] P. W. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in *35th Annual Symposium on Foundations of Computer Science*, pp. 124–134, 1994.

[29] T. D. Wu *et al.*, "The secure remote password protocol," in *NDSS*, vol. 98, pp. 97–111, 1998.

# Biography

**Xinwei Gao** received B.S. degree from Beijing Jiaotong University in 2014. He is a Ph.D. student at Beijing Key Laboratory of Security and Privacy in Intelligent Transportation at Beijing Jiaotong University. He is currently a visiting student at University of Cincinnati with sponsorship from China Scholarship Council. His research interest include post-quantum cryptography and RLWE-based key exchange.

**Jintai Ding** received Ph.D. degree in Yale in 1995. He is currently a professor of Mathematics at the University of Cincinnati. He was Humboldt fellow and visiting professor at TU Darmstadt in 2006-2007. He received the Zhong Jia Qing Prize from the Chinese math society in 1990. His research interest lies in post-quantum cryptography (PQC). He was a co-chair of the second international workshop on PQC. He and his colleagues invented LWE and RLWE-based key exchange protocols, Rainbow signature, "GUI" HFEV- signature and Simple Matrix encryption.

**Lin Li** received Ph.D. degree from Shandong University in 2007. She is currently an assistant professor at Beijing Key Laboratory of Security and Privacy in Intelligent Transportation at Beijing Jiaotong University. Her current research interests include cryptography and privacy preserving.

**Saraswathy RV** received bachelor degree in mathematics from University of Madras, India in 2006. She worked as a software quality assurance engineer in information technology industry for a few years till 2011 and is now a Ph.D. candidate in mathematics at University of Cincinnati. Her current research interests include lattice based cryptography, Learning with Errors and key exchange using RLWE.

**Jiqiang Liu** received B.S. and Ph.D. degree from Beijing Normal University in 1994 and 1999 respectively. He is currently a professor at the Beijing Key Laboratory of Security and Privacy in Intelligent Transportation at Beijing Jiaotong University. He is also the vice dean of graduation school of Beijing Jiaotong University. His research interests include security protocols, trusted computing and privacy preserving.

# A Fast Scalar Multiplication Algorithm Based on Alternate-Zeckendorf Representation

Shuang-Gen Liu and Xue-Jing Sun
(Corresponding author: Shuang-Gen Liu)

Department of Information and Communication Engineering, Xi'an University of Posts and Telecommunications
W Chang'an Ave, ChangAnQu XiBu DaXueCheng ShangQuan, Changan Qu, Xi'an, Shaanxi 710121, China
(Email: liusgxupt@163.com )

(Received May 15, 2017; revised and accepted Sept. 25, 2017)

## Abstract

This paper proposed a new method to point scalar multiplication on elliptic curve and it is defined in a finite field with a characteristic greater than 3. It is based on the Transformed Fibonacci type sequence like (2P + Q) and it can resist the Simple Power Attack (SPA). Although the sequence quite easier to calculate, expressing any $k$ using the sequence remains a very difficult problem, so we proposed the Alternate-Zeckendorf representation and given the proof of this view.The NewADD algorithm is also added to the new algorithm and in meanwhile we also listed (2P + Q) results as a table to reduce the computation cost.The performance comparisons show that our algorithm is less costly than other algorithms 12.7 % to 27.9 % at least.

*Keywords: Addition Chain; Elliptic Curve; Fibonacci Sequence; Scalar Multiplication; Simple Power Attack*

## 1 Introduction

Since the Koblit [8] and Miller [15] firstly applied the elliptic curve in the encryption system, Elliptic Curve Cryptography (ECC) has received more and more attention. It gradually become a mainly standard in public key cryptography and widely used in various areas of information security, such as message encryption, authentication and digital signatures [2, 6, 9, 16, 17, 21]. The literature [20] proposed a combination of RSA and ECC. Compared with RSA, the advantages and benefits of the ECC as follows:

- High security. RSA's security level is sub-exponential and ECC is exponential.
- Shorter key. At the same level of security, the secret key length of ECC is much smaller than that of RSA and ElGamal, which makes the ECC is applied in the storage-constrained environments.
- Small storage space,lower bandwidth requirements. This advantage allows ECC have a good prospect in many limited areas.
- The faster computational speed. Due to the small size of the finite field bottom of ECC and the deepening of related research, it is much faster than RSA's computational speed.

The SET agreement which is introduced by the Visa and Master card has set its default public key cryptography algorithm to ECC. That means with the development of ECC it will be gradually replace the status of RSA mainstream application algorithms.

Although the existing ECC is much faster than RSA, with the development of information technology, the existing computing speed has been unable to meet people's need [22]. So it is imperative to improve the computational efficiency of ECC. However, In elliptic curve calculation the most basic and time-consuming operation is the elliptic curve scalar multiplication (ECSM), which calculate the [$k$]P, *i.e.*, the computation of the point $kP = P + P + \cdots + P$, where the integer $k$ is in the known domain, the P is a point on the elliptic curve. The literature [19, 23] introduced a fast scalar multiplication algorithm. In fact, the entire process of computing ECSM on two levels: top level and bottom level. Among them, the top level operation refers to the scalar multiplication is converted to the double and addition operation on the elliptic curve. On the other hand, the bottom level operation refers to through the multiplication, square, inversion, addition operation to achieve double and addition operation on top level. Therefore, there are two aspects of the research of scalar multiplication: The top level seeks the efficient representation of scalar $k$, and the bottom level find the methods achieved the fast calculation of point doubling and addition. Obviously the top level is the operation on the elliptic curve E and the common methods are double-and-add, NAF, sliding window method and so on. The literature [5, 11] were introduced separately improving Miller's Algorithm using NAF, Window NAF algorithm and extend Φ-NAF algorithm. The literature [10] introduced the bottom algorithm on Jacobian coordinates. Our algorithm select the former to research.

In this paper, we will devote to the two levels. Firstly, we proposed a new addition chain based on the deformation of the Fibonacci sequence [12, 13] and prove the theory that any integer can be expressed by the sum of this sequence. Secondly, we presented a new scalar multiplication algorithm based on the transformed Fibonacci sequence and the algorithm form is like (2P + Q). By the method of combining the NewADD algorithm with new algorithm and calculating the results of (2P + Q) form the table we improved the efficiency of the calculation greatly. In addition to that, our algorithm can resist Simple Power Attack (SPA) [7, 18] naturally, SPA is a type of side channel attack discovered by Kocher *et al.* Literature [1] introduced a method that resist SPA using the addition chain.

The paper is organized as follows. Section 2 gives some relevant conception and definition for the ECC. Section 3 is about addition chain definition. Section 4 introduce our new algorithm and some examples to description. Section 5 we compare and analysis with the other algorithms. And the summary is given in Section 6.

# 2 Elliptic Curve Arithmetic

## 2.1 Definition of Elliptic Curve

An elliptic curve E over a finite field $K$ is defined by the equation:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6. \tag{1}$$

where $a_1, a_2, a_3, a_4, a_6 \in K$, and $\Delta \neq 0$. It is defined as

$$\begin{aligned}
\Delta &= d_2^2 d_8 - 8d_4^3 - 27d_6^2 + 9d_2d_4d_6 \\
d_2 &= a_1^2 + 2a_2 \\
d_4 &= 2a_4 + a_1a_3 \\
d_6 &= a_3^2 + 4a_6 \\
d_8 &= a_1^2a_6 + 4a_2a_6 - 4a_1a_3a_4 + a_3a_2^2 - a_4^2.
\end{aligned}$$

When the characteristic of $K$ is not equal 2 or 3, the equation can be written in another form, such that

$$y^2 = x^3 + a_4x + a_6. \tag{2}$$

where $a_4, a_6 \in K$, and $\Delta = -16(4a_4^3 + 27a_6^2) \neq 0$.

In practice, we simplified the formula to

$$y^2 = x^3 + ax + b. \tag{3}$$

where $a, b \in K$ and $4a^3 + 27b^2 \neq 0$, over the characteristic greater than 3. The set of points of $E(K)$ is an Abelian group.

## 2.2 Addition on Elliptic Curve

People proposed several algorithms to compute the addition of two points and also found the many coordinate systems to improve the addition efficiency, you can refer to [4]. Algorithm 1 is executed in the Jacobian coordinates.

**Algorithm 1** EACCD

```
Input: $P = (X_1, Y_1, Z_1)$ , $Q = (X_2, Y_2, Z_2)$
Output:$P + Q$
$A \leftarrow X_1Z_2^2, B \leftarrow X_2Z_1^2, C \leftarrow Y_1Z_2^2, D \leftarrow Y_2Z_1^2$
$E \leftarrow B - A, F \leftarrow D - C$
$X_3 \leftarrow F^2 - E^3 - 2AE^2$
$Y_3 \leftarrow F(AE^2 - X^3 - CE^3)$
$Z_3 \leftarrow Z_1Z_2E$
Return$(X_3, Y_3, Z_3)$
```

If one of the point given the form like $(X, Y, 1)$, we can obtain the addition cost is 12 multiplications(M) and 4 square(S), the cost of the doubling is 4 multiplications(M) and 6 square(S).

However, if we assume that P and Q sharing the same Z-coordinate, P $= (X_1, Y_1, Z)$ and Q $= (X_2, Y_2, Z)$. Then P+Q$=(X_3, Y_3, Z_3) = (X_3^{'}Z^6, Y_3^{'}Z^9, Z3^{'}Z^3) \sim (X_3^{'}, Y_3^{'}, Z_3^{'})$. Where $A = (X_2 - X_1)^2, B = X_1A, C = X_2A, D = (Y_2 - Y_1)^2$, therefore

$$\begin{aligned}
X_3^{'} &= D - B - C \\
Y_3^{'} &= (Y_2 - Y_1)(B - X_3) - Y_1(C - B) \\
Z_3^{'} &= Z(X_2 - X_1)
\end{aligned}$$

The cost is reduced to 5M+2S.

# 3 Addition Chain Theory

**Addition Chain:** The addition chain is defined as a sequence $v = (v_1, \cdots, v_l)$, where $v_1 = 1, v_l = k, v_i = v_{i-1} + v_{i-2} (1 \leq i \leq l)$. And the l is a length of the addition chain [3].

Euclidean Addition Chain: The Euclidean addition chain (EAC) of $k$ is defined as an addition chain which satisfies $v_1 = 1, v_2 = 2, v_3 = v_1 + v_2$ and $\forall 3 \leq i \leq s - 1$, if $v_i = v_{i-1} + v_j$, some $j < i - 1$, then $v_{i+1} = v_i + v_{i-1}$ or $v_{i+1} = v_i + v_j$ [14].

**Theorem 1.** *We denotes the average number of steps to compute gcd*(k, g) *using the subtraction Euclid's algorithm when* g *is uniformly distributed in the range* $1 \leq \mathrm{g} \leq \mathrm{k}$ *[14].*

$$S(\mathrm{k}) = 6\pi^{-2}(ln\mathrm{k})^2 + O(log\mathrm{k}(logk)^2). \tag{4}$$

This formula shown that the chain length of the EAC is determined by random $g$ and the chain length is about $(lnk)^2$, the chain length is too long to practical application. Even if we choose the $g$ close to the $k/\phi$, where $\phi = (1 + \sqrt{5}) \div 2$ is the golden section, the average chain length is still nearly 1100. So we need to find the appropriate addition chain, however how to seek the shortest addition chain is NP complete problem and solving this problem is very difficult for us. In order to avoid the effect of $g$, we choose the Fibonacci sequence and the definition is shown as follows.

**Fibonacci Sequence:** The Fibonacci sequence is defined as $F_0 = 0, F_1 = 1, F_i = F_{i-2} + F_{i-1}$ [13].

**Fibonacci Type Sequence:** All sequences satisfied the Fibonacci condition $F_i = F_{i-2} + F_{i-1}$ collectively called Fibonacci type sequence.

We use the example to illustrate the difference between the Fibonacci Sequence and the Fibonacci Type Sequence, the sequence $\{0, 1, 1, 2, 3, 5, 8, \cdots\}$ is the Fibonacci sequence, but the Fibonacci type sequence like $\{4, 5, 9, 14, 23, \cdots\}$, this means that the Fibonacci sequence is a special form of the Fibonacci type sequence. Fibonacci type sequences have a greater range and form. Thus we can compute the Fibonacci type sequence of arbitrary integers.

# 4 A New Algorithm about 2P+Q

In previous section we have presented various addition chain forms. Notice that the Fibonacci type sequence is a special EAC and no gap in the application, and inherited all the advantages of EAC. Based on this idea, we proposed a new Transformed Fibonacci type sequence in this section and introduced how to apply this algorithm in the elliptic curve.

## 4.1 Transformed Fibonacci Type Addition Chain

Transformed Fibonacci Type Addition Chain: A sequence which satisfied the formula $T_i = 2T_{i-2} + T_{i-1}$ $(4 \leq i \leq l)$, where $T_1 = 1$, $T_2 = 2$, $T_3 = 3$. We called this sequence as Transformed Fibonacci Type Addition Chain(TFTAC), marked as T. For example $T = \{1, 2, 3, 7, 13, \cdots\}$.

Alternate-Zeckendorf Representation: Let $k$ be an integer and $T_i (i \geq 0)$ is the Transformed Fibonacci Type (TFT) sequence, the $k$ can be written in the form:

$$k = \Sigma d_i [2T_{i-2} + T_{i-1}]. \tag{5}$$

with $d_i \in \{0, 1\}$. This representation is similar to that the Zeckendorf representation, we called it Alternate-Zeckendorf (A-Z) representation method.

*Proof.* We use the inductive deduction to prove it. For the $i = 1, 2, 3$, there is no doubt that it is correct since these are Transformed Fibonacci definition numbers, for $i = 4$ we get $7 = 2 * 2 + 3$. Now suppose each $i \leq k$ can be represented correctly. If $k+1$ is a Transformed Fibonacci number then we're done. Otherwise hypothesis there exists $j$ and $T_j < k+1 < T_{j+1}$, making the $a = k + 1 - T_j$. Since $a < k$ and $a$ can be represented, so $k + 1 = a + T_j$. At the same time $T_j + a < T_{j+1}$, $a < T_{j+1} - T_j = 2T_{j-1}$. Taking into account the relationship between $2T_{j-1}$ and $T_j$ we obtained that the $a \leq T_j$, now consider the following two cases:

**Case 1:** where $a = T_j$, $a$ is represented as $a = 2T_{j-2} + T_{j-1}$, the $k+1$ can be represented as $a$ and $T_j$.

**Case 2:** where $a < T_j$, so representation $a$ does not contain $T_j$. As a result, $k+1$ can be represented as the sum of $T_j$ and representation $a$.

□

In summary, $k+1$ can be represented with this method. Therefore, for any $k$, it can be represented using A-Z representation.

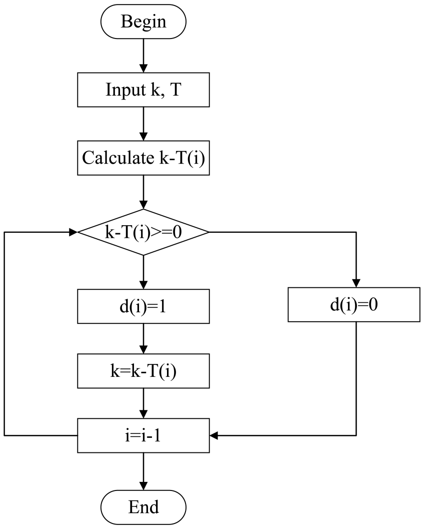


Figure 1: The flow chart of the addition chain

## 4.2 Specific Algorithm

### 4.2.1 Generated the A-Z Representation Sequence

Now, we discuss how to generate a sequence represented by A-Z. Firstly, we set the third number of the sequence is equal to the second number plus the doubling of the first number , in this way we can gain the next result until all the T number are calculated. Then we calculate the addition chain of the $k$. For any integer $k$, we calculate the $k - T_i$, where $i$ is the last element of T. Case 1: if $k - T_i \geq 0$ , then set $k = k - T_i$, mark $d_i = 1$, until the $k = 0$. Case 2: if $k - T_i < 0$, mark $d_i = 0$ then calculate $k - T_{i-1}$, till the $k - T_i \geq 0$ using Case 1. The flow chart is shown in Figure 1.

Algorithm 2 is a specific TFTAC algorithm. Through a large number of experiments we found that the number of T is generally related to the number of the binary bits of $k$. For example, if we select 160-bits integer to

**Algorithm 2** Transformed Fibonacci Addition Chain

```
Input: A positive integer $k$ and $n$
Output: d=$\{d_1, d_2, \cdots, d_n\}$.
$T = \{1, 2, 3\}$, d=$\{\}$,Set $i = 4$
while $i \leq n$ do
    $T_i = T_{i-2} + T_{i-1}$
    Output T.
end while
while $j \geq 1$ and $j \leq n$ do
    if $k - T_n < 0$ then
        Output n is not appropriate, return 1.
    end if
    if $k - T_j \geq 0$ then
        $d_j = 1$
        $d = d \bigcup d_j$
        $k = k - T_j$
        $j = j - 1$
    end if
    if $k - T_j < 0$ then
        $d_j = 0$
        $d = d \bigcup d_j$
        $j = j - 1$
    end if
end while
Output d.
```

execute our experiment, the number of the T is 160 as well. That means the chain length of the $k$ obtained by this method is the same as the binary length. But if we use the Zenkendorf method to represent the $k$, the chain length is about 230. It needs 44 % more digits than the A-Z representation and the binary representation. In fact, we can see that the definition of the A-Z representation is similar to the binary method. Here we use an example to illustrate our algorithm in Example 1.

**Example 1.** $k$ = 567, n = 10, T ={1,2,3}

$T_4 = 2T_2 + T_3 = 7$
$T_5 = 2T_3 + T_4 = 13$
$T_6 = 2T_4 + T_5 = 27$
$T_7 = 2T_5 + T_6 = 53$
$T_8 = 2T_6 + T_7 = 107$
$T_9 = 2T_7 + T_8 = 213$
$T_{10} = 2T_8 + T_9 = 427$
We get the $T = \{1, 2, 3, 27, 53, 107, 217, 427\}$ , then
$k - T_{10} = 567 - 427 > 0, d_{10} = 1, k = 567 - 427 = 140$
$k - T_9 = 140 - 213 < 0, d_9 = 0$
$k - T_8 = 140 - 107 > 0, d_8 = 1, k = 140 - 107 = 33$
$k - T_7 = 33 - 53 < 0, d_7 = 0$
$k - T_6 = 33 - 27 > 0, d_6 = 1, k = 33 - 27 = 6$
$k - T_5 = 6 - 13 < 0, d_5 = 0$
$k - T_4 = 6 - 7 < 0, d_4 = 0$
$k - T_3 = 6 - 3 > 0, d_3 = 1, k = 6 - 3 = 3$
$k - T_2 = 3 - 2 > 0, d_2 = 1, k = 3 - 2 = 1$
$k - T_1 = 1 - 1 = 0, d_1 = 1, k = 1 - 1 = 0$
**Output** $d = \{1010100111\}$

### 4.2.2 Application of Elliptic Curve

In previous section, we discussed how to compute the addition chain sequence of $k$, and now we describe the application of this algorithm on elliptic curves. It is shown in Algorithm 3.

**Algorithm 3** Scalar Multiplication Algorithm using TF-TAC

```
Input: $P \in E(K)$, $k = (d_l, \cdots, d_1)$
Output: W = $[k]P \in E(K)$
Begin W=0
if $d_1 = 1$ then
    W ← W + P
end if
(U, V)←(2P, 3P)
for $i = 2, \cdots, l$ do
    if $d_i = 1$ then
        W←W + U
    end if
    (U,V)←(V, 2U+V)
end for
Return W
```

In Section 2 we know that the NewADD algorithm can be used as long as two points sharing the same Z-coordinates. In the later content, we will use this theory and the specific algorithm is shown in Algorithm 4.

**Algorithm 4** Scalar Multiplication Algorithm using NewADD

```
Input: $P \in E(K)$, $k = (d_l, \cdots, d_1)$
Output: W = $[k]P \in E(K)$
Begin W=0
if $d_1 = 1$ then
    W← P
end if
(U, V)←(2P, 3P)
for $i = 2, \cdots, l$ do
    if $d_i = 1$ then
        upgrade W
        ($\cdots$,W)← NewADD(W,U)
    end if
    Calculate 2U
    upgrade V
    (U,V)← NewADD(2U,V)
end for
($\cdots$,W)← NewADD(W,U)
Return W
```

Now we need to analyze the cost of the algorithm. Supposing the U= $(X_U, Y_U, Z)$ , V= $(X_V, Y_V, Z)$ and P= $(x, y, 1)$. When $d_1$ =1, W= $(x, y, 1)$, upgrade the W= $(xZ^2, yZ^3, Z)$ and the cost is 3M+S. At this time we can add W and U using the NewADD algorithm with the cost is 5M+2S. And then doubling the U needs 4M+4S. Since the $Z_{2U} = 2Y_U Z$, thus the $X_V^{'} = (X_V(2Y_U)^2$, $Y_V(2Y_U)^3, 2Y_U Z)$ only costs 2M. Because of the density of "1" is about 0.5, the add step cost is 4M+1.5S. Finally, the total cost is 15M+7.5S.

**Example 2.** Computation of $[39]P = 27+7+3+2 = (101110)_{A-Z}$:

initialization : W= 0
$d_2 = 1$ : W = 0 + 2P = 2P, (U, V) ← (3P, 7P)
$d_3 = 1$ : W = 2P + 3P = 5P, (U, V) ← (7P, 13P)
$d_4 = 1$ : W = 5P + 7P = 12P, (U, V) ← (13P, 27P)
$d_5 = 0$ : (U, V) ← (27P, 53P)
$d_6 = 1$ : W = 12P + 27P = 39P
Return W = [39]P

**Example 3.** Computation of $[67]P = 53+13+1 = (1010001)_{A-Z}$:

initialization : W= 0
$d_1 = 1$ : W = 0 + P = P
$d_2 = 0$ : (U, V)← (3P, 7P)
$d_3 = 0$ : (U, V)← (7P, 13P)
$d_4 = 0$ : (U, V)← (13P, 27P)
$d_5 = 1$ : W = P + 13P = 14P , (U, V)← (27P, 53P)
$d_6 = 0$ : (U, V)← (53P, 107P)
$d_7 = 1$ : W = 14P + 53P = 69P
Return W = [67]P

Example 2 and Example 3 are illustrated for Algorithm 4. In Algorithm 4, the elliptic curve scalar multiplication process has a double and a addition each time.

#### 4.2.3 Improve Algorithm

By the large of statistical analysis, we get the density of "1" in the A-Z representation is approximate 0.5. That means representing a 160-bits integer needs 80 Transformed Fibonacci number and the total multiplication is about 3360. In this part we will introduce how to improve Algorithm 4 to reduce the cost of multiplication. We known that each iteration of the algorithm needs to calculate 2P + Q, so we can calculate the results of 2P + Q in advance and recorded it as a table. That means we do not need to calculate 2P + Q in the algorithm. The part table shown in Table 1.

Table 1: Transformed fibonacci type number

| Tth num. | $T_1$ | $T_2$ | $T_3$ | $T_4$ | $T_5$ |
|---|---|---|---|---|---|
| TFT num. | 1 | 2 | 3 | 7 | 13 |
| Tth num. | $T_6$ | $T_7$ | $T_8$ | $T_9$ | $\cdots$ |
| TFT num. | 27 | 53 | 107 | 213 | $\cdots$ |

Now, we recalculate Example 3 using improved Algorithm. As you can see in Example 4 only the $d_i = 1$ is performed, and the amount of computation is reduced by half.

**Example 4.** Computation of $[67]P = 53+13+1 = (1010001)_{A-Z}$:

initialization : W= 0
$d_1 = 1$ : TFT num.=1, then W = 0 + P = P
$d_2 = 0$ : Next
$d_3 = 0$ : Next
$d_4 = 0$ : Next
$d_5 = 1$ : TFT num.=13, then W = P + 13P = 14P
$d_6 = 0$ : Next
$d_7 = 1$ : TFT num.=53, then W = 14P + 53P = 69P
return W = [67]P

## 5 Comparison with Other Algorithms

In this section we will compare with other algorithms and give some practical results about new algorithm. We choose mixed coordinates to calculate the cost of the TFTAC and the cost of the various mixed coordinates is shown in Table 2 [4]. In addition, we will also introduce how to resist the SPA attacks using the new algorithm.

Where $A$ is the Affine coordinates, the $J$ is Jacobian, $J^c$ is Chudnovsky Jacobian, $J^m$ is modified Jacobian, the $P$ is Projective coordinate.

### 5.1 Scalar Multiplication Analysis

In Table 3, We compared our algorithm with others, such as NAF, 4-NAF and Double-and-add and so on on mixed coordinates. In Table 4, we compared several algorithms which used the NewADD. Table 5 shown the chain length of several algorithms.

In Section 3, we know that there is unnecessary to calculate the cost of (2P+Q). What we just need to do is to calculate the remaining addition when 1 appears. So the average cost is 4M+1.5S, the final multiplications are 832. However, the number of occurrences of 1 dependent on the specific number, in order to illustrate the advantages of the TFTAC algorithm clearly, we choose the largest number to do comparison and the cost is 8M+3S, the final field multiplications is 1664. On mixed coordinates, TFTAC is faster at least 6.5 % than NAF, 20.9 % than double-and-add, 12.7 % than GRAC-258 , 21 % DFAC-160. Although slower 3.8 % than 4-NAF, most of the TFT numbers are faster than it.

Table 4 shown the comparisons algorithm using the NewADD algorithm. From 27.9 % more than the Fibonacci-and-add, 20.3 % more than the Signed Fib-and-add and 15.1 % more than the Window Fib-and-add. So we can see the new algorithm significantly reduces the addition computation.

Table 5 shown the chain length comparison of the various algorithms with the 160-bits, the A-Z representation chain length is 160 and it is same as the binary counterpart. And compared with other algorithm chain length the A-Z representation chain length is quite shorter than

Table 2: The cost of various mixed coordinate

| doubling | |
|---|---|
| **operation** | **costs** |
| $2A = J$ | $2[M] + 4[S]$ |
| $2J^m = J$ | $3[M] + 4[S]$ |
| $2A = J^m$ | $3[M] + 4[S]$ |
| $2J^m$ | $3[M] + 5[S]$ |
| $2J^m = J^c$ | $4[M] + 4[S]$ |
| $2J$ | $4[M] + 6[S]$ |
| $2J^c$ | $5[M] + 6[S]$ |
| $2P$ | $7[M] + 5[S]$ |
| **addition** | |
| **operation** | **costs** |
| $P + P$ | $12[M] + 2[S]$ |
| $J^m + J^m$ | $13[M] + 6[S]$ |
| $J + A$ | $8[M] + 3[S]$ |
| $J^m + A = J^m$ | $9[M] + 5[S]$ |
| $J^m + A = J$ | $8[M] + 3[S]$ |
| $J^c + J = J$ | $11[M] + 3[S]$ |
| $J^c + J^c = J^m$ | $11[M] + 4[S]$ |
| $J^c + J^c = J$ | $10[M] + 2[S]$ |
| $J^c + J^c$ | $11[M] + 3[S]$ |
| $J^c + A = J^m$ | $8[M] + 4[S]$ |
| $J^c + A = J^c$ | $8[M] + 3[S]$ |
| $J + A = J^m$ | $9[M] + 5[S]$ |
| $A + A = J^m$ | $5[M] + 4[S]$ |
| $A + A = J^c$ | $5[M] + 3[S]$ |
| $J + J$ | $12[M] + 4[S]$ |
| $J^c + J = J^m$ | $12[M] + 5[S]$ |
| $J^m + J^c = J^m$ | $12[M] + 5[S]$ |

Table 3: Comparison the classical algorithm for the 160-bits

| Algorithm | Costs |
|---|---|
| 4-NAF | 1600 |
| NAF | 1780 |
| Double-and-add | 2104 |
| GRAC-258 | 1907 |
| DFAC-160 | 2016 |

Table 4: Comparisons with different algorithm using the NewADD

| Algorithm | Chain length |
|---|---|
| Fibonacci-and-add | 2311 |
| Signed Fib-and-add | 2088 |
| Window Fib-and-add | 1960 |
| TFTAC | 1664 |

Table 5: Comparisons with the chain length for 160-bit

| Algorithm | Chain length |
|---|---|
| Fibonacci-and-add | 358 |
| Signed Fib-and-add | 322 |
| Window Fib-and-add | 292 |
| Binary representation | 160 |
| Zenkendorf representation | 230 |
| TFTAC | 160 |

others. It is 55 % shorter than the Fibonacci-and-add, 50.3 % shorter than Signed Fib-and-add, 45 % shorter than Window Fib-and-add, and 44 % shorter than Zenkendorf representation.

### 5.2 SPA Analysis

SPA is a technology which is a direct interpretation of energy consumption measured value. The system consumption of energy is different that mainly depending on the instructions executed by the microprocessor. When the microprocessor operation performed at different part of the encryption algorithm, some of the energy consumption of the system is very obvious. With this feature, the attacker can distinguish a single instruction to achieve the purpose of breaking the algorithm.

We already know that the attacker obtained the key information by observing the energy curve changing. Therefore we can utilize the method of fixed sequence to resist SPA. Our algorithm based on the Transformed Fibonacci sequence and it is a fixed sequence, therefore it can resist the SPA as well.

## 6 Conclusion

In this paper we proposed a new algorithm TFTAC which is based on the Fibonacci deformation sequence, the new algorithm combined with the advantages of NewADD and through the method of generating tables reduced the cost of scalar multiplication significantly. This method utilized the space exchange for time achieve the purpose of saving resources effectively. Among them the multiplication is less at least 12.7 % than others, the chain length also reduced from 45 % to 55 %. In addition, the algorithm against the SPA as well.

## Acknowledgments

The support of NSFC (National Natural Science Foundation of China, No.61272525), Jiangxi Natural Science Foundation (No.2009GQN0094Scientific Research program funded by Shaanxi Provincial Education Department (No.2013JK1097) and the Youth Foundation of Xi'an University of Posts and Telecommunications (No.103-0454) are gratefully acknowledged.

## References

[1] A. Byrne, N. Meloni, F. Crowe, W. P. Marnane, A. Tisserand, and E. M. Popovici, "Spa resistant elliptic curve cryptosystem using addition chains," in *Fourth International Conference on Information Technology (ITNG'07)*, pp. 995–1000, 2007.

[2] K. Chatterjee, A. De, and D. Gupta, "Mutual authentication protocol using hyperelliptic curve cryptosystem in constrained devices," *International Journal of Network Security*, vol. 15, no. 1, pp. 9–15, 2013.

[3] D. V. Chudnosky and G. V. Chudnovsky, "Sequences of numbers generated by addition in formal groups and new primarily and factorization tests," *Advances in Applied Mathematics*, vol. 7, no. 4, pp. 385–434, 1986.

[4] H. Cohen, A. Miyaji, and T. Ono, "Efficient elliptic curve exponentiation using mixed coordinates," *Lecture Notes in Computing Science*, vol. 1514, pp. 51–65, 1998.

[5] S. Ezzouak, M. E. Amrani, and A. Azizi, "Improving miller's algorithm using the naf and the window NAF," *Lecture Notes in Computer Science*, vol. 7853, pp. 279–283, 2013.

[6] G. Hou and Z. Wang, "A robust and efficient remote authentication scheme from elliptic curve cryptosystem," *International Journal of Network Security*, vol. 19, no. 6, pp. 904–911, 2017.

[7] F. Jia and D. Xie, "A unified method based on spa and timing attacks on the improved RSA," *China Communications*, vol. 13, no. 4, 2016.

[8] N. Koblitz, "Elliptic curve cryptosystems," *Mathematics of Computation*, vol. 48, no. 177, pp. 203–209, 1987.

[9] F. Laguillaumie and D. Vergnaud, "Time-selective convertible undeniable signatures with short conversion receipts," *Information Sciences*, vol. 180, no. 12, pp. 2458–2475, 2010.

[10] Z. X. Lai and Z. J. Zhang, "Research on elliptic curve bottom level algorithm in jacobian coordinate system," *Bulletin of Science and Technology*, vol. 31, no. 10, pp. 244–248, 2015.

[11] T. C. Lin, "Algorithms on elliptic curves over fields of characteristic two with non-adjacent forms," *International Journal of Network Security*, vol. 9, no. 2, pp. 117–120, 2009.

[12] S. G. Liu and Y. P. Hu, "Fast and secure elliptic curve scalar multiplication algorithm based on special addition chains," *Journal of Southeast University*, vol. 24, no. 1, pp. 29–32, 2008.

[13] S. G. Liu, G. L. Qi, and X. A. Wang, "Fast and secure elliptic curve scalar multiplication algorithm based on a kind of deformed fibonacci-type series," in *10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC'15)*, pp. 398–402, 2015.

[14] N. Meloni, "New point addition formulae for ecc applications," *Lecture Notes in Computing Science*, vol. 4547, 2007.

[15] V. S. Miller, "Use of elliptic curves in cryptography," *Lecture Notes in Computer Science*, vol. 218, pp. 417–426, 1986.

[16] V. S. Naresh and N. V. E. S. Murthy, "Elliptic curve based dynamic contributory group key agreement protocol for secure group communication over ad-hoc networks," *International Journal of Network Security*, vol. 17, no. 3, pp. 328–339, 2015.

[17] B. Rashidi, S. M. Sayedi, and R. R. Farashahi, "Efficient and low-complexity hardware architecture of gaussian normal basis multiplication over GF(2m) for elliptic curve cryptosystems," *IET Circuits, Devices and Systems*, vol. 11, no. 2, 2017.

[18] Reddy and E. Kesavulu, "Elliptic curve cryptosystems and side-channel attacks," *International Journal of Network Security*, vol. 12, no. 3, pp. 151–158, 2011.

[19] Y. Sakemi, T. Izu, and M. Shirase, "Faster scalar multiplication for elliptic curve cryptosystems," in *16th International Conference on Network-Based Information Systems*, pp. 523–527, Sep. 2013.

[20] A. Thomas and M. Manuel, "Embedment of montgomery algorithm on elliptic curve cryptography over rsa public key cryptography," *International Conference on Emerging Trends in Engineering, Science and Technology*, vol. 24, pp. 911–917, 2016.

[21] M. Toorani and A. A. B. Shirazi, "Cryptanalysis of an elliptic curve-based signcryption scheme," *International Journal of Network Security*, vol. 10, no. 1, pp. 51–56, 2010.

[22] S. F. Tzeng, M. S. Hwang, "Digital signature with message recovery and its variants based on elliptic curve discrete logarithm problem", *Computer Standards & Interfaces*, vol. 26, no. 2, pp. 61–71, 2004.

[23] D. Yong, Y. F. Hong, W. T. Wang, Y. Y. Zhou, and X. Y. Zhao, "Speeding scalar multiplication of elliptic curve over GF(2)," *International Journal of Network Security*, vol. 11, no. 10, pp. 70–77, 2010.

## Biography

**Shuang-Gen Liu** was born in 1979, associate professor. He graduated from Xidian University in 2008 with a major in cryptography, PhD, a member of the Chinese Institute of computer science, and a member of the Chinese code society.

**Xue-Jing Sun** is a graduate student of Xi'an University of post and telecommunications. She is mainly engaged in the research of elliptic curve cryptosystem.

# A Wireless Key Generation Algorithm for RFID System Based on Bit Operation

Rui Xie[1,2], Jie Ling[1], and Dao-Wei Liu[1]

(Corresponding author: Rui Xie)

School of Computers, Guangdong University of Technology[1]
No. 100, Waihuan Xi Road, Panyu District, Guangzhou 510006, China
School of Automation, Guangdong University of Technology, Guangzhou 510006, China[2]
(Email: xr_1977@126.com)

(Received Oct. 19, 2017; Revised and Accepted Mar. 18, 2018)

## Abstract

Radio Frequency Identification (RFID) system consists of three parts: tags, reader and database. Because of the advantages of low cost, long life and easy deployment, RFID technology has been widely used in logistics, supply chain management system and other fields. With the widespread use of RFID technology, a number of security problems occur. The prerequisite for further promotion of RFID technology is to ensure the security of shared keys stored between legitimate readers and tags. However, because the tags and reader communicate in wireless channel, the shared key written directly to the tag by the reader is easily intercepted by the attacker. Because the computing power and the storage space of tags in RFID are limited, the key generation protocols based on cryptographic can not be used in RFID systems. In addition, the shared key written by the tag manufacturer will result in a shared key escrow problem, and the user can not customize the shared key as well. For the above reasons, the secure generation of the shared key in RFID system is somewhat difficult. The paper abandons the general way that the shared private key has been set to the tag before leaving the factory, then uses the method of dynamically generating the shared key in the application between the tag and the reader through the wireless communication, and proposes a wireless generation algorithm WKGA-BO (Wireless Key Generation Algorithm for RFID System Based on Bit Operation) for private key in RFID system based on bitwise operation. WKGA-BO uses the bit replacement operation and the self-combinatorial cross-bit operation to encrypt the transmission information. The random number is generated by the reader to ensure the freshness of the shared key, and the methods of generating the random number on the tag side is discarded, thus reducing the computational complexity. Finally, a comprehensive security analysis demonstrates the security and reliability of WKGA-BO. A comprehensive performance analysis shows the feasibility of WKGA-BO in the existing RFID system.

*Keywords: RFID; Sac Internet of Things; Wireless Key Generation*

## 1 Introduction

RFID is a technology that uses non-physical contact to achieve object recognition and data exchange. It emerged in the last century, and had been widely promoted and applied in the late nineties of the $20^{th}$ century [16, 20]. Radio Frequency Identification (RFID) system consists of three parts: tags, reader and database. Because of the advantages of low cost, long life and easy deployment, RFID technology has been widely used in logistics, supply chain management system and other fields [17, 21].

RFID tags are composed of coupling elements and chips, and are embedded in the object, which is used to identify the target object. According to the tag's own characteristics and operation modes, it can be divided into two categories: one is active tag; another is passive tag [1, 27]. Active tag generally carries its own batteries, so it doesn't need to use the energy of the reader. It can also take the initiative to the reader to send information, to complete some of the work independently. Compared to the passive tag, the active tag works much larger. However, because the active tag carry its own power, there are disadvantages of high production cost, too large volume, heavy quality, excessive power consumption and so on, which restricts the development of such tags [6, 18]. Because the passive tag doesn't carry its own power, and the energy required in the work process can be only obtained through the carrier sent by the reader, so the work scope of this tag is very small [19, 23].

In the existing RFID systems, the most commonly used tags are passive tags. Based on the above description, it is known that the way that the passive tag works will easily expose the information stored in the tags and even expose the privacy information stored on the reader. In order to solve the above problems, privacy protection certification has become the most commonly used security mechanism between the tag and reader, whose main idea is using a shared private key for bidirectional authentication and identification between the tag and the reader [7, 22, 30].

Privacy protection is mainly based on the security and reliability of the shared private key, but the traditional method of generating the shared private key has some flaws. The traditional method of generating the shared private key is that the producer defaults the shared key to the tag, but it'll cause the tag escrow problems, and the user can not generate his own trusted shared private key according to his own needs [9, 28]. It is very challenging to securely generate keys on RFID tags. In addition to the reasons described above, there are the following factors. First, passive tag does not have a physical interface, so it can not be connected with other devices, and it can not be used to generate the shared private key through the physical connection [29]. Second, if the reader directly write the shared private key to the tag through the wireless, then because the reader can read and write a larger scope, it's easier for attackers to obtain the shared private key by the means of eavesdropping [12]. Third, the computing ability of the tag is very limited, so the more complex traditional cryptographic calculations are disabled, which makes the existing shared private key generation protocol based on cryptographic unable to be applied to the tag [14].

In this paper, we innovatively propose an algorithm WKGA-BO to generate the shared private key wirelessly based on bitwise operation in the RFID system for the problems described above. The basic idea of WKGA-BO is as follows. The tag which does not initialize the shared private key sends the fragment of shared private key to the reader. When the tag receives another part of the shared private key fragment generated by the reader, the tag can generate the shared private key by mixing the key fragments. The information about the shared private key fragment between the tag and the reader is transmitted over the ultra-lightweight bitwise operation, and the tag does not only superimpose the shared private key fragments simply to obtain the final shared private key. So even if the attacker obtains the shared private key fragments' information exchanged between the tag and the reader through eavesdropping, it is impossible to restore or derive the shared private key, thus realizing the wireless generation of the shared private key between the tag and the reader.

The first chapter of this article is the introduction about the background of RFID technology and its existing problems, which leads to the focus of this study. The second chapter provides a comprehensive introduction to the current work in the shared private key generation in the RFID system. The third chapter introduces the mathematical knowledge and the meanings of the symbols that are used in the WKGA-BO design process. The fourth chapter describes the detailed design steps of the WKGA-BO systematically. The fifth chapter analyzes the security of WKGA-BO from two aspects: active attack and passive attack. In the sixth chapter, the WKGA-BO is rigorously proved and deduced by GNY formal logic. In the seventh chapter, the performance of WKGA-BO is analyzed in detail from the aspects of calculation, storage space and communication traffic. The eighth chapter summarizes the whole paper, and gives the next research direction.

## 2 Related Works

There are many ways to generate the shared private key in the existing wireless systems and wireless devices. We can classify the existing methods into two categories from the point of view of usefulness and cryptography. One is based on the cryptography method; the other is based on the non-cryptographic method. The advantages and disadvantages of the existing shared private key generation methods from these two types of methods are described as follows.

The method based on cryptography – this type of shared private key generation method is mainly based on the public key cryptography, such as Diffie-Hellman key agreement protocol [8]. Although this type of method can solve the problem of shared private key generation, but it needs more complex mathematical operations (mostly based on the mathematical problem of the decomposition of large number when using cryptography), so that it can not be used very well in the RFID system tags whose resources are severely limited, such as weak calculation capacity, small storage space and so on.

The method based on non-cryptographic – this type of shared private key generation method is different from the cryptographic-based method, but it uses physical methods or the characteristics of the communication channels to generate shared private key, so it does not require complicated calculation [24]. In the following, we'll describe the two types of methods in detail.

The physical methods of generating the shared private key involve two different ways. One way is to generate the shared private key by physical isolation; the other is to write the shared private key over a wired connection link. The physical isolation method generates the shared private key primarily by using a Faraday cage container to protect the communication channel between the wireless device or the wireless system from being intercepted by the attacker [15]. From the Reference[19] we know that the Faraday cage is a container made of metal mesh for shielding the wireless signal transmission. The wireless signal in the cage is shielded by the Faraday cage (the wireless signal outside the cage is not in the Faraday cage category), so that any two wireless devices within the Faraday cage and wireless systems can communicate in plain text. However, this type of the shared private key generation method must be carried out in the space of the Faraday cage, so it is limited by the size of Faraday cage space, and it can not be promoted in large-scale in the practical application of RFID. For example, Faraday cage containers can not be accommodated in the RFID tags of the large objects such as trains, containers. At the same time, users may not be willing to use RFID tags before spending too much manpower and financial resources to build a large Faraday cage. In the case of the method of generating the shared private key over a wired connection link, the two communicating devices require a hardware interface to establish a physical electrical signal connection so that the shared private key can be generated or exchanged [25]. However, in the existing RFID systems, the tag does not have a physical interface, so it can not support this kind of circuit connection communication, and it does not apply to the existing RFID systems and equipment.

Generating the shared private key based on the characteristics of the communication channel: In Reference [4] it is found that the shared private key is generated between two wireless devices with limited CPU resources through an anonymous communication channel. However, it has been analyzed that although the computation of the method is small, it is necessary to generate a packet for each secret bit, so that the communication traffic is large. Meanwhile, large-scale passive tag can only reflect the reader's signal as a response, and can not produce random data packets, so the scheme proposed in Reference [4] does not apply to the existing RFID systems. Reference [2] has proposed a scheme for generating the shared private key by measuring a random variable in a wireless communication channel, which can be used to generate the shared private key with high entropy in two wireless devices. However, the above-mentioned scheme requires the wireless device to autonomously measure the data of random variables (such as signal strength) in the communication channel, but the tag in the existing RFID systems does not have this capability, so the proposed scheme can not be applied on a large scale.

In view of the shortcomings of the existing schemes, this paper proposes an algorithm WKGA-BO of generating the shared private key wirelessly based on bitwise operation in the RFID system. By dynamically generating the shared private key wirelessly, WKGA-BO resolves the problem that the shared private key between the tag and the reader must be pre-set in the RFID system, and the user can not customize it. Compared with other existing solutions, WKGA-BO has the following main advantages. Firstly, WKGA-BO doesn't need to provide additional physical hardware interface, and doesn't require complex cryptographic algorithms, which is able to minimize the amount of computation. Secondly, WKGA-BO dynamically generates the shared private key wirelessly, which solves the problem of key escrow caused by the factory setting for the shared private key. Thirdly, WKGA-BO uses bit operations for encryption. Because the ultra-lightweight characteristic of bit operations is able to meet the requirements of low cost, WKGA-BO can be implemented and scaled up in real RFID systems.

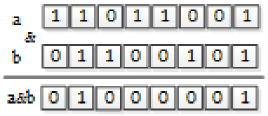


Figure 1: Bitwise AND Operation Flow Chart

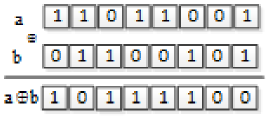


Figure 2: Bitwise XOR Operation Flow Chart

# 3 Related Knowledge Introduction

## 3.1 Bitwise AND Operation

If the two values a, b are not both 1, then the result is0; if the two values a, b are both 1, then the result is 1, where $a\in\{0,1\}^l$, $b\in\{0,1\}^l$, meaning that a and b are both binary numbers with the length of l bits. In order to facilitate the description of the symbol, we use the symbol "&" to represent bitwise AND operation [13].

The bitwise AND operation rules are described below: *0&0=0*, *1&0=0*, *0&1=0*, *1&1=1*. Only when the values of a and b are both 1, the result of the operation is 1; otherwise, the result of the operation is0. For example, if *l=8, a=11011001, b=01100101,* then *a & b=01000001.* The specific process is shown in Figure 1.

## 3.2 Bitwise XOR Operation

If the two values *a,b* are not the same, the result is 1; if the two values a, b are the same, the result is0, where $a\in\{0,1\}^l$, $b\in\{0,1\}^l$, meaning that $a$ and $b$ are both binary numbers with the length of $l$ bits. In order to facilitate the description of the symbol, we use the symbol "$\oplus$" to represent bitwise XOR operation [13].

The bitwise XOR operation rules are described below: $0\oplus 0=0$, $1\oplus 0=1$, $0\oplus 1=1$, $1\oplus 1=0$. Only when the values of a and b are the same, the result of the operation is0; otherwise, the result of the operation is 1. For example, if *l=8, a=11011001, b=01100101,* then $a \oplus b$*=10111100.* The specific process is shown in Figure 2.

Bitwise XOR operation has a clever use, that is, $a \oplus b \oplus b=a$. For example, if *l=8, a=11011001, b=01100101,* then $a \oplus b$*=10111100*, $a \oplus b \oplus b$*=11011001*, so $a \oplus b \oplus b=a$. Because the bitwise XOR operation has this feature, so the WKGA-BO design process in the fourth chapter uses the bitwise XOR operation to transmit the information, and then through this feature the original information can be obtained by the corresponding encrypted information.

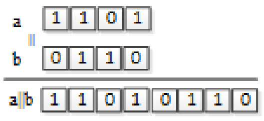


Figure 3: a || b Bitwise Concatenation Operation Flow Chart

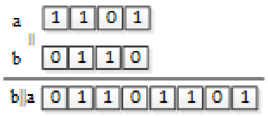


Figure 4: $b||a$ Bitwise Concatenation Operation Flow Chart

## 3.3 Bitwise Concatenation Operation

If $a∈\{0,1\}^l$, $b∈\{0,1\}^l$, $a$ and $b$ are both binary numbers with the length of $l$ bits, then $a||b$ or $b||a$ means concatenating the two numbers to a new binary number with the length of *2l* bits, but the results of $a||b$ or $b||a$ are different. In order to facilitate the description of the symbol, we use the symbol "||" to represent bitwise concatenation operation [5]. For example, if *l=4, a=1101, b=0110,* then *a || b=11010110, b || a=01101101.*The specific process of $a||b$ is shown in Figure 3, and the specific process of $b||a$ is shown in Figure 4.

## 3.4 Bitwise Substitution Operation

In order to facilitate the description of the symbol, we use the symbol "*Sub(X, Y)*" to represent bitwise substitution operation. The definition of *Sub(X, Y)* is as follows. Let $X$ and $Y$ be two binary numbers with the length of $l$ bits, $X = x_1x_2x_3 \cdots x_L$, $Y = y_1y_2y_3...y_L$, where $X∈\{0,1\}^l$, $Y∈\{0,1\}^l$. We get each bit which is one in the binary number $Y$, complement the corresponding bit in the binary number $X$ and substitute it. In the binary number $X$, totally $n = wt(Y)$ bits are substituted, where $wt(Y)$ is the Hamming weight of $Y$.

When the bitwise substitution operation is implemented in the tag, using the pointer form proposed in Reference [26] will be more efficient than using a logic gate directly. Two pointers are introduced, one for $P_X$ and the other for $P_Y$, where $P_X$ points to the binary number $X$ and pointer $P_Y$ points to the binary number $Y$. When the pointer $P_X$ starts traversing from the most significant bit of the binary number $X$, the pointer $P_Y$ starts traversing from the most significant bit of the binary number $Y$ at the same time. When the pointer $P_Y$ points to the zero bit of binary number $Y$, the bit that the pointer $P_X$ points to in the binary number $X$ does not change. When the pointer $P_Y$ points to the one bit of the binary number $Y$, the bit that the pointer $P_X$ points to in the binary number $X$ is substituted with its complement. Finally, *Sub(X, Y)* is the binary number $X$ after substituted. For example, if $l = 8, X = 11011001, Y = 01100101$, then $Sub(X, Y) = 10111100$. The specific process is shown in Figure 5.

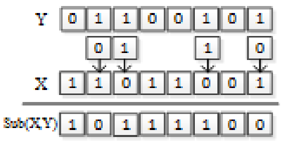


Figure 5: Bitwise Substitution Operation Flow Chart

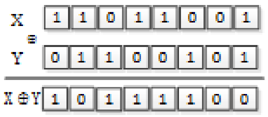


Figure 6: Bitwise Substitution Operation Flow Chart

## 3.5 Self-assembling Cross-bit Operation

In order to facilitate the description of the symbol, we use the symbol "$Sac(Z)$" to represent self-assembling cross-bit operation. Let $X, Y, Z$ be three binary numbers with even length $l$ bits, $X = x_1x_2x_3...x_L, Y = y_1y_2y_3...y_L, Z = z_1z_2z_3...z_L$, where $X∈\{0,1\}^l$, $Y∈\{0,1\}^l$, $Z∈\{0,1\}^l$. $X$ makes a bitwise XOR operation to $Y$ and then we get the result $Z$. The operation $Sac(Z)$ is a new binary number $W$ with the even length of $l$ bits formed by the combination of the high and low bits of $Z$, that is, $Sac(Z) = z_1z_{L/2+1}z_2z_{L/2+2} \cdots z_{L/2}z_L$.

The self-assembling cross-bit operation can be implemented in the tag and reader as described below. Introduce two pointers, one for $P_1$ and the other for $P_2$, where the pointer $P_1$ points to the head of the binary number $Z$, and the pointer $P_2$ points to the end of the binary number $Z$. When the pointer $P_1$ traverses from the head of the binary number $Z$, the pointer $P_2$ starts traversing from the end of the binary number $Z$ at the same time. The numbers traversed by the pointer $P_1$ are sequentially placed in the odd bits of the new binary number $W$, and the numbers traversed by the pointer $P_2$ are sequentially placed in the even bits of the new binary number $W$. Then through the final combination we can get the new binary number $W$, that is $Sac(Z)$ [10].

The self-assembling cross-bit operation only needs the shift and the bitwise OR operation, and makes the final combination, thereby reducing system throughput and storage capacity, to achieve an ultra-lightweight level. According to the different order of the assignment of the pointers, it will be combined to obtain different values, thereby increasing the difficulty of cracking. For example, if $l = 8, X = 11011001, Y = 01100101$, then $X \oplus Y = Z, Sac(Z) = 11011010$. The specific process is shown in Figure 6.

# 4 WKGA-BO Design

In this section, WKGA-BO is designed for the three cases of the shared private key wireless generation in the practical application: (1) Generating the shared private key for a single tag; (2) Generating each individual shared private key separately for a batch of tags at the same time; (3) Generating a shared private key for a batch of tags at the same time.

RFID system generally consists of three parts: the tag, reader, and database. Because the reader and the database communicate through the wired link, the general researches point out that the communication between the two is safe and reliable. So we regard the reader and the database as a whole, and the reader has a strong search ability.

As is the same to other related protocols,we assume that the communication between the reader and database is safe and reliable. We also suppose that the communication between the reader and the tag is unreliable and is easy to be eavesdropped by the attacker, and that the privacy information shared between the reader and the tag is safe and is what the attacker does not know in advance (such as the shared $ID_L$ *etc.*).

## 4.1 Single Tag's Shared Private Key Generation

Before the start of the single tag's shared private key generation (SPKG) algorithm, the tag $T_i$ stores the information $(ID_{i_{t_R}}, ID_{i_{t_L}})$, and the reader R stores the information $(ID_{i_{t_R}}, ID_{i_{t_L}})$ of all the tags.

The meaning of each symbol in this algorithm is given as Table 1.

Table 1: Symbol used in single tag's SPKG algorithm

| Symbol | Description |
|---|---|
| $T$ | A tag |
| $R$ | A reader |
| $DB$ | A database |
| $ID_t$ | The tag identifier ID |
| $ID_{t_L}$ | The left half of the tag identifier ID |
| $ID_{t_R}$ | The right half of the tag identifier ID |
| $Key$ | The generated shared private key |
| $r$ | The random number generated by the reader |
| $r_L$ | The left half of r |
| $r_R$ | The right half of r |
| $\oplus$ | Bitwise XOR operation |
| $\|\|$ | Bitwise concatenation operation |
| $\&$ | Bitwise AND operation |
| $Sub(X,Y)$ | Bitwise substitution operation |

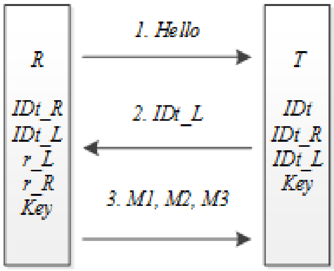


Figure 7: Single Tag's Shared Private Key Generation Flow Chart

The flow chart of wirelessly generating the single tag's shared private key is shown in Figure 7. Table 2 shows the operation formulas that appear in this algorithm.

Table 2: Formula used in single tag's SPKG algorithm

| Symbol | Description |
|---|---|
| $M1$ | $r_L \oplus ID_{t_R}$ |
| $M2$ | $r_R \oplus ID_{t_R}$ |
| $M3$ | $Sub(r_L\|\|ID_{t_R}, ID_{t_R}\|\|r_R)$ |
| $r'_L$ | $M1 \oplus ID_{t_R}$ |
| $r'_R$ | $M2 \oplus ID_{t_R}$ |
| $M3'$ | $Sub(r'_L\|\|ID_{t_R}, ID_{t_R}\|\|r'_R)$ |
| $Key$ | $Sub(r_L\&ID_{t_R}, ID_{t_R}\&r_R)$ |

The single tag's shared private key generation algorithm in the WKGA-BO protocol consists of four steps, as shown in Figure 7.

**Step 1:** The reader sends a *Hello* message to the tag, and informs the tag to start the private key generation process.

**Step 2:** The tag sends $ID_{t_L}$ as the response message to the reader.

**Step 3:** The reader searches the database for the result of whether $ID_{t_L}$ exists or not. If the result exists, the reader generates a random number $r \in \{0,1\}^l (r_L$ means the left half of $r$, and $r_R$ means the right half of $r$), then calculates the values of $M1, M2, M3$, and enumerates the shared private key $Key$. Finally, it sends $M1, M2, M3$ to the tag. If it does not exist, $WKGA-BO$ terminates. The calculations of $M1, M2, M3, Key$ are described above.

**Step 4:** The tag calculates the value of $r'_L, r'_R$, then calculates $M3'$, and compares the value of $M3$ and $M3'$. If $M3$ is equal to $M3'$, the tag successfully verifies the reader, which meanwhile indicates that $r'_L = r_L$, $r'_R = r_R$, and then the tag starts calculating the shared private key $Key$. If $M3$ is not equal to $M3'$, $WKGA-BO$ terminates. The calculations of $r'_L, r'_R, M3', Key$ are described above.

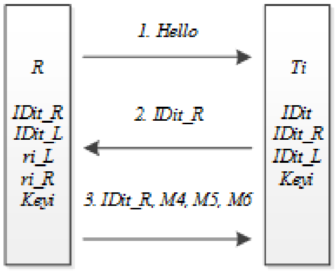


Figure 8: A Batch of Shared Private Keys Generation Flow Chart

## 4.2 A Batch of Shared Private Keys Generation

Before the start of a batch of SPKG algorithm, the tag $T_i$ stores the information ($ID_{i_{t_R}}$, $ID_{i_{t_L}}$), and the reader $R$ stores the information ($ID_{i_{t_R}}$, $ID_{i_{t_L}}$) of all the tags.

The meaning of each symbol in this algorithm is given as Tables 1 and 3.

Table 3: Symbol used in batch of SPKG algorithm

| Symbol | Description |
|---|---|
| $T_i$ | The $i$-th tag |
| $ID_{i_t}$ | The $i$-th tag identifier ID |
| $ID_{i_{t_L}}$ | The left half of $ID_{i_t}$ |
| $ID_{i_{t_R}}$ | The right half of $ID_{i_t}$ |
| $Key_i$ | The shared private key generated between the $i$-th tag and the reader |
| $r_i$ | The random number generated by the reader for the $i$-th tag |
| $r_{i_L}$ | The left half of $r_i$ |
| $r_{i_R}$ | The right half of $r_i$ |
| $Sac(X)$ | Self-assembling cross-bit operation |

The flow chart of wirelessly generating a batch of shared private keys is shown in Figure 8. Table 4 shows the operation formulas that appear in this algorithm.

Table 4: Formula used in batch of SPKG algorithm

| Symbol | Description |
|---|---|
| $M4$ | $r_{i_L} \oplus ID_{i_{t_L}}$ |
| $M5$ | $r_{i_R} \oplus ID_{i_{t_L}}$ |
| $M6$ | $Sac((r_{i_L} \vert\vert ID_{i_{t_L}})\ \&\ (ID_{i_{t_L}} \vert\vert r_{i_R}))$ |
| $r'_{i_L}$ | $M4 \oplus ID_{i_{t_L}}$ |
| $r'_{i_R}$ | $M5 \oplus ID_{i_{t_L}}$ |
| $M6'$ | $Sac((r'_{i_L} \vert\vert ID_{i_{t_L}})\ \&\ (ID_{i_{t_L}} \vert\vert r'_{i_R}))$ |
| $Key_i$ | $Sac((r_{i_L} \& ID_{i_{t_L}}) \oplus (ID_{i_{t_L}} \& r_{i_R}))$ |

This algorithm enables the reader in the RFID system to generate individual shared private keys for a large number of different tags at the same time, and *WKGA-BO* protocol consists of four steps, as shown in Figure 8.

**Step 1:** The reader sends a *Hello* message to the tag, and informs the tag to start the private key generation process.

**Step 2:** The tag $T_i$ sends $ID_{i_{t_R}}$ as the response message to the reader.

**Step 3:** The reader searches the database for the result of whether $ID_{i_{t_R}}$ exists or not. If the result exists, the reader generates a random number $r_i \in \{0,1\}^l (r_{i_L}$ means the left half of $r_i$; $r_{i_R}$ means the right half of $r_i$; $r_i$ is applied to calculate the random number used by the shared private key between the reader and the tag $T_i$), then calculates the values of $M4, M5, M6$, and generates the shared private key $Key_i$. Finally, it sends $M4, M5, M6$, $ID_{i_{t_R}}$ to the tag $T_i$. If it does not exist, $WKGA-BO$ terminates. The calculations of $M4, M5, M6, Key_i$ are described above.

**Step 4:** The tag $T_i$ compares the received value $ID_{i_{t_R}}$ with its own value $ID_{t_R}$. If $ID_{i_{t_R}}$ is not equal to $ID_{t_R}$, the tag $T_i$ discards the message. Otherwise, the tag $T_i$ calculates the value of $r'_L, r'_R$, then calculates $M6'$, and compares the value of $M6$ and $M6'$. If $M6$ is equal to $M6'$, the tag $T_i$ successfully verifies the reader, which meanwhile indicates that $r'_{i_L} = r_{i_L}$, $r'_{i_R} = r_{i_R}$, and then the tag starts calculating the shared private key $Key_i$. If $M6$ is not equal to $M6'$, $WKGA - BO$ terminates. The calculations of $r'_{i_L}, r'_{i_R}, M6', Key_i$ are described above.

## 4.3 Group-based Shared Private Keys Generation

Before the start of group-based SPKG algorithm, the tag $T_i$ stores the information ($ID_{i_{t_R}}$, $ID_{i_{t_L}}$), and the reader $R$ stores the information ($ID_{i_{t_R}}$, $ID_{i_{t_L}}$) of all the tags.

The meaning of each symbol in this algorithm is given as Tables 1 and 3.

The flow chart of wirelessly generating a group of shared private keys is shown in Figure 9. Table 5 shows the operation formulas that appear in this algorithm.

This algorithm needs to generate a unique shared private key between the reader $R$ and a group of tags $T1, T2$, $\cdots$, $T_i$, $\cdots$, $Tn$, where $n$ is the total number of tags in the group. $WKGA - BO$ protocol consists of four steps, as shown in Figure 9.

**Step 1:** The reader sends a *Hello* message to the tag group, and informs all the tags to start the private key generation process.

**Step 2:** The tag $T_i$ sends $ID_{i_{t_L}}$ as the response message to the reader.

Table 5: Formula used in group-based SPKG algorithm

| Symbol | Description |
|---|---|
| $M7$ | $r_{i_L} \oplus ID_{i_{t_L}}$ |
| $M8$ | $r_{i_R} \oplus ID_{i_{t_L}}$ |
| $M9$ | $Sac((r_{i_L} \,\&\, r_{i_R}) \oplus ID_{i_{t_R}})$ |
| $M10$ | $Key \oplus ID_{i_{t_R}}$ |
| $r'_{i_L}$ | $M7 \oplus ID_{i_{t_R}}$ |
| $r'_{i_R}$ | $M8 \oplus ID_{i_{t_R}}$ |
| $M9'$ | $Sac((r'_{i_L} \,\&\, r'_{i_R}) \oplus ID_{i_{t_R}})$ |
| $Key$ | $Sac(ID_{1_{t_R}} \oplus ID_{2_{t_R}} \oplus ID_{i_{t_R}} \oplus \cdots \oplus ID_{n_{t_R}}$ |

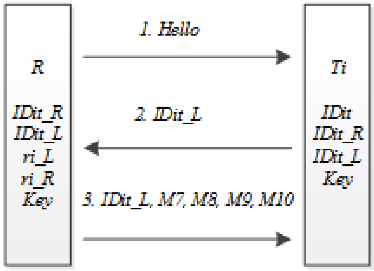


Figure 9: Group-based Shared Private Keys Generation Flow Chart

**Step 3:** After the reader receives the responses of all the tags, $R$ compares all the received tags $ID_{1_{t_L}}, ID_{2_{t_L}}, \cdots, ID_{i_{t_L}}, \cdots, ID_{n_{t_L}}$ with its own tag $ID_{t_L}$ and determines whether they are consistent or not. If they're not consistent, which indicates that some tags do not answer the reader, then $R$ re-sends a $Hello$ message to restart the shared private key generation. If they're consistent, it indicates that all the tags have answered the reader, and that the shared private key can be generated. The reader generates a random number $r_i \in \{0,1\}^l (r_{i_L}$ means the left half of $r_i$; $r_{i_R}$ means the right half of $r_i$; $r_i$ is the random number used by the tag $T_i$ to verify the reader), then calculates the value of the shared private key $Key$ and the values of $M7, M8, M9, M10$. Finally, it sends $M7, M8, M9, M10, ID_{i_{t_L}}$ to the tag $T_i$. The calculations of $M7, M8, M9, M10, Key$ are described above.

**Step 4:** The tag $T_i$ compares the received value $ID_{i_{t_L}}$ with its own value $ID_{t_L}$. If $ID_{i_{t_L}}$ is not equal to $ID_{t_L}$, the tag $T_i$ discards the message. Otherwise, the tag $T_i$ calculates the value of $r'_{i_L}, r'_{i_R}$, then calculates $M9'$, and compares the value of $M9$ and $M9'$. If $M9$ is equal to $M9'$, the tag $T_i$ successfully verifies the reader, which meanwhile indicates that $r'_{i_L} = r_{i_L}, r'_{i_R} = r_{i_R}$, and then the tag Ti starts calculating the shared private key $Key = M10 \oplus ID_{i_{t_R}}$. If $M9$ is not equal to $M9'$, $WKGA-BO$ terminates. The calculations of $r'_{i_L}, r'_{i_R}, M9'$, Key are described above.

The $WKGA-BO$ proposed in this paper has the following advantages. Firstly, the shared private key between the reader and the tag need not be pre-set and can be dynamically generated when being used. Secondly, the user can customize the shared private key according to his preference and requirement. Thirdly, the $WKGA-BO$ uses the ultra-lightweight bit operation to encrypt the transmitted information so that it can achieve the ultra-lightweight level, which is able to reduce the amount of calculation and achieve the goal of low cost. It's suitable for the existing RFID systems. Fourthly, the corresponding algorithms for the shared private key generation are given in different scenarios, and $WKGA-BO$ is no longer limited in real application.

## 5 Security

This section first gives the target of the security of $WKGA-BO$, and then gives a specific security analysis process.

### 5.1 Security Target

The communication channel between the tag and the reader in the RFID system can be divided into a forward channel and a backward channel. The forward channel refers to the channel from the reader to the tag; the backward channel refers to the channel from the tag to the reader. Research scholars generally believe that the forward and backward channels are not safe. The attacker can gain access to the communication between the tag and the reader by tapping the two channels. The attacker deduces the privacy information stored in the tag or reader according to what he eavesdrops. One of the WKGA-BO's security target is to ensure that the attacker can not infer any useful privacy information from the eavesdropping information.

1) Resist passive attack:
Before the shared private key is not generated, there is no trusted private key between the tag and the reader to protect the communication between the two messages, so $WKGA-BO$ must be able to resist the attacker eavesdropping, and make the attacker unable to eavesdrop the information to infer the generated shared private key. Even if the attacker can eavesdrop on the forward and backward channels, $WKGA-BO$ should make the attacker unable to infer the generated shared private key based on the information obtained from what he eavesdrops.

2) Resist active attack:
The attacker can inject the false information into the forward channel and backward channel. When the attacker injects or modifies the information into the

forward channel, the attacker is equivalent to disguise himself as the reader to send messages to the tag. This kind of attack is called the reader impersonation attack. What's more, when the attacker injects or modifies the information into the backward channel, the attacker is equivalent to disguise himself as the tag to send messages to the reader. This kind of attack is called the tag impersonation attack. WKGA-BO should be able to resist the two kinds of impersonation attacks.

## 5.2 Security Analysis

This section will analyze the security of $WKGA-BO$ from the aspects of resisting the passive attack and the active attack.

### 5.2.1 Passive Attack

The ability of $WKGA-BO$ to resist the passive attack depends on the characteristic of the random number generated by the reader and the security of $ID_{i_{t_R}}$, $ID_{i_{t_L}}$ shared between the reader and the tag. $WKGA-BO$ can resist the passive attack, if the probability of each bit (0 or 1) in the random number generated by the reader is the same, and if $ID_{i_{t_R}}$, $ID_{i_{t_L}}$ shared between the reader and the tag are secure. The *EPC Gen* 2 air interface standard has defined 16-bits pseudo-random number generators embedded in RFID tags to meet this characteristic. At the same time $WKGA-BO$ algorithm has a certain application of the scene, this section has been set in the applicable scene between the label and the reader at the time of sharing the information is safe and reliable. At the same time $WKGA-BO$ algorithm has a certain application of the scene. In this section the shared information that has been set between the tag and the reader before leaving factory is safe and reliable in the applicable scene.

In order to obtain the generated shared key, the attacker can eavesdrop on the communication between the tag and the reader. In the following we uses the single tag shared private key generation algorithm as an example, to analyze that the $WKGA-BO$ can resist the passive attack.

First, the random number are used in the calculations of$M1, M2, M3$, and the random numbers in each round of the random generation of the shared private key are different, so as to ensure that $M1, M2, M3$ in each round are various. Second, the attacker can obtain $ID_{t_L}$, but the calculations of $M1, M2, M3$ do not use $ID_{t_L}$, but the $ID_{t_R}$, and there is no association between $ID_{t_R}$ and $ID_{t_L}$, so that the attacker can not infer $ID_{t_R}$ from the eavesdropping $ID_{t_L}$. Third, there are at least two values that the attacker doesn't know in the calculations of $M1, M2, M3$, so that the attacker can not exhaust private information such as the shared private key and so on. Based on the above description, $WKGA-BO$ can resist the passive attack.

### 5.2.2 Active Attack

The attacker can inject the error messages into the forward channel and the backward channel, and deceive the reader and the tag, which makes them generate the wrong shared private key. So In this section we will analyze the attack from two aspects – the attacker disguises as the reader, and the attacker disguises as the tag.

1) The attacker disguises as the reader:

   The attacker can inject the error messages to the forward channel before/while/after the WKGA-BO is executed at the legal reader and the legal tag, which can be seen as the attacker disguises as the legal reader and executes $WKGA-BO$ algorithm with the tag. In the following we uses the single tag shared private key generation algorithm as an example, to analyze that $WKGA-BO$ can resist the active attacks.

   The attacker chooses to attack before the $WKGA-BO$ is executed. The attacker disguises as the reader to send information to the tag. Even if the attacker can pass the certifications of the first two steps, the attacker can not correctly execute the third step and the fourth step. In the third step, the attacker can not correctly calculate $M1, M2, M3$ by the unknown $ID_{t_R}$. In the fourth step, the tag can confirm the authenticity of $M1, M2$ and $M3$ to determine the authenticity of the reader quickly. The specific determination process is as follows.

   Next the tag uses the received $M1$, its own $ID_{t_R}$ to calculate $r'_L = M1 \oplus ID_{t_R}$, uses the received $M2$, its own $ID_{t_R}$ to calculate $r'_R = M2 \oplus ID_{t_R}$, uses $r'_L, r'_R$ to calculate $M3' = Sub(r'_L||ID_{t_R}, ID_{t_R}||r'_R)$, and then compares $M3'$ with $M3$. The attacker does not know the value of $ID_{t_R}$, so he can only choose a random number instead of $ID_{t_R}$ to calculate and obtain the incorrect value, which results in the fourth step in which the $ID_{t_R}$ used by the tag is not consistent with the $ID_{t_R}$ used by the attacker. So the tag can determine that the reader is forged by the attacker, and $WKGA-BO$ terminates.

   The attacker chooses to attack while the $WKGA-BO$ is executed. At this time the legal reader and tag are in the process of generating the shared private key. Because there is no initial shared private key, the tag can not distinguish the source of the message in the forward channel (that is, the tag can not determine that the message sent by the forward channel at this time is derived from the legal reader or from the attacker). Even if the attacker injects the wrong data, the tag can also distinguish the authenticity of the reader in the fourth step, and the specific analysis process is as described above.

   The attacker chooses to attack after the $WKGA-BO$ is executed. At this time the tag and the legal reader has generated the shared private key through executing $WKGA-BO$. Then the tag uses the generated shared private key to bidirectionally authenticate the reader (such as using the Hash-Lock protocol), and the attacker fails because the attacker can not obtain the shared private key.

In summary, when the attacker disguises as the legal reader to inject the message in the forward channel, the tag can always distinguish the authenticity of the reader, so $WKGA-BO$ can resist the impersonation attack that is caused by the attacker disguising as the reader.

2) The attacker disguises as the tag:
In this attack mode, the attacker sends information to a legal reader by actively injecting a message to the backward channel, in order to destroy the generation process of the shared private key or to obtain the shared private key between the reader and the tag, which is equivalent to the attacker deceiving the tag and the legal reader between the implementation of $WKGA-BO$ to generate the shared private key. Because before $WKGA-BO$ is executed, there is no shared private key between the tag and the reader, the legal reader can not distinguish the message sent by the legal tag or the tag that is disguised by the attacker. In the following we uses the single tag shared private key generation algorithm as an example, to analyze that $WKGA-BO$ can resist the active attacks.

The attacker can obtain the information $ID_{t_L}, M1, M2, M3$ through eavesdropping a complete communication process between the legal tag and the legal reader. The attacker tries to send the $ID_{t_L}$ to the legal reader and wants to get the correct shared private key, but the attacker can not succeed. After the reader receives the $ID_{t_L}$ sent by the attacker, although the $ID_{t_L}$ can be found, the calculations of $M1, M2, M3$ do not use $ID_{t_L}$, but the $ID_{t_R}$, and there is no association between $ID_{t_R}$ and $ID_{t_L}$. The random number are used in the calculations of $M1, M2, M3$, and the random numbers in each round of the random generation of the shared private key are different, so as to ensure that $M1, M2, M3$ in each round are various. Although the tag disguised by the attacker can obtain $M1, M2, M3$ sent by the legal reader, but because the attacker does not know $ID_{t_R}$, he can not calculate the random number generated by the legal reader, whick makes the attack unable to infer the shared private key from the received information. Before the subsequent tags communicate with the reader, there will be a bidirectional authentication process (such as Hash-Lock protocol). Through this process, the reader can argue that the two can not generate a unified shared private key, and then the reader and the tag re-execute $WKGA-BO$ to generate the shared private key.

Based on the above description, when the attacker disguises as the tag to inject information into the backward channel, although the $WKGA-BO$ can not generate a unified legal shared private key, the legal reader can still distinguish the authenticity of the tag. So $WKGA-BO$ can resist the impersonation attack that is caused by the attacker disguising as the tag.

# 6 GNY Logic Formal Proof

That the security of a complete protocol can be analyzed in words is far from enough. It can also be proved by the rigorous mathematical formulas. Based on this thought, in 1989 *Burrows* et al proposed a BAN formal logic analysis method, which was regarded as a milestone in the analysis of security protocols [3]. BAN logic is only concerned with the part of the protocol that is directly related to the authentication logic, and the rest is not a concern. It uses the rigorous mathematical rules to formalize the analysis and proof of the certification of the protocol. It also derives the target authentication step from the initialized hypothesis step of the protocol.

Because BAN form logic analysis has certain limitations, Gongli, etc. in 1990 put forward the GNY formal logic analysis method [11]. GNY formal logic analysis method is an expansion for BAN formal logic analysis method. GNY formal logic analysis method is more comprehensive than BAN logic analysis method, mainly in expanding the type and scope of analyzing the protocol. In this paper, the formal analysis and proof of $WKGA-BO$ protocol are carried out by using GNY formal logic analysis method.

$WKGA-BO$ gives three different scenarios to generate the shared private key. Because of the small differences among the three kind of the shared private key generation processes and the limited paper space, we choose to use the single tag's shared private key generation algorithm as an example to analyze and prove the protocol by GNY formal logic analysis method.

## 6.1 Formal Description of the Protocol

In order to make the $WKGA-BO$ protocol easy to be described by GNY formal logic, we use $R$ to represent the reader, and T for the tag. The process of WKGA-BO protocol is as follows.

$$Msg1 : R \rightarrow T : Hello;$$
$$Msg2 : T \rightarrow R : ID_{t_L};$$
$$Msg3 : R \rightarrow T : M1 = r_L \oplus ID_{t_R},$$
$$M3 = Sub(r_L || ID_{t_R}, ID_{t_R} || r_R), M2 = r_R \oplus ID_{t_R}.$$

Using GNY formal logic to standardize the above protocol, it can be described as follows.

$$Msg1 : T < *Hello;$$

$$Msg2 : T < *ID_{t_L};$$

$$Msg3: T < *M1 = r_L \oplus ID_{t_R},$$
$$M3 = Sub(r_L||ID_{t_R}, ID_{t_R}||r_R), M2 = r_R \oplus ID_{t_R}.$$

### 6.2 Initialization Hypothesis of the Protocol

The $WKGA-BO$ protocol assumes that $R$ and $T$ represent the entities, that is, $R$ for the reader, and $T$ for the tag. Hypothesis 1-3 represent what the tag and the reader have. Hypothesis 4 represents the trust of the reader and the tag on the freshness of the random number. Hypothesis 5-8 represent $ID_{t_L}, ID_{t_R}$ are shared between the reader and the tag.

$$Sub\ 1: T \ni (ID_{t_R}, ID_{t_L})$$
$$Sub\ 2: R \ni (ID_{i_{t_R}}, ID_{i_{t_L}})$$
$$Sub\ 3: R \ni (r_L, r_R)$$
$$Sub\ 4: R| \equiv \#(r_L, r_R);$$
$$Sub\ 5: T| \equiv R \overset{ID_{t_R}}{\leftrightarrow} T;$$
$$Sub\ 6: R| \equiv T \overset{ID_{t_R}}{\leftrightarrow} R;$$
$$Sub\ 7: T| \equiv R \overset{ID_{t_L}}{\leftrightarrow} T;$$
$$Sub\ 8: R| \equiv T \overset{ID_{t_L}}{\leftrightarrow} R.$$

### 6.3 Proof Target of the Protocol

There are 3 proof targets in $WKGA-BO$, mainly in the trust of the reader and the tag on the freshness of the interaction information.

The formulas of the targets are as follows.

$$Goal\ 1: T| \equiv R| \sim \#(M1 = r_L \oplus ID_{t_R});$$
$$Goal\ 2: T| \equiv R| \sim \#(M2 = r_R \oplus ID_{t_R});$$
$$Goal\ 3: T| \equiv R| \sim \#(M3 = Sub(r_L||ID_{t_R}, ID_{t_R}||r_R)).$$

### 6.4 Proof Process of the Protocol

The proof of the $WKGA-BO$ protocol is based on the initialization hypothesis. The proof process follows the logical reasoning rules, being-told rules, freshness rules and possession rules in Reference[30].The message interpretation rules follow the written form of the GNY logical reasoning rule in Reference[30], which are represented by $T, P, F, I$ respectively.

This process $Goal\ 2: T| \equiv R| \sim \#(M2 = r_R \oplus ID_{t_R})$, $Goal\ 3: T| \equiv R| \sim \#(M3 = Sub(r_L||ID_{t_R}, ID_{t_R}||r_R))$ is similar to the process of the proof target $Goal\ 1: T| \equiv R| \sim \#(M1 = r_L \oplus ID_{t_R})$. Therefore, this chapter takes the process of the proof target $Goal\ 1: T| \equiv R| \sim \#(M1 = r_L \oplus ID_{t_R})$ as an example, which is described as follows.

1) Because $Rule\ P1: \frac{P<X}{P \ni X}$ and $Msg3: T* < \{M1 = r_L \oplus ID_{t_R}\}$, therefore $T \ni \{M1 = r_L \oplus ID_{t_R}\}$.

2) Because $Rule\ F1: \frac{P|\equiv(x)}{P|\equiv(x,y), p|\equiv \#F(x)}$ and $Sup4: R| \equiv \#(r_L, r_R)$, therefore $T = \#\{M1 = r_L \oplus ID_{t_R}\}$.

3) Because $Rule\ P2: \frac{P \ni X, P \ni Y}{P \ni (X,Y), P \ni F(X,Y)}$, $Sup1: T \ni (ID_{t_R}, ID_{t_L})$ and $Sup3: R \ni (r_L, r_R)$, therefore $T \ni \{M1 = r_L \oplus ID_{t_R}\}$.

4) Because $Rule\ F10: \frac{P|\equiv(X), P \ni X}{P|\equiv \#(H(X))}$ and the inferred rule $T = \#\{M1 = r_L \oplus ID_{t_R}\}$, $T \ni \{M1 = r_L \oplus ID_{t_R}\}$, therefore $T| \equiv \#\{M1 = r_L \oplus ID_{t_R}\}$.

5) Because $Rule\ I3: \frac{P<H(X,<S>)>, P \ni (X,S), P|\equiv \#(X,S)}{P|\equiv Q|\sim(X,S), P|\equiv Q \sim H(X,<S>)}$, $Sup5: T| \equiv R \overset{ID_{t_R}}{\leftrightarrow} T$, $Sup6: R| \equiv T \overset{ID_{t_R}}{\leftrightarrow} R$ and $Msg3: T* < \{M1 = r_L \oplus ID_{t_R}\}$, therefore $T| = R \sim \{M1 = r_L \oplus ID_{t_R}\}$.

6) Because The definition of freshness and the inferred $T = \#\{M1 = r_L \oplus ID_{t_R}\}, T| = R \sim \{M1 = r_L \oplus ID_{t_R}\}$, therefore $Goal\ 1: T| \equiv R| \sim \#(M1 = r_L \oplus ID_{t_R})$.

The protocol is proved.

## 7 Performance Analysis

RFID system consists of three parts: tag, reader, and database. In this paper, the shared private key generation process is mainly completed in the tag and the reader, and the reader and the database are viewed as a whole, so the performance analysis is more concerned about the tag. In this section, we analyze the advantages and disadvantages of the WKGA-BO from the following three aspects: the storage space of the tag, the calculation of the tag, the communication cost of $WKGA-BO$.

Table 6: The Single Tag's Shared Private Key Generation Algorithm

| Storage space | Calculation | Communication cost |
|---|---|---|
| $3l$ | $2XOR + 2AND$<br>$+2OR + 2Sub()$ | $6l$ |

The data in Table 7 is analyzed for the single tag's shared private key generation algorithm. The tag stores three kinds of data structures: $ID_{t_R}, ID_{t_L}, Key$. We set the length of each data structure is $l$ bit, so the storage space of the tag is $3l$. XOR means bitwise XOR operation; $AND$ means bitwise AND operation; $OR$ means bitwise concatenation operation; $Sub()$ means bitwise substitution operation. The above four operations are bitwise operations, which can share part of the circuit to a certain extent, and they belong to super-lightweight operation, which can achieve the goal of reducing the cost of the tag. A complete communication process in the single tag's shared private key generation algorithm requires the transmission of the following five data: $Hello, ID_{t_L}, M1, M2, M3$. According to the above analysis, the lengths of $Hello, ID_{t_L}, M1, M2$ are $l$ bits, and the length of $M3$ is $2l$ bits, so the communication cost is $6l$.

Table 7: A Batch of Shared Private Keys Generation Algorithm

| Storage space | Calculation | Communication cost |
|---|---|---|
| $3l$ | $3XOR + 3AND$ <br> $+2OR + 2Sac()$ | $7l$ |

The data in Table 8 is analyzed for a batch of shared private keys generation algorithm. The tag stores three kinds of data structures: $ID_{i_{t_R}}$, $ID_{i_{t_L}}$, $Key_i$. We set the length of each data structure is $l$ bit, so the storage space of the tag is $3l$. $XOR$ means bitwise XOR operation; $AND$ means bitwise AND operation; $OR$ means bitwise concatenation operation; $Sac()$ means self-assembling cross-bit operation. The above four operations are bitwise operations, which can share part of the circuit to a certain extent, and they belong to super-lightweight operation, which can achieve the goal of reducing the cost of the tag. A complete communication process in a batch of shared private keys generation algorithm requires the transmission of the following five data: $Hello$, $ID_{i_{t_R}}$, $M4, M5, M6$. According to the above analysis, the lengths of $Hello$, $ID_{i_{t_R}}$, $M4, M5$ are $l$ bits, and the length of $M6$ is $2l$ bits. $ID_{i_{t_R}}$ is transmitted twice, so the communication cost is $7l$.

Table 8: Group-based Shared Private Keys Generation

| Storage space | Calculation | Communication cost |
|---|---|---|
| $3l$ | $4XOR + 1AND$ <br> $+1Sac()$ | $7l$ |

The data in Table 9 is analyzed for the group-based shared private keys generation algorithm. The tag stores three kinds of data structures: $ID_{i_{t_R}}$, $ID_{i_{t_L}}$, $Key$. We set the length of each data structure is $l$ bit, so the storage space of the tag is $3l$. $XOR$ means bitwise XOR operation; $AND$ means bitwise AND operation; $Sac()$ means self-assembling cross-bit operation. The above three operations are bitwise operations, which can share part of the circuit to a certain extent, and they belong to super-lightweight operation, which can achieve the goal of reducing the cost of the tag. A complete communication process in the group-based shared private keys generation algorithm requires the transmission of the following six data: $Hello$, $ID_{i_{t_L}}$, $M7, M8, M9, M10$. According to the above analysis, the lengths of $Hello$, $ID_{i_{t_L}}$, $M7, M8, M9, M10$ are $l$ bits, and the length of $M6$ is $2l$ bits. $ID_{i_{t_L}}$ is transmitted twice, so the communication cost is $7l$.

# 8 Conclusions

This paper proposes a wireless generation algorithm $WKGA - BO$ for private key in RFID system based on bitwise operation. This algorithm mainly solves the problems that the shared private key between the tag and the reader must be pre-set, and it can not be customized by the user. $WKGA - BO$ can not only be used in the single tag's shared private key generation, but also for group tag's shared private key generation, which makes $WKGA - BO$ more widely used. The security analysis shows that $WKGA - BO$ can resist passive attacks and active attacks. GNY formal logic is used for the rigorous mathematical derivation of $WKGA - BO$. The performance analysis shows that $WKGA - BO$ is able to generate the shared private key on the tag and the reader in RFID systems. In summary, $WKGA - BO$ is suitable for use in existing RFID systems. The next step for the paper is to implement an RFID system that uses $WKGA - BO$, and to research the total number of required gate circuits, a complete communication time and so on, with the combination of theory and practice.

# Acknowledgments

This work was supported in part by the Natural Science Foundation of China under Grant 61472090, Grant 61472089 and Grant 61672169, in part by the Science and Technology Project of Guangdong Province under Grant 2015B090906015 and Grant 2017B090906003, in part by the Science and Technology Planning Project of Guangzhou under Grant 201707010492, Grant 201604016003, Grant 201604016067 and Grant 201604016041.

# References

[1] T. Agrawal, P. K. Biswas, and A. D. Raoot, "Optimum frame size evaluation framework for efficient tag identification in passive RFID systems," in *2013 IEEE Wireless Power Transfer*, pp. 48–51, 2013.

[2] P. Bellot and M. Dang, "Secret key agreement over a non-authenticated channel," *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 48–55, 2003.

[3] M. Burrows, M. Abadi, and R. M. Needham, "A logic of authentication," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, pp. 233–271, 1989.

[4] C. Castelluccia and P. Mutaf, "Shake them up! a movement based pairing protocol for cpu-constrained

devices," in *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services*, pp. 51–64, 2005.

[5] C. L. Chen, Y. L. Lai, and C. C. Chen, "RFID ownership transfer authorization systems conforming epc global class-1 generation-2 standards," *International Journal of Network Security*, vol. 13, no. 1, pp. 41–48, 2011.

[6] W. T. Chen, "A feasible and easy-to-implement anticollision algorithm for the epc global uhf class-1 generation-2 RFID protocol," *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 2, pp. 485–491, 2014.

[7] Y. C. Chen, W. L. Wang, M. S. Hwang, "RFID authentication protocol for anti-counterfeiting and privacy protection", in *The 9th International Conference on Advanced Communication Technology*, pp. 255–259, 2007.

[8] W. Diffie and M. E. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, vol. 22, no. 6, pp. 644–654, 1976.

[9] R. Doss, W. Zhou, and S. Yu, "Secure RFID tag ownership transfer based on quadratic residues," *IEEE Trans on Information Forensics and Security*, vol. 8, no. 2, pp. 390–401, 2013.

[10] Z. Y. Du, G. A. Zhang, and H. L. Yuan, "Crossover based ultra-lightweight RFID authentication protocol," *Computer Science*, vol. 40, no. 14, pp. 35–42, 2013.

[11] L. Gong, R. Needham, and R. Yahalom, "Reasoning about belief in cryptographic protocols," *IEEE Computer Society Symposium in Security and Privacy*, pp. 234–248, 1990.

[12] B. Jian and D. Liu, "Wireless key generation algorithm for RFID system based on bit operation," *Computer Engineering and Applications*, vol. 53, no. 16, 2017.

[13] Y. Jin and Q. Wu, "RFID lightweight authentication protocol based on prf," *Journal of computer Research and Development*, vol. 51, no. 7, pp. 1506–1514, 2014.

[14] G. Kapoor and S. Piramuthu, "Single RFID tag ownership transfer protocols," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 42, no. 2, pp. 164–173, 2012.

[15] C. Kuo, M. Luk, and R. Negi, "Message-in-a-bottle: User-friendly and secure key deployment for sensor nodes," in *Proceedings of the 5th International Conference on Embedded Networked Sensor Systems*, pp. 223–246, 2007.

[16] Y. C. Lai, L. Y. Hsiao, and H. J. Chen, "A novel query tree protocol with bit tracking in RFID tag identification," *IEEE Transactions on Mobile Computing*, vol. 12, no. 10, pp. 2063–2075, 2013.

[17] H. Lalndaluce, A. Perallos, and I. J. G. Zuazola, "A fast RFID identification with low tag complexity," *IEEE Communications Letters*, vol. 17, no. 9, pp. 1704–1706, 2013.

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

[19] L. Lu, "Wireless key generation for RFID systems," *Chinese Journal of Computers*, vol. 38, no. 4, pp. 822–832, 2015.

[20] C. S. Ma, "Low cost RFID authentication protocol with forward privacy," *Chinese Journal of Computers*, vol. 34, no. 8, pp. 1388–1398, 2011.

[21] Y. Ma and D. Liu, "Improved mutual authentication with backward security for RFID protocols," *Computer Engineering and Applications*, 2017.

[22] H. Mala, M. Dakhilalian, and M. Shakiba, "Cryptanalysis of mcrypton-a lightweight block cipher for security of RFID tags and sensors," *International Journal of Communication Systems*, vol. 25, no. 4, pp. 415–426, 2012.

[23] D. Moriyama, S. Matsuo, and M. Ohkubo, "Relations among notions of privacy for RFID authentication protocols," *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 97, no. 1, pp. 225–235, 2014.

[24] J. Si, B. Y. Yang, and D. Liu, "Wireless key generation algorithm for RFID system," *Computer Engineering and Design*, vol. 38, no. 9, 2017.

[25] F. Stajano and R. Anderson, "The resurrecting duckling: Security issues for ubiquitous computing," *Computer*, vol. 35, no. 4, pp. 22–26, 2002.

[26] Y. Tian, G. Chen, and J. Li, "A new ultralightweight RFID authentication protocol with permutation," *IEEE Communications Letters*, vol. 16, no. 5, pp. 702–705, 2012.

[27] S. Wang, S. Liu, and D. Chen, "Scalable RFID mutual authentication protocol with backward privacy," *Journal of computer Research and Development*, vol. 5, no. 6, pp. 1276–1284, 2013.

[28] C. H. Wei, M. S. Hwang, A. Y. H. Chin, "A mutual authentication protocol for RFID", *IEEE IT Professional*, vol. 13, no. 2, pp. 20–24, Mar. 2011.

[29] M. H. Yang, "Secure multiple group ownership transfer protocol for mobile RFID," *Electronic Commerce Research and Applications*, vol. 11, no. 4, pp. 361–373, 2012.

[30] Z. Zhang, Y. Liu, and D. Liu, "Based on tag's id wireless key generation for RFID system algorithm," *Application Research of Computers*, vol. 34, no. 1, pp. 261–263, 2017.

# Biography

**Rui Xie** received his B.S. in electrical engineering and automation from Dalian Maritime University in 2000, and the M.Sc. in Computer Science from Guangdong University of Technology in 2003. He is currently a Ph.D Candidates in Guangdong University of Technology. His research interests cover a variety of different topics including netwok security, machine learning, cloud computing, data mining and their applications.

**Jie Ling** received his Ph.D degree in computation mathematics from Sun Yat-sen University (China) in June 1998. He is a professor in computer science in Guangdong University of Technology. His current research interest fields include computer applications and intelligent video processing technology.

**Dao-wei Liu** received a master's degree in School of Computers from Guangdong University of Technology (China) in June 2016. His current research interest fields include information security.

# Perseverance of Uncertainty in Cloud Storage Services through Reputation Based Trust

Vegi Srinivas[1], Vatsavayi Valli Kumari[2], and KVSVN Raju[3]

*(Corresponding author: Vegi Srinivas)*

Department of Computer Science and Engineering, Dadi Institute of Engineering and Technology[1]
DIET, Anakapalle-531002, Andhra Pradesh, India
(Email: srini.vegi@gmail.com)
Department of Computer Science and Systems Engineering, College of Engineering[2]
Andhra University, Visakhapatnam-530003, Andhra Pradesh, India.
Research and Development Cell, Anil Neerukonda Institute of Technology and Sciences[3]
Sangivalasa-531162, Visakhapatnam, Andhra Pradesh, India.

*(Received Feb. 8, 2017; revised and accepted July 13, 2017)*

## Abstract

The overwhelming technology in the world of computing is Cloud Data Storage and it is also a significant approach to making the highest level of durability, availability, and performance of the services to the users. A sudden and significant change over the data at cloud storage is much more problematic to find the risks. Security plays a vital role in cloud computing and trust is one of the most fascinating and promising factors to prevent uncertainty. Since the organizational hand over direct control over the data, it trusts on the provider to keep that data in a protected way. Clients make sure that the service provider protects data confidentiality by using reputed results to send and storing static data. In this paper, a new approach for reputation based interactions is proposed that are characterized by the trust which is critical for the cloud data persistence and the promise of gaining the advantage in a competitive market.

*Keywords: Cloud Computing; Cloud Data Storage; Credibility; Inconsistency; Reputation Based Trust*

## 1 Introduction

The environment of cloud computing offers two basic types of functions: computing and data storage. In the cloud computing environment, consumers of cloud services do not need anything and they can get access to their data and finish their computing tasks just through the Internet connectivity. During the access to the data and computing, the clients do not even know where the data are stored and which machines execute the computing tasks. The data in the cloud storage to be revealed by the user if the provider is considered trustworthy in the field of cloud computing.

There are some questions can be arising in the area of cloud data storage. Where my data is residing in the network? What types of vulnerabilities are exist in Cloud? In the Cloud Data Storage, Data Segregation and Accountability issues are one of the major problems. The cloud storage solution for a specific application or service may change based on many factors, such as Maturity, Performance, Compliance, Risk, Location Demands, Security, Technology Changes, and Changing Business Requirements [5]. To support enterprise customers with a solution flexible enough to meet their application requirements, cloud service providers must offer a broad range of cloud capabilities that falsification the lines between types of cloud infrastructure [11].

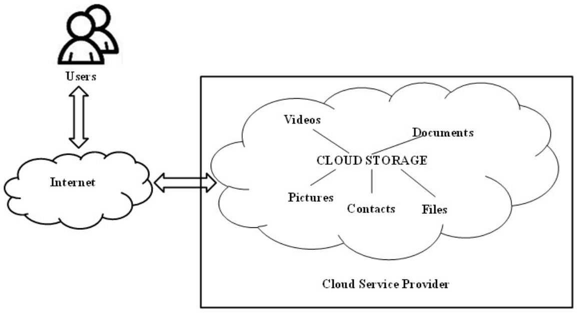


Figure 1: Cloud storage architecture

As shown in Figure 1, the cloud storage architecture will consist of several amenities such as videos, documents, pictures, files, *etc.* The derived trust values or reputation scores must be transparent to and clear enough for the consumers so that they can easily and confidently make the trust-based decision. Users are willing to disclose their data to the cloud provider if the provider is deemed to be competent, to be of integrity, and to be benevolent. Realizing that trustworthy, high-quality providers will not risk their reputation, users will be less concerned to use the cloud storage service [25].

Cloud computing is one of the emerging fields which replaces the burden of IT industry from spending huge expenditure on resources such as storage and network. Remote storage and easy accessibility of data combined with characteristics such as on-demand self-service, broad network access, resource pooling, rapid elasticity and measured services [24]. Virtualization provides a practical vehicle for transferring compute environments and sharing physical compute resources in the cloud. This approach has been used successfully by financial institutions and the life sciences to solve heavy compute models. It is expensive to run data centers full of servers ready to run complex mathematical models [13].

The computing resources that are provided by cloud service provider's (CSP's) shared to serve all consumers using a multi-tenant form, with different physical and virtual resources dynamically assigned according to consumer demand. The customer generally has no control of the location of the allocated resources. As a result, establishing accountability in distributed and layered architecture is an issue [4].

Trust Foundation has two stage forms; at first, it is trailed by irregular trust upgrade. Next, after the preliminary verification, the security properties of every user established occasionally for conformance with predefined security arrangements. Guarantee the service provider secures data confidentiality by utilizing encryption to transmit information, and utilizing it when putting away static information. The stronger the security, the greater the consumption of computing, memory, and bandwidth resources and the more difficult the service is to use, requiring manual configuration of security mechanism parameters [3].

The management and containment issues with rapid resource pooling are the main drawbacks in the cloud environment. Cloud computing differs from previously studied products and services in the way that it introduces a continuous uncertainty into the relationship between the provider and the user. Although the user depends on the cloud service provider at all time, he has only limited information about the providers's qualities, intentions, and actions [25].

The rest of the paper is organized as follows. In Section 2, we present the trust management and their techniques. The related work is discussed in Section 3. Section 4 describes the reputation based trust models so as to minimize the drawbacks of the existing models and enhance the trust values. The system implementation and experimental results are demonstrated at Section 5, Section 6 shows a small case study on file sharing in the cloud environment and finally; the paper is concluded in Section 7.

## 2 Trust Management

There are two ways to model the trust or distrust among peers: namely trust and reputation. Trust can be transitive but not necessarily symmetric between two parties. The combined trust model is the combination of three popular models such as identity-based trust, capability-based trust, and behavior-based trust. Lacking trust between service providers and cloud users has delayed the universal acceptance of cloud computing as a service on demand. As a virtual environment, the cloud poses new security threats that are more difficult to contain than traditional client and server configurations. In many cases, one can extend the trust models for P2P networks and grid systems to protect clouds and data centers.

### 2.1 Trust Management Techniques

Trust can be transitive yet not so much symmetric between two parties. The joined trust model is the blend of three mainstream models, for example, identity-based trust, capability based trust, and behavior based trust. Lacking trust between service provider and cloud consumer has overdue the comprehensive acknowledgment of distributed computing as an administration on interest. Trust management service is a difficult problem due to a unpredictable number of consumers and the highly dynamic nature of the cloud services. The trust management service should be flexible and extremely scalable to be practical in cloud environments.

#### 2.1.1 Trust Models

There are several trust models were already proposed by several researchers, each one having their own advantages and disadvantages. Which models are appropriate based on their security and trust requirements and the systems they need to interface it. Some of the trust models we have discussed in this paper are, Public Key Infrastructure (PKI)-based trust model, Feedback credibility-based trust model, Behavioral-based trust model, Subjective trust model, and Domain-based trust model [10]. Most of the trust models are subject to different kinds of attacks, while a few of them are resistant to particular attacks like false praise or accusation (FPA), Sybil and white washing attacks.

The PKI-based trust model depends on a few leader nodes to secure the whole system. This model may cause uneven load or a single point of failure since it relies on leader nodes too much. Behavioral-based trust model uses history trade records to compute trust. Subjective trust is a personal choice about the definite level of entity's particular characters or behaviors. The Domain-based trust model is mostly used in Grid computing which divides into two kinds of trust; one is in-domain trust relationship and the other is inter-domain trust relationship.

### 2.1.2 Taxonomy of Trust

A trust metric is a measure of how a member of a group is trusted by the other members of the group. Trust metric can be classified into local trust metric and global trust metric; a local trust metric predicts trust scores that are personalized from the point of view of every single user. On the other hand, a global trust metric computes a single global trust value for every single user. A trust management technique for direct and indirect trust can be calculated and it is defined in [2] as,

$$Direct\,Trust, DT^{A,B} = C^{A,B}(\sum_{i=1}^{p} W_i * T_i^{A,B}) \quad (1)$$

Where $C^{A,B}$ is the confidence factor calculated as a function of collected direct measurements, $W_i$ is the weighting factor for each one of the $p$ event types, $T_i^{A,B}$ is node A trust value of event $i$ regarding node B.

$$Indirect\,Trust, IT^{A,B} = \sum_{j=1}^{n} W(DT^{A,N_j})DT^{N_j,B} \quad (2)$$

Where $n$ is the number of neighboring nodes A, $N_j$ are the neighboring nodes to A, $DT^{A,N_j}$ is node $N_j$ reputation value of node B, $W(DT^{A,N_j})$ is a weighting factor reflecting node A direct trust value of node $N_j$.

### 2.1.3 Trust Evaluation Models

The trust evaluation models are different from the trust models. Firdhous *et al.*, in [7] had discussed about these models: Cuboid Trust, Eigen Trust, Bayesian Network Based Trust Management (BNBTM), GroupRep, AntRep, Semantic Web, Global Trust, Peer Trust, comPrehensive repuTation-based TRust mOdeL (PATROL-F), Trust Evaluation, Time-based Dynamic Trust Model (TDTM), Trust Ant Colony System (TACS).

Cuboid trust represents global reputation trust model which precedes three factors namely, peer's trustworthiness in giving feedback, a contribution of the peer to the system and quality of resources. Eigen trust assigns each peer a unique global trust value in a P2P file sharing network, based on the peer history of upload. BNBTM uses multidimensional applications specific trust values and each domain is evaluated using a single Bayesian network. GroupRep is a group based trust management system.

## 2.2 Trust Assessment

Integration of security measures, accreditation, bandwidth or customer support are the complex challenges regarding computation of trust. Another issue that is relevant when selecting or designing of trust or reputation mechanism relates to how much customization should be supported and where should be trusted values is aggregated [8]. Trust assessment should be based on not only experiences and user interactions but they also depend on other trustworthy communities.

### 2.2.1 Feedback

Trust feedback is used for getting the evolution of trust results, depends on the consistency and reliability of the services. The trust system explicitly depends on the credibility of the feedback of the users and their potential behaviors. Cloud consumers either give feedback regarding the trustworthiness of a particular cloud service or request trust assessment for the service. Let $j$ and $k$ be any two peers, then the feedback $f$ about $j$ given by $k$ is represented by $f_{jk}$ and is computed as given below.

$$f_{jk} = \frac{\sum_{i=1}^{n} S_{jk_i}}{t} \quad (3)$$

where $t$ is the total number of transactions performed by $k$ with $j$. $S_{jk_i}$ represents the satisfaction of $k$ on $j$ in $i^{th}$ transaction and its value is always assumed to be between 0 (not satisfied) and 1 (completely satisfied).

### 2.2.2 Reputation

Reputation clearly is an important aspect of trust establishment, a fact evident in the numerous reputation-based computational trust models in existence. The quality of the reputation system is primarily indicated by its accuracy and effectiveness in updating periodically. It is the one the important technique in trust because the feedback of the various cloud service consumers give the reputation of the service either positively or negatively.

### 2.2.3 Quality of Service

The Quality of Service (QoS) evaluation based on recommended trust is about feedback information of clients after executing service. Total set of QoS attributes is Q = T (Execution Time), D (Reliability), U (Availability), H (Throughput), and R (Comprehensive Evaluation) [12]. Eigen Trust algorithm is based on the notion of Transitive Trust. Each peer calculates the local trust value. Trust is stored in opinions, which are a 4-tuple (b,d,u,a): b-belief, d-disbelief, u-uncertainty, a-a-priori trust, where (b+d+u)=1.0 and a=[0, 1]. Eigen Trust requires the inclusion of pre-trusted users to get good performance. The longer time pasts, the more the trust degree reduces.

The importance of Eigen Trust in this paper is to get a global trust value with more weight given to pre-trusted peers. The main advantage of Eigen Trust is scalable computation and the trust does not weaken via transitivity. At the time of joining the new peer in the network and does not so far know anyone; the peer uses the perception of the network provided by the pre-trusted peers, from whom it can learn who else to trust.

## 2.3 Cloud Trust Models and Their Limitations

In the cloud computing environment, there are several trust models have been defined so far. But, none of these trust models satisfies the qualitative service provided to the user. So, our approach is to provide a qualitative service to the user from the cloud provider through recommender based trust. Some of the existing trust models in cloud computing are discussed below:

A cloud trust model which holds two layers of trust called inside trust layer and contracted trust layer is proposed by Sato *et al.* in [19]. Both the layers, however, give trust in a layered way yet the trust figured are inside to the affiliation. The Cloud Service Provider (CSP) has nothing to do with the advantages' security. So the affiliation needs to have a private cloud to secure its information which is unfeasible with minimal/medium affiliations.

Shen *et al.* in [20] and Shen and Tong [21] have proposed trusted processing development for trust appraisal. The principal shortcoming of this model is that the fundamental basic arranging relies on upon Trusted Computing Platform [TCP] which is difficult to facilitate with isolated registering concerning equipment.

Alhamad *et al.* in [1] have proposed Service Level Agreement (SLA) based trust show just and no utilization or evaluation has been created or depicted. This model is notoriety based trust that has a disadvantage that customer with high scores for reputation can cheat customer in a couple of trades regardless of the way that they get negative criticism. This model has a concentrated development displaying, so every one of the organizations and reputation information has the single purpose of disappointment.

In Role Based Trust show the trust relies on upon the parts, ID used for TCP, standard confirmation for affirmation. The gear keeps up a specialist key for each machine and it utilizes the master key to delivering unique subkey for every setup of the machine. The data mixed for one setup can't be decoded in another outline of the same machine. In case the machine's outline changes the session key of the adjacent machine won't be significant.

The Active Bundle Scheme [16] proposed in perspective of Identity Management model approach is free of an outcast, it is less disposed to assault as it lessens the threat of association ambushes and side channel assaults, on the other hand, it is slanted to foreswearing of organization as dynamic gathering may in like manner be not executed at all in the remote host.

# 3 Related Work

Cloud computing services are continually evolving and providers are offering new options, but it is not always in their best interests to enable data mobility. As IT organizations assume a greater role as service broker to the business, they must take ownership of ensuring that technologies from multiple vendors integrate seamlessly [14]. The cloud service provider's reputation reverts the overall view of a community towards that provider; therefore it is more useful for the cloud users (mostly individual users) in choosing a cloud service from many options without particular requirements.

## 3.1 Sources of Uncertainty

The data in the cloud is not always reliable and it is not under control by the provider. Also, applications that are hosted by the provider may not be available all the time, and/or they may not present the latest versions of these applications. As such, the client may encounter uncertainties with respect to these applications or the results that are delivered by these applications [6]. Most of the existing sources of uncertainty are:

1) Missing information.

2) Trusting the available information.

3) Inconsistency of available information.

4) Irrelevant information.

5) Interpretable information.

## 3.2 Modeling Uncertainty

There are various qualitative and quantitative approaches to model uncertainty. Uncertainty alone (without the consideration of Trust aspect of the information source) can be modeled as one of the following ways:

1) Probabilistic logic: This is the most common and widely used way of representing uncertainty.

2) Fuzzy logic: This approach allows to classify data into different classes called Fuzzy Sets, depending upon their relevance or closeness to the set.

3) Dempster - Shafer belief theory: It basically deals with measures of two main aspects belief and credibility.

4) Subjective logic: Based on probabilistic logic and Dempster-Shafer evidence Theory (DST), this approach has come up as one of the important ways to model uncertainty.

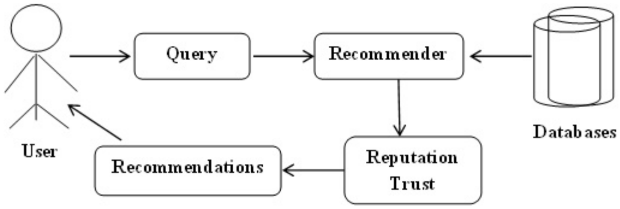


Figure 2: Reputation based trust organizational design

A data source may have different levels of trust at different times and different contexts. Uncertainty on a data source which has provided accurate measurements fairly regularly is less than compared to a new data source [17].

# 4 Reputation based Trust

Trust and reputation are related but different. Mostly, trust is between two entities; but the reputation of an entity is the collective opinion of a community towards that entity. Usually, an entity that has a high reputation is trusted by many entities in that community; an entity, who needs to make trust judgment on a trustee, may use the reputation to calculate or estimate the trust level of that trustee [9].

The reputation of a cloud service provider follows the overall view of a community against that provider, therefore it is more useful for the cloud users in choosing a cloud service from many options without particular requirements. The user always puts a request(query) to the recommender about the reliable services from the cloud databases and get several recommendations for the services based on trust and reputation as presented in Figure 2.

Trust in cloud computing services is based on several recommendations provided by numerous researchers. Nowadays recommender-based trust models are used in several e-commerce business enterprises, like Amazon and E-Bay. In recommender systems, it is clear based on the other users' ability to provide valuable recommendations. Since the number of direct interactions of the users is very small, so the number of direct relationships plays a minor role in the process of recommendation. The trust relationships between the users are not static but dynamically change over time which may lead to change the recommendation results [26].

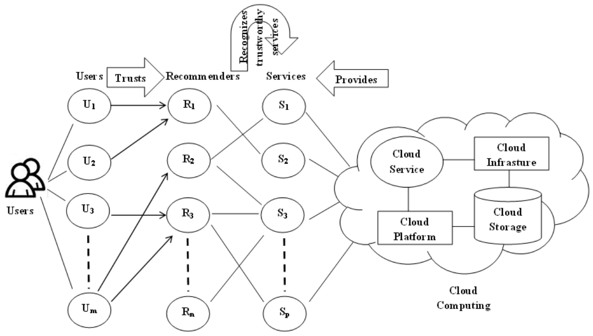


Figure 3: Relationship between users, recommenders and services

In Figure 3, $U_1$, $U_2$, $U_3$, $\cdots$, $U_m$ are the users, $R_1$, $R_2$, $R_3$, $\cdots$, $R_n$ are the recommenders and $S_1$, $S_2$, $S_3$, $S_p$ are the services provided by cloud service providers. Several users can put requests for different services like an infrastructure or software or a database. Each user in this environment gets the feedback from the recommenders on a particular service over a cloud service provider. Talal *et al.*, in [15] discussed trustworthiness of a certain cloud service $s$, and then the trust result,

$$T_r(s) = \frac{\sum\limits_{l=1}^{|V(s)|} F_c(l,s)}{|V(s)|} \tag{4}$$

Where $V(s)$ is the all trust feedback given to the cloud service $s, |V(s)|$ represents the length of $V(s)$. $F_c(l,s)$ are trust feedbacks from the $l^{th}$ cloud consumer weighted by the credibility. The weights can be calculated based on the consumer experience and satisfaction on the cloud services.

A cloud user is an individual or an organization that has a formal contract or arrangement with a cloud provider to use several resources made available by the cloud provider. Whereas, the cloud provider is an organization which provides cloud-based resources to the consumer. Finally, the recommender is also an individual/organization which analyzes the feedbacks coming from several attributes about the services in the cloud and also recommends to the cloud user whether he/she remain or terminate the services from the cloud providers.

Reputation-based trust is eventually assessed through several trust feedback mechanisms. Each one is having their own advantages and disadvantages during the process of trust evaluation. The users can select a particular service based on their preferences from the cloud service provider; meanwhile, the users can have a direct interaction with the recommenders to get the trust feedback. Non-negative weight is added to the feedback based on recent transactions. We also consider the old transactions so as to give specific weightage to the recommenders to calculate the trust value for the service providers.

# 5 Implementation and Experimental Results

In this research, a new approach of reputation-based trust evaluation was proposed which is based on weightage given to each and every transaction of the service for the cloud storage to minimize the uncertainty. Our projected trust model helps both the recommender and cloud user, where the user can make a decision on whether to continue or discontinue the service with the service provider.

Trust facilitates users to select the best available service in a diverse cloud infrastructure. Trust value is calculated using three parameters; capability, behavior, and feedback. A more serious type of attack is when malicious peers exploit file sharing networks to distribute viruses and Trojan horses. Peers also need to detect inauthentic file attacks, in which corrupted or blank files are passed off as legitimate files. Before going to undertake a transaction, peers should decide who to trust based on the reputation system which helps to address this need by establishing a trust mechanism [23].

Malicious peers can weaken the reputation system by assigning underprivileged reputation ratings to honest peers and privileged ratings to other malicious peers. Most of the existing reputation systems absorb into their trust model in view of the correlated trust, to deal with malicious feedback: peers reputed to provide trustworthy service, in general, will likely provide trustworthy feedback.

In Equation (3), an equal importance is given to the satisfaction values due to the most recent transactions as well as the oldest transactions. While in [22] different weights were attached to the satisfaction values, we suggest another addition to the above equation show the difference between the most recently performed the transaction and not so recently performed a transaction. So, that we should have to minimize the responses from malicious peers [18]. Assume $int$ is an interval representing a set of transactions performed during a time period. Let $int = 0$ represent the most recent period and int=1 be the next recent period. Assume the $i^{th}$ transaction was performed in an interval $int$. Then its corresponding $S_{jk_i}$ is adjusted according to the following equation.

$$f_{jk} = \frac{\sum_{i=1,n\ \ int=0,|tf|} \left[\frac{(ts-2^{int})+1}{ts}\right] * S_{jk_i}}{n} \qquad (5)$$

where $ts$ stands for timestamp which represents the exact time taken when the transaction was performed, $tf$ for timeframe where the considerable past time is categorized into intervals $int$ numbered from 0 to $|tf|$ onwards. Equation (2) allows graceful reduction of feedback ratings as they get old. Figure 2(b) shows how a satisfaction rating fades with time. The significance is that the recent ratings overweigh the past ratings. The advantages are twofold:

1) The recent feedbacks are given more importance and hence;

2) Reputation computation gets more dynamic.

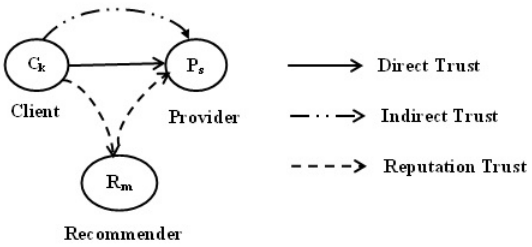


Figure 4: Reputed trust through recommendations

As in Figure 4, $C_k$ is the service request by the customer to the provider through direct interaction, $P_s$ is the various services provided by the CSP to the customers, and $R_m$ is the Recommender used for giving the feedback to the customers about the trusted services. So, the weightage to the specified service for the user is:

$$w = (tv)^p \qquad (6)$$

$0 \leq tv \leq 1$, where $tv$ is a single value for local trust which is suggested by the recommender and $p$ is the time period in which the transaction is done between the user and the service provider. The local trust value can be projected based on feedback given by the trustworthy users to the recommenders.

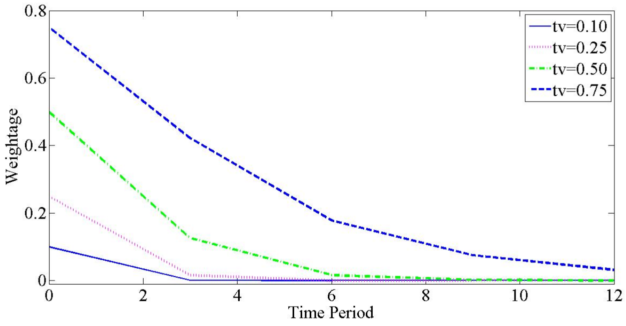


Figure 5: Weightage to the cloud services based on time and trust values

Since, the trust values of the services always lie in between 0 and 1, and then in Fig.5 shows that if the trust value is 0.1, the weightage given to the services is computed based on the time period. For the longest time period and low trust value, the weightage is below 0.1. If the trust value is 0.25 and the weightage is below 0.2. If the trust value is 0.50 and the weightage is below 0.4. Finally, if the trust value is 0.75, then the weightage is nearest to 0.5.

So, now we have to calculate the reputation based trust value for the specified service for a particular user in a specific time period with the given weightage is:

$$RT\,(Q,S) = G_t.\sqrt(w) \qquad (7)$$

Where, $Q$ is a service requester, $S$ is a service provider, $G_t$ is a global trust value on a particular service and $w$ is weightage which is already computed in the equation 4. Here, the square root is used for increasing the weightage so as to give the preference to the recent transactions. Based on the above equation, we calculate the reputed trust to each and every transaction between the service requester and service provider, in order to minimize the uncertainty about the services.

The algorithms that are used for the above computations are presented as follows.

Algorithm 1 is used to calculate the difference between the most recent transaction and not so recently performed a transaction. Here $S$ is the satisfaction value, $N$ is the number of Common Vendors and $f$ is the feedback.

Algorithm 2 is used to calculate the weightage to the specified service. Here $tv$ is local trust value, $w$ is the weightage and $p$ is time period.

**Algorithm 1** Age of transaction

```
Begin
Input: $S, N$ Output: $f$
Let $ts$ be the timestamp
Let $tv$ is the transaction value
Time is categorized into intervals, $int \ \epsilon \ tf$
If $int$ is in between 0 and $tf$, then $tv := tv + (ts - 2^{int}) + 1/ts, where\ ts \ \epsilon \ tf$
Compute the feedback
End
```

**Algorithm 2** Weightage

```
Begin
Input: $tv$, $p$ Output: $w$
Let $p$ is the time period
Calculate the local trust $tv$ based on feedback $f_{jk}$
Compute the weighting factor $(w) = (tv)^p$
End
```

Algorithm 3 is used to calculate the reputed trust between the service requester and service provider for each and every transaction. Here $Q$ is a service requester, $S$ is service provider, $G_t$ is a global trust value on a particular service.

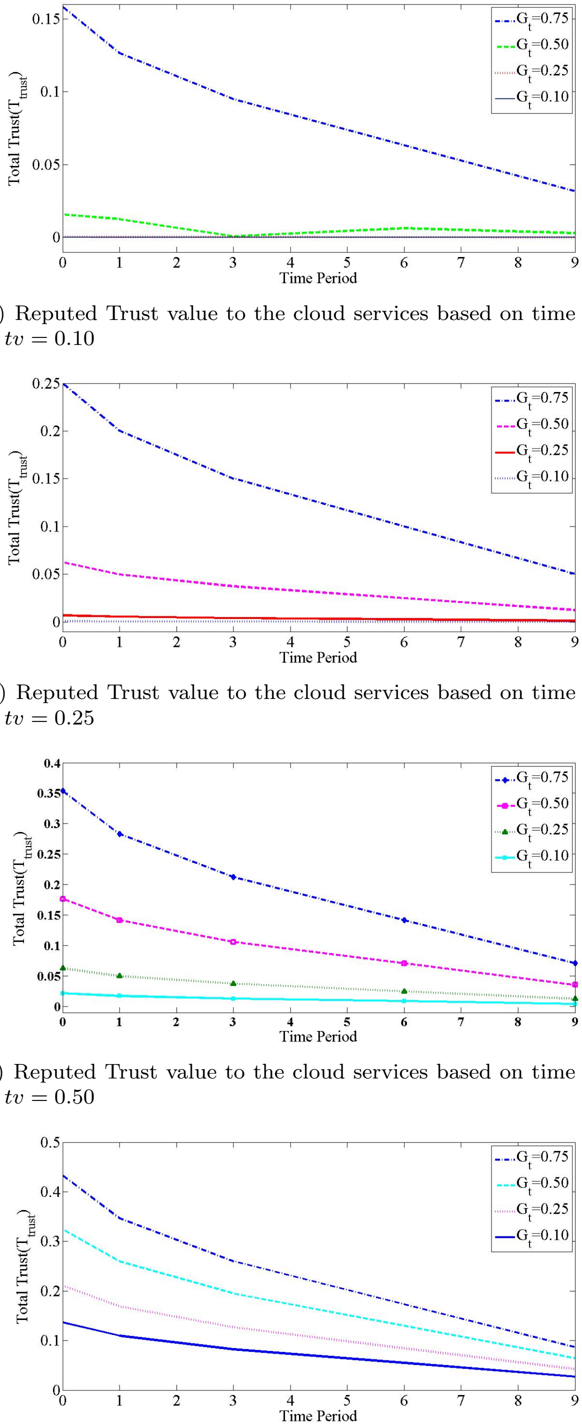
**Algorithm 3** Reputed Trust

```
Begin
Input: $G$, $w$ Output: $RT$
Let $G$ is the Global Trust
Give more weightage to the recent transactions using $\sqrt(w)$
Compute the Reputed Trust using $G_t.\sqrt(w)$
End
```

## 6 Case Study

In this section, a case study related to the distributed file sharing service has been represented under SaaS in a cloud environment and trustworthiness of the related entities have been evaluated based on the proposed trust management model. In the cloud environment, let a specific service of distributing files sharing, where the files have a desired distribution and availability. When any entity wants to share a file in cloud environment then first it needs to ensure that whether a node or entity is trustworthy or not. The trustworthiness can be decided based on service level agreement (SLA) like processing capacity, recovery time, connectivity, peak-load performance, and availability.

In the Reputed Trust Model (RTM), let the service provider be the vendor $v$ and the trust relationship is established using trust degree based on a request sent to other entities in the cloud. Each entity will maintain two trust tables: direct trust table and the recommended list table. If an entity wants to calculate the trust degree of



(a) Reputed Trust value to the cloud services based on time period at $tv = 0.10$

(b) Reputed Trust value to the cloud services based on time period at $tv = 0.25$

(c) Reputed Trust value to the cloud services based on time period at $tv = 0.50$

(d) Reputed Trust value to the cloud services based on time period at $tv = 0.75$

Figure 6: Reputed Trust($RT$) values to the cloud services

another entity then it first checks the direct trust table. If the trust degree value for the entity exists then it will guarantee for last communication time and then calculate the decay function using Equation (5).

After calculating decay function, reputation based trust for the specified service can be calculated using Equation (7) and where the weightage factor also will be considered. The reputation computation is more dynamic according to decay function effect. Also, the comparative review between the proposed method and the related work is shown in Table 1.

Table 1: Comparative study of proposed method with methods of the related work

| Mechanisms | Trust | Reputation | Feedback | Accessibility |
|---|---|---|---|---|
| Proposed Method | ✓ | ✓ | ✓ | ✓ |
| Dipen *et al.*[5] | ✓ | × | × | × |
| Ayesha *et al.*[7] | ✓ | ✓ | × | × |
| Shaik *et al.*[8] | ✓ | ✓ | × | ✓ |
| Firdhous *et al.*[9] | ✓ | ✓ | ✓ | × |
| Mahbub *et al.*[10] | ✓ | ✓ | × | × |
| Alhamad *et al.*[15] | ✓ | ✓ | ✓ | × |
| Huang *et al.*[20] | ✓ | ✓ | × | × |

As shown in Table 1, in most of the related work, just some options in the field of the trusted service description are studied. For example, Dipen *et al.* [6] have considered only the trust, Ayesha *et al.* [8] and Huang *et al.* [23] have considered the trust and reputation. Also, the results show that the provided method acts well than the other related work.

# 7 Conclusions

- Service availability is one of the significant challenges in the cloud storage to predict the number of requests by several users for the service has to handle at a single point in time. Even though it achieves high availability of services but faces the uncertainties of reliable transactions between the cloud users and providers. Achieving trustworthy services is possible through a reputation-based trust as we are here presented in this research. The mechanism suggested in this paper consider the several communities in general, and allows reputation correction based on the type of community the particular peer belongs to. The simulation results that support our claims have been presented.
- In this research, a number of results can be considered based on the total trust values provided by the recommenders. Even though our proposed system will improve the availability of services by minimizing the malicious peers, but still there is some limitation in our new approach. The proposed system can not be addressed the vendor lock-in; migration of user data and service from one vendor to other is nearly impossible. Future improvements that need to be addressed are how to combine trust and clustering relationships to improve the algorithm performance, and performance of the services in the cloud.

# Acknowledgments

The authors gratefully acknowledge the anonymous reviewers for their valuable comments which helped us to improve the quality and presentation of this paper.

# References

[1] M. Alhamad, T. Dillon, E. Chang, "SLA-based trust model for cloud computing," in *13th International Conference on Network-Based Information Systems*, pp. 321-324, 2010.

[2] S. S. Babu, A. Raha, M. K. Naskar, "Trust evaluation based on node's characteristics and neighbouring node's recommendations for WSN," *Journal of Wireless Sensor Network*, pp. 157-172, vol. 6, no. 8, 2014.

[3] J. Chen, Y. Wang, X. Wang, "On-demand security architecture for cloud computing," *Computer*, pp. 73-78, July 2012.

[4] D. Contractor, D. Patel, "Accountability in cloud computing by means of chain of trust," *International Journal of Network Security*, vol. 19, no. 2, pp. 251-259, Mar. 2017. (DOI: 10.6633/IJNS.201703.19(2).10)

[5] Dimension Data, *Service Providers Need Flexible Cloud Services to Compete*, White Paper, 2012. (www.dimensiondata.com/onecloud)

[6] B. N. Farah, "A model for managing uncertainty on the cloud," *Journal of Management Policy and Practice*, vol. 14, no. 6, 2013.

[7] M. Firdhous, O. Ghazali, S. Hassan, "Trust management in cloud computing: A critical review," *International Journal on Advances in ICT for Emerging Regions*, pp. 24-36, vol. 4, no. 2, 2011.

[8] S. M. Habib, S. Hauke, S. Ries and M. Muhlhauser, "Trust as a facilitator in cloud computing: A survey," *Journal of Cloud Computing: Advances, Systems and Applications*, pp. 1-19, 2012.

[9] J. Huang, D. M. Nicol, "Trust mechanisms for cloud computing," *Journal of Cloud Computing: Advances, Systems and Applications*, pp. 2-9, 2013.

[10] A. Kanwal, R. Masood, U. E. Ghazia, M. A. Shibli and A. G. Abbasi, "Assessment criteria for trust models in cloud computing," in *IEEE International*

*Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*, pp. 254-261, 2013.

[11] C. Lan, H. Li, S. Yin, L. Teng, "A new security cloud storage data encryption scheme based on identity proxy re-encryption," *International Journal of Network Security*, vol. 19, no. 5, pp. 804-810, Sep. 2017. (DOI: 10.6633/IJNS.201709.19(5).18)

[12] W. Lu, X. Hu, S. Wang, X. Li, "A multi-criteria QoS-aware trust service composition algorithm in cloud computing environments," *International Journal of Grid and Distributed Computing*, vol. 7, no. 1, pp. 77-88, 2014.

[13] Nasuni, *Understanding Security in Cloud Storage*, White Paper, 2010. (www.nasuni.com)

[14] NetApp,*Data Fabric: Realize the Full Potential of the Hybrid Cloud*, White Paper, 2017. (www.netapp.com/datafabric)

[15] T. H. Noor, Q. Z. Sheng, "Trust as a service: A framework for trust management in cloud environments," in *Web Information System Engineering (WISE'11)*, pp. 314-321, 2011.

[16] R. Ranchal, B. Bhargava, "Protection of identity information in cloud computing without trusted third party," *29th IEEE International Symposium on Reliable Distributed Systems*, pp. 1060-9857, 2010.

[17] M. Ravi, Y. Demazeau, F. Ramparany, "Managing trust and uncertainty for distributed AI systems," *RJCIA*, 2014. (https://rjcia2014.greyc.fr/sites/rjcia2014.greyc.fr/files/rjcia2014_submission_3.pdf)

[18] P. RVVSV, V. Srinivas, V. V. Kumari, R. KVSVN, "An effective calculation of reputation in P2P networks," *Journal of Networks*, vol. 4, no. 5, July, 2009.

[19] H. Sato, A. Kanai, S. Tanimoto, "A cloud trust model in a security aware cloud," *10th IEEE/IPSJ International Symposium on Applications and the Internet (SAINT'10)*, pp. 121-124, 2010.

[20] Z. Shen, L. Li, F. Yan, X. Wu, "Cloud computing system based on trusted computing platform," in *International Conference on Intelligent Computation Technology and Automation (ICICTA'10)*, vol. 1, pp. 942-945, 2010.

[21] Z. Shen, Q. Tong, "The security of cloud computing system enabled by trusted computing technology," in *2nd International Conference on Signal Processing Systems (ICSPS'10)*, vol. 2, pp. 11-15, 2010.

[22] M. Srivatsa, L. Xiong, L. Liu, "Trustguard: Countering vulnerabilities in reputation management for decentralized overlay networks," in *Proceedings of the 14th International Conference of World Wide Web*, pp. 422-431, 2005.

[23] G. Swamynathan, B. Y. Zhao, K. C. Almeroth, "Decoupling service and feedback trust in a Peer-to-Peer reputation system," *International Symposium on Parallel and Distributed Processing and Applications (ISPA'05)*, pp. 82-90, 2005.

[24] D. Thiyagarajan, R. Ganesan, "Cryptographically imposed model for efficient multiple keyword-based search over encrypted data in cloud by secure index using bloom filter and false random bit generator," *International Journal of Network Security*, vol. 19, no. 3, pp .413-420, May 2017. (DOI: 10.6633/IJNS.201703.19(3).10)

[25] M. Trenz, J. C. Huntgeburth, D. J. Veit, "The role of uncertainty in cloud computing continuance: Antecendents, mitigators, and consequences," in *Proceedings of the 21st European Conference on Information Systems (ECIS'13)*, pp. 147, 2013.

[26] J. Yuan, L. Li, "Recommendation based on trust diffusion model," *Research Article, The Scientific World Journal*, June, 2014.

# Biography

**Vegi Srinivas** received his M.Sc. in Computer Science and M.Tech. in Computer Science and Engineering from Andhra University. He is a Research Scholar in the Department of Computer Science & Engineering, JNTUK, Kakinada. He is currently working as an Associate Professor in Dadi Institute of Engineering and Technology, Anakapalli, Visakhapatnam, India. His main areas of interests are Cloud computing, Security, Privacy and Trusted Computing. He is a member of IEEE, ACM, CSTA and Life Member of CSI & ISTE.

**Vatsavayi Valli Kumari** received her B.E. in Electronics and Communication Engineering and M.Tech. and PhD in Computer Science and Engineering all from Andhra University, India and is currently working as Professor in the same department. Her research interests include Security and Privacy issues in Data Engineering, Network Security and E-Commerce. She is a member of IEEE and ACM and is a fellow of IETE.

**KVSVN Raju** received the B.E. in Electrical Engineering from Government College of Engineering, Kakinada, India and M.E. in Control Systems from Andhra University, India and obtained the Ph.D. in Computer Science and Technology from IIT, Kharagpur, India. He is currently working as Director, R & D Cell, Anil Neerukonda Institute of Technology & Sciences (ANITS). He is a retired Professor in the Department of Computer Science and Systems Engineering at A.U. College of Engineering, Visakhapatnam, India. His research interests include Data Engineering, Security Engineering and Software Engineering.

# An Efficient Compromised Nodes Detection System in Wireless Sensor Networks

Xiaolong Xu[1], Zhonghe Gao[2], Lijuan Han[1]
*(Corresponding author: Xiaolong Xu)*

Experiment Teaching Center, Qufu Normal University[1]
Rizhao Shandong 276826, China
(Email: xiaolongxu@foxmail.com)
Institute of Software, Qufu Normal University[2]
Qufu Shandong 273165, China

*(Received Jan. 26, 2018; revised and accepted Apr. 28, 2018)*

## Abstract

Wireless sensor networks have limited resources and are deployed in an open environment, this makes it vulnerable to attacks. The CND method we proposed can accurately detect the compromised nodes in wireless sensor networks. Experimental results show that the CND system has the following advantages: The detection rate and false positive rate are better than the existing compromised node detection methods; It is not vulnerable to slander attacks; It can run in most wireless sensor networks and can automatically adjust the detection behavior according to the network transmission; It requires only small memory and low communication overhead, so it can be applied to large-scale networks.

*Keywords: Attack Detection; CND; Compromised Node; Wireless Sensor Network*

## 1 Introduction

Sensor nodes are cheap and autonomous, this extends wireless sensor networks to several applications, including environmental monitoring, medical care, smart home, traffic control, and so on. However, the application of wireless sensor networks has been extended to many security fields, and information security has become an important aspect of people's concern. In hostile environments, it is impossible to trust reports from wireless sensor networks without information security.

However, sensor nodes only have limited resources, such as the limited computing power, memory and battery life, and are usually deployed in an open environment, so they are vulnerable to attacks, and attackers can control some nodes [14]. If not detected, a compromised node is considered a network's authorized participant, which can use its own authority to launch an internal attack.

Therefore, security oriented wireless sensor networks must take measures to prevent node compromise. Generally, security policy can be divided into prevention, detection and recovery. Because of the small size and low cost of micro sensor nodes, limited resource limits the effectiveness of the defense mechanisms [15]. Attackers can use more powerful machines, such as laptops, to capture nodes. So defensive measures can only delay an attacker's attack.

Attacks on wireless sensor networks can be divided into external attacks and internal attacks. The external attacker is located outside the wireless sensor network, and the internal attacker is the authorized user of the wireless sensor network. Both external attackers and internal attackers can capture sensor nodes and make them compromised nodes. Malicious codes are running on compromised nodes, so they become nodes controlled by attackers, which seriously threaten the security of wireless sensor networks. Compromised nodes detection is of great importance for ensuring the security of wireless sensor networks [10].

Detection mechanism is an active measure to prevent node compromise. Once the compromised node is detected, appropriate measures are taken to reduce the loss caused by compromise. At present, there are many kinds of compromised node detection methods, but they have various disadvantages. In addition, most of the detection methods are for specific situations, and they do not perform well in other cases. For example, most of the detection systems do not consider lossy environments, and packet loss is more common in wireless sensor networks. The packet loss rate of 20% reduces the detection rate by more than 50% and leads to false positive rates of over 90% [5].

The intrusion detection system CND (Compromised Nodes Detection) is proposed to identify the compromised nodes in wireless sensor networks. This detection method has the following advantages:

**Accuracy:** CND can detect the compromised nodes in wireless sensor networks timely and accurately. Specifically, it requires a high detection rate and low false alarm rate and short detection time. High detection rate means that the vast majority of compromise behaviors can be detected. Low false positive rate means that most reports of compromised nodes are accurate, so you can rest assured that such nodes should be taken corresponding measures. Finally, shorter detection time limits the amount of malicious activities that a compromised node can perform before being detected.

**Flexibility:** CND does not change for a particular application or deployment because it limits its scope of application. It considers the underlying network as little as possible and can be used in most situations to detect compromise behaviors.

**Robustness:** Compromised nodes may try to damage detection system through malicious behaviors, such as slander attack. They will send false information so that the legitimate node is mistaken for a compromised node. CND must be able to prevent such malicious behaviors. Even with full knowledge of CND, attackers can't use it to control the rest of the network.

**Extendibility:** Because micro sensor nodes only have limited resources, some high cost applications will interfere with other applications and reduce the lifetime of sensor nodes [12]. CND has very low overhead, so it has only a minor influence on other applications deployed in the network.

The main goal of CND is to provide a system to identify compromised nodes accurately, so as to improve the overall security of wireless sensor networks. CND uses a lightweight distributed architecture that can be deployed in resource limited devices, such as wireless sensor nodes. It has little influence on other applications and network lifetime. Through many experiments, we find that CND has more accurate detection ability than other similar systems, and can be extended to tens of thousands of nodes.

The rest of this paper is structured as follows: The first chapter reviews the previous research of compromised nodes detection in wireless sensor networks. The second section discusses the proposed system and threat model. From Sections 3 to 6, the design, implementation and evaluation of CND are introduced respectively. Section 7 is the conclusion and future work.

## 2 Literature Review

Most detection methods of wireless sensor networks mainly focus on some specific attacks, such as node replication attack, wormhole attack, sybil attack, etc [9]. Although they may detect compromised nodes indirectly, attackers can escape detection by avoiding target attacks.

The traditional method of detecting compromised nodes is authentication. The authentication method is to check the changes of node memory to find out whether the modified code is running. The advantage of this approach is the ability to detect compromised nodes that do not perform destructive activities. A variety of software based authentication techniques have been proposed for wireless sensor networks [3], but software based security authentication has not been implemented in wireless sensor networks yet.

Other programmes focus on monitoring suspicious communication behaviors. Suspicious behavior can be confirmed by anomaly detection or rule based detection. Anomaly detection establishes a baseline of normal behaviors and considers a behavior abnormal when detected beyond baseline. For example, intrusion detection system proposed by Onat and Miri mainly monitors two features - packet arrival rate and received power [13]. The detection nodes continuously monitor these two features from adjacent nodes and are considered abnormal if new data is found to deviate from the established baseline. Malicious behaviors that can cause changes in these two features will be detected, such as replay attacks. Rule based detection judges a behavior as malicious when the behavior is found to consistent with the rules set earlier. For example, the COOL system is an intrusion detection system that detects the compromised nodes using the relationship between incoming and outgoing messages [16]. The COOL system is based on the idea that the vast majority of outgoing messages should be forwarded to incoming messages. When a node sends more information than it receives and reaches a threshold, it is considered a compromised node.

Compared with the existing methods, the proposed method is more flexible, robust and scalable. CND can accurately detect compromised nodes in the presence of packet loss, without being affected by other applications running on the sensor nodes. It is very effective in combating large-scale slander attacks. In this attack, the compromised nodes hinder the detection process. In addition, it has the advantage of low overhead. This allows it to be deployed in a wireless sensor network with thousands of nodes without affecting other applications deployed, without significantly shortening the lifetime of the network.

## 3 System and Threat Model

A CND system is designed based on the following common features of wireless sensor networks and compromised nodes.

1) The sensor nodes are densely deployed in the network, so that the sensor nodes have overlapping perception range. Thus, an event may be detected by multiple nodes at the same time. Because of range overlap, one sensor node can monitor the behavior of its neighbors.

2) Sensor nodes have limited energy, calculation ability, and communication capability. For example, the Mica2 micro sensor uses an Atmel microprocessor with a main frequency of 4 MHz and a word length of 8 bits, it is equipped with 128 KB instruction memory and 4 KB RAM.

3) A routing protocol that forwards messages between the base station and the node is required.

4) A base station is a higher order device, such as a computer placed in a secure location.

5) The sensor node has a unique identifier that enables the base station to know which node corresponds to the reported compromise behavior.

6) All messages have time stamps.

7) Attackers can capture nodes either by physical capture or by means of wireless communication channels. Once a node is compromised, all the information, including the key, is acquired by the attacker.

8) Although compromised nodes can perform any number of attacks to reduce the security of network, this paper focuses on compromised nodes that perform malicious behaviors, such as forges and tampers with data.

CND can make use of these characteristics to achieve accurate identification of compromised nodes, and only a small amount of overhead is needed.

# 4 System Architecture

When designing CND, you must determine whether to use a distributed, centralized, or hybrid architecture. Building a pure distributed intrusion detection system is very challenging because the limited resources of sensor nodes restrict the use of complex algorithms. For example, traditional security protocols, such as modulo operations used by RSA, run more difficult on 8 bit node processors. Complex computations can be distributed over a number of sensor nodes, but nodes that engage in critical operations can become compromised nodes, resulting in spurious results. In contrast, centralized solutions do not have these problems because the base station has more resources and is more secure. However, the data received by the base station from compromised nodes may be false. Therefore, the network needs to have some degree of additional functions to detect false data from the nodes.

Thus, CND takes a hybrid approach, as shown in Figure 1. The system consists of two parts: a distributed system running on each node in the network and a centralized system running on the base station.

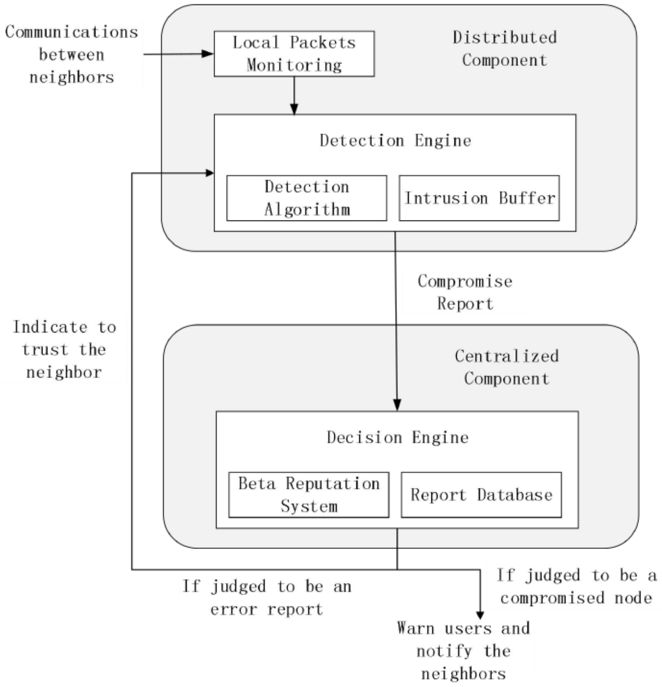


Figure 1: The framework of CND

**Distributed component:** Copies of this component run on each sensor node and run concurrently with applications, routing protocols, and so on. Each copy is responsible for detecting possible compromise behavior among adjacent nodes and reporting to the base station. This detection relies on adjacent node monitoring, and each node records and analyzes the behavior of its neighboring nodes. Because of the broadcast characteristics of wireless sensor networks, this does not cause excessive communication overhead.

**Centralized component:** A base station is a higher level device, so CND uses it to perform complex analysis to determine whether a reported compromise is correct. The base station collects data from the entire network, which is what the sensor nodes cannot do with their own local views and limited resources.

After the network deployment, there is an initial setup phase. During this phase, nodes establish their neighbor lists, routes to base stations, and so on. The network is safe for some time after the initial deployment, and this phase does not introduce any vulnerabilities because the attacker cannot immediately capture a node after the network has just been deployed. If this requirement cannot be reached, then the information must be preprogrammed to each node before the network is deployed.

## 4.1 Distributed Component

Each sensor node has a distributed component running on it that will record data from neighboring nodes and establish baselines based on these records. The baseline indicates the normal behavior of the nodes, and the behavior that deviates from the baseline will be considered an abnormal behavior. If a neighbor node continues to perform an abnormal behavior, it will be identified as a compromised node and reported to the base station.

Considering instantaneous errors, such as collisions or other unmalicious behaviors, CND is flexible in determining a node as a compromised node and can tolerate certain abnormal behavior. When judging whether a neighbor node is abnormal, there is no cooperation between nodes. This independent decision process implies that the compromised nodes can not affect the perspective of legitimate neighbor nodes.

#### 4.1.1 Monitoring Features

The first step in designing any security system based on detection is to select the system features to be monitored. To support most wireless sensor networks, CND monitors only common features of wireless sensor networks.

1) Sensor reading: By monitoring sensor readings, attacks attempting to distort the collected information can be detected [4].

2) Received power: In a static network, the received power should remain constant. Fluctuations may be caused by changes in the location of communication hardware or corresponding nodes.

3) Sending rate: Most applications read sensor readings and periodically send them. Routing packets are also sent periodically. Therefore, the rate at which packets are sent by nodes should follow a consistent pattern. Most attacks, such as selective forwarding, sybil attack, replay attack, etc., can cause metric deviation. In addition, a sudden idle period may be caused by opponent's rewriting node program.

4) Receiving rate: The ratio of incoming and outgoing packets should be constant, because the outgoing packets can only be those routed or generated by nodes. A neighbor node whose receiving rate has changed, but its sending rate does not change, such a node may be a compromised node. It should be noted that, regardless of whether the data is encrypted, the header of a packet is usually visible to all nodes.

Because most wireless sensor networks have these characteristics, CND has a wide range of applicability. However, these features may not be appropriate for two scenarios: (1) Packets can only be decrypted by base stations; (2) Applications rarely communicate with base stations.

The first scenario will appear when the confidentiality of the information is very important. Since compromised nodes cannot be detected immediately and blocked, some sent messages may be tapped by compromised nodes. Therefore, packets can be encrypted and only base stations can decrypt them. Under such conditions, the number of monitored neighbors can be increased to make up for defects that cannot monitor sensor readings, thus enabling CND to achieve appropriate performance by occupying a little more memory.

The second scenario is caused by applications that are not periodically communicated. For example, wireless sensor networks in a demilitarization zone send messages only when an attack is detected, and they do not communicate in a secure environment. Due to insufficient monitoring information, the baseline cannot be established for most features. CND compensates by making the node send its unique identifier at a certain speed. Long silence will cause the overcome nodes cannot be found, therefore a certain amount of communication overhead is needed. This behavior pattern is used only when the amount of communication in the application is small.

#### 4.1.2 Detection Algorithm

There are two kinds of algorithms for detecting abnormal behaviors: anomaly detection and rule based detection. They all use records of monitoring system characteristics. The anomaly detection algorithm uses the existing record to establish the baseline, and any new record that deviates from the baseline to a certain extent is considered to be an abnormal behavior. On the contrary, rule based detection should establish a specific standard. For example, any two packets have the same header means a replay attack occurred. In CND system, the main attention is paid to anomaly detection algorithms to meet the requirements of CND for flexibility. Rule based algorithms aim at special situations, and the rules must be updated for each new situation.

The distributed component of CND can be divided into five algorithms for detecting attack behaviors: The first four algorithms are anomaly detection algorithm, which uses network features such as sensor reading, received power, sending rate and receiving rate; The fifth algorithm is a rule based detection algorithm.

For rule based algorithms, if a node detects a new neighbor that conforms to the characteristics of the previously predefined rule base, it is considered that the new neighbor is a compromised node.

These rules can prevent compromised nodes and external attackers masquerading as normal nodes without being discovered, and Figure 2 illustrates this nature. Suppose that node A is compromised and want to impersonate another node, if node D does not detect new neighbors, it cannot impersonate B or C; if node B and C do not detect new neighbors, it cannot impersonate D; if node B, C, D do not detect new neighbors, it cannot impersonate any other node. Therefore, if there are enough neighbors to monitor each other, any attack and impersonation can be detected.

All anomaly detection algorithms follow a similar approach. Each node sets two buffers for each monitored neighbor: a packet buffer and an abnormal behavior buffer. All anomaly detection algorithms share the buffer and use the sliding window mechanism. It stores the last

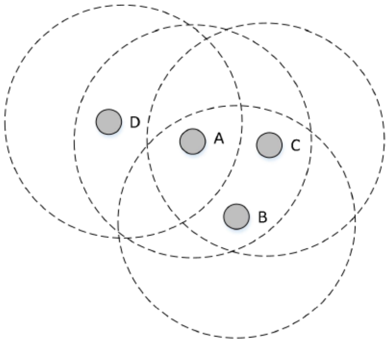


Figure 2: Node A must be detected by neighbor nodes if it wants to impersonate other nodes

N packets or reports of the corresponding neighbors. The data stored in the packet buffer is used to compute the baseline of the neighbors. The new packet is compared with the baseline, and any packet that deviates from the baseline beyond a certain threshold is considered to be abnormal. The abnormal packet indicates the intrusion behavior, and causes the detection node to produce abnormal behavior report. All the reports are added to the abnormal behavior buffer. When the cumulative report number in the abnormal behavior buffer exceeds the threshold, the node will report the corresponding neighbor to the base station as a compromised node.

An overview of the algorithm that uses the received power is shown in Figure 3, and the corresponding equation is:

$$power_{new} - power_{max} > T, \quad \text{if } power_{new} > power_{max}$$
$$power_{min} - power_{new} > T, \quad \text{if } power_{new} < power_{min}$$

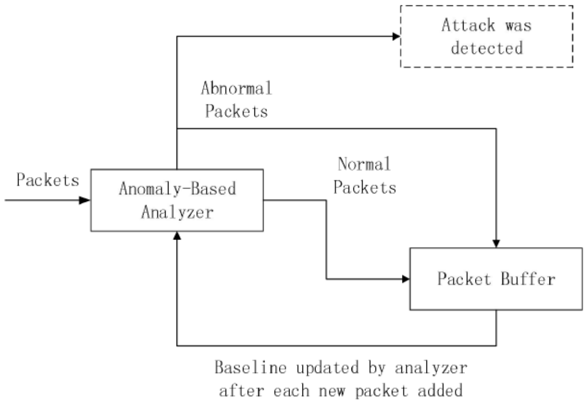


Figure 3: Overview of the detection algorithm using received power and sensor readings

This algorithm calculates the maximum and minimum values of packet received power in the packet buffer. If the received power of a new packet is lower than the minimum value of T or higher than the maximum value of T, it is considered to be abnormal. Abnormal data packets are added to the packet buffer so that anomalies caused by environmental changes can be taken into account when calculating baselines in the future.

The sensor reading algorithm is almost the same as the algorithm using the received power, and the only difference is the use of sensor readings from nodes and neighbor nodes instead of the received power.

Figure 4 shows an overview of the algorithm that uses the sending rate. It calculates two rates: The sending rate of the last $N_2$ packets $rate_{N_2}$ and the sending rate of the last $N$ packets $rate_N (N > N_2)$. If the ratio of these two rates is higher than the threshold $K$, the corresponding neighbor nodes are considered to be compromised nodes.

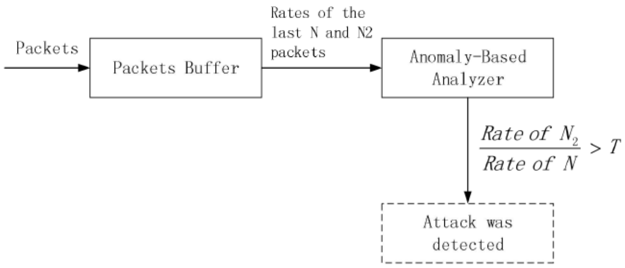

Packets
Packets Buffer
Rates of the last N and N2 packets
Anomaly-Based Analyzer
$\frac{Rate\ of\ N_2}{Rate\ of\ N} > T$
Attack was detected

Figure 4: Overview of the detection algorithm using sending rate and receiving rate

The algorithm that uses receiving rate differs only in two ways. First, instead of calculating packets sent by neighbors, the data packets received by neighbors are calculated. Second, the rate is replaced by the ratio of the sending rate and receiving rate, that is, $rate_{N_2}$ becomes $rate_{sent_{N_2}}/rate_{rec_{N_2}}$, and $rate_N$ becomes $rate_{sent_N}/rate_{rec_N}$.

All illegal behaviors detected by anomaly detection algorithm are stored in a shared illegal behavior buffer. Each reported illegal behavior is assigned a weight based on the detection time $t_{stamp}$ and the current time $t_{current}$. When a neighbor's illegal behavior is detected, the weight of its illegal behavior is calculated:

$$\sum_M (t_{current} - t_{stamp}) + 0.3 \sum_m (t - current - t_{stamp}). \quad (1)$$

Where $M$ represents all detected illegal behaviors of the same type, and $m$ represents all other types of illegal behavior. When the result of Equation (1) exceeds the threshold $T_M$, the corresponding neighbors are considered to be compromised. After thousands of simulations, the weight 0.3, the optimal value of the threshold and other parameters of the equation can be determined.

Once a node determines that a neighbor A is compromised, it sends three reports about this node to the base station. Each of these reports has three domains, which are reporter, reported node, and illegal behavior type. Since then, the reporter will continue to record information from node A, but will no longer detect abnormal behavior unless instructed by the base station.

### 4.2 Centralized Component

A centralized component runs on a base station to determine whether a reported node is really compromised based on data from other nodes. If the reported node is a compromised node, the base station will notify the user and execute the recovery process, such as ignoring all the messages of the node. If the reported node is not compromised, the base station will notify the reporter to treat it as a non compromised node and continue to monitor it.

Users will receive notifications for all new neighbor reports. If an actual node is added to the network, the user can notify the network that the node is not malicious. For other cases, the base station will process data based on reports from other nodes. In order to determine whether a reported node is compromised, CND uses the beta reputation system [12]. Research shows that Beta reputation system can accurately detect illegal behavior and reduce false positives rate based on a large number of reports. Because the system takes historical factors into account, in order to successfully hide a compromised node A, the average 72% of the neighbors of node A need to become compromised nodes. This is better than other programs (such as Majority Voting) of 33% 50% [7]. Beta reputation system uses probability density function and multi source feedback to determine reputation rating. For this paper, reputation rating is to judge whether a node beyond the threshold is a compromised node.

In Beta reputation system, the probability is $\rho$, each reported event is given two parameters $\alpha$ and $\beta$ of beta distribution. $f(\rho|\alpha,\beta)$ can be represented by $\Gamma$ function:

$$\begin{aligned} f(\rho|\alpha,\beta) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)}\rho^{\alpha-1}(1-\rho)^{\beta-1}, \\ & \qquad 0 \leq \rho \leq 1, \alpha > 0, \beta > 0. \end{aligned}$$

The parameter $\alpha$ and $\beta$ denote the weighted sum of all previous reports of the reported nodes and the number of compromised nodes within two hops of the reported node. This allows the network topology and past reports to influence the final decision on whether a reported node is a compromised node. The initial baseline value is determined during installation.

The base station knows the neighbor information of each node, and this information is collected during the network installation. If there are more than one neighbor reporting that $A$ is a compromised node, then there is a higher likelihood that it is right. The longer the history of the report, the more compromised nodes there may be.

On the other hand, if the reported node is unlikely to be a compromised node, then the report node $B$ may be a compromised node and initiate a slander attack on node $A$. In this case, the base station will notify the other nodes that node $B$ is a compromised node, and alerts the user, and starts the recovery program.

This method can prevent attackers from using CND to attack the network without being detected. If a compromised node poses as a base station, the nodes near the base station on the routing path will detect messages from the wrong direction and alert the base station. Therefore, once an attacker is posing as a base station, it will be detected immediately.

CND is not vulnerable to slander attacks. As mentioned earlier, masquerading as other nodes will be detected by neighbor nodes, and the nodes do not affect each other. Suppose a compromised node $C$ wants to slander its neighbor node $D$, because the base station knows the neighbor of node $C$, so reporting a non neighbor node can also lead to detection. The only possible slander attack is that node $C$ affects base stations by sending reports on compromised neighbors. If there is no support report from the neighbor of node $D$, the base station will not consider node $D$ to be a compromised node. Slander attack only leads to node $C$ being considered a compromised node.

## 5 System Implementation

This paper uses TinyOS operating system to implement CND [1]. There are two key issues that might be applicable to other implementations of CND:

**First,** each node has an illegal behavior buffer to store the illegal behaviors of neighbor nodes. The format and size of the buffer are related to the specific implementation. For example, it depends on the required accuracy and the performance of wireless sensor networks. The report in the buffer must contain the following domains: alarm time, illegal behavior type and source.

**Second,** monitoring all neighbors makes the buffer require a higher memory overhead. Memory overhead is exponentially increased by the number of direct neighbors, so the overhead for high-density networks is very large. To solve this problem, CND nodes can select a subset of neighbor nodes to monitor, and the selection can be random or in accordance with other protocols. For example, in the random pairwise key distribution protocol [8], some keys are generated before deployment, and each node is assigned a random key. After deployment, it is possible that two neighbor nodes have compatible keys and can communicate with each other. The number of neighbors monitored by each node is controlled by the density of the network, so that each neighbor can communicate with it. This paper will show in the following chapters that when the number of monitored neighbors reaches a certain value, the performance of CND will reach its maximum. So, in dense networks, there is no need to monitor all neighbors.

## 6 System Performance

In order to analyze the performance of CND, a series of experiments were carried out using SenSec [17]. SenSec is an evaluation tool that enables people to imitate and analyze various attacks in wireless sensor networks. The validity of CND in different parameters is quantitatively analyzed.

In this paper, the standard performance indicators of the detection system are as follows:

1) Detection rate. This indicator is a percentage of the actual compromised behaviors detected by the system. However, even if the detection rate is 100%, the accuracy of the system can not be determined without considering the false positives.

2) False positive rate. Legitimate nodes can be erroneously reported as compromised nodes, and these reports are called false positives. The detection rate is not inversely proportional to the false positive rate. The system with high false positive rate is inaccurate, because most of the reported compromised behaviors are false.

3) Detection time. Before determining whether a node is compromised, the detection mechanism takes time to process the collected data. Detection time refers to the time that a compromised node keeps the state of being not detected.

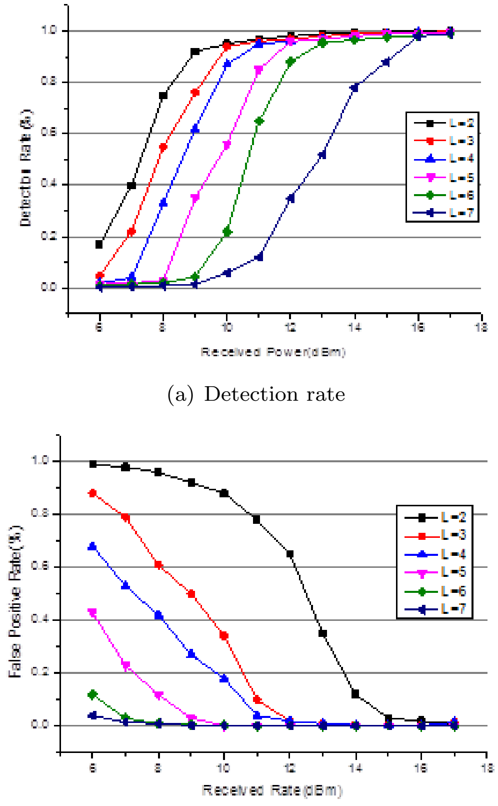
The performance of the distributed component and the overall performance of the CND will be discussed below.

## 6.1 Performance of Distributed Component

When modeling the compromised nodes in network, a gradient based model proposed by Chen *et al.* is used [2]. The model is based on the perspective of the spatial locality of the compromised node. For example, if a node is close to a compromised node, it will be more likely to become a compromised node. Therefore, the probability of a node being conquered forms a gradient, the closer to the compromised node, the more likely it will be conquered.

The experimental network topology consists of 100 analog nodes randomly deployed in a $100m \times 100m$ area. The node has a wireless transmission device with a transmission power of 5 dBm, and runs a universal sensor application, reads the sensor readings every second, and routes them to a base station at any edge of the network. The tree routing protocol and CSMA protocol are used in the experiment. First, the system is set up. At a random time after the setup stage, a random node in the analog wireless sensor network is conquered every 10 simulated minutes, and a series of attacks on the network are launched. Attacks initiated by compromised node are provided by SenSec, such as replay attacks, witch attacks, wormhole attacks, pulse delays, selective forwarding, and so on. In the simulation, each node runs a real TinyOS application with a sensor readings every 0.1 s. Each experiment includes 50 runs, and each run lasts for 1 simulated hours.

Figure 5(a) and 5(b) show the experimental results, which can be used to evaluate the performance of different received power detection algorithms. For the original packets, the constant 5 dBm transmission power is used in this paper, while the received power is simulated according to physical topology, and then the level of transmission power increases gradually.



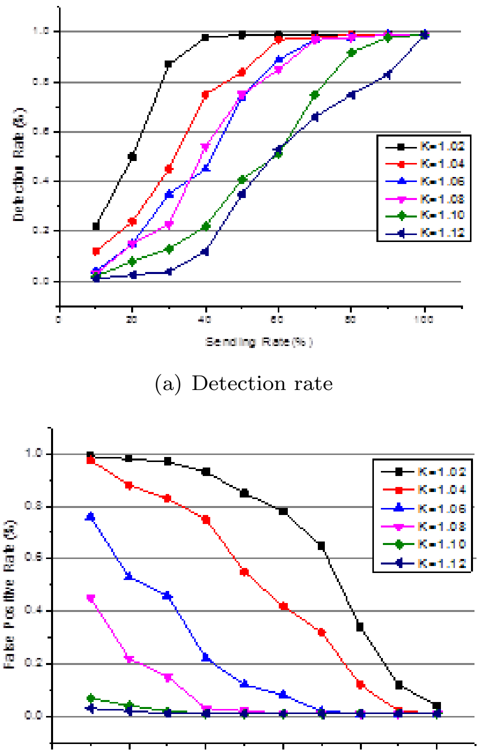
(a) Detection rate

(b) False positive rate

Figure 5: Performance of detection algorithm based on received power change

As you can see from Figure 5, smaller packet buffer requires smaller received power changes when compromised nodes are detected. A buffer with a length of 2 can reach a positive rate of 95% with the minimum change in received power. However, the false positive rate will be higher. For example, a buffer with a length of 2 has a false positive rate of 95%. These results can be explained by setting up a baseline with past data. A smaller buffer means that the algorithm is more sensitive to small changes, whether the change is caused by a compromise or a temporary change in the environment.

For the algorithm using the transmission rate, the buffer length $L$ is 6 and the intercepted length $L_2$ is 2. Figure 6(a) and 6(b) show the performance of the algorithm when the received power is changed according to a certain percentage and threshold $K$. The result of this experiment is consistent with the results of previous experiments: smaller threshold and buffer length will make the algorithm more sensitive and provide higher detection rate at the expense of higher false positive rate. For example, if the value of $K$ is 1.02, an increase of 30% of the transmission rate will make the detection rate up to 90%, but the false positive rate is 97%. The detection time is only dependent on $K$ and remains unchanged when the transmission rate changes.



(a) Detection rate

(b) False positive rate

Figure 6: Performance of detection algorithm based on sending rate change

The purpose of these experiments is to analyze the performance of the detection algorithms deployed on each node, and the actual parameters should be adjusted according to the needs of the application security. However, these results show that the algorithm can detect the compromised node with a detection rate of over 98% and a false positive rate below 5%.

In different environments, the performance of the algorithm will be quite different, because a single node with limited resources can not achieve high accuracy in any case. The function of the detection algorithm is to notify the base station of possible compromise behaviors. The base station determines whether a report is correct by collecting reports from multiple nodes, which will partly compensate for the limitations of the sensor nodes.

## 6.2 Overall Performance of CND

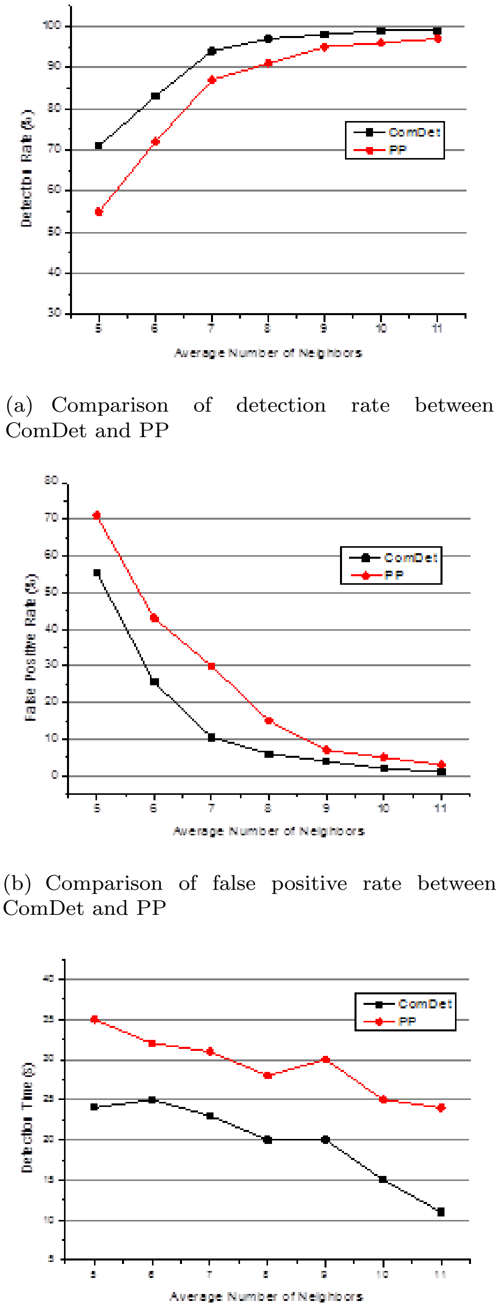
The ComDet system adopts a hybrid architecture consisting of two parts: distributed components and centralized components. Distributed components running on sensor nodes can detect compromised nodes and report them to the base station. Centralized components are used to perform complex analysis to determine whether the report of compromise is correct. The combination of these two aspects can effectively improve the efficiency and accuracy of detection. Document [6] proposes an intrusion detection scheme based on projection pursuit algorithm for wireless sensor networks, in which the proposed algorithm is called PP algorithm. Figure 7 shows a comparison between the ComDet algorithm and the PP algorithm in terms of detection rate, false positive rate and detection time when the packet loss rate is 15%. Through comparison, we can see that ComDet algorithm has higher detection efficiency and accuracy than BP algorithm.

In this paper, several experiments have been carried out to make a quantitative analysis of the performance of CND. The experimental setup is the same as that before, and 100 nodes that run TinyOS application are deployed randomly.

Figure 8 shows the performance evaluation results of CND with various packet loss rates and multiple monitoring neighbors. Figure 8(a) and 8(b) show that the algorithm can compensate for high packet loss rates when multiple nodes are monitored each other. When the packet loss rate is 30%, if each node monitors an average of 9 neighbors, it can reach a positive rate of 99% and a false positive rate of 2%. The high packet loss rate has a higher impact on the detection rate than the false positives, because the loss of the report makes the real compromise appear to be an instantaneous error. However, in most cases, the compromise can be detected before the damage is caused. As shown in figure 8(c), the higher the packet loss rate is, the longer the detection time is.

At the same time, the number of packets sent is also measured to be related to the operation of CND, as shown in Figure 8(d). As expected, the high packet loss rate will cause more packets to be sent because of retransmission. However, increasing the number of monitoring neighbors does not increase the number of packets sent. In most cases, the number of sending packets does not change significantly, and in some cases, as each neighbor is added, the number of actual packets sent by each neighbor is reduced by 5%. Careful observation shows that when the number of monitoring neighbors increases, more neighbors will send reports on the same attack, which will lead to increased communication overhead. However, more reports will make the base station detect compromised nodes faster when some reports are missing. The malicious behavior of the detected compromised node will no longer generate reports, which will reduce the communication overhead. The actual result is that, when the packet loss rate is greater than 15%, the communication overhead will decline or remain unchanged with each additional neighbor.

In addition, the energy consumption of CND is also measured. The energy consumption of wireless transmission accounts for the vast majority of the total energy consumption, which is consistent with the previous conclusion, that is, the total energy consumption is proportional to the communication overhead [11].



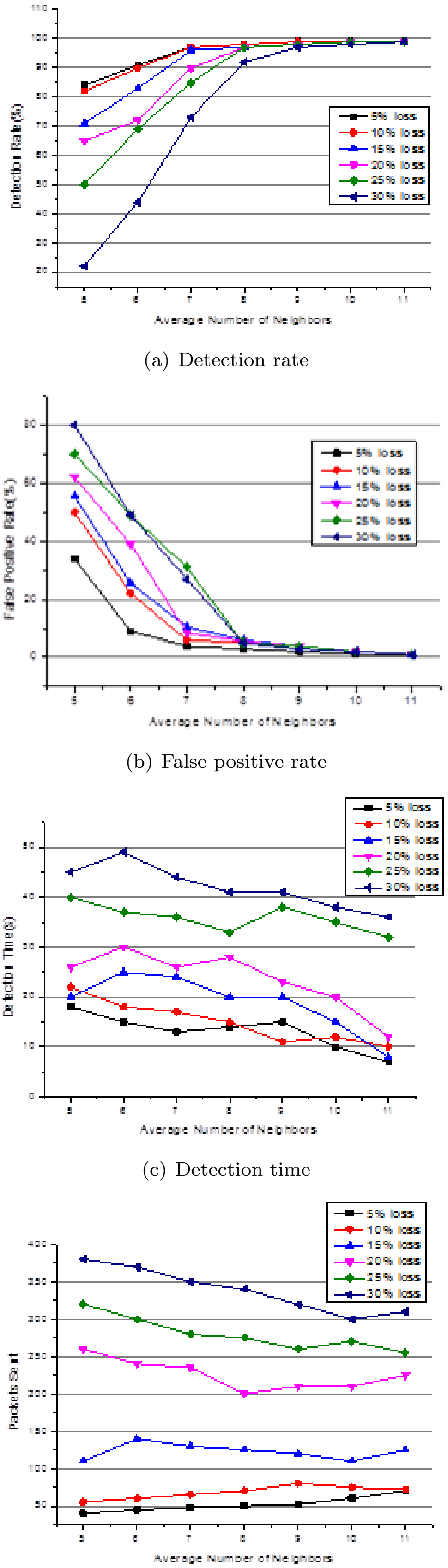
(a) Comparison of detection rate between ComDet and PP

(b) Comparison of false positive rate between ComDet and PP

(c) Comparison of detection time between ComDet and PP

Figure 7: Performance comparison between ComDet and PP

These results show that CND can provide accurate detection of compromised nodes and can be extended to large networks. Although similar systems can achieve almost the same performance without losing packets, but when the packet loss rate reaches 30%, the highest detection rate is reduced to 14%, and the false positive rate is as high as 99%. However, in the case of the packet loss rate of 30%, CND can reach a detection rate of 99% and a false positive rate of 2%. In addition, for a larger net-



(a) Detection rate

(b) False positive rate

(c) Detection time

(d) Communication overhead

Figure 8: Performance of CND under various packet loss rates and neighbors

work of density and size, its overhead does not increase significantly.

# 7 Conclusion

In wireless sensor networks, compromised nodes can destroy data integrity by sending false reports, injecting erroneous data and interfering data transmission. Because encryption is not enough to prevent these attacks, CND is proposed to detect compromised nodes in wireless sensor networks. A series of experiments show that CND can reach a 99% detection rate and a false positive rate of less than 2% when the packet loss rate is 30%. CND can run in most wireless sensor networks, because it uses common application features and adjusts detection behavior when there is no periodic transmissions or lack of communications between nodes. It has smaller memory and lower computing and communication overhead, which enable it to be extended to large networks with thousands of nodes.

The goal of future work is to create a response system and a challenge system. CND provides a mean to identify compromised nodes in the network, but it does not provide a way to deal with attacks. The basic method is to isolate compromised nodes, but it is not suitable for all occasions. Besides, once a node is determined to be a compromised node, it should be allowed to prove that it is not a compromised node. This can further improve the accuracy of the detection.

# Acknowledgments

This work was supported by the Experimental Technology Research Project Fund of Qufu Normal University (No. SJ201719), MOE (Ministry of Education in China) Project of Humanities and Social Sciences under Grant [No. 17YJCZH026].

# References

[1] M. Amjad, M. Sharif, M. K. Afzal, S. W. Kim, "TinyOS - New trends, comparative views, and supported sensing applications: A review," *IEEE Sensors Journal*, vol. 16, no. 9, pp. 2865–2889, 2016.

[2] X. Chen, K. Makki, Y. Kang, and N. Pissinou, "Node compromise modeling and its applications in sensor networks," in *IEEE Symposium on Computers and Communications*, pp. 575–582, 2007.

[3] Y. Chen, and Q. Ye, "Summaryon security authentication scheme for wireless sensor networks," *Computer & Digital Engineering*, vol. 42, no. 2, pp. 261–266, 2014.

[4] A. R. Dhakne, P. N. Chatur, "Detailed survey on attacks in wireless sensor network," in *Proceedings of the International Conference on Data Engineering and Communication Technology*, Springer Singapore, 2017.

[5] Y. Gao, P. Zeng, K. K. R. Choo, and S. Fu, "An improved online/offline identity-based signature scheme for WSNs," *International Journal of Network Security*, vol. 18, no. 6, pp. 1143–1151, 2016.

[6] X. Ge, L.Wang, and X. Guo, "Intrusion detection model for WSNs based on projection pursuit," *Transducer & Microsystem Technologies*, vol. 34, no. 9, pp. 24–26, 2015.

[7] S. Javanmardi, A. Barati, S. J. Dastgheib, and I. Attarzadeh, "A novel approach for faulty node detection with the aid of fuzzy theory and majority voting in wireless sensor networks," *International Journal of Advanced Smart Sensor Network Systems*, vol. 2, no. 4, pp. 1–10, 2012.

[8] Z. Ji, X. Du, L. Xu, and J. Lin, "Security key predistribution scheme for wireless sensor networks," *Journal of Computer Applications*, vol. 33, no. 7, pp. 1851–1853, 2013.

[9] M. Kumar, K. Dutta, "Impact of wormhole attack on data aggregation in hierarchical WSN," *International Journal of Electronics & Information Engineering*, vol. 1, no. 2, pp. 70–77, 2014.

[10] C. T. Li, M. S. Hwang and Y. P. Chu, "An efficient sensor-to-sensor authenticated path-key establishment scheme for secure communications in wireless sensor networks", *International Journal of Innovative Computing, Information and Control*, vol. 5, no. 8, pp. 2107-2124, Aug. 2009.

[11] M. J. Miller and N. H. Vaidya, "A MAC protocol to reduce sensor network energy consumption using a wakeup radio," *IEEE Transactions on Mobile Computing*, vol. 4, no. 3, pp. 228–242, 2005.

[12] A. Pananjady, V. K. Bagaria, R. Vaze, "Optimally approximating the coverage lifetime of wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 25, no. 1, pp. 98–111, 2017.

[13] Y. B. Saied, A. Olivereau, "A lightweight threat detection system for industrial wireless sensor networks," *International Journal of Network Security*, vol. 18, no. 5, pp. 842–854, 2016.

[14] R. Singh, M. S. Manu, "An energy efficient grid based static node deployment strategy for wireless sensor networks," *International Journal of Electronics & Information Engineering*, vol. 7, no. 1, pp. 32–40, 2017.

[15] L. Tan, and C. Pei, "Research of wireless sensors network security algorithms," *Computer Science*, vol. 42, no. S1, pp. 438–443, 2015.

[16] Y. Zhang, J. Yang, W. Li, and L. Jin, "An authentication scheme for locating compromised sensor nodes in WSNs," *Journal of Network & Computer Applications*, vol. 33, no. 1, pp. 50–62, 2010.

[17] J. Zhu, P. Wu, X. Wang, and J. Zhang, "SenSec: Mobile security through passive sensing," in *IEEE International Conference on Computing, Networking and Communications*, pp. 1128–1133, 2013.

# Biography

**Xiaolong Xu,** born in 1977, he is now a experimenter of Experiment Teaching Center, Qufu Normal University. He obtained his bachelor's degree of computer science from Qufu Normal University in 1999, and master's degree of computer application technology from Qufu Normal University in 2006. He has published more than 20 papers. His major research interests include network security and wireless sensor network.

**Zhonghe Gao,** born in 1961. He is a professor and postgraduate tutor. His major research interests include computer network and mobile communication technology.

**Lijuan Han,** born in 1976. She got the master degree and now is a associate professor. Her major research interests include artificial intelligence and digital signal processing.

# Optimal Control Analysis of Computer Virus Transmission

Okosun Kazeem[1], Ogunlade Samson[2], Bonyah Ebenezer[3]
(Corresponding author: Okosun Kazeem)

Department of Mathematics, Vaal University of Technology, South Africa[1]
Andries Potgieter Blvd, Vanderbijlpark 1900, South Africa
(Email: stogunlade@gmail.com)
Mathematical Sciences Department, Federal University of Technology[2]
PMB, 704, FUTA Rd, Akure, Nigeria
Department of Information Technology Education, University of Education Winneba[3]
Kumasi Campus, P. O Box 1277, Kumasi, Ghana

(Received June 13, 2017; revised and accepted Oct. 21, 2017)

## Abstract

Computer virus has become a global problem and affecting many industries both developed and the developing countries.In this study, a deterministic computer virus model is formulated incorporating removal devices. The basic properties of the model is studied and the reproduction number is calculated. The steady states are studied and found to be stable. We analyze different properties with parameter change by carrying out the sensitivity analysis of the model. Time optimal control is included and Pontryagin's Maximum Principle is used to characterize the all necessary condition for controlling the spread of computer virus. The most effective strategy for controlling computer virus is the combination of all the three controls. Graphical illustrations are presented to show the effects.

*Keywords: Computer Virus; Optimal Control; Pontryagin's Maximum Principle; Removable Devices; Sensitivity Analysis*

## 1 Introduction

The study of Computer Virus and its control has been a challenge over the years. The Computer virus is defined as a piece of software that contains malicious code, which could propagate, be installed and cause damage to computer data without the authorised permission or knowledge of the user [15, 16, 33].

In the industries, we observe that though the impact of the virus and the damages it has caused have been reduced moderately due to the increasing awareness by the public and high technological inventions. Despite this, the problem still persists [3, 13, 19, 24, 26, 35]. In general, the importance of mathematical models is that it helps to analyse the steady states, minimize the disease spread, control the disease outbreak and characterize the model propagation [17, 37].

Also, our modern financial institutions, culture, infrastructures and information and communication technology now depends mostly on computer networks and internet connectivities [27]. As the rate of dependence on computer networks increase, we observe that cyber attacks are also on the increase [31]. In order to avert this, or reduce considerably the cyber attack rate, there is a need to formulate deterministic models which would capture if not most but at least very important parameters such that there would be a control measure to which the computer virus could be spread [26].

Mathematical modeling in recent times has played a vital role of providing important insights into many processes including population behavior and their controls. Again, for some year now, it has also become an indispensable vehicle with which dynamical behaviors of many systems are understood such as computers virus so that the appropriate decision concerning the right interventions are undertaken.

For instance, Jin investigated the significance of applying epidemiological models in computer virus protection and prevention, and discussed their implication in developing anti-virus technologies and policies [15]. Also Lopez and Cipolatti introduced a simplified theoretical model to describe a virtual virus propagation process in a set of interacting computers. They also considered the propagation mechanisms which are those related to the reception of messages through internet as well as the ones concerning the simple exchange of files using recording devices as compact disks or the commonly used floppy disks [10]. Zhu *et al.* considered the effect of removable devices on the transmission of computer virus [37]. Chen *et al.*, presented a mathematical model, referred to as the Analytical Active Worm Propagation (AAWP) model, which characterized the propagation of worms that employ random scanning. They compared the model with the Epidemiological model and Weaver's simulator. Their results showed that the model characterized the spread of worms effectively [3]. Furthermore, Omair and Samir also analyzed the efficiency of antivirus software and crashing of the nodes due to virus attack [26].

Our main goal is to construct a mathematical model on computer virus transmission and incorporating removal devices with some control strategies. The paper is arranged as follows, in Section 2, we present the model formulation and carry out the stability analysis of the model. In Section 3, we perform sensitivity analysis of parameters. In Section 4, time dependent control is incorporated in the model and analytical solution of the controls. The numerical solutions are presented in Section 5. Finally conclusion is drawn in Section 6.

## 2 The Model

The model sub-divides the total Computer population, denoted by $N$, into sub-populations of Susceptible computers ($S$), Exposed computers ($E$), Infected computers ($I$), Recovered computers ($R$). We assume that computers can be infected through electronic mails and internet access. But computers are not contributing or infected internet network. Let Susceptible removable device be $D_S$ and Infected removable device be $D_I$. So that $N = S + E + I + R$ and $D_N = D_S + D_I$. The diagram and differential equations are given in Figure 1 and the following:

$$\begin{aligned}
\frac{d}{dt}S &= \Lambda - \frac{\beta_2 D_I S}{D_N} - \beta_1 SI - dS + \eta R \\
\frac{d}{dt}E &= \frac{\beta_2 D_I S}{D_N} + \beta_1 SI - (d+\mu)E \\
\frac{d}{dt}I &= \mu E - (d+\gamma+\alpha)I \\
\frac{d}{dt}R &= \gamma I - (d+\eta)R \qquad (1)\\
\frac{d}{dt}D_S &= \Lambda_d - \frac{\beta_2 D_S I}{N} + \sigma D_I - d_2 D_S, \\
\frac{d}{dt}D_I &= \frac{\beta_2 D_S I}{N} - (d_2+\sigma)D_I.
\end{aligned}$$

The $\beta_1, \beta_2$ are the transmission probabilities of computer virus, while the terms $\sigma$ is the ingestion rate and $\gamma$ is the recovery rate while $\alpha$ virus induced computer death. Computer recruitment rate is $\Lambda$, $\mu$ is progression from exposed to infected class.

**Lemma 1.** *The closed set*

$$D = (S, E, I, R, D_S, D_I) \in \mathbb{R}^6_+ : N \leq \frac{\Lambda}{d}, D_N \leq \frac{\Lambda_d}{d_2}$$

*is positively invariant and attracting for the model [8].*

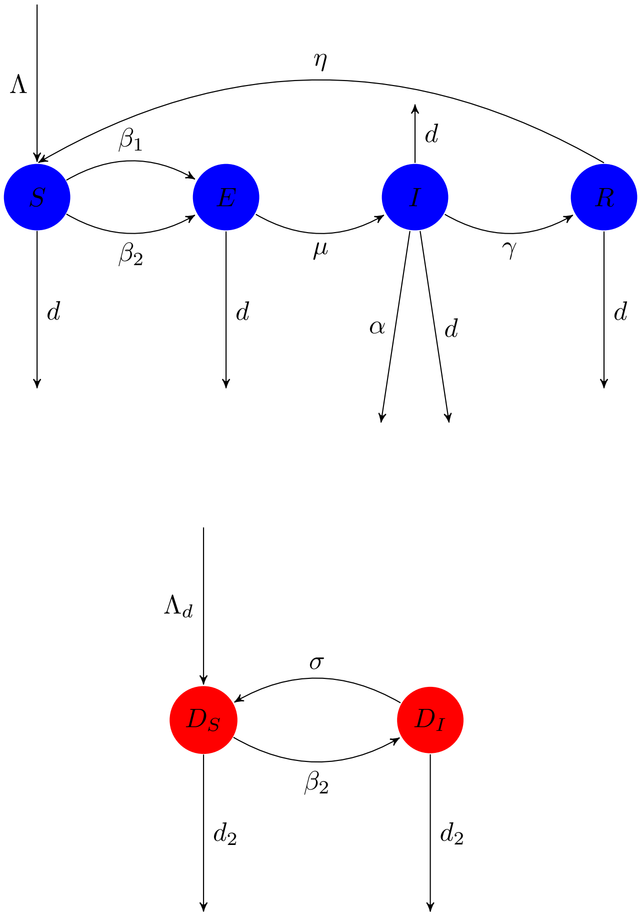


Figure 1: The computer virus transmission model diagram

Table 1: Computer virus transmission model notations

| Parameters | Description |
|---|---|
| $S$ | The number of Susceptible Computers. |
| $E$ | The number of Exposed Computer. |
| $I$ | The number of Infected Computers. |
| $R$ | The number of Recovered Computers. |
| $D_S$ | The number of Susceptible removable device. |
| $D_I$ | The number of Infected Removable device. |
| $\Lambda$ | Computer recruitment |
| $\Lambda_d$ | Removable device recruitment |
| $\beta_1$ | The infectivity contact rate at with network attacks occur. |
| $\beta_2$ | The infectivity contact rate at with virus attacks removable device. |
| $d$ | Natural death of Computer System. |
| $d_2$ | Death rate of removable device. |
| $\alpha$ | Rate of Virus induced death. |
| $\gamma$ | Recovery rate. |
| $\eta$ | Waning rate of Computer. |
| $\mu$ | The Computer exposed rate. |
| $\sigma$ | Ingestion rate. |

*Proof.* Adding the first four equations and then the last two equations of the model, we have:

$$\begin{aligned}\frac{dN}{dt} &= \Lambda - dN - \alpha I\\ \frac{dD_N}{dt} &= \Lambda_d - d_2 D_N\end{aligned}$$

Since $\frac{dN}{dt} \leq \Lambda - dN$ and $\frac{dD_N}{dt} \leq \Lambda_d - d_2 D_N$, it follows that $\frac{dN}{dt} \leq 0$ and $\frac{dD_N}{dt} \leq 0$ if $N \geq \frac{\Lambda}{d}$ and $D_S \geq \frac{\Lambda_d}{d_2}$ respectively.

Hence by Comparison theorem in [12],

$$\begin{aligned}\frac{dN}{dt} + dN &\leq \Lambda\\ \frac{d}{dt}(Ne^{dt}) &\leq \Lambda e^{dt}\end{aligned}$$

$\Longleftrightarrow$

$$\begin{aligned}N(t) &\leq N(0)e^{-dt} + \frac{\Lambda}{d}(1 - e^{-dt})\\ D_N(t) &\leq D(0)e^{-d_2 t} + \frac{\Lambda_d}{d_2}(1 - e^{-d_2 t})\end{aligned}$$

Therefore, $N(t) \leq \frac{\Lambda}{d}$ if $N(0) \leq \frac{\Lambda}{d}$ and $D_S(t) \leq \frac{\Lambda_d}{d_2}$ if $D_S(0) \leq \frac{\Lambda_d}{d_2}$. Thus, the region D is positively invariant for the model.

Furthermore, if $N(0) > \frac{\Lambda}{d}$ and $D_S(0) > \frac{\Lambda_d}{d_2}$, then, one of the solution enters $D$ in finite time i.e. $N(t) \longrightarrow \frac{\Lambda}{d}$ and $D_S(t) \longrightarrow \frac{\Lambda_d}{d_2}$ as $t \longrightarrow \infty$. We conclude that the region $D$ attracts all solution in $\mathbb{R}^6_+$. □

# 3 Analysis of Steady States

## 3.1 Basic Reproduction Number $R_0$ of Model

The Virus-Free Equilibrium (VFE) of Equation (1) is computed as

$$\begin{aligned}\varepsilon_0 &= (S^*, E^*, I^*, R^*, D_S^*, D_I^*)\\ &= \left(\frac{\Lambda}{d}, 0, 0, 0, \frac{\Lambda_d}{d_2}, 0\right) \qquad (2)\end{aligned}$$

By the Van den Driessche and Watmough [5], the basic reproduction number $R_0$ of the computer-virus model is computed by using the Next Generation Matrix Method. It is given by:

$$R_0 = r(FV^{-1})$$

where $r(.)$ is the spectral radius. Therefore,

$$F = \begin{pmatrix} 0 & \frac{\Lambda d_2 \beta_1}{d\Lambda_d} & \frac{\Lambda\beta_3}{d}\\ 0 & 0 & 0\\ 0 & \frac{d\beta_3\Lambda_d}{\Lambda d_2} & 0\end{pmatrix}$$

$$V^{-1} = \begin{pmatrix} \frac{1}{d+\mu} & 0 & 0\\ \frac{\mu\sigma+\mu d_2}{(d+\alpha+\gamma)(d+\mu)(\sigma+d_2)} & \frac{1}{d+\alpha+\gamma} & 0\\ 0 & 0 & \frac{1}{\sigma+d_2}\end{pmatrix}$$

$$FV^{-1} = \begin{pmatrix} \frac{\Lambda d_2(\mu\sigma+\mu d_2)\beta_1}{d(d+\alpha+\gamma)(d+\mu)(\sigma+d_2)\Lambda_d} & \frac{\Lambda d_2\beta_1}{d(d+\alpha+\gamma)\Lambda_d} & \frac{\Lambda\beta_3}{d(\sigma+d_2)}\\ 0 & 0 & 0\\ \frac{d(\mu\sigma+\mu d_2)\beta_3\Lambda_d}{(d+\alpha+\gamma)\Lambda(d+\mu)d_2(\sigma+d_2)} & \frac{d\beta_3\Lambda_d}{(d+\alpha+\gamma)\Lambda d_2} & 0\end{pmatrix} \qquad (3)$$

Since the second row in Equation (3) has zero entries, then we reduce the above matrix to:

$$FV^{-1} = \begin{pmatrix} \frac{\Lambda d_2(\mu\sigma+\mu d_2)\beta_1}{d(d+\alpha+\gamma)(d+\mu)(\sigma+d_2)\Lambda_d} & \frac{\Lambda d_2\beta_1}{d(d+\alpha+\gamma)\Lambda_d}\\ \frac{d(\mu\sigma+\mu d_2)\beta_3\Lambda_d}{(d+\alpha+\gamma)\Lambda(d+\mu)d_2(\sigma+d_2)} & \frac{d\beta_3\Lambda_d}{(d+\alpha+\gamma)\Lambda d_2}\end{pmatrix}$$

Now, we have that by the largest eigenvalue of the above matrix, the basic reproduction number for the model is given by:

$$R_0 = \frac{\Lambda^2\mu d_2^2\beta_1 + d^3\beta_2\Lambda_d^2 + d^2\mu\beta_2\Lambda_d^2}{d(d+\alpha+\gamma)\Lambda(d+\mu)d_2\Lambda_d}$$

Hence, we will establish the local and global stability of the VFE.

## 3.2 Local Stability of Virus-Free Equilibruim

**Theorem 1.** *The virus-free equilibrium $\varepsilon_0$ exists for all $R_0$ and is locally asymptotically stable if $R_0 < 1$ and unstable if $R_0 > 1$.*

*Proof.* We compute the Jacobian matrix and evaluate it at VFE. Therefore we have:

$$J_{\varepsilon_0} = \begin{pmatrix} -d & 0 & -\beta_1\frac{\Lambda}{d} & \eta & 0 & -\beta_2\frac{\Lambda d_2}{\Lambda_d d}\\ 0 & -(d+\mu) & \beta_1\frac{\Lambda}{d} & 0 & 0 & \beta_2\frac{\Lambda d_2}{\Lambda_d d}\\ 0 & \mu & -(d+\gamma+\alpha) & 0 & 0 & 0\\ 0 & 0 & \gamma & -(d+\eta) & 0 & 0\\ 0 & 0 & -\beta_2\frac{\Lambda_d d}{\Lambda d_2} & 0 & -d_2 & \sigma\\ 0 & 0 & \beta_2\frac{\Lambda_d d}{\Lambda d_2} & 0 & 0 & -(d_2+\sigma)\end{pmatrix} \qquad (4)$$

We now compute the eigenvalues of the Jacobian matrix. From Matrix (4), we observed that the first and the fifth columns contain only the diagonal terms $-d$ and $-d_2$, which are the first two eigenvalues. To obtain the other eigenvalues, the Jacobian matrix (4) is reduced to a submatrix (5) as follows:

$$J'_{\varepsilon_0} = \begin{pmatrix} -(d+\mu) & \beta_1\frac{\Lambda}{d} & 0 & \beta_2\frac{\Lambda d_2}{\Lambda_d d}\\ \mu & -(d+\gamma+\alpha) & 0 & 0\\ 0 & \gamma & -(d+\eta) & 0\\ 0 & \beta_2\frac{\Lambda_d d}{\Lambda d_2} & 0 & -(d_2+\sigma)\end{pmatrix} \qquad (5)$$

Also in the new matrix (5), we see that, the third column contains a diagonal term $-(d+\eta)$ which is the third eigenvalue. Therefore we reduce Matrix (5) to:

$$J''_{\varepsilon_0} = \begin{pmatrix} -(d+\mu) & \beta_1\frac{\Lambda}{d} & \beta_2\frac{\Lambda d_2}{\Lambda_d d}\\ \mu & -(d+\gamma+\alpha) & 0\\ 0 & \beta_2\frac{\Lambda_d d}{\Lambda d_2} & -(d_2+\sigma)\end{pmatrix} \qquad (6)$$

Now let $P_i$, $i = 1, 2, 3$, be the eigenvalues of Matrix (6), we have

$$\begin{aligned} & Det(J''_{\varepsilon_0} - PI) \\ = & \begin{vmatrix} -(d+\mu)-P & \beta_1 \frac{\Lambda}{d} & \beta_2 \frac{\Lambda d_2}{\Lambda_d d} \\ \mu & -(d+\gamma+\alpha)-P & 0 \\ 0 & \beta_2 \frac{\Lambda_d d}{\Lambda d_2} & -(d_2+\sigma)-P \end{vmatrix} \\ = & 0. \end{aligned} \tag{7}$$

The eigenvalues of the characteristic Equation (7) are the zeros which satisfies the following equation below:

$$P^3 + P^2 F_2 + P^1 F_1 + F_0 = 0 \tag{8}$$

Where

$$\begin{aligned} F_2 &= 2d+\alpha+\gamma+\mu+\sigma+d_2 \\ F_1 &= \Lambda d_2(d+\alpha+\gamma)\Lambda_d(d+\mu)[P-R_0] \\ F_0 &= \Lambda_d d_2(d+\alpha+\gamma)(d+\mu)\Lambda[Q-R_0] \end{aligned} \tag{9}$$

with

$$\begin{aligned} P = & \ (d(2d+\alpha+\gamma+\mu)d_2 - \Lambda\mu\beta_1 + \Lambda^2\mu d_2^2\beta_1 \\ & +d((d+\alpha+\gamma)(d+\mu) + (2d+\alpha+\gamma+\mu)\sigma \\ & +d(d+\mu)\beta_3\Lambda_d^2))/(d(d+\alpha+\gamma)\Lambda(d+\mu)d_2\Lambda_d), \\ Q = & \ (\Lambda^2\mu d_2^2\beta_1 + (\sigma+d_2)(d(d+\alpha+\gamma)(d+\mu) - \Lambda\mu\beta_1) \\ & -d\mu\beta_2^2 + d^3\beta_2\Lambda_d^2 + d^2\mu\beta_2\Lambda_d^2)/(d(d+\alpha+\gamma) \\ & \Lambda(d+\mu)d_2\Lambda_d). \end{aligned}$$

Now, from the basic reproduction number deduced, we can make the following observations for $F_1$ and $F_0$, that is

$$F_1 = \begin{cases} > 0 & when \quad R_0 < P < 1 \quad or \quad 1 < R_0 < P \\ < 0 & when \quad P < R_0 < 1 \quad or \quad 1 < P < R_0, \end{cases}$$

$$F_0 = \begin{cases} > 0 & when \quad R_0 < Q < 1 \quad or \quad 1 < R_0 < Q \\ < 0 & when \quad Q < R_0 < 1 \quad or \quad 1 < Q < R_0, \end{cases}$$

Therefore, we see that when

$R_0 < 1$ provided that $R_0 < P$ and $R_0 < Q$, the virus-free equilibrium is locally and asymptotically stable, otherwise it is unstable.

The requirement of the real and negative eigenvalues ensuring stability is clearly satisfied by $P$.

Now, for the roots of the polynomial equation (8) by which the eigenvalues are obtained, we therefore make the following analysis based on the Routh-Hurwitz stability criteria [12],

Firstly, the coefficients $F_0, F_2 > 0$, that is, must be positive.

Secondly, for the eigenvalues to have real negative parts, i.e. $F_2F_1 > F_0$.

It is obvious that the coefficients $F_0 > 0$ and $F_2 > 0$. Also,

$$\begin{aligned} F_2F_1 - F_0 = & (d+\alpha+\gamma)\Lambda(d+\mu)d_2((2d+\alpha+\gamma+\mu \\ & +\sigma+d_2)(P-R_0) + (R_0-Q)) > 0, \end{aligned}$$

Therefore, by the Routh-Hurwitz criteria for stability, we conclude that the virus free equilibrium is locally asymptotically stable whenever $R_0 < 1$. □

## 3.3 Endemic Equillibrium

Endemic Equilibrium: In order to obtain the endemic equilibrium of the model *i. e.* the equilibrium where at least one of the infected components of the model is non-zero [25], we solve the system of equations at steady states and obtain:

Let $p = (d+\alpha+\gamma)$, $q = (d+\mu)$, $r = (d+\eta)$, $v = (d+\alpha)$ and $w = (\sigma + d_2)$.

$$\begin{aligned} S^* &= d\gamma p\Lambda^2\mu q d_2\Lambda_d - d^2q(dpr + (dp+v\eta)\mu)\beta_2\Lambda_d^2 \\ &\quad - \Lambda^2\mu(dpr + (dp+v\eta)\mu)d_2^2\beta_1/d^2\gamma p\Lambda\mu q d_2\Lambda_d \\ E^* &= \frac{r\left(\Lambda^2\mu d_2^2\beta_1 + d^3\beta_2\Lambda_d^2 + d^2\mu\beta_2\Lambda_d^2\right)}{d\gamma\Lambda\mu q d_2\Lambda_d} \\ I^* &= \frac{r\left(\Lambda^2\mu d_2^2\beta_1 + d^3\beta_2\Lambda_d^2 + d^2\mu\beta_2\Lambda_d^2\right)}{d\gamma p\Lambda q d_2\Lambda_d} \\ R^* &= \frac{\Lambda^2\mu d_2^2\beta_1 + d^3\beta_2\Lambda_d^2 + d^2\mu\beta_2\Lambda_d^2}{dp\Lambda q d_2\Lambda_d} \\ D_S^* &= \frac{w\Lambda_d(\alpha r\Lambda^2\mu d_2^2\beta_1 - d\gamma p\Lambda^2 q d_2\Lambda_d + d^2\alpha r q\beta_2\Lambda_d^2)}{r\Lambda^2\mu d_2^2\beta_1(\alpha w - d\beta_2) - d\gamma p\Lambda^2 q d_2 w\Lambda_d + d^2 rq(\alpha w - d\beta_2)\beta_2\Lambda_d^2} \\ D_I^* &= \frac{dr\beta_2\Lambda_d(\Lambda^2\mu d_2^2\beta_1 + d^2q\beta_2\Lambda_d^2)}{r\Lambda^2\mu d_2^2\beta_1(-\alpha w + d\beta_2) + d\gamma p\Lambda^2 q d_2 w\Lambda_d + d^2 rq(-\alpha w + d\beta_2)\beta_2\Lambda_d^2} \end{aligned}$$

**Theorem 2.** *The unique Endemic Equilibrium of the model* (1) *is globally asymptotically stable whenever* $R_0 > 1$.

*Proof.* If $R_0 > 1$, then there exist a unique Endemic equilibrium. We therefore consider the non-linear Lyapunov function $\mathcal{V}$ such that

$$\begin{aligned} \mathcal{V} =& S^*\left[\frac{S}{S^*} - \ln\frac{S}{S^*}\right] + E^*\left[\frac{E}{E^*} - \ln\frac{E}{E^*}\right] + I^*\left[\frac{I}{I^*} - \ln\frac{I}{I^*}\right] \\ & + R^*\left[\frac{R}{R^*} - \ln\frac{R}{R^*}\right] + D_S^*\left[\frac{D_S}{D_S^*} - \ln\frac{D_S}{D_S^*}\right] + D_I^*\left[\frac{D_I}{D_I^*} - \ln\frac{D_I}{D_I^*}\right], \\ \dot{\mathcal{V}} =& \left[1 - \frac{S^*}{S}\right]\dot{S} + \left[1 - \frac{E^*}{E}\right]\dot{E} + \left[1 - \frac{I^*}{I}\right]\dot{I} + \left[1 - \frac{R^*}{R}\right]\dot{R} \\ & + \left[1 - \frac{D_S^*}{D_S}\right]\dot{D_S} + \left[1 - \frac{D_I^*}{D_I}\right]\dot{D_I}, \\ \dot{\mathcal{V}} =& \left[1 - \frac{S^*}{S}\right]\left[\Lambda - \frac{\beta_2 D_I S}{D_N} - \beta_1 SI - dS + \eta R\right] \\ & + \left[1 - \frac{E^*}{E}\right]\left[\frac{\beta_2 D_I S}{D_N} + \beta_1 SI - h_1 E\right] \\ & + \left[1 - \frac{I^*}{I}\right][\mu E - h_2 I] + \left[1 - \frac{R^*}{R}\right][\gamma I - h_3 R] \\ & + \left[1 - \frac{D_S^*}{D_S}\right]\left[\Lambda_d - \frac{\beta_2 D_S I}{N} + \sigma D_I - d_2 D_S\right] \\ & + \left[1 - \frac{D_I^*}{D_I}\right]\left[\frac{\beta_2 D_S I}{N} - h_4 D_I\right], \end{aligned} \tag{10}$$

where

$$\begin{aligned} h_1 &= d + \mu, \\ h_2 &= d + \gamma + \alpha, \\ h_3 &= d + \eta, \\ h_4 &= d_2 + \sigma. \end{aligned}$$

By expanding (10), and rearranging the expressions, we have

$$\begin{aligned}\dot{\mathcal{V}} =& dS^* \left[1 - \frac{S}{S^*} + \frac{\beta_2}{d}\left[\frac{D_S I}{D_I^* N} - \frac{D_S I}{D_I N}\right] + \frac{\Lambda}{d}\left[\frac{1}{S^*} - \frac{1}{S}\right]\right] \\ &+ h_1 E^* \left[1 - \frac{E}{E^*}\left[1 - \frac{\mu}{h_1}\left[1 - \frac{I^*}{I}\right]\right] + \frac{\beta_2}{h_1}\left[\frac{D_I S}{D_N E^*} - \frac{D_I S}{D_N E}\right]\right] \\ &+ h_2 I^* \left[1 - \frac{I}{I^*}\left[1 - \frac{\gamma}{h_2}\left[1 - \frac{R^*}{R}\right]\right] + \frac{\beta_1}{h_2} I \left[\frac{S^*}{I^*} - \frac{E^* S}{E I^*}\right]\right] \\ &+ h_3 R^* \left[1 - \frac{R}{R^*}\left[1 - \frac{\eta}{h_3}\left[1 - \frac{S^*}{S}\right]\right]\right] \\ &+ d_2 D_S^* \left[1 - \frac{D_S}{D_S^*} + \frac{\beta_2}{d_2}\left[\frac{I}{N}\left[1 - \frac{D_S}{D_S^*}\right]\right] + \frac{\Lambda_d}{d_2}\left[\frac{1}{D_S^*} - \frac{1}{D_S}\right]\right] \\ &+ h_4 D_I^* \left[1 - \frac{D_I}{D_I^*}\left[1 - \frac{\sigma}{h_4}\left[1 - \frac{D_S^*}{D_S}\right]\right] + \frac{\beta_2}{h_4}\left[\frac{I}{N}\left[\frac{D_S}{D_I^*} - \frac{D_S}{D_I}\right]\right]\right]\end{aligned}$$

We now have that, since the arithmetic mean exceeds the geametric mean value [9, 30], then

$$\begin{aligned}\left[1 - \frac{S}{S^*} + \frac{\beta_2}{d}\left[\frac{D_S I}{D_I^* N} - \frac{D_S I}{D_I N}\right] + \frac{\Lambda}{d}\left[\frac{1}{S^*} - \frac{1}{S}\right]\right] &\le 0 \\ \left[1 - \frac{E}{E^*}\left[1 - \frac{\mu}{h_1}\left[1 - \frac{I^*}{I}\right]\right] + \frac{\beta_2}{h_1}\left[\frac{D_I S}{D_N E^*} - \frac{D_I S}{D_N E}\right]\right] &\le 0, \\ \left[1 - \frac{I}{I^*}\left[1 - \frac{\gamma}{h_2}\left[1 - \frac{R^*}{R}\right]\right] + \frac{\beta_1}{h_2} I \left[\frac{S^*}{I^*} - \frac{E^* S}{E I^*}\right]\right] &\le 0, \\ \left[1 - \frac{R}{R^*}\left[1 - \frac{\eta}{h_3}\left[1 - \frac{S^*}{S}\right]\right]\right] &\le 0, \\ \left[1 - \frac{D_S}{D_S^*} + \frac{\beta_2}{d_2}\left[\frac{I}{N}\left[1 - \frac{D_S}{D_S^*}\right]\right] + \frac{\Lambda_d}{d_2}\left[\frac{1}{D_S^*} - \frac{1}{D_S}\right]\right] &\le 0, \\ \left[1 - \frac{D_I}{D_I^*}\left[1 - \frac{\sigma}{h_4}\left[1 - \frac{D_S^*}{D_S}\right]\right] + \frac{\beta_2}{h_4}\left[\frac{I}{N}\left[\frac{D_S}{D_I^*} - \frac{D_S}{D_I}\right]\right]\right] &\le 0.\end{aligned}$$

Since the parameters of the model (1) are greater then or equal to zero, therefore we have that $\dot{\mathcal{V}} \le 0$ for $R_0 > 1$. Hence it follows from LaSalle's Invariance Principle [17] that every solution of the equation in the model (1) approaches the Endemic Equilibrium as $t \longrightarrow \infty$ whenever $R_0 > 1$. □

# 4 Sensitivity Analysis

In order to investigate the above model robustness, due to uncertainties associated with the estimation of certain parameter values, it is important and useful to carry out a sensitivity analysis to investigate how sensitive the basic reproduction number is with respect to these parameters. It will also give us insight to know the parameters that have high impact or cause most reduction on the virus transmission, that is, in $R_0$ and therefore determine the control measure that is most effective in the control of the Computer virus transmission [25, 32].

To carry out this analysis, we compute the normalized forward sensitivity index of the reproduction number with respect to these parameters. This is also referred to as the ratio of the relative change in the variable change in the parameter [4, 25].

**Definition 1.** *The normalized forward sensitivity index of a variable h, that depends differentially on a parameter m, is defined as:*

$$\mathit{\Pi}_m := \frac{\partial h}{\partial m} \times \frac{m}{h}.$$

## 4.1 Sensitivity Indices of $R_0$

We derive the sensitivity of $R_0$ corresponding to the following parameters:

$$\begin{aligned}\mathit{\Pi}_\alpha &:= -\frac{\alpha}{d + \alpha + \gamma}, \\ \mathit{\Pi}_\gamma &:= -\frac{\gamma}{d + \alpha + \gamma}, \\ \mathit{\Pi}_{\beta_1} &:= \frac{\Lambda^2 \mu d_2^2 \beta_1}{\Lambda^2 \mu d_2^2 \beta_1 + d^2 (d + \mu) \beta_2 \Lambda_d^2}, \\ \mathit{\Pi}_{\beta_2} &:= \frac{d^2 (d + \mu) \beta_2 \Lambda_d^2}{\Lambda^2 \mu d_2^2 \beta_1 + d^2 (d + \mu) \beta_2 \Lambda_d^2}, \\ \mathit{\Pi}_\mu &:= \frac{d \Lambda^2 \mu d_2^2 \beta_1}{(d + \mu)(\Lambda^2 \mu d_2^2 \beta_1 + d^2 (d + \mu) \beta_2 \Lambda_d^2)}, \\ \mathit{\Pi}_\eta &:= \mathit{\Pi}_\sigma := 0,\end{aligned}$$

Using parameter values from Table 2, (it should be stated that these parameters are chosen for illustrative purpose only, and may not necessarily be realistic in terms of epidemiological interpretations), we calculate the sensitivity indices of $R_0$ based on the following parameters $\mu, \beta_1, \beta_2, \alpha, \eta, \sigma, \gamma$. The parameters are therefore, arranged from the most sensitive to least. The most sensitive parameter is proportion of the natural death rate $\beta_2 = 0.8901$. While the least of the sensitivity parameters are the $\eta$ and $\sigma = 0.0000$. An increase (or decrease) in $\beta_2$ by 10% increases (or decreases) the $R_0$ by 8.91%. Similarly increasing (or decreasing) the rate of recovery $\gamma$ by 10% decreases (or increases) the $R_0$ by 3.33%. From the sensitivity analysis, it is clear that control efforts should be targeted towards the rate at which the infectivity contact rate at which the virus attacks removable device ($\beta_2$).

# 5 Simulations for Computer Virus Model

In this section, we illustrated the effects of the changes of some basic parameters that may influence the transmission dynamics of the Computer virus model. In order to investigate the graphical trend of these changes of parameters in the model (1), we illustrate these by focusing on the transmission dynamics of the each sub-class of the model with respect to changes in some of its basic parameter values such as $\beta_1$ and $\mu$.

In the course of this investigations, we studied the dynamical flow of the trend of the following graphs below. And Hence, we draw some conclusions based on the result obtained under the graphs.

Table 2: Sensitivity analysis of $R_0$

| Parameters | Descriptions | Sensitivity |
|---|---|---|
| $\beta_2$ | The infectivity contact rate at which virus attacks removable device | 0.8901 |
| $\gamma$ | Rate of Recovery | $-0.3333$ |
| $\alpha$ | Rate of Virus induced Computer death | $-0.1667$ |
| $\beta_1$ | The infectivity contact rate at with network attacks occur | 0.1099 |
| $\mu$ | Exposed rate | 0.0824 |
| $\eta$ | Waning rate of Computer | 0.0000 |
| $\sigma$ | Rate of Ingestion | 0.0000 |

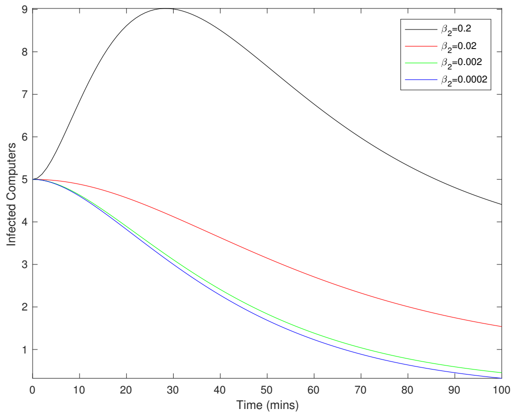


Figure 2: Model (1) Effects caused by changes in parameter $\beta_2$ in Infected class for $\Lambda = 0.8,$, $\beta_1 = 0.0002$, $\beta_2 = 0.2$, $\alpha = 0.01$, $\eta = 0.03$, $\gamma = 0.02$, $\mu = 0.01$, $\Lambda_d = 0.6$, $d = 0.03$, $d_2 = 0.005$, $\sigma = 0.002$

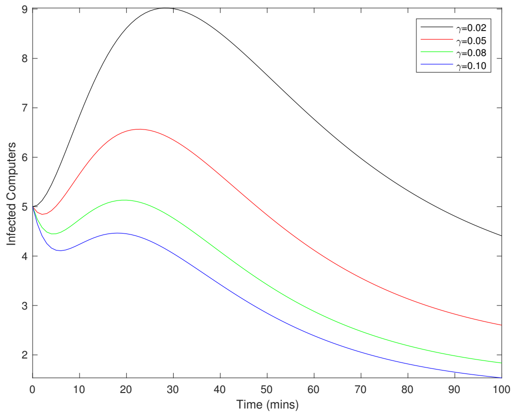


Figure 3: Model (1) Effects caused by changes in parameter $\gamma$ in Infected class for $\Lambda = 0.8,$, $\beta_1 = 0.0002$, $\beta_2 = 0.2$, $\alpha = 0.01$, $\eta = 0.03$, $\gamma = 0.02$, $\mu = 0.01$, $\Lambda_d = 0.6$, $d = 0.03$, $d_2 = 0.005$, $\sigma = 0.002$

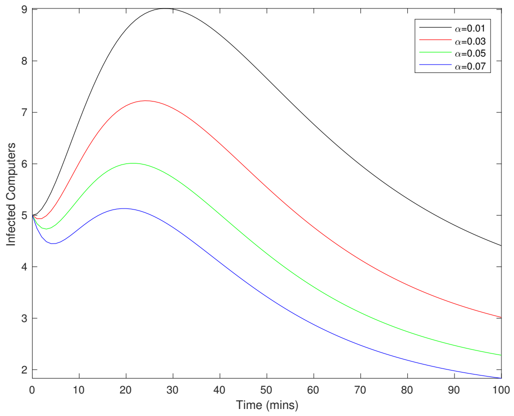


Figure 4: Model (1) Effects caused by changes in parameter $\alpha$ in Infected class for $\Lambda = 0.8,$, $\beta_1 = 0.0002$, $\beta_2 = 0.2$, $\alpha = 0.01$, $\eta = 0.03$, $\gamma = 0.02$, $\mu = 0.01$, $\Lambda_d = 0.6$, $d = 0.03$, $d_2 = 0.005$, $\sigma = 0.002$

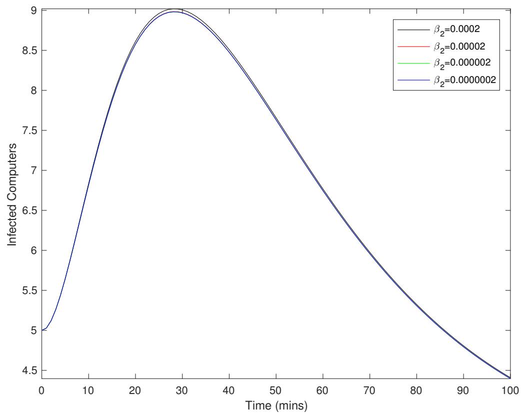


Figure 5: Model (1) Effects caused by changes in parameter $\beta_1$ in Infected class for $\Lambda = 0.8,$, $\beta_1 = 0.0002$, $\beta_2 = 0.2$, $\alpha = 0.01$, $\eta = 0.03$, $\gamma = 0.02$, $\mu = 0.01$, $\Lambda_d = 0.6$, $d = 0.03$, $d_2 = 0.005$, $\sigma = 0.002$

Studying the behavior of the mathematical model (1) for the transmission and spread of Computer virus through numerical simulations illustrated above, it was observed that the most sensitive parameter in the control of the virus spread are the infectivity contact rate, the virus attacks removable device $\beta_2$ rate. Although some other parameters such as $\gamma$, the rate of recovery, $\alpha$, virus induced computer death rate and $\beta_1$ the infectivity contact rate at with network attacks occur, are also quite sensitive to the model control. These graphs obtained also confirm the results in the sensitivity analysis.

From Figure 2, it is shown that, decrease in the rate at which the virus attacks removable device, decreases the rate at which the computer systems get infected with time and as a result decreases $R_0$. Figures 3 and 4 showed decrease in the infected class as a result of increase in the recovery rate and virus induced computer deaths, which is $\gamma$ and $\alpha$ respectively, which also led to a reduction in $R_0$. It is also shown in Figure 5 that the reduction of the infectivity contact rate at which network attacks occur $\beta_1$ slightly affects the transmission dynamics of the model by reducing the Infected class which also ascertain the sensitivity analysis.

By observing the numerical simulations, the control strategy should be targeted at decreasing $\beta_2$ by properly doing a thorough system scanning using an antivirus software before the system can be used. Adequate and frequent checks should be done always and ensure infected external devices such as flash drives should not be used on susceptible computers. By doing this, $R_0$ is greatly reduced and hence the virus drastically dies out.

# 6 Analysis of Optimal Control

In this section, we make use of Pontryagin's Maximum Principle so that we can obtain the essential conditions for the optimal control of the for computer virus model. Time dependent controls are incorporated the model in order to determine the best strategy for controlling the computer virus.For this reason, we consider the following objective functionals,

$$J(u_1, u_2, u_3) = \int_0^{t_F} \left( b_1 E + b_2 I + b_3 D_I + \frac{c_1}{2} u_1^2 + \frac{c_2}{2} u_2^2 + \frac{c_3}{2} u_3^2 \right) dt. \tag{11}$$

where $c_1, c_2$ and $c_3$ deal with the weighting constants for providing intensive public education on the use of removable device(prevention), effort on public campaign of how to maintain virus free on computers use (prevention) and treatment of infected computers with viruses (treatment) respectively. The cost corresponding to with prevention and treatment mechanisms are assumed to have a nonlinear character. Hence, we explore an optimal control $u_1^*, u_2^*$ and $u_3^*$ in a way that, $J(u_1^*,\ u_2^*, u_3^*) = \min J(u_1,\ u_2, u_3), \Gamma = \{(u_1,\ u_2, u_3)|0 \leq u_i \leq 1,\ i = 1, 2\}$.

$$\begin{aligned}
\frac{dS}{dt} &= \Lambda - (1-u_1)\frac{\beta_2 D_1 S}{D_N} - (1-u_2)\beta_1 SI - dS + \eta R \\
\frac{dE}{dt} &= (1-u_1)\frac{\beta_2 D_1 S}{D_N} + (1-u_2)\beta_1 SI - (d+\mu)E \\
\frac{dI}{dt} &= \mu E - (d + u_3\gamma + \alpha)I
\end{aligned}$$

$$\begin{aligned}
\frac{dR}{dt} &= u_3\gamma I - (d+\eta)R \\
\frac{dD_S}{dt} &= \Lambda_d - (1-u_1)\frac{\beta_2 D_S I}{N} + \sigma D_1 - d_2 D_S \\
\frac{dD_I}{dt} &= (1-u_1)\frac{\beta_2 D_S I}{N} - (d_2+\sigma)D_I. \qquad (12)
\end{aligned}$$

The necessary conditions that an optimal solution has to satisfy emanate from Pontryagin Maximum Principle [2, 25]. The principle actually converts (11)-(12) into a type of problem which principally aimed at minimizing pointwise a Hamiltonian $H$, with respect to $u_1$, $u_2$ and $u_3$.

$$\begin{aligned}
H &= b_1 E + b_2 I + b_3 D_I + c_1 u_1^2 + c_2 u_2^2 + c_2 u_3^2 \\
&\quad + N_S\{\Lambda - (1-u_1)\frac{\beta_2 D_1 S}{D_N} - (1-u_2)\beta_1 SI - dS + \eta R\} \\
&\quad + N_E\{(1-u_1)\frac{\beta_2 D_1 S}{D_N} + (1-u_2)\beta_1 SI - (d+\mu)E\} \\
&\quad + N_I\{\mu E - (d + u_3\gamma + \alpha)I\} \\
&\quad + N_R\{u_3\gamma I - (d+\eta)R\} \\
&\quad + N_{D_S}\{\Lambda_d - (1-u_1)\frac{\beta_2 D_S I}{N} + \sigma D_1 - d_2 D_S\} \\
&\quad + N_{D_I}\{(1-u_1)\frac{\beta_2 D_S I}{N} - (d_2+\sigma)D_I\} \qquad (13)
\end{aligned}$$

where $N_S, N_E, N_I, N_R, N_{D_S}$ and $N_{D_I}$ denote the adjoint variables or also referred to as co-state variables. The system of equations are arrived at by considering the right partial derivatives of the Hamiltonian (13) with respect to the associated state variable.

**Theorem 3.** *Given optimal controls $u_1^*, u_2^*, u_3^*$ and solutions $S, E, I, R, D_S, D_I$ of the associated state system (11)-(12) that minimize $J(u_1, u_2, u_3)$ over $\Gamma$. Thus, there exists adjoint variables*
*$N_S, N_E, N_I, N_R, N_{D_S}, N_{D_S}$ satisfying*

$$\frac{-d\lambda_i}{dt} = \frac{\partial H}{\partial i}$$

*where $i = S, E, I, R, D_S, D_I$ and with transversality conditions*

$$\begin{aligned}
N_S(t_f) &= N_E(t_f) = N_T(t_f) \\
&= N_R(t_f) = N_{D_S}(t_f) = N_{D_I} = 0 \qquad (14)
\end{aligned}$$

*and*

$$\begin{aligned}
u_1^* &= \min\left\{1, \max\left(0, \frac{\frac{\beta_2 D_1 S}{D_N}(N_E - N_S) + \frac{\beta_2 D_1 S}{D_N}(N_{D_I} - N_{D_S})}{2c_1}\right)\right\}, \\
u_2^* &= \min\left\{1, \max\left(0, \frac{\beta_1 SI(N_E - N_S)}{2c_2}\right)\right\}, \qquad (15) \\
u_3^* &= \min\left\{1, \max\left(0, \frac{\gamma(N_I - N_R)}{2c_3}\right)\right\} \qquad (16)
\end{aligned}$$

*Proof.* Corollary 4.1 of Fleming and Rishel [7] gives the appropriate condition of possible existence of an optimal control as a result of convexity of the integrand of $J$ with respect to $u_1, u_2$ and $u_3$, a *priori* boundedness of the state control solutions, and the *Lipschitz* characteristics of the state system with regard to the state variables. The Hamiltonian function computed at the optimal control provides the governing adjoint variables. Therefore, the adjoint equations can be rearrange as

$$\begin{aligned}
-\frac{dN_S}{dt} &= dN_S + (1-u_1)\frac{\beta_2 D_1}{D_N}(N_S - N_E) \\
&\quad +(1-u_1)\frac{\beta_2 D_S I}{N^2}(N_{D_I} - N_{D_S}) \\
&\quad +(1-u_2)\beta_1 I(N_S - N_E) \\
-\frac{dN_E}{dt} &= -b_1 + (d+\mu)N_E - \mu N_I \\
&\quad +(1-u_1)\frac{\beta_2 D_S I}{N^2}(N_{D_I} - N_{D_S}) \\
-\frac{dN_I}{dt} &= -b_2 + (d+\alpha)N_I + (1-u_2)\beta_1 S(N_S - N_E) \\
&\quad +u_3\gamma(N_I - N_R) \\
&\quad +(1-u_1)\frac{\beta_2 D_S(N-I)}{N^2}(N_{D_S} - N_{D_I}) \\
-\frac{dN_R}{dt} &= dN_R + \eta(N_R - N_S) \\
&\quad +(1-u_1)\beta_2 D_S I(N_{D_1} - N_{D_S}) \\
-\frac{dN_{D_S}}{dt} &= d_2 N_{D_S} + (1-u_1)\frac{\beta_2 D_1 S}{D_N^2}(N_E - N_S) \\
&\quad +\frac{\beta_2 I}{N}(N_{D_S} - N_{D_I}) \\
-\frac{dN_{D_I}}{dt} &= -b_3 + (d_2+\sigma)N_{D_I} \\
&\quad +(1-u_1)\frac{\beta_2 S(D_N - D_I)}{D_N^2}(N_S - N_E) \\
&\quad \frac{\frac{\beta_2 D_1 S}{D_N}(N_E - N_S) + \frac{\beta_2 D_1 S}{D_N}(N_{D_I} - N_{D_S})}{2c_1}
\end{aligned}$$

Solving for the values of $u_1^*, u_2^*$ and $u_3^*$ with respect to the constraints, the characterization (15-16) can be arrived at as

$$\begin{aligned}
0 &= \frac{\partial H}{\partial u_1} \\
&= -2c_1 + \frac{\beta_2 D_1 S}{D_N}(N_E - N_S) + \frac{\beta_2 D_1 S}{D_N}(N_{D_I} - N_{D_S}) \\
0 &= \frac{\partial H}{\partial u_2} = -2c_2 + \beta_1 SI(N_E - N_S) \\
0 &= \frac{\partial H}{\partial u_3} = -2c_3 + \gamma(N_I - N_R).
\end{aligned}$$

Thus, we have (see for example Lenhart and Workman [18])

$$\begin{aligned}
u_1^* &= \frac{\frac{\beta_2 D_1 S}{D_N}(N_E - N_S) + \frac{\beta_2 D_1 S}{D_N}(N_{D_I} - N_{D_S})}{2c_1} \\
u_2^* &= \frac{\beta_1 SI(N_E - N_S)}{2c_2} \\
u_3^* &= \frac{\gamma(N_I - N_R)}{2c_3}.
\end{aligned}$$

Making use of standard control arguments which considers the bounds on the controls, we make the conclusion that

$$\begin{cases} 0 \quad \text{If} \quad \xi_{\mathrm{i}}^* \leq 0 \\ \xi_{\mathrm{i}}^* \quad \text{If} \quad 0 < \xi_{\mathrm{i}}^* < 0 \\ 1 \quad \text{If} \quad 0 < \xi_{\mathrm{i}}^* \quad \geq 0 \end{cases}$$

For $i \in 1, 2, 3$ and where

$$\begin{aligned}
\xi_1^* &= \frac{\frac{\beta_2 D_1 S}{D_N}(N_E - N_S) + \frac{\beta_2 D_1 S}{D_N}(N_{D_I} - N_{D_S})}{2c_1} \\
\xi_2^* &= \frac{\beta_1 SI(N_E - N_S)}{2c_2} \\
\xi_3^* &= \frac{\gamma(N_I - N_R)}{2c_3}
\end{aligned}$$

□

The next section shall be focused on the detailed discussion on the numerical simulation solution results which is hinged on the optimality of the model taking into account of different kind of strategies of the optimal controls $u_1 u_2$ and $u_3$, the parameter selections and meanings from various strategies.

## 7 Numerical Simulations

The numerical simulation solutions for this model is undertaken using MATLAB 10.0 version. The optimality system, which comprise the state system and the adjoint system, was worked out to obtain the optimal control solution. The optimality system solution was calculated using a fourth-order Runge-Kutta iterative scheme. The adjoint equations were also worked out by the backward fourth-order Runge-Kutta scheme applying the preceding solutions of the state equations hinged on the transversality conditions Equation (14). The controls results obtained were updated using a convex combination of the previous controls and the value obtained from the characterizations. This activity was carried on and the iterations were terminated if the values of the unknowns at the former iterations were similar to the ones arrive at the current iteration [2, 18]. Table 1 shows parameters and values used in the numerical simulation of the computer virus model. The following weight constants were considered: $b_1 = 10, b_2 = 30, b_3 = 90$ and $c_1 = 100, c_2 = 120, c_3 = 300$.

### 7.1 Prevention $(u_1, u_2)$ of Computer Virus and Removable Infected Device

The prevention control $u_1$ on the intensive public education on the use of removable device (intensive public education) and the prevention control $u_2$ (public campaign on computer virus free)are used to optimise the objective function J, and at the same time the control $(u_3)$ is set to zero. Figure 6(a) depicts that the number of removable infected device infected $D_I$ is significant different from application of control and without control presence. This control strategy only brings the number of infected removable device down but cannot control it as the number of

Table 3: Description of variables and parameters of the model

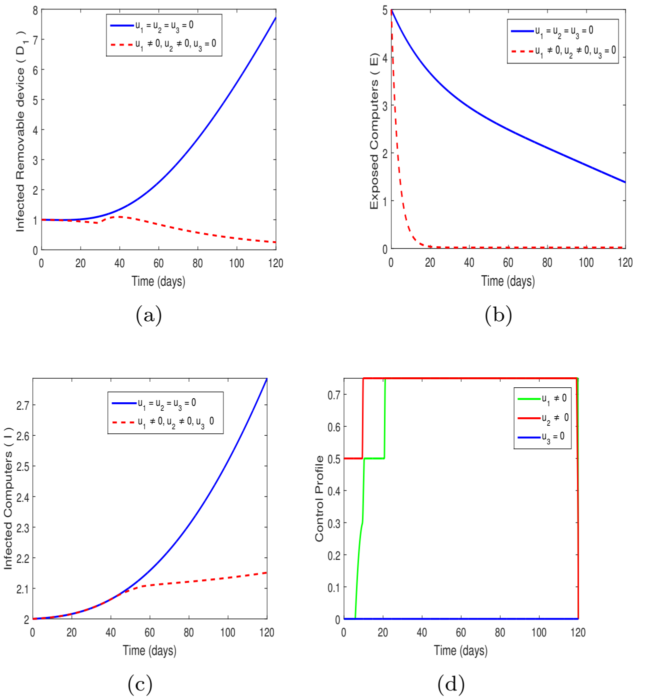
| Parameter | Value | Ref. |
|---|---|---|
| $\beta$ | 0.65 $day^{-1}$ | Assumed |
| $\sigma$ | 0.95 | Assumed |
| $\eta_1$ | 1/(365x60) $day^{-1}$ | [6] |
| $\eta_2$ | 0.44 | [11] |
| $\eta_3$ | 0.6 | [6] |
| $\pi$ | 1/14 | [6] |
| $\gamma$ | 1/14 | [6] |
| $\rho_1$ | 0.75 | [6] |
| $\rho_2$ | 0.4 | [11] |
| $\rho_3$ | 0.95 $day^{-1}$ | assumed |
| $\delta$ | 0.0085 $day^{-1}$ | assumed |
| $\mu$ | 0.055 $day^{-1}$ | assumed |

infected removable device increases after the intervention as shown in Figure 6(a). In Figure 6(b) there is a substantial difference also exists between the case controlled and without controlled case.

The positive impact on the control strategy suggests that giving intensive public campaign on the computer virus free is effective however, it does not completely controlled the number of exposed computers $E$. The Figure 6(c) shows the number of infected computer with virus and this strategy suggests that both the controlled and without controlled case are rising after the intervention. This condition is expected since there is no effective strategy on treatment of infected computers as shown in Figure 6(c). The control profile is depicted in Figure 6(d) as control $u3$ is set to zero.The control $u1$ is initially set to 50% for 20 days which is then increased to 100% for the rest of the intervention. While control $u2$ is initially also set to 50% for 6 days then rise up to 100% for the rest of the intervention as shown in Figure 6(c).

### 7.2 Prevention ($u_1$) and Treatment ($u_3$) of Computer Virus and Removable Device

Prevention and treatment control $u_1, u_3$ (intensive public education on the use of removable device and treatment of infected computers) are used to optimize the objective function J, and while control ($u_2$) is set to zero. Figure 7(a) shows a significant difference between the use of control and that of without control. This strategy suggests that the number of infected removal devices are minimized but infected removable device $D_1$ increases. In Figure 7(b) there is a negative impact on the control strategy since the presence of the control has no effect on reducing the number of exposed computers $E$ to infected computers. There is no control strategy designed to effectively ensure that intensive public education on computer virus are carried on. Figure 7(b) depicts the number of infected computers $I$ and there is a significant difference



(a) (b)

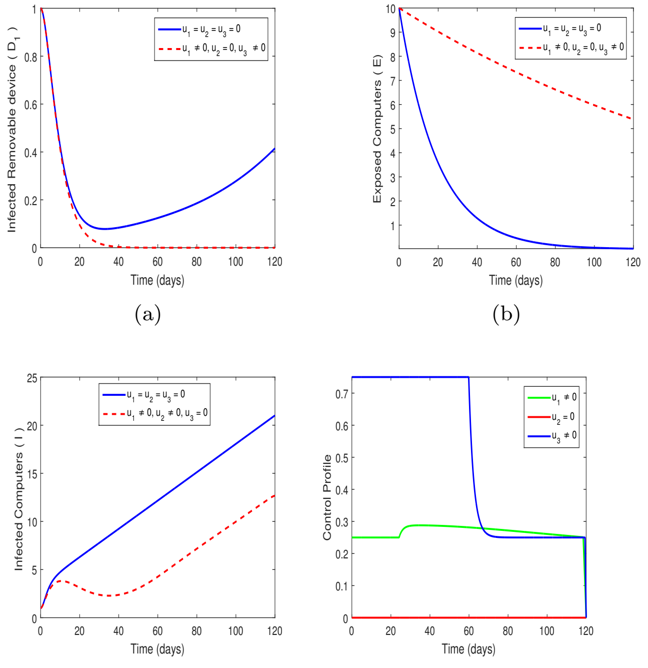
(c) (d)

Figure 6: Simulations of the model showing the effect of computer virus and infected removable device prevention on transmission

between the controlled case and without the application of control. However, the presence of the control strategy has a minimal effect since both controlled and without control increasing after the intervention. In Figure 7(c) the control $u_2$ is set to zero while control $u_1$ is set to 25% for the beginning and relatively kept same for the entire intervention. The control $u_3$ is also for a start is kept at 100% for 60 day and then reduce to 24% which is then maintained for the rest of the intervention.

### 7.3 Prevention ($u_2$) and Treatment ($u_3$) of Computer Virus

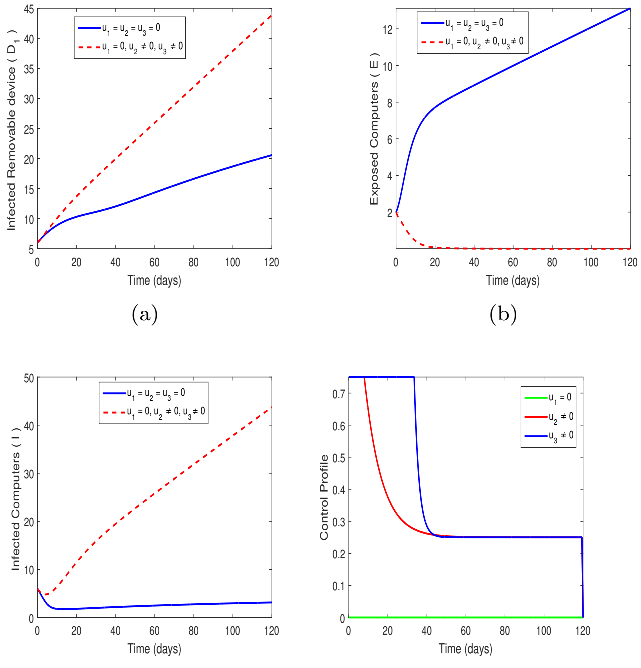
In this strategy, the prevention control strategy $u_2$ and treatment control $u_3$ (intensive public campaign on computer virus free and treatment) are used to optimize the objective function J. Figure 8(a) shows the number of infected removable device $D_I$ and there is a significant difference between the controlled and without controlled. The results in Figure 8(a) suggests that there is a negative impact on the control strategy since controlled case is higher than without controlled case. Therefore, this control mechanism has no effect on controlling the number of infected removal devices $D_I$. Again, in Figure 8(b) there is a substantial difference between the application of control and without the use of control. The positive effect of this strategy suggests that the control strategy is effective during the entire intervention and is able to control the number of exposed computer $E$to virus. However, the uncontrolled case rise up at the end of the intervention as



(a) (b)

(c) (d)

Figure 7: Simulations of the model showing the effect of computer virus and infected removal device prevention and treatment on transmission



(a) (b)

(c) (d)

Figure 8: Simulations of the model showing the effect of computer virus and infected removal devices prevention and treatment on transmission

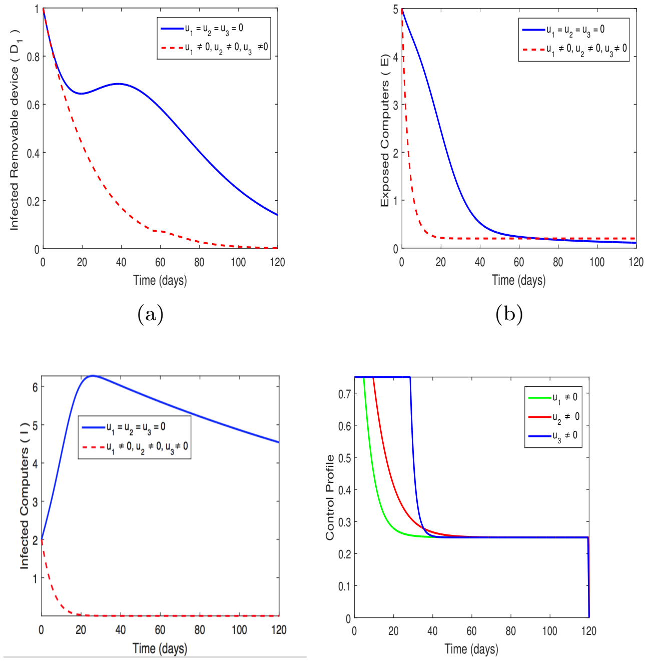
shown in Figure 8(b). In Figure 8(c) there is a significant difference between the use of control and without control application. The negative impact as shown in Figure 8(b) suggests that this control strategy is not very robust in reducing the the number of infected computers. The control profile of this strategy is shown in Figure 8(d) and control $u_1$ is set to zero. The control $u_2$ is initially set to 100% for 18 days and gradually reduce to 27% which is then maintain throughout the rest of the intervention. Similarly, the control $u_3$ is at 100% for 34 days and gradually minimize to 27% for the rest of the intervention.

### 7.4 Prevention $(u_1)$,$(u_2)$ and Treatment $(u_3)$ of Computer Virus and Infected Removal Devices

In this strategy all the control strategies $u_1, u_2, u_3$ ( intensive education on infected removal devices, campaign on computer virus free and treatment of infected computers) are used simultaneously to optimize the objective function J. Figure 9(a), shows that there is significantly different between the controlled case and without controlled case. This positive impact suggests that the control mechanism is very effective and robust in controlling the number of infected removal devices. This maybe attributable to the combination of the other controls for their effectiveness. A similar pattern is depicted in Figure 9(b) as there is a vast positive difference between controlled case and without application of control. This also infers that the strategy is effective in controlling the number of exposed computers $E$. In fact, the number of exposed computers are totally brought under control with the combination of all the controls as shown in Figure 9(b). Figure 9(c) shows the number of infected computers $I$ and there is a significant difference between the use of control and without control. The positive effect indicates that the control strategy is effective as both controlled and without controlled are brought under effective control as shown in Figure 9(c). The control profile for this strategy is depicted in Figure 9(d) as all the controls are used at the same time. Control $u_1$ is initially kept at 100% for 6 days then reduce to 28% which is maintained throughout the rest of the intervention. Similarly, control $u_2$ is also kept at 100% for 10 days for the beginning and then reduce to 28% which is kept for the rest of the intervention. While control $u_3$ is maintained at 100% for 25 days which is also reduce to 28 day and kept constantly for the rest of the intervention.

## 8 Conclusion

In this work, a computer virus model incorporating a removal device of deterministic type was formulated. The basic properties of the model was investigated then the stability analysis of the model was studied. The steady states found to be stable.The reproduction number $R_0$ was calculated. Time dependent controls was incorporated into the model and Pontryagin's Maximum Princi-



(a) (b)

(c) (d)

Figure 9: Simulations of the model showing the effect of infected computer virus and infected removal devices prevention and treatment only on transmission

ple was used to determine all the necessary conditions for controlling the spread of computer virus. The numerical simulation carried out on the control suggests that the best strategy in controlling the spread of computer virus is the use of all the three controls at the same time. The implication of the result seems to suggest that all effort must given to all the strategies designed without relaxing.

# References

[1] G. Abramson, "Mathematical modeling of the spread of infectious diseases," *A Series of Lectures Given at PANDA*, vol. 92, no. 1, pp. 33V42, 2009.

[2] F. B. Agusto, "Optimal chemoprophylaxis and treatment control strategies of a tuberculosis transmission model," *World Journal of Modelling and Simulation*, vol. 5, no. 3, pp. 163–173, 2009.

[3] Z. Chen, L. Gao, K. Kwiat, "Modeling the spread of active worms", in *IEEE Societies of Twenty-Second Annual Joint Conference of the IEEE Computer and Communications (INFOCOM'03)*, vol. 3, pp. 1890–1900, 2003.

[4] N. Chitnis, J. M. Hyman, J. M. Cushing, "Determining important parameters in the spread of malaria through the sensitivity analysis of a mathematical model," *Bulletin of Mathematical Biology*, vol. 70, no. 5, pp. 1272–1296, 2008.

[5] P. V. Driessche, J. Watmough, "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission," *Mathematical Biosciences*, vol. 180, no. 1, pp. 29–48, 2002.

[6] M. H. Farahi, M Shirazian, "Optimal control strategy for a fully determined hiv model," *Intelligent Control and Automation*, vol. 1, no. 1, pp. 15–19, 2010.

[7] W. H. Fleming, R. W. Rishel, "Deterministic and stochastic optimal control,"*Application of Mathematics*, vol. 1, 1975.

[8] S. M. Garba, M. R. A. Bakar, A. B. Gumel, "Backward bifurcations in dengue transmission dynamics," *Mathematical Biosciences*, vol. 215, no. 1, pp. 11–25, 2008.

[9] P. Georgescu, H. Zhang, "A lyapunov functional for a SIRI model with nonlinear incidence of infection and relapse," *Applied Mathematics and Computation*, vol. 219, no. 16, pp. 8496–8507, 2013.

[10] J. L. Gondar, R. Cipolatti, "A mathematical model for virus infection in a system of interacting computers," *Computational & Applied Mathematics*, vol. 22, no. 2, pp. 209–231, 2003.

[11] S. Gourari J. Karrakchou, M. Rachik, "Optimal control and infectiology application to an HIV/AIDS model," *Applied Mathematics and Computation*, vol. 177, no. 2, pp. 807–818, 2006.

[12] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Review*, vol. 42, no. 4, pp. 599–653, 2000.

[13] S. M. Hashemi, J. He, "An evolutionary multi-objective approach for modelling network security," *International Journal of Network Security*, vol. 19 no. 4, pp. 528–536, 2017.

[14] M. Javidi, N. Nyamorady, "Stability analysis of a novel (VEISV) propagation model of computer worm attacks," *World Journal of Modelling and Simulation*, vol. 10, no. 3, pp. 163–174, 2014.

[15] W. Jin, "Applying epidemiology in computer virus prevention: Prospects and limitations," *Department of Computer Science, University of Auckland*, pp. 1–13, 1996.

[16] M. Kumar, K. Dutta, I. Chopra, "Impact of wormhole attack on data aggregation in hieracrchical WSN," *International Journal of Electronics and Information Engineering*, vol. 1, no. 2, pp. 70–77, 2014.

[17] R. S. Lebelo, M. Mukamuri, S. T. Ogunlade, K. O. Okosun, "Optimal control analysis of cryptosporidiosis disease," *Global Journal of Pure and Applied Mathematics*, vol. 12, no. 6, pp. 4959–4989, 2016.

[18] S. Lenhart, J. T. Workman, *Optimal control applied to biological models*, Chapman and Hall/CRC, 2007.

[19] B. K. Mishra, D. Saini, "Mathematical models on computer viruses," *Applied Mathematics and Computation*, vol. 187, no. 2, pp. 929–936, 2007.

[20] B. K. Mishra, S. K. Pandey, "Dynamic model of worms with vertical transmission in computer network," *Applied Mathematics and Computation*, vol. 217, no. 21, pp. 8438–8446, 2011.

[21] B. K. Mishra, N. Jha, "SEIQRS model for the transmission of malicious objects in computer network," *Applied Mathematical Modelling*, vol. 34, no. 3, pp. 710–715, 2010.

[22] B. K. Mishra, S. K. Pandey, "Fuzzy epidemic model for the transmission of worms in computer network," *Nonlinear Analysis: Real World Applications*, vol. 11, no. 5, pp. 4335–4341, 2010.

[23] B. K. Mishra, A. Prajapati, "Dynamic Model on the transmission of malicious codes in network," *International Journal of Computer Network and Information Security*, vol. 5, no. 10, pp. 17-23, 2013.

[24] H. Okamura, T. Dohi, "Estimating computer virus propagation based on markovian arrival processes," *IEEE 16th Pacific Rim International Symposium on Dependable Computing (PRDC'10)*, pp. 199–206, 2010.

[25] K. O. Okosun, O. D. Makinde, I. Takaidza, "Impact of optimal control on the treatment of HIV/AIDS and screening of unaware infectives," *Applied Mathematical Modelling*, vol. 37, no. 6, pp. 3802–3820, 2013.

[26] S. M. Omair, S. K. Pandey, "e-Epidemic model on the computer viruses in the network," *European Journal of Advances in Engineering and Technology*, vol. 2, no. 9, pp. 78–82, 2015.

[27] E. U. Opara, O. A. Soluade, "Straddling the next cyber frontier: The empirical analysis on network security, exploits, and vulnerabilities," *International Journal of Electronics and Information Engineering*, vol. 3, no. 1, pp. 10–18, 2015.

[28] M. Peng, X. He, J. Huang, T. Dong, "Modeling computer virus and its dynamics," *Mathematical Problems in Engineering* vol. 2013, no. 17, pp. 1–5 , 2013.

[29] J. R. C. Piqueira, V. O. Araujo, A. A. de Vasconcelos, C. E. C. J. Gabriel, "Dynamic models for computer viruses," *Computers & Security*, vol. 27, no. 7-8, pp. 355–359, 2008.

[30] M. A. Safi, S. M. Garba, "Global stability analysis of SEIR model with holling type II incidence function," *Computational and Mathematical Methods in Medicine*, vol. 8, no. 4, pp. 1–8, 2012.

[31] J. Singh, "Cyber-attacks in cloud computing: A case study," *International Journal of Electronics and Information Engineering*, vol. 1, no. 2, pp. 78–87, 2014.

[32] B. Song, J. P. Aparicio, C. Castillo-Chavez, "Tuberculosis models with fast and slow dynamics: The role of close and casual contacts," *Mathematical Biosciences*, vol. 180, no. 1, pp. 187–205, 2002.

[33] M. Styugin, "Establishing systems secure from research with implementation in encryption algorithms", *International Journal of Network Security* vol. 20, no. 1, pp. 35–40, 2017.

[34] H. Toyoizumi, K. Kaiwa, Y. Kobayashi, J. Shitozawa, "Stochastic features of computer viruses: Towards theoretical analysis and simulation," *The 5th St. Petersburg Workshop on Simulation*, pp. 695–702, 2005.

[35] X. Yang, L. Yang, "Towards the epidemiological modeling of computer viruses," *Discrete Dynamics in Nature and Society*, vol. 2012, pp. 1–11, 2012.

[36] H. Zheng, Z. Gao, D. Li, "An epidemic model of mobile phone virus, pervasive computing and applications," *First International Symposium on Pervasive Computing and Applications*, pp. 1–5, 2006.

[37] Q. Zhu, X. Yang, J. Ren, "Modeling and analysis of the spread of computer virus," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 12, pp. 5117–5124, 2012.

# Biography

**Okosun Kazeem** is a Professor of Mathematics in Vaal University of Technology, Vanderbijlpark, South Africa. He received his Ph.D degree in Applied Mathematics, University of the WesternCape, South Africa in 2010. He has written several articles. As a result of this, he is currently rates as NRF researcher. His present interests include Mathematical modeling, Numerical analysis and Optimization.

**Ogunlade Samson** is a lecturer in the Department of Mathematical Sciences, Federal University of Technology, Akure. He received his masters in 2015 at AIMS-Senegal and is currently undergoing his PhD in UNSW, Australia. His research interests are Biomathematics and Modelling.

**Bonyah Ebenezer** is a Senior Lecturer in Mathematics and Statistics Department, Kumasi Technical University, Ghana. He has several publications to his credit and he majors in Mathematical modelling and Optimization.

# A Detection and Defense Technology for Information-stealing and Deceitful Trojan Viruses Based on Behavioral Features

Li Lin
*(Corresponding author: Li Lin)*

Department of Information Engineering, Huanghuai University
Zhumadian, Henan, 463000, China
(Email: linli2016_li@126.com)

*(Received Dec. 12, 2017; revised and accepted Apr. 28, 2018)*

## Abstract

With the rapid development of Internet technology, computer networks play an increasingly important role in improving the living standard of people. However, it also induces events which threaten the security of Internet. Information stealing and network intrusion are always threatening network information security. In this study, the behavioral features of Trojan were analyzed, and a detection and defense system for information-stealing and deceitful Trojan virus based on behavioral features was designed and tested by simulation experiments. The simulation results demonstrated that the system has favorable accuracy and practicability, and its performance can satisfy practical applications. The system installed at a network gateway can monitor network flows and detect whether there is invasion of information-stealing and deceitful Trojan viruses, which is of great significance to the defense of network security.

*Keywords: Behavioral Features; Cyber Theft; Network Information Security; Trojan Virus*

## 1 Introduction

Network information security is an important cornerstone for the development of a modern network [6,9]. A large amount of valuable resources and information stored in the computer network make criminals on the Internet eager to get all these valuable resources by using various hacking technologies [12]. Trojan virus is the most common technology. Information-stealing Trojan virus can acquire computer information resources by invading computers [3]. Developing Trojan virus detection and defense technology has been one of priorities today. Many scholars have made deep study [1].

Ni *et al.* [8] put forward a wavelet transform based a noise optimization method through detecting hardware Trojan viruses using a back propagation neural network technology. The experimental results suggested that the wavelet transform-based noise optimization method could eliminate high-frequency noises and make the sensitivity of detecting hardware Trojan viruses based on neural network increase from 92.2% to 99.2%. In a study of Xi [14], several defense algorithms were proposed based on the summary of the study status and key technologies of Address Resolution Protocol (ARP) and analysis of the formation mechanism of ARP bugs, and the advantages of the improved defense algorithms were introduced; finally the optimal defense algorithm was determined after testing every improved algorithm.

In this study, the behavioral features of Trojan virus were analyzed, and then an information-stealing Trojan virus detection and defense system was designed based on the behavioral features of Trojan virus. The test suggested that the system could effectively test information-stealing and deceitful Trojan virus, which has great significance to the protection of network information security.

## 2 Analysis on Behavioral Features of Trojan Virus

Trojan virus invades systems by concealing itself. Many behavioral features will be presented though it can hide its tracks left in the targeted system [10]. Next is the analysis of behavioral features of Trojan virus invasion at different stages.

1) Stage of Trojan virus implantation:
When invading a system, Trojan virus will deceive users to gain their trust using illegal means and then enter the target [5]. In this stage, the behavioral features of Trojan virus include attacking via bugs in software and systems, rogue programs, ports, foreign unknown e-mails, and unknown network links.

Table 1: Main behavioral features of Trojan virus and division of risk levels

| Moderate | High |
|---|---|
| 1. Replicate or create files in catalog;<br>2. Self-starting;<br>3. Call cmd process;<br>4. Bind monitor port. | 1. Automatically delete files;<br>2. Automatically compress or decompress files;<br>3. Revise system time;<br>4. Creating file associations;<br>5. Do Internet Control Messages Protocol communication;<br>6. Conceal process;<br>7. Close corresponding process;<br>8. Automatically send e-mail outward;<br>9. Disguise system process or communication path;<br>10. Entering system process or IE process. |

2) Stage of Trojan virus installation:
There are some special behavioral features when Trojan virus is installed. In this stage, the main action object of Trojan virus is itself. Therefore, it is a good stage to detect Trojan virus. The behavioral features of Trojan virus in this stage include compressing or decompressing files automatically, deleting some files automatically, automatically restarting, automatically system timing, automatically closing or opening programs, revising system configuration files, and revising the system relevance.

3) Stage of startup operation:
When Trojan virus is successfully installed in the target system, operation modes such as process and thread are needed. Generally, processes can be observed, but threads cannot [4]. Therefore, Trojan virus has to conceal itself. In this stage, the behavioral features of Trojan virus include concealing processes, calling cmd processes, closing specific processes. and transferring to other processes via a remote thread technology.

4) Stage of network communication:

   a. After a system is controlled by Trojan virus, information communication is needed. Trojan virus will receive information from outside such as the new commands of the controller and transmit information to the controlling end via special communication modes [7]. It can control the target system and steal system information via those means.

   b. Main behavioral features of Trojan virus and division of risk levels A corresponding database of Trojan virus behavioral features was established based on Table 1. A large number of Trojan virus features were collected and applied in the detection and defense of Trojan virus.

# 3 Design an Information-Stealing and Deceitful Trojan Virus Detection System

## 3.1 Framework of Trojan Virus Detection Model

A behavioral features-based Trojan virus detection model can classify programs based on behaviors of programs and the aforementioned database of Trojan virus behavioral features by using Bayes classifier [13]. The model designed in this study was composed of behavioral extraction, behavioral features database, a program behavioral analyzer, Trojan virus processor and user assistance.

Behavioral extraction aims at monitoring suspicious behaviors in the system and sending them to the behavioral analyzer [15]. Database of Trojan virus features included a large amount of Trojan virus behaviors, action objects of behaviors, descriptive information of behaviors and basic probability information. The program behavioral analyzer was the most important in the model. It can classify programs using Bayes classifier and perfect its classification processing ability through communicating with the features of database. User assistance could achieve human-computer interactions by sending contents which cannot be determined by the analyzer to users and present data to users.

## 3.2 The Main Module of the Proposed System

### 3.2.1 Design of Behavioral Extraction Module

To realize monitoring and capture of all program behaviors, a behavior monitoring module should be installed in the core of the operation system to monitor file system, registry, processes, storages, and communication. API-HOOK [11] was used to intercept a program system called in this study. The implementation of the behavioral monitoring system was introduced by taking the monitoring of registry as an example.

Table 2: The features - behavioral database

| | **Feature set database** | | | | **Feedback self-adaption database** | |
|---|---|---|---|---|---|---|
| 2-7 Data sheet | Feature set sheet | Test feature sheet | Parameter sheet | Feature benchmark sheet | Classification performance sheet | System parameter sheet |
| Description | Storing sample feature vectors which need examination | Storing testing data of normal programs and Trojan virus | Storing API parameters | Storing various data of behavioral features after classification | Storing system classification performance description | Storing system analysis risk coefficient and classification width |

Registry, the core database of an operation system, can store the setting information of systems and application programs. SSDT HOOK technology is needed in the monitoring of registry. All the operations associated to registry can be monitored under the assistance of SSDT HOOK. Registry operations in the monitoring module include creating, revising and deleting registry key items, and creating, revising and deleting registry key values. In the monitoring based on SSDT HOOK technology [16], SSDT entry should be backed up before revision, and the monitored function should be replaced by a self-defined monitoring function.

Besides API HOOK technology, static analysis of PE was used to acquire behavioral characteristic vectors. PE is a file format which can be transplanted and executed in window systems. It can store data set codes in a linear address space and analyze PE files using BK-50 scanner to acquire the features-behavioral vector of programs.

Static analysis of PE files is ineffective to programs which apply a code obfuscation technology. Therefore, API HOOK was the main body, and static analysis of PE was the assist in the design of the behavioral extraction module.

### 3.2.2 Features-Behavioral Database

The common behaviors of Trojan virus have been introduced in Section 2. The features- behavioral database contained various data, which could provide a reference for classification of the classifier. The behavioral features of Trojan virus include normal programs such as characteristic weight a, risk coefficient S, width of classifier *theta* and basic conditional probability. The features-behavioral database could be divided into characteristic set database and feedback self-adaption database. The classification of the behavioral characteristics database is shown in Table 2.

### 3.2.3 Behavioral Analysis Module

Behavioral analysis module, an important part of the Trojan virus detection model, was mainly composed of data preprocessing, program classification and feedback study. The algorithm was described as follows.

Data preprocessing included redundancy elimination, feature vector independent processing and weight calculation.

1) In redundancy processing, feature set and test feature set sheets were input. Through calculation of feature CRR and deletion of features with low relevancy, a non-redundant feature set sheet was output.

2) In feature vector independent processing, a non-redundant feature set was input. Then behavioral attributes were merged using SNCB model. Finally, an irrelevant non-redundant feature set was output.

3) In weight calculation, an irrelevant non-redundant feature set was input; then impact factors were assigned, scales were obtained, and weights were calculated. Finally, feature weights were output.

4) In classification calculation, sample feature sets, classification algorithms, unknown examples and a system parameter sheet were input; then features of examples were extracted by the classifier which has been regulated by statistical calculation of sample features; after classification, the classification results were output.

5) In classifier learning, a classification performance sheet was input, and feature set and system parameter sheets were output through incremental learning and re-learning.

### 3.2.4 System Response Module

After classification on running programs, corresponding operations needed to be done according to the classification results. For example, the discovered Trojan virus program needed to be cleared, and the classification conditions were fed back to users. After classification, the system stopped monitoring normal programs, and Trojan virus programs and programs which could not be determined were provided to users for processing. The system

Table 3: The detection results of the Trojan virus samples

| | Huigezi Trojan virus | Dark distant control Trojan virus | Byshell Trojan virus |
|---|---|---|---|
| Number of samples | 20 | 20 | 20 |
| Correct detection number | 18 | 19 | 19 |
| False detection number | 2 | 1 | 1 |
| Detection rate | 90% | 95% | 95% |
| False alarm rate | 10% | 5% | 5% |

Table 4: The testing results of thread performance

| | | | | | |
|---|---|---|---|---|---|
| Performance test on single thread | | 500 Mbps | 800 Mbps | 1 Gbps | 1.5 Gbps |
| | Peak CPU | 87.1% | 91.6% | 89.3% | 91.4% |
| | Memory usage | 9.6% | 11.6% | 12.0% | 13.1% |
| | Thread | 1 | 1 | 1 | 1 |
| | Packet loss probability | 0% | 0% | 0.14% | 1.6% |
| Performance test on multi-thread | | 2 Gbps | 4 Gbps | 6 Gbps | 8 Gbps |
| | Peak CPU | 114% | 154% | 234% | 301% |
| | Memory usage | 13.4% | 15.7% | 19.8% | 25.6% |
| | Thread | 2 | 4 | 6 | 8 |
| | Packet loss probability | 0% | 0% | 0.05% | 0.02% |

deleted Trojan virus programs once discovered and transmitted the information of Trojan virus to users. When it was difficult to determine whether a program was normal or not, the system would isolate the program and display the condition to users for determination. After determination, the system deleted or restored it according to actual conditions.

## 4 System Test

The detection and false alarm rates of the Trojan virus detection system were tested. The purpose of the test was to test the accuracy of the Trojan virus detection system. Rules were written according to the Trojan virus features-behavioral database.

The test process was as follows: A Trojan virus controlled the terminal was implanted into a virtual host. Then the control terminal was installed in an external computer. The categories and a number of Trojan virus samples were controlled through the computer. Detection rules were formulated according to the Trojan virus samples and features-behavioral database. Then the Trojan virus detection system was started and controlled to communicate with the host. The total number and valid number of alarms were accounted. Finally, the detection and false alarm rates of the system were calculated.

The expected detection rate was not lower than 80%, and the expected false alarm rate was not higher than 10%.

In the test, Huigezi Trojan virus, dark distant control Trojan virus, and Byshell Trojan virus were selected for testing. During testing, the three viruses were installed in virtual hosts, and an abnormal communication flow was generated. In the testing cluster, there were 100 hosts, and every host was installed with Huigezi Trojan viruses, dark distant control Trojan viruses, and Bysell Trojan viruses, containing 20 viruses for each kind of above viruses. Then the step number was detected at the cluster exit. After repeating tests following the above procedures, the results obtained were in following tables.

It is seen from Table 3 that the detection rates of those three Trojan viruses were 90%, 95%, and 95%, respectively, and the corresponding false alarm rates were 10%, 5% and 5%, respectively. The results suggest that the behavioral features-based, information-stealing, and deceitful Trojan virus detection system could detect Trojan viruses included in the features- behavioral database. The false alarm might happen because the users presented some behavioral features similar to Trojan viruses under a certain condition. The detection and false alarm rates satisfied the aforementioned expectations.

In Table 4, it is demonstrated that the packet loss gradually happened with the increase of the flow in a single-thread operation; the higher the flow, the severer the pack loss phenomenon. Thus, it could be concluded that the calculation capability of a single thread was not suitable for calculating a large flow, as an excessively large data packet could lead to overflow of bottom data, leading to a packet loss. Therefore, a single thread was only suitable for processing the data flow between 500 M and 1 G. Flow packages could increase continuously in the process of multi-thread processing. In Table 2, it is demonstrated that CPU occupancy rate and memory utilization rate are significantly improved during the multi-thread processing; however, the packet loss gradually appeared with the increase of flow. Hence, the peak flow supported by the designed system was 10 G.

## 5 Conclusion

The development of Internet facilitates worldwide information sharing, but it also brings huge challenges to network security. Network development results in the spread of information stealing events. Some lawbreakers develop many information-stealing and deceitful Trojan viruses to steal information. In this study, the behavioral features of information-stealing and deceitful Trojan viruses were analyzed, then an information-stealing and deceitful Trojan virus detection system was developed based on the behavioral features, and the feasibility and performance of the system were tested. The test results demonstrated that the detection rate and false alarm rate of the system satisfied the standards, and the supportable peak flow was 10 G, which lays a reference for behavioral features-based Trojan virus detection technologies.

## References

[1] A. A. Al-khatib, W. A. Hammood, "Mobile malware and defending systems: Comparison study," *International Journal of Electronics and Information Engineering*, vol. 6, no. 2, pp. 116–123, 2017.

[2] A. Anurag, "Network neutrality: Developing business model and evidence based net neutrality regulation," *International Journal of Electronics and Information Engineering*, vol. 3, no. 1, pp. 1–9, 2015.

[3] E. C. Dunn, "Trojan pig: Paradoxes of food safety regulation," *Environment & Planning A*, vol. 35, no. 8, pp. 1493–1511, 2015.

[4] F. Farahmandi, Y. Huang, P. Mishra, "Trojan localization using symbolic algebra," in *Design Automation Conference*, pp. 591–597, 2017.

[5] I. Hsiao, Y. K. Hsieh, C. F. Wang, *et al.*, "Trojan-horse mechanism in the cellular uptake of silver nanoparticles verified by direct intra- and extracellular silver speciation analysis," *Environmental Science & Technology*, vol. 49, no. 6, pp. 3813–3821, 2015.

[6] S. Islam, H. Ali, A. Habib, N. Nobi, M. Alam, and D. Hossain, "Threat minimization by design and deployment of secured networking model," *International Journal of Electronics and Information Engineering*, vol. 8, no. 2, pp. 135–144, 2018.

[7] X. W. Li, X. H. Wang, Y. Zhang, K. Y. Chen, L. Xu, "A new hardware trojan detection method based on kernel maximum margin criterion," *Acta Electronica Sinica*, vol. 45, no. 3, pp. 656–661, 2017.

[8] L. Ni, J. Li, S. Lin, *et al.*, "A method of noise optimization for Hardware Trojans detection based on BP neural network," in *IEEE International Conference on Computer and Communications*, pp. 2800–2804, 2017.

[9] E. U. Opara, O. A. Soluade, "Straddling the next cyber frontier: The empirical analysis on network security, exploits, and vulnerabilities," *International Journal of Electronics and Information Engineering*, vol. 3, no. 1, pp. 10–18, 2015.

[10] A. R. Ribeiro, S. Gemini-Piperni, R. Travassos, *et al.*, *Trojan-like Internalization of Anatase Titanium Dioxide Nanoparticles by Human Osteoblast Cells*, Scientific Reports, 6:23615, 2016.

[11] Y. Song, Y. Shen, G. Zhang, "The new INLINE hook technology combination of hard-code technology and independent code injection," in *IEEE International Conference on Software Engineering and Service Science*, pp. 521–525, 2017.

[12] A. Tayal, N. Mishra and S. Sharma, "Active monitoring & postmortem forensic analysis of network threats: A survey," *International Journal of Electronics and Information Engineering*, vol. 6, no. 1, pp. 49–59, 2017.

[13] M. Turkoglu, D. Hanbay, "Classification of the grape varieties based on leaf recognition by using SVM classifier," in *Signal Processing and Communications Applications Conference*, pp. 2674–2677, 2015.

[14] H. Xi, "Research and application of ARP protocol vulnerability attack and defense technology based on trusted network," in *Advances in Materials, Machinery, Electronics. Advances in Materials, Machinery, Electronics (AMME'17)*, pp. 090019.1–090019.7, 2017.

[15] M. Xue, A. Hu, W. Liu, *et al.*, "Detecting hardware Trojan through feature extraction in subspace domain," *Journal of Southeast University (Natural Science Edition)*, vol. 44, no. 3, pp. 457–461, 2014.

[16] Y. Zhang, H. Bi, "Anti-rootkit technology of kernel integrity detection and restoration," in *International Conference on Network Computing and Information Security*, pp. 276–278, 2011.

## Biography

**Li Lin** who has gained the master's degree is a lecturer in the college of information engineering of Huanghuai University, Henan, China. Her interests of research include computer application, data mining and big data application. She has engaged in the teaching of computer course for 13 years. Moreover she has participated in multiple provincial-level, city-level and school-level projects, published more than 10 academic papers (more than half in Chinese core periodicals), participated in the editing of a textbook, and gained multiple honors from the school for her works.

# Exploring a Novel Cryptographic Solution for Securing Small Satellite Communications

Samuel Jackson[1], Jeremy Straub[2], and Scott Kerlin[3]
*(Corresponding author: Jeremy Straub)*

Department of Computer Science, Oklahoma State University[1]
219 MSCS; Stillwater, Oklahoma 74078, USA
Department of Computer Science, North Dakota State University[2]
1320 Albrecht Blvd, Room 258; NDSU Dept 2740; Fargo, ND 58102, USA
(Email: jeremy.straub@ndsu.edu)
Department of Computer Science, University of North Dakota[3]
3950 Campus Road, Grand Forks, ND 58202, USA

*(Received May 1, 2017; revised and accepted Aug. 26, 2017)*

## Abstract

The need for encryption for use onboard satellites is a growing issue. While larger and modern satellites may have the hardware capabilities to use common terrestrial encryption schemes, smaller and older satellites may lack such capabilities. Small satellites are becoming increasingly important, and securing their communications a necessity. The Department of Defense, for example, is developing CubeSats, a type of small satellite, for use in operations. Despite the growing need, currently, there is no agreed upon encryption algorithm for these devices, which have limited hardware capabilities and physical space. This paper presents and characterizes a novel algorithm that uses chaos theory to encrypt data. Specifically, a proof-of-concept of this novel algorithm is presented. It is then compared against AES and SPECK in terms of speed of encryption and decryption.

*Keywords: Chaotic Cryptosystem; CubeSat; Small Spacecraft*

## 1 Introduction

Encryption for use onboard satellites is an open research problem. While larger, modern satellites may be able to employ common terrestrial encryption schemes, small satellites and older larger satellites may not have the requisite hardware capabilities. Because of the radio-based transmission medium used, spacecraft communications have no inherent physical security mechanism and thus have a particular need for other security mechanisms. The Advanced Encryption Standard (AES), a standard technique used by the NSA and others for securing data transmissions, could potentially be brought to bear on this challenge.

Muhaya [21], in particular, considered the use of multiple standard techniques. While determining that AES was a component of the solution, he demonstrated the utility of and discussed the need for enhancing the AES cryptographic algorithm with a chaotic pixel shuffling mechanism. However, this prior work failed to evaluate the computational costs of typical approaches and the proposed hybrid approach. Knowledge of this is, of course, critical to determining their suitability for use on a small or older spacecraft.

The lack of a standard, usable cryptographic algorithm to secure communications to and from spacecraft is a growing concern, as programs such as NASA's Educational Launch of Nanosatellites [25] program, the University NanoSat Program [15] and the European Space Agency's Fly Your Satellite [11] program promote the building and launching of CubeSats. While not all CubeSats require encryption (and some may be precluded from its use due to FCC restrictions on amateur licenses [27]), many future CubeSats plan to incorporate propulsion (see, *e.g.*, [20]), making their potential comprise problematic and driving a need for encryption. The growth of the use of small satellites for military applications (see, *e.g.*, [1, 29]) also drives the need for suitable cyptographic technologies for use on small satellites.

The challenge of securing communications is not limited to small satellites. In 2014, NOAA's Satellite network was hacked, shutting down the network for two days. This satellite data is vital to many different applications, ranging from providing weather forecasts to national security applications. The same lightweight approaches used for small spacecraft may also be appropriate for older satellites, which may have limited capabilities (as compared to newer models).

This paper investigates a potential solution that would allow CubeSats, other small spacecraft, and older spacecraft with limited resources to encrypt and decrypt data within their hardware capabilities or transmission time requirements. Specifically, this paper presents a proof-of-concept for a novel cryptographic algorithm, implemented in software, which uses chaos-theory to encrypt data. The algorithm, invented by Huang, Ye, and Wong [13] for the use of encrypting images, has been modified to turn it into a block cipher that can be used to encrypt any data. The algorithm has been tested and compared against other, more common algorithms in terms of encryption and decryption speed. In line with Muhaya's prior work [21], the use of the proposed algorithm on its own as well as its use (and the use of SPECK) to augment AES are also considered.

This paper continues with a discussion on the challenges of the space environment that prompt the need for this technology. Then, prior work on block ciphers and encryption for use onboard spacecraft is reviewed. This is followed by the presentation of the proposed algorithm. The experiments used to validate the algorithms functionality, characterize its performance across multiple block sizes, and compare it to alternate techniques are then presented. Finally, this data is analyzed, before concluding.

# 2 Background

This section provides background information related to prior work in the fields of small satellites and cryptography. First, an overview of challenges to cryptography specific to operating in space is provided. This is followed by an overview of block ciphers. Finally, prior work on encryption for use in space is discussed.

## 2.1 Challenges of the Space Environment

The space environment presents a number of challenges relative to security. The first is the lack of any appreciable physical security for ground-to-space and space-to-ground communications. All of the foregoing occurs over radio frequency transmissions and, thus, can easily be intercepted and possibly jammed or even manipulated by an adversary. State actors, in particular, may have the capability of placing another craft (either aerial or in a lower orbit) in between the transmitting and receiving station, allowing them to perform a man-in-the-middle style attack. A variety of security techniques are needed to protect against this and similar scenarios (see, *e.g.*, [28]). Encryption is only a small portion of this challenge; however, it is critical in protecting sensitive commands (which might reveal the tactics or other plans of a spacecraft controller) and data.

Spacecraft, and in particular small spacecraft, have limited mass and volume available. Wertz *et al.* [30] and Fortescue *et al.* [12] discuss these constraints and the deliberate trade-off process that is made to determine what will be included and what must be excluded from a spacecraft design. These constraints may make a dedicated encryption hardware system impractical and also limit the computational capabilities available onboard the spacecraft.

Beyond the limitations posed to what hardware can be included, additional restrictions exist. Both power and communications time are at a premium. Thus, encryption techniques must be as fast and lightweight as possible to maximize the available processing time and capabilities. The amount of data overhead imposed by the technique must also be kept to a minimum.

## 2.2 Prior Work on Block Ciphers

The Data Encryption Standard (DES) was one of the first block ciphers widely used and accepted as a standard. It was created by IBM, in cooperation with the National Security Agency (NSA), and published in 1975 [26]. This algorithm has been adapted multiple times as computers have become more powerful and avenues of attack were discovered.

In 2001, an Advanced Encryption Standard (AES) was published and superseded DES [26]. It was created by Vincent Rijmen and Joan Daemen, and is currently listed in NSA Suite B of cryptographic algorithms for use in encryption. It is a block cipher with a block size of 16 bytes, and variable key sizes ranging from 16, 24, or 32 bytes. AES is a symmetric standard, which uses the same keys to encrypt and decrypt data. Prior work has evaluated the performance of AES and other symmetric algorithms [10] and its power consumption when processing different types of data [19].

More recently, a new field of cryptographic algorithms has been investigated, namely lightweight cryptographic algorithms. While these may be less secure than other, heavier algorithms, they are created with the goal of being used on systems with limited resources, such as sensor networks or RFID tags. A family of lightweight algorithms was recently published by the NSA in 2013 [3]. Known as the SIMON and SPECK family of algorithms, SIMON was optimized for hardware and SPECK for software. Both families have been proposed as the standard for lightweight algorithms, although more testing is needed to prove their security.

## 2.3 Prior Work on Encryption for Space

Quantum Cryptography is a field of cryptographic research that is being investigated for use in satellites. Hughes *et al.* [14] investigated the technique of using Quantum cryptography to securely generate keys for use on either ground station/satellite communications or satellite/satellite communications. Their tests were successful. Rarity, Tapster, Gorman and Knight [23] also investigated the possibility of using Quantum Cryptography to create a secure key exchange technique between a ground station and a satellite, and their work suggests that there are no technical obstacles to building such a system, within their technical specifications.

In 2011, Challa, Bhat and Mcnair [4] proposed a security solution for CubeSats called CubeSec and GndSec. Their proposed scheme involved using AES and DES encryption operating in Galois/Counter mode on hardware that supports AES and DES encryption. Specifically, they tested their algorithm on an ATXMega128 microcontroller, and achieved throughput of between 43 KBps and 256Kbps depending on the configuration. However, they did not investigate a solution based solely on software implementation.

As mentioned previously, Muhaya investigated the possibility of combining AES with a chaotic encryption scheme [21]. In his work, Muhaya looked at the possibility of using an Arnold Cat Map [22] to shuffle pixel values, and using a Chaotic Henon Map [7] to generate a random sequence of key values for the AES algorithm. However, while the results listed were promising, no analysis was performed looking at either the performance or security of AES or the chaotic algorithms operating separately.

Encryption techniques used in space are well-served by making use of ongoing development of terrestrial protocols and advancements. Attackers are, similarly, able to make use of Earth-based advances in protocol-cracking techniques, driving a need for spacecraft developers to stay current. Advances may come from areas such as sensor networks [18], the encryption of specific types of data (such as images [9, 16]) or supporting technologies (such as seed generators [2]).

# 3 Proposed Technique

This section presents the proposed novel algorithm, based on the prior work of [13]. First, a general overview is presented. Then, specific elements of the approach are discussed, including the use of the Lorenz system, key generation, diagonal and anti-diagonal permutation, and block-based diffusion.

## 3.1 Overview

The first step of the proposed algorithm, which is based on prior work [13], is to read the data into an $n \times n$ matrix. The data is then permuted along the diagonal and anti-diagonal lines. This randomizes the location of each pixel throughout the matrix. Then, block-based diffusion is performed on the matrix. As adjacent pixels normally have a high correlation with their neighbors, block-based diffusion helps remove this correlation. A general diagram of the novel algorithm is shown below.

## 3.2 Lorenz System and Key Generation

The algorithm uses the Lorenz system of equations (see [6] for a more detailed discussion of the Lorenz system) to generate several values used in both the Diagonal/Anti-diagonal permutation step and the Block-based diffusion step. The Lorenz system of equations is as shown below.

$$\begin{aligned} \dot{x} &= m(y-x)+u_1 \\ \dot{y} &= rx-y-xz+u_2 \\ \dot{z} &= xy-bz+u_3 \end{aligned} \tag{1}$$

Where $m, r$, and $b$ are constants and $x, y$, and $z$ are initial values. For values $m = 10$, $r = 28$, $b = 8/3$, and $u_1 = u_2 = u_3 = 0$, the Lorenz system is in a chaotic state, which in essence means that it will diverge from another Lorenz System with similar starting values. In other words, given a set of $\{x_1, y_1, z_1\}$ similar to $\{x, y, z\}$, the two systems will soon look completely different. Note that in [13], they use a Time-delay Lorenz System, as seen below:

$$\begin{aligned} \dot{x} &= m(y-x)+u_1 \\ \dot{y} &= rx-y-xz+u_2 \\ \dot{z} &= xy-bz(t-)+u_3 \end{aligned} \tag{2}$$

For the sake of simplicity, in this paper the algorithm uses the regular set of Lorenz Equations shown in Equation (1). Using this simpler set of equations preserves the essence of the original algorithm while making it easier for the reader to understand. It also offers prospective speed benefits, as well as making the implementation and testing of the proposed algorithm simpler.

Given the Lorenz system, initial values $x, y$, and $z$ are chosen (as discussed in [13]). These are the secret key set. As mentioned above, given enough time, the Lorenz system will diverge from Lorenz systems with similar starting values. Hence, the system of equations is iterated some number $p$ times to preserve this quality and to give sufficient time for the system of equations to diverge. In [13], $p$ was chosen to be equal to 30. In this paper, $p$ was similarly chosen to be 30. Next, the system is iterated an additional $n$ amount of times, where $n$ is equal to the length of one side of the $n \times n$ data matrix. This generates the values used in the Diagonal/Anti-diagonal permutation step and the Block-based diffusion step.

Each $x, y$, and $z$ value set generated is stored in an array with each other $x, y$, or $z$ value. At the end of this process, the algorithm has generated three arrays of length $n$, each array storing either the $x$ values, the $y$ values, or the $z$ values. As each $x, y$, and $z$ values in a given successive sequence are generally increasing or decreasing, each $x, y$, and $z$ value are processed via the equation:

$$m = abs(m * 10^3) - floor[abs(m * 10^3)] \tag{3}$$

Where abs(a) returns the absolute value of a, and floor(b) returns the nearest integer less than or equal to m. Note that this differs from the original equation found in [6], which is:

$$m = abs(m * 10^{14}) - floor[abs(m * 10^{14})]. \tag{4}$$

At this point, as was done in [13], the array that holds the x values are copied into a second array and sorted.

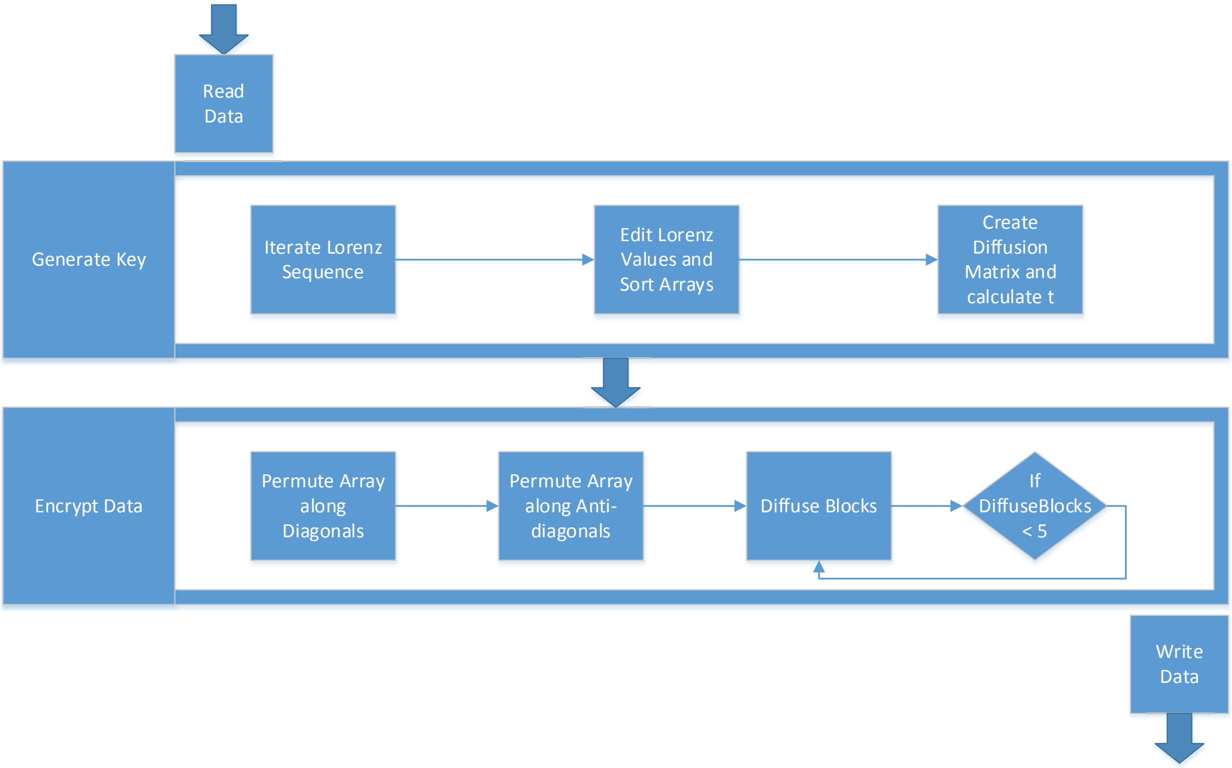

Read Data
Generate Key
Iterate Lorenz Sequence
Edit Lorenz Values and Sort Arrays
Create Diffusion Matrix and calculate t
Encrypt Data
Permute Array along Diagonals
Permute Array along Anti-diagonals
Diffuse Blocks
If DiffuseBlocks < 5
Write Data

Figure 1: A general outline of the steps taken during the encrypting of data

This then produces two arrays, one with the $x$ values in their original positions, and one with the $x$ values in sorted order. Both of these arrays will be used in the diagonal/anti-diagonal permutation steps. Similarly, the y values are processed so that there are now two arrays, one which holds the y values in their original position and one which holds the y values in sorted order.

### 3.3 Diagonal and Anti-diagonal Permutation

The algorithm, as was performed in [13], now starts the process of permuting the matrix of data along the diagonal and anti-diagonal lines. First, it is important to note that each diagonal line in the matrix has a different number of elements. For example, the diagonal which reaches from the top left corner down to the bottom right corner holds $n$ elements, while the diagonal to the right holds $n-1$ elements. This can be fixed by patching each diagonal with another diagonal so that, together, they have $n$ elements. In the above example, the diagonal with $n-1$ elements is patched with either the element in the top right or bottom left corner (which is in a diagonal of 1). Thus, together they make a diagonal with $n$ elements. By doing this to each diagonal, it is now possible to permute each diagonal symmetrically.

To permute the diagonal line, the proposed algorithm uses the two arrays of $x$ values (sorted and unsorted) produced in the key generation step. Consider a given element in the unsorted array. There is a similar element in the sorted array, however, the position is different. By looking at the difference in positions, it is possible to generate a key that will permute a given diagonal. Consider the case where each pixel on a given diagonal is mapped to an element in the unsorted array of $x$ values. For example, the first pixel is mapped to the first element of the unsorted array of $x$ values, the second pixel is mapped to the second element, and so on. With this mapping complete, it becomes possible to permute the pixels by simply matching them with the position of their assigned value of $x$ in the array of sorted $x$ values. Permutation in the anti-diagonal direction is performed similarly, except the arrays of sorted and unsorted y values are used.

During the block-based diffusion step, the original values of each pixel are modified using the diffusion matrix calculated from the values iterated from the Lorenz sequence. The goal of these steps are to harden the data against known plaintext attacks, as well as reduce correlation between pixels that were adjacent in the source image [13].

### 3.4 Block Based Diffusion

This step, based on [13], begins by breaking the $n \times n$ matrix into two matrices of size $n/2$ by $n$, which will be called matrices A and B. If $n$ is an odd number, the

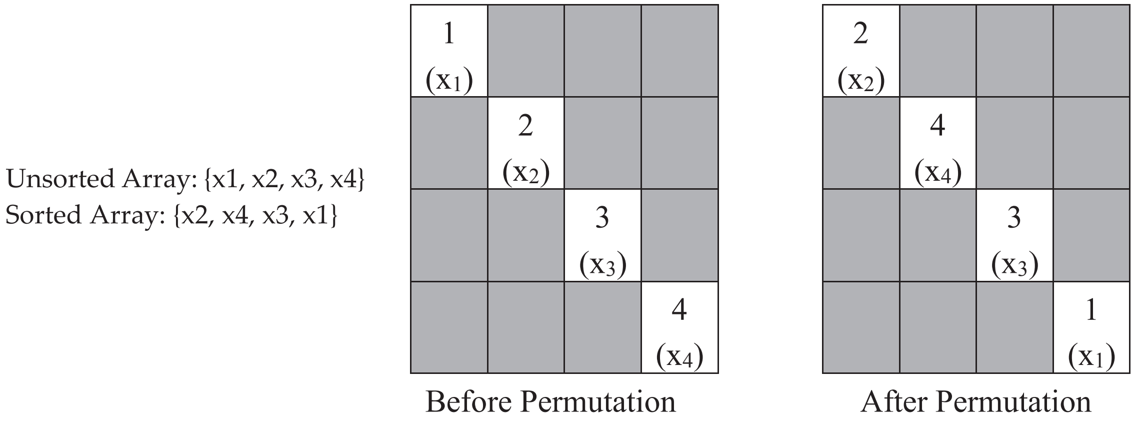


Figure 2: An example of performing permutations along a diagonal line [17]

data will need to be padded by adding one more row to make $n$ even. Next, $n^2/2$ values must be selected from the arrays of X, Y, and Z values randomly. These will form a diffusion matrix which will be applied to $a$. The diffusion matrix will be called D. It is important to note that for $n > 6$, more values must be selected than there are unique elements in X, Y, and Z. In addition to creating the diffusion matrix, a value t is calculated:

$$\begin{aligned} t = (\Sigma B) \mod (\frac{n}{2}) + (\Sigma B) \mod (n) \\ +(\Sigma B) \mod (\frac{n^2}{2}) + 1. \end{aligned} \tag{5}$$

where $\Sigma$ B is the sum of the elements in B, and (a) mod (b) applies the modulo(b) operator to $a$. The matrix A is then diffused via the following equation, where G is the resulting matrix and $G_{i,j}$ is a position in matrix G:

$$G_{i,j} = (A_{i,j} + (t * W_{i,j})) \mod (256). \tag{6}$$

The matrix B is diffused with the following equation, where H is the resulting matrix and $H_{i,j}$ is a position in matrix H:

$$H_{i,j} = (B_{i,j} + G_{i,j}) \mod (256). \tag{7}$$

While the algorithm discussed in [13] was originally created to encrypt pictures, and hence deals with pixels and pixel values the proposed algorithm has been extended to support any type of data that can be placed in an $n \times n$ matrix. This makes it possible to use the algorithm as a general block-based cipher, and allows for easier comparison with common algorithms.

## 4 Experimental Methods and Results

The process of testing the proposed cryptographic algorithm and characterizing its performance relative to other, more common algorithms is now presented. The goal of this work is to determine what techniques would be well-suited for use in the domain of small satellites. As such, there are several constraints that limit possible solutions, which were discussed previously. It is important that algorithms not be inefficient in terms of processing time or overall data size. Thus, algorithms that double or triple the size of the ciphertext, as compared to the plaintext, would be inefficient in this environment.

Two other cryptographic algorithms were chosen to compare the proposed algorithm to. The first algorithm chosen was the Advanced Encryption Standard (AES). Specifically, AES operating in Electronic Codebook (ECB) mode was chosen to compare against the novel algorithm for several reasons. Firstly, being a block cipher with a block size of 16 bytes, it is easy to compare against the proposed algorithm. In addition, it is a common protocol for terrestrial computing and arguably could be a first choice for use on most platforms, including small satellites. ECB mode of operation was chosen because it appeared to operate most closely to the proposed algorithm, and thus made for balanced testing. Note that AES operating in ECB mode of operation is known to be insecure [6]. However, similar techniques to those used in the other forms of AES could prospectively be used (with similar performance and other impacts) to increase the security of the proposed algorithm.

The other algorithm that was chosen was SPECK. SPECK is a relatively new lightweight encryption algorithm that was created and proposed by the National Security Agency. While new, it appears to be secure, and may be a seen as a possible standard Lightweight Encryption approach [6]. Specifically, Speck (128/128) was chosen for testing, for several reasons. Firstly, SPECK (128/128) is also a block cipher with a block size of 16 bytes. In addition, lightweight algorithms are increasing in popularity and also have the potential to be used in the domain of small satellites. Hence, it is important to consider how the novel algorithm compares to a lightweight algorithm.

The tests were run on a Raspberry Pi Model B. The code of the proposed algorithm, as well as the cryptographic algorithms with which the proposed algorithm was compared to, were written in the C++ programming language. AES code was implemented with the use of the Crypto++ Library version 5.6.2.

### 4.1 Experimental Setup

This section provides an overview of the configuration and design choices made in implementing the experimental design. Several areas are now discussed.

In the block-based diffusion step, a number of elements iterated from the Lorenz System must be chosen to fill a diffusion matrix. In the implementation presented in this paper, this was achieved by using a pseudo-random number generator with a seed based on the current time. To generate the seed, both the encryption and decryption algorithms have a file that stores the current seed. When the algorithm encrypts a file, it encrypts the current time as well and sends it with the ciphertext. It also stores this new value in the local file, overwriting the previous seed. When the decryption program decrypts the ciphertext, it then reads the new random number seed that was sent and overwrites the old seed. Both algorithms use the last known random seed when encrypting and decrypting, and with each transmission creates a new seed to use. It is important to note that by using this approach, the seed is overwritten with each transmission. Hence, if one transmission is not sent or received properly, it may create a scenario where all future transmissions are unrecoverable. This is an issue that will need to be addressed to facilitate the use of this algorithm in a real-world environment.

Additionally, the internal clock on a Raspberry Pi resets every time the system restarts. Even if the system is designed to never shut down, unexpected restarts may happen. This means that random number seeds may not be unique, given enough time. One potential solution to this is to set the current time from an external source, such as a GPS, or to send an updated time via the ground station during transmission, minimizing the amount of time that the Pi is running off of an old time.

### 4.2 Results

In this paper, six files were encrypted and decrypted to test the throughput of each algorithm. One text file, three JPEG pictures of varying size, and two QuickTime movie files of different size were used for testing. The picture and movie files were sources from NASA imagery and would be typical of images that could be transmitted from a spacecraft. The text file contains a short paragraph. This sort of a text file might be generated after the satellite runs a debug or system check. While it is uncommon for small satellites to transmit movies (due to bandwidth and communication window limitations), as small spacecraft capabilities advance it is not impossible for this to occur in the near future.

The file sizes of the test data are as follows. The text file was 67,655 bytes. The three picture files had sizes of 148,497 bytes (small) 409,306 bytes (medium) and 538,156 bytes (large). The small movie was 5,493,374 bytes, and the large movie was 9,689,032 bytes. Each file was encrypted and decrypted a total of 150 times, in groups of 50 operations of encryption and decryption at a time. In addition, the initial key values used in our novel algorithm testing are the same values used in [13]: x=-0.175, y=0.216, and z=-0.811. Multiple block sizes of the proposed algorithm were tested. Namely, $n = 4$, 8, 16, 32, 64, 128, and 256, where $n$ is one side of a square matrix. Block sizes for the algorithm are respectively 16, 64, 256, 1024, 4096, 16384, and 65536 bytes. Results from these experiments are presented in Tables 1 to 4.

As this testing occurred on mission-realistic hardware (and the Raspberry Pi units are single core computers, meaning that system processes could impair a given test), noise was introduced into the data set. For this reason, both the mean (which is more impacted by the interference, but would be an accurate measure of performing multiple encryptions/decryptions over time) and the median values (which provide a better view of the single file encryption cost) are presented. For sake of consistency, the median values are used as the comparison metric.

## 5 Analysis of Results

As shown in Tables 1 and 3, AES encrypted and decrypted the files faster than both SPECK and the proposed algorithm. The processing times for SPECK and the proposed algorithm are comparable, however SPECK performed better than the proposed algorithm across all scenarios. It is important to note, however, that the Raspberry Pi has built-in hardware support for AES which the Crypto++ Library has been built to take advantage of this feature [8]. Previous testing [24] has shown that (hardware aided) AES-NI is several times faster than AES encryption without the hardware optimization.

Note that times were also compared where SPECK and the novel algorithm were both added to AES encryption. While some nominal overhead is expected, these values demonstrate to how the algorithms would perform if they were combined with AES (mirroring the work done by [21]). While SPECK + AES was faster than the proposed algorithm combined with AES by a reasonable margin, it is possible that, with further optimization, the proposed algorithm combined with AES could produce similar result times. This will serve as a prospective topic for future work.

A comparison of the performance of the algorithms under the various conditions also produces data of some interest. Considering Table 2, for example, it is clear that encryption times for the Test file vary, with the 256 and 1025 byte block size encrypting in the least amount of

Table 1: Mean and media, showing comparison between encryption times of AES, SPECK, and the novel algorithm with a block size 16 bytes. Times shown are in milliseconds.

| | | Text File | Small Picture | Medium Picture | Large Picture | Small Movie | Large Movie |
|---|---|---|---|---|---|---|---|
| AES | Mean | 122.21 | 155.67 | 208.36 | 218.67 | 1896.95 | 3369.14 |
| | Median | 26.44 | 59.07 | 167.11 | 192.11 | 1820.96 | 3307.72 |
| SPECK | Mean | 133.48 | 229.60 | 563.94 | 768.55 | 7565.59 | 13180.98 |
| | Median | 92.53 | 199.89 | 545.28 | 717.17 | 7558.81 | 13131.20 |
| 16 bytes | Mean | 247.08 | 321.14 | 721.23 | 938.66 | 9376.04 | 16380.62 |
| | Median | 116.39 | 249.53 | 677.43 | 890.02 | 9310.50 | 16276.60 |
| AES + SPECK | | 118.96 | 258.96 | 712.40 | 909.27 | 9379.77 | 16438.92 |
| AES + 16 bytes | | 142.83 | 308.59 | 844.54 | 1082.13 | 11131.46 | 19584.32 |

Table 2: Mean and median times of encryption for each file, showing data from the novel algorithm for block sizes of 16, 64, 256, 1024, 4096, 16384, and 65536 bytes. Times shown are in milliseconds.

| | | Text File | Small Picture | Medium Picture | Large Picture | Small Movie | Large Movie |
|---|---|---|---|---|---|---|---|
| 16 bytes | Mean | 247.08 | 321.14 | 721.23 | 938.66 | 9376.04 | 16380.62 |
| | Median | 116.39 | 249.53 | 677.43 | 890.02 | 9310.50 | 16276.60 |
| 64 bytes | Mean | 235.02 | 280.08 | 630.58 | 820.12 | 8162.89 | 14142.78 |
| | Median | 181.70 | 214.81 | 584.57 | 768.05 | 8079.60 | 14058.80 |
| 256 bytes | Mean | 294.86 | 376.05 | 615.42 | 770.55 | 7944.10 | 13910.27 |
| | Median | 94.40 | 202.35 | 549.64 | 721.93 | 7707.64 | 13622.30 |
| 1024 bytes | Mean | 216.49 | 261.31 | 588.56 | 771.28 | 7735.40 | 13503.79 |
| | Median | 96.97 | 206.13 | 556.55 | 730.85 | 7686.17 | 13365.95 |
| 4096 bytes | Mean | 611.32 | 396.03 | 641.75 | 811.43 | 7726.50 | 13618.76 |
| | Median | 586.87 | 205.07 | 543.04 | 714.36 | 7490.01 | 13251.30 |
| 16384 bytes | Mean | 524.18 | 539.65 | 590.94 | 780.36 | 8044.89 | 14309.31 |
| | Median | 551.14 | 343.87 | 555.32 | 730.54 | 7804.75 | 14235.70 |
| 65536 bytes | Mean | 370.34 | 475.52 | 961.50 | 1217.91 | 10073.63 | 14139.14 |
| | Median | 294.81 | 306.69 | 679.57 | 865.36 | 8117.33 | 14030.40 |

Table 3: Comparison between the mean and median decryption time data results between AES, SPECK, and the novel algorithm with a block size of 16 bytes. Times shown are in milliseconds.

| | | Text File | Small Picture | Medium Picture | Large Picture | Small Movie | Large Movie |
|---|---|---|---|---|---|---|---|
| AES | Mean | 103.43 | 141.83 | 183.75 | 196.29 | 1800.80 | 3171.62 |
| | Median | 24.46 | 51.70 | 137.42 | 170.73 | 1775.65 | 3093.66 |
| SPECK | Mean | 122.53 | 225.49 | 573.42 | 748.60 | 7624.45 | 13264.09 |
| | Median | 92.25 | 201.15 | 548.90 | 722.06 | 7608.54 | 13212.20 |
| 16 bytes | Mean | 230.21 | 326.44 | 770.75 | 990.79 | 10051.66 | 17379.17 |
| | Median | 123.84 | 265.01 | 718.77 | 945.63 | 9938.16 | 17243.20 |
| AES + SPECK | | 116.70 | 252.85 | 686.32 | 892.79 | 9384.18 | 16305.86 |
| AES + 16 bytes | | 148.30 | 316.71 | 856.19 | 1116.36 | 11713.80 | 20336.86 |

Table 4: Comparison between the mean and median decryption time data results for the novel algorithm for block sizes of 16, 64, 256, 1024, 4096, 16384, and 65536 bytes. Times shown are in milliseconds.

| | | Text File | Small Picture | Medium Picture | Large Picture | Small Movie | Large Movie |
|---|---|---|---|---|---|---|---|
| 16 bytes | Mean | 230.21 | 161.50 | 770.75 | 990.79 | 10051.66 | 17379.17 |
| | Median | 123.84 | 265.01 | 718.77 | 945.63 | 9938.16 | 17243.20 |
| 64 bytes | Mean | 228.58 | 309.22 | 676.09 | 880.58 | 8695.97 | 15322.08 |
| | Median | 107.81 | 232.37 | 628.41 | 824.79 | 8656.10 | 15104.05 |
| 256 bytes | Mean | 228.81 | 400.91 | 680.78 | 820.42 | 8457.53 | 14827.25 |
| | Median | 101.27 | 218.52 | 591.57 | 777.03 | 8119.46 | 14551.20 |
| 1024 bytes | Mean | 213.35 | 309.07 | 636.41 | 851.08 | 8328.38 | 14533.98 |
| | Median | 105.23 | 222.94 | 600.10 | 788.16 | 8269.76 | 14382.65 |
| 4096 bytes | Mean | 482.16 | 479.97 | 672.05 | 843.53 | 8427.94 | 14958.49 |
| | Median | 105.81 | 222.31 | 588.58 | 776.07 | 8064.88 | 14661.90 |
| 16384 bytes | Mean | 512.90 | 599.42 | 687.79 | 829.91 | 8575.43 | 15253.62 |
| | Median | 586.59 | 463.81 | 600.59 | 788.52 | 8162.46 | 14906.35 |
| 65536 bytes | Mean | 350.06 | 486.52 | 1034.06 | 1315.57 | 10737.79 | 15190.99 |
| | Median | 286.71 | 328.53 | 727.43 | 930.95 | 8681.52 | 15062.95 |

time. For the small picture, the 256, 1025, and 4096 block sizes are all comparable, and all slightly faster than the 16 byte block size. For larger file sizes, the larger block sizes appear to function faster than the smaller block sizes. The medium picture was best encrypted by the 256, 1024, 4096, and 16384 block sizes, with the 64 block size only marginally slower. The same trend occurs when looking at the large picture. The small picture was best encrypted by the 4096 block size, with 256, 1024, and 16384 slightly slower. It is important to note that, as we look at these large file sizes, the 16 byte encryption algorithm is performing the slowest. Finally, the large movie was best encrypted by the 4096 block size, with 256 and 1024 block sizes performing slightly slower.

Reviewing the data in Table 4 shows that the results for the text file and the small picture are similar, with block sizes of 64, 256, 1024, and 4096 all performing best and within very close limits to one another, with 16 block encryption close behind. The medium picture is also similar, with 245, 1024, and 4096 performing optimally and similarly. However, the 16384 block size performs similarly to the other three, and the 64 block size is operating slightly slower. For the large picture, again, block sizes of 256, 1024, 4096, and 16384 perform best. The small movie results are similar to the large picture. Finally, the large movie was best encrypted by block sizes of 256, 1024, and 4096, with 64, 16384 and 65536 falling not unbelievably far behind.

The foregoing demonstrates that, in addition to the selection of an algorithm for use on the spacecraft, it may be necessary to use multiple block sizes. Alternately, system developers might choose to project the types of data that will be sent in order to optimize the block size selected.

## 6 Conclusions and Future Work

The work performed has shown that, in the context of the use of a single algorithm on the Raspberry Pi hardware, AES seems to be a better choice for encrypting and decrypting data on small satellites than the proposed algorithm when hardware optimization is present. Further testing is needed to consider how the proposed algorithm compares to AES in a non-optimized environment. The potential of implementing hardware enhancement for the proposed algorithm could also be considered.

In addition, while the SPECK algorithm implementation performed better than the 16 byte block size version of the proposed algorithm, the results are similar enough that further optimization the proposed algorithm could potentially increase the speed to the point of matching or outperforming SPECK. The potential to hardware optimize both would also bear consideration.

Finally, this paper has demonstrated the importance of block size selection for optimization of the performance of the proposed algorithm. It has provided data that may aid block size selection, based on the size of the data being encrypted and decrypted.

## Acknowledgements

This research was funded by the U.S. National Science Foundation (NSF Award # 1359224) with support from the U.S. Department of Defense. This paper builds on prior work presented in [17].

## References

[1] L. R. Abramowitz, "US air force SMC/XR sense nanoSat program," in *Aiaa Space 2011 Conference*

*& Exposition*, 2011. (https://doi.org/10.2514/6.2011-7332)

[2] Y. Asimi, A. Amghar, A. Asimi and Y. Sadqi, "New random generator of a safe cryptographic salt per session," *International Journal of Network Security*, pp. 445-453, 2016.

[3] R. Beaulieu, D. Shors, J. Smith, S. Treatman-Clark, B. Weeks and L. Wingers, "The SIMON and SPECK families of lightweight block ciphers," *IACR Cryptology ePrint Archive*, pp. 404, 2013.

[4] O. Challa, G. Bhat and J. Mcnair, "CubeSec and GndSec: A lightweight security solution for CubeSat communications," in *Small Satellite Conference*, pp. 8, 2012.

[5] G. Chen, Y. Mao and C. K. Chui, "A symmetric image encryption scheme based on 3D chaotic cat maps," *Chaos, Solitons & Fractals*, pp. 749-761, 2004.

[6] J. H. Curry, "A generalized lorenz system," *Communications in Mathematical Physics*, vol. 60, no. 3, pp. 193-204, 1978.

[7] P. Cvitanovi, G. H. Gunaratne and I. Procaccia, "Topological and metric properties of hnon-type strange attractors," *Physical Review A*, pp. 1503, 1988.

[8] W. Dai, *Crypto++ Library 5.6.2*, May 25, 2018. (http://www.cryptopp.com/)

[9] N. El-Fishawy and O. M. A. Zaid, "Quality of encryption measurement of bitmap images with RC6, MRC6 and rijndael block cipher algorithms," *International Journal of Network Security*, pp. 241-251, 2007.

[10] D. S. A. Elminaam, H. M. A. Kader and M. M. Hadhound, "Evaluating the performance of symmetric encryption algorithms," *International Journal of Network Security*, pp. 213-219, 2010.

[11] European Space Ahgency, *CubeSats - Fly Your Satellite!* May 25, 2018. (https://www.esa.int/Education/CubeSats_-_Fly_Your_Satellite)

[12] P. Fortescue, G. Swinerd and J. Stark, "Spacecraft systems engineering, 4th edition," *Wiley*, pp. 724, 2011. ISBN: 978-0-470-75012-4.

[13] X. Huang, G. Ye and K. Wong, "Chaotic image encryption algorithm based on circulant operation," *Presented at Abstract and Applied Analysis*, vol. 2013, pp. 8, 2013.

[14] R. J. Hughes, W. T. Buttler, P. G. Kwiat, S. K. Lamoreuax, G. L. Morgan, J. E. Nordholt and C. G. Peterson, "Quantum cryptography for secure satellite communications," in *Proceedings of the IEEE Aerospace Conference*, 2000. (DOI: 10.1109/AERO.2000.879387)

[15] G. Hunyadi, J. Ganley, A. Peffer and M. Kumashiro, "The university nanosat program: An adaptable, responsive and realistic capability demonstration vehicle," in *Proceedings of the IEEE Aerospace Conference*, vol. 5, 2004.

[16] S. P. Indrakanti and P. S. Avadhani, "Permutation based image encryption technique," *International Journal of Computer Applications*, pp. 45-47, 2011.

[17] S. Jackson, S. Kerlin and J. Straub, "Implementing and testing a novel chaotic cryptosystem for use in small satellites," in *Proceedings of the ACM Conference on Computer and Communications Security*, pp. 1638-1640, 2015.

[18] C. H. Ling, C. C. Lee, C. C. Yanh and M. S. Hwang, "A secure and efficient one-time password authentication scheme for WSN," *International Journal of Network Security*, 2017. (DOI: 10.6633/IJNS.201703.19(2).02)

[19] D. S. A. Minaam, H. M. Abdul-Kader and M. M. Hadhound, "Evaluating the effects of symmetric cryptography algorithms on power consumption for different data types," *International Journal of Network Security*, pp. 78-87, 2010.

[20] J. Mueller, R. Hofer and J. Ziemer, "Survey of propulsion technologies applicable to cube-Sats," in *NASA Techical Reports Server*, 2010. hdl:2014/41627.

[21] F. T. B. Muhaya, "Chaotic and AES cryptosystem for satellite imagery," *Telecommunication Systems*, pp. 573-581, 2013.

[22] G. Peterson, "Arnold cat map," *Math45-Linear Algebra*, 1997. (http://pages.physics.cornell.edu/~sethna/teaching/562_S03/HW/pset02_dir/catmap.pdf)

[23] J. Rarity, P. Tapster, P. Gorman and P. Knight, "Ground to satellite secure key exchange using quantum cryptography," *New Journal of Physics*, pp. 82, 2002.

[24] P. Schmid and A. Roos, "AES-NI performance analyzed; Limited To 32nm core i5 CPUs," *Conclusion*, 2010. (http://www.tomshardware.com/reviews/clarkdale-aes-ni-encryption,2538.html.)

[25] G. Skrobot and R. Coelho, "ELaNaeducational launch of nanosatellite: Providing routine RideShare opportunities," in *Presented at Proceeding SmallSat Conference*, 2012. (https://ntrs.nasa.gov/search.jsp?R=20120015542)

[26] D. Stinson, "Cryptography: Theory and Practice," *Cryptography & Coding Theory*, pp. 616, 2006.

[27] J. Straub and J. Vacek, "Escaping earths orbit but not earthly regulations: A discussion of the implications of ITAR, EAR, FCC regulations and title VII on interplanetary CubeSats and CubeSat programs," *Interplanetary CubeSat Workshop*, 2013. (https://icubesat.files.wordpress.com/2013/06/icubesat-org_2013-b-3-3-regulations_straub_201305291815.pdf)

[28] J. Straub, "CyberSecurity for aerospace autonomous systems," in *Presented at SPIE Defense Security*, 2015. (DOI:10.1117/12.2177959)

[29] D. Weeks, A. B. Marley and J. London III, "SMDC-ONE: An army nanosatellite technol-

ogy demonstration," in *Small Satellite Conference*, 2009. (https://digitalcommons.usu.edu/smallsat/2009/all2009/62/)

[30] J. R. Wertz, D. F. Everett and J. J. Puschell, "Space mission engineering: The new smad," *Amazon*, 2011. (https://www.amazon.co.uk/Space-Mission-Engineering-New-Smad/dp/1881883159)

# Biography

**Samuel Jackson** is an undergraduate student studying Computer Science at the Oklahoma State University in Stillwater, OK. In the summer of 2015, he participated in the small satellite research experience for undergraduates program at the University of North Dakota.

**Jeremy Straub** is an assistant professor in the Department of Computer Science at the North Dakota State University. He holds a PhD in Scientific Computing, an MS and an MBA and two BS degrees. He has published over 40 journal articles and over 120 full conference papers, in addition to making numerous other conference presentations. His research spans the gauntlet between technology, commercialization and technology policy. He serves as the associate director of the NDSU Institute for Cyber Security Education and Research.

**Scott Kerlin** graduated with an MS in Computer Science in 2010. He is currently Undergraduate Director for the University of North Dakota Computer Science Department. He specializes in programming, Computer Science education, cyber security and information assurance.

# A Survey of E-book Digital Right Management

Cheng-Yi Tsai[1], Cheng-Ying Yang[2], Iuon-Chang Lin[3], and Min-Shiang Hwang[1,4],

(Corresponding author: Min-Shiang Hwang)

Department of Computer Science and Information Engineering, Asia University[1]
No. 500, Lioufeng Raod, Wufeng Shiang, Taichung 41354, Taiwan, R.O.C.
Department of Computer Science, University of Taipei[2]
Taipei 10048, Taiwan, R.O.C.
Department of Management Information Systems, National Chung Hsing University[3]
145 Xingda Rd., South Dist., Taichung City 402, Taiwan
Department of Medical Research, China Medical University Hospital, China Medical University[4]
No.91, Hsueh-Shih Road, Taichung 40402, Taiwan
(Email: mshwang@asia.edu.tw)

(Invited Jan. 31, 2018)

## Abstract

In past two decade, with the mobile communication technology rises rapidly and as well as the popularity of smart phones, tablets and mobile devices. Human's daily life and habits had a completely changed. Information comes from with many different ways as before. People read paper books or newspapers to get information in the past. Shift to reading from electronic text files or Internet information. Therefore, making the demand for electronic books greatly increased. In this situation, traditional publishers forced to face huge impact and challenges. How to manage the digital rights of e-books effectively has become a serious issue. In this paper, we propose some key issues of of e-book digital copyright management for interesting researchers.

*Keywords: Blockchain; Digital Right Management; E-book*

## 1 Introduction

Digitization and Internet are the greatest impact on the text and books, that is, making books lose its own physical properties. A permanent solution to the book's issues like heavy weight, difficulty to carry, taking up space, and difficulty to maintain. Before the digital publication appears, the proportion of books that readers bring around is not much. With the advent of the digital age, a variety of traditional publications have been digitized and uploaded to the Internet for sales. Consumers can easily store digital publications in digital devices that they carry with them. Such convenience promotes the thriving digital industry.

However, as same as all digital content products, e-books can be copied without any loss of quality [5] and be distributed at costs close to zero [1, 9]. Like most traditional industries, the book industry generates most of its profit from direct books sales, not from advertising in Internet. Hence, illegal mass-scale low-price duplication and distribution of copied content should constitute serious issues, which are similar to the music, video and other digital content industries [10].

Since e-books can be copied and transmitted by anyone very easily and quickly, for consumers likely without the knowledge of the situation, they may become a victim of piracy and participants inadvertently. That also hurts authors, publishers and distributors. In reality, when consumers buy e-books, there is nothing that can be referred as the basis to confirm whether the purchased goods have legal copyrights. A large number of e-books are stored in mobile devices or home computers, so it also makes it difficult for copyright owners to trace and confirm. This situation also reduces the willingness of publishers and creators to issue e-books. Therefore, a reliable e-book digital rights management mechanism is needed to help the entire e-book industry.

## 2 Current Status of E-Book Publishing Industry

In the past 20 years, the rapid development of the Internet as well as the popularization of digital cloud application system has changed the consumer's reading behavior gradually and offered more different business opportunities and challenges to the publishing industry. E-books market has three aspects: Firstly, the content provider, including authors, publishers, authority, and digital right management. Secondly, the online service platform, including online shopping book stores, the pricing mechanisms of e-books. Thirdly, the e-books reading devices, including desktop computers, notebooks, smart phones and tablets.

Sony was the pioneer of the e-book reading device. It has started the development as early as 2004 and announced its first-generation e-book reader device in 2006. In 2007, Amazon introduced a device named Kindle, that enables users to browse, buy, download, and read e-books, newspapers, magazines and other digital media via networking. Amazon becomes a pioneer of reading device with e-books. After that, Apple presented iPad in 2010, selling e-books from his online i-Bookstore. In addition, Google of search engines has opened an online e-book store named e-Bookstore.

From the publishing industry's perspective, digitizing books is an inevitable trend. However, the e-book sales have not grown significantly, which is a fact to be faced. Publishers after years of trying to change and longtime observation of consumers have found that at the beginning publishers worry that e-books will strongly affect the sales of paper books and even accelerate the decline of the paper book, but later, publishers have generally agreed that e-books and paper books have their preferred user groups, respectively, instead of being regarded as two opposing groups. Regardless e-books or paper books, readers never disappear, but reading habits have changed. In diversified reading models which are so dispersed, the readers' reading behaviors can be still seen.

In the U.S. book market, the American Association of Publishers statistics report on 1,209 publishers in the nation pointed out that the overall net revenue from the book industry in the United States grew 4.1% (from $357.1 millions to$ 371.79 millions) from January to July 2014 compared with the same period in 2013, with "Adult Books" slumped 2.2% and "Religious Publications" increased by 1.9%. The best performing section was "Child and Juvenile Books", substantial increase by25.8%. The dazzling performance of both categories of e-books grew by 25.8% and 59%, respectively [3].

In the United Kingdom book market, according to the Publishers Association, total revenue from British books and journals amounted to 4.7 billion pounds in 2013, of which 29% (1.5 billion pounds) from digital services (including e-books, e-journals, etc.). Over the past five years, overall book sales (excluding journals) in the United Kingdom alone for physical books and e-books have risen 6% to 3.4 billion pounds, along with physical books sales down 6% while e-books sales Up 305%. Especially in the past two years, e-books sales have doubled to 590 million pounds, of which sales in textbooks and novels are the best, accounting for 42% and 39% of the total, respectively. In terms of periodicals, periodicals have sales of 1.3 billion pounds, of which 850 million pounds come from digital services. Although physical books sales decline, the growth of e-books makes up for the decline in physical books sales [7].

In terms of book publishing format, downloadable audio books are still the fastest growing data format. From January to July 2014 increased by 26.2% over the same period in 2013; followed by e-books, over up 7.5% over the same period of last year. Revenues from paperback books edged up 5.3% over the same period of last year, while hardback books declined 0.3% [3]. Overall, books sales in the United States increased slightly in the first half of this year, while the kinetic energy of growth mainly came from non-paper books such as audio books and e-books.

According to a survey of the U.S. market conducted by Pew Research in January 2014, the ratio of dedicated hand-held reading devices (tablet PC or e-reader) owned by American rose from 7% in 2010 to 43% in 2013, further to 50% by 2014. That means, about half of Americans have iPads Tablet PCs, Kindles, or Nook reader. In reading habits, 69% of the respondents in this survey indicated that they would read paper books, while the proportion of readers reading paper books did not changed much in the past three years. However, readers who read e-books have risen sharply from 17% in 2011 to 28% in 2014.With the popularity of dedicated reading devices, the proportion of readers of the aforementioned devices has also grown. The percentage of readers on e-readers has risen from 41% in 2011 to 57% in 2014;at the same period, reading on the tablet computers increased from 23% to 55% [11]. As of the 2016 survey, the percentage of readings on tablets and smartphones were15% and 13%, respectively, higher than 8% of dedicated reading devices. The reading rates of all reading devices increased. This means that e-book reading penetration continues to rise [6].

## 3 Digital Right Management Model in E-Books Publish Industry

There are many members to work together in an e-book supply chain, including authors, publishers, online book stores, book distributors, and reading device manufacturers. Although e-books have been around for more than two decades, at the very beginning, publishers do not take the e-book market seriously because there are no appropriate industrial support measures. The overall e-books industry is quite deserted. Until 2007, Amazon introduced its first-generation of Kindle e-book reader, consumers can buy an e-book from the Internet (Amazon.com) and read an e-book on the Kindle. Amazon has integrated the e-book sales platform with consumers' reading devices successfully. The first release of Kindles are sold out just in 5.5 hours [2]. This is not only a successful business model, but also a change in consumers' reading behavior.

Copyright management of the e-book industry includes author authorizations, publisher agreements, e-book formats, digital restrictions management (DRM), pricing, transaction service platforms, reading devices and programs. The overall e-book distribution processes can be

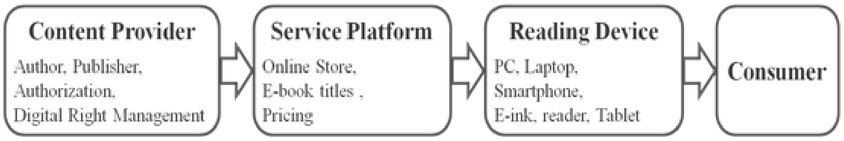


Figure 1: E-book publishing process

classified into three elements: content providers, service platforms, and reading devices (see Figure 1). At the end, consumers buy an e-book from an Internet service platform, and the book file is downloaded into a reader device. Then, e-books are read by reading software from a reader device. These three elements are the key factors influencing the success of e-book industry [4].

However, the above three components of the e-books industry, both in the industrial or technical environments, have been roughly mature. Only the digital rights management of e-books still lacks a mechanism that can be trusted by authors, publishers and distributors altogether. The distribution of e-books is different from that of paper books. Paper books have a specific number of printing, but the e-books only have an e-book digital file released. Each online sales platform holds an e-book copy file for sale. When consumers make purchases through the online sales platform, the e-book digital file is again copied and downloaded to the consumer's storage device. When the original consumers want to transfer their holdings of e-books, the e-book file is just copied again to another user to complete the transaction (see Figure 2). However, it is hard to grasp whether the original consumers have actually deleted the e-book files. Although e-books can be quickly copied and transferred, for the e-book digital rights holders, it has the difficulty to manage copyrights of e-books. From authors to consumers throughout the industry cycle, no matter which one of the roles, it is difficult to control the current circulation of the actual number of e-books on the market. Even online platform providers are difficult to confirm whether the sale of e-books has legal digital rights.

At present, there are mainly two models for the management mechanism of e-books digital right. The first model is online shopping and reading cloud management model. In this model, consumers buy e-books on online sales platform, but will not download e-book files. Consumers obtain access to the online reading of the e-book, and then dealers confirm users' identity and read permission for specific e-books by managing users' accounts. The advantage of this type of management is that the overall number of e-book files does not increase with the number of sales. Consumers can read through any Internet-enabled device. The disadvantage is that consumers can't read e-books offline, so needing technical reference properties of books will cause inconvenience to users. The dealer actually manages the number of accounts rather than the actual number of readers (see Figure 3). It may happen that many people share an account. For the author of e-books, this management model can't effectively get the actual number of sales, unless the dealer returns the sales data from time to time. For consumers, what they buy is e-books reading authorization rather than actually e-book files. Therefore, consumers will only agree to pay a lower fees rather than actually fees. When consumers no longer use the e-book, but the e-book can't be transferred to another people, this motivates consumers to move their accounts to others.

The second mode is to restrict the user's reading device management mode. In this model, consumers must install specific application software before purchasing or using e-books, and purchase an e-book from the online e-book platform by the application installed. And then thee-book's file will be downloaded to the user's reading device (see Figure 4). After the download of the e-book file has been completed, the installed application will encrypt the downloaded e-book file or add some tags to the downloaded e-book file. Users can only read e-books by the installed application on this particular device. When the file is transferred to another device, it cannot be read. The advantage of this mode is that the user can read an e-books offline and obtain the actual e-book digital file, which increases the willingness of the user to purchase and the selling price. The disadvantage is that the user can only read e-books on the particular device. When the consumer's devices is damaged, lost or replaced, it also loses all previously purchased e-books at the same time. For the author of e-books, this management model also can't effectively get the actual number of sales.

Even we do not consider the commercial interests, no matter what kind of choice from the above two management models, it is a challenge for e-books holders, users or consumers to make sure whether those e-books that they are holding, using or purchasing have a legal digital right or not. It is also difficult to help digital right holders (authors or publishers) try to verify whether other people's e-books have legally digital rights.

## 4 Digital Rights Manage Problems in E-Book Industry and Research Purposes

As mentioned in chapter 1.3 of this dissertation, the digital right management model of e-books industry lack a reliable mechanism that can validate the e-books held and used. If the publisher deliberately conceals the information, then the authors or the authorizers can't know the actual number of e-books published. Similarly, if a sales platform distributor deliberately conceals sales figures, it

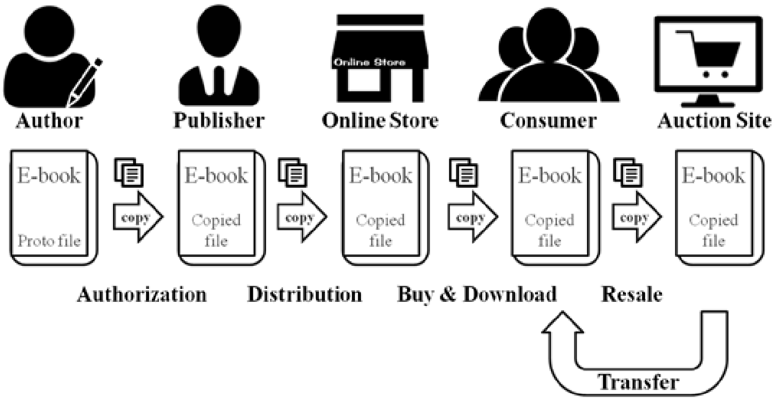


Figure 2: E-book industry product circulation diagram

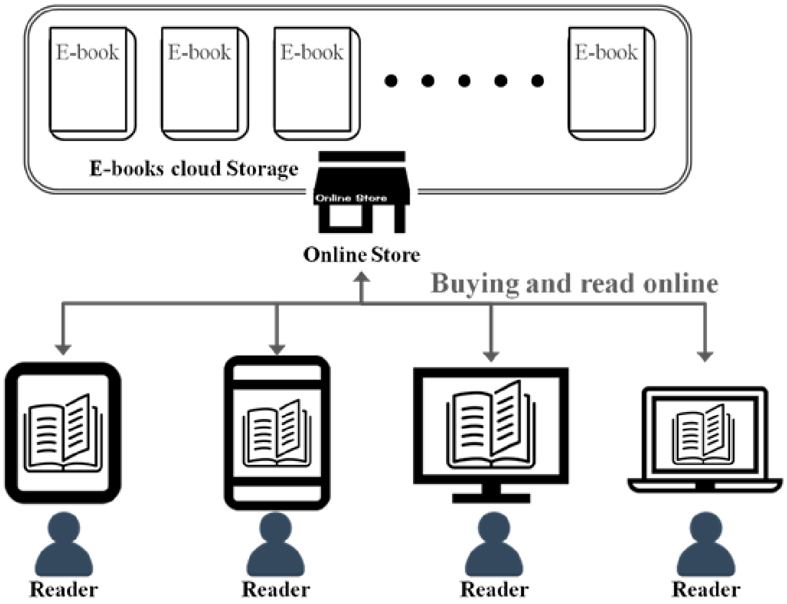


Figure 3: Buying and read online management mode

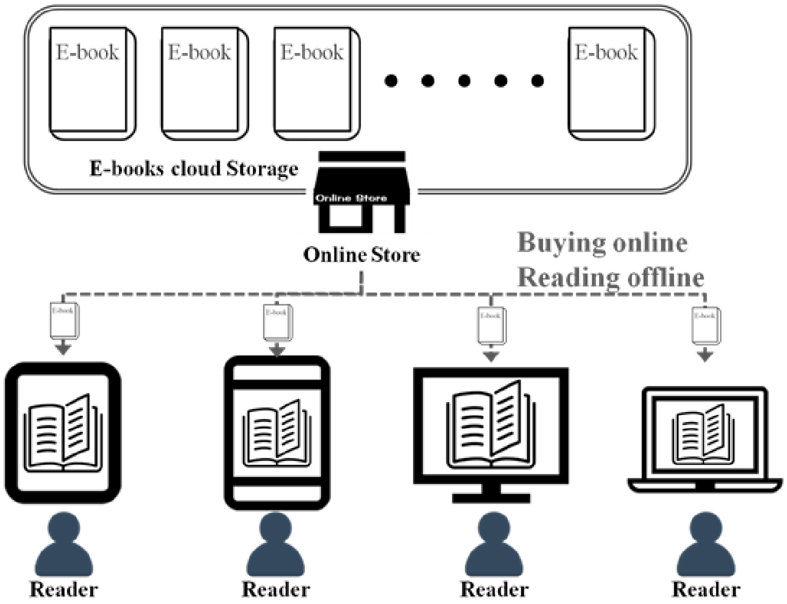


Figure 4: Buying online and read offline management mode

is difficult for the publisher to know the actual number of e-books sold. It is much more difficult for consumers to try identifying an e-book's digital right which they hold and used. The criminals can copy the files of e-books, sell for profit, or illegally distribute by this loophole. This situation has caused considerable economic losses to authors, publishers and sales platforms distributors, and consumers may become victims of piracy without their knowledge.

These problems not only cause economic damage to the e-book industry, but also drag down the growth of the whole industry and impede the dissemination of information. Because of such concerns, it will make authors, authorizers, and publishers reduce the willingness to put out or publish e-books, and create barriers for e-book consumers to access such information. For sales platform distributors as well as consumers, they may violate the law without expectation. Therefore, establishing an open and transparent e-books digital rights management platform is needed to provide e-book holders to confirm whether the e-books purchased or held have legal digital rights, and provide that authors, authorizers, and the publishers can quickly and easily obtain e-books circulation quantity. For the e-book industry, it has become an important project.

The current e-book sales model is: consumers must buy all chapters in a packaged model. In actual demand, consumers may only need some chapters or pages in an e-book. However, such a sales model isn't fair for consumers and would also affect consumers' willingness to buy e-books. At the same time, in the existing e-book management model, a lot number of e-books held by consumers cannot be re-sold or transferred. This has caused consumers' e-book assets to be frozen, and that is a big disadvantage compared with paper books. Therefore, the digital rights of the e-books, need to be refined when an e-book is being sold, so the consumer may only choose to purchase the required chapters or pages so that the consumers no longer need to pay for the unwanted content. Moreover, the consumers could transfer the digital rights of e-books to another consumer when they don't need them anymore. That should reduce the motivation for demand for e-book piracy and accelerate to the e-book industry growth.

In this research, I will base on the blockchain technology and present a model for managing e-book digital rights by creating a blockchain network. All e-books from publishing, sales, purchase to transfer processes, any participating users of this blockchain can use this e-book blockchain network to verify the digital rights of an e-book. To help every participant in the eBook industry, regardless selling, purchasing or holding, you can be guaranteed on the e-books copyright. Digital rights of e-books are also refined, allowing consumers to purchase only the chapters or pages that they actually need, and to permit the transfer of digital rights held in the e-books, reducing the user's motivation to violate the digital rights of the e-book, and making a reference for digital rights management of the e-book industry.

## 5 Key Issues of E-book Digital Right Management

In this section, we propose some key issues of of e-book digital right management for interesting researchers.

### 5.1 Establish a Digital Right Blockchain Model of E-books

The blockchain technology has high reliability for the storage, application, transparency, privacy and security of data, which is appropriate [12] for data recording, data verification, data tracking, and anonymity in business activities. There are five main features of blockchain technologies: Decentralization, transparency, independence, immutability and anonymity [8]. In this issue, The researchers need to apply blockchain technologies to establish an e-book digital right management model.

### 5.2 E-Books Digital Right Publishing and Trading

In today's e-book industry, the author completes an e-book writing, and authorizes the publishers to publish the e-book, which is uploaded by the publishers to the online sales platform. After the consumer pays for the purchase of an e-book via the online platform, he/she can download the digital file of the e-book and read it offline through a device installed with the e-book reading software. When the original purchase user is no longer in use, this e-book file may be resold, or copied directly to other people. Therefore, there is no proper way for readers to confirm the legitimacy of digital rights of e-books.

In another type of e-books digital right publishing process, after the consumer pays for the purchase of an e-book via the online platform. The online platform provides online reading permission for consumers to read online through an internet browser. When the original purchase user is no longer in use, due to the reading permission, this e-book can't be resold, resulting an idle asset. Otherwise, the user may tell other people his account and password, so the multiple people use the same account.

In this issue, the researchers need to develop a secure scheme to protect e-books digital right's publishing (generation), trading (transfer), and asset management in the e-book digital right blockchain.

### 5.3 E-Books Digital Right Refinement and Distribution Statistics

In the market of traditional paper-books, a paper-book is completely bound before sales. Therefore, either in a physical bookstore or an online bookstore is sold on the whole book. While consumers are accustomed to such selling patterns, it often happens that consumers only need some of the information in a book they buy, but they still have to buy a whole book. This sales model forces consumers to buy a lot of unwanted parts. On one hand it will reduce the willingness of consumers to buy, and on the other hand it also virtually increase the intent of piracy.

In this issue, the researchers need to design an efficient scheme to refine the digital right distribution, sale and transfer of e-books digital rights. The scheme also provides authors, authorizers or publishers the statistical data of e-books in the market circulation.

## 6 Conclusion

Digitization has brought many opportunities and shocks to traditional paper books. The tendency of digital publishing has gradually become a mainstream trend. Although there still have many paper-books lovers, the demand for undeniable e-books is constantly growing. How to establish a reliable digital right management mechanism for e-books is precisely the topic of concern to this research. In this paper, we have proposed some key issues of of e-book digital right management for interesting researchers.

## Acknowledgment

This research was partially supported by the Ministry Of Science and Technology, Taiwan (ROC), under contract no.: MOST 103-2632-E-324-001-MY3.

## References

[1] M. Bichler and C. Loebbecke, "Pricing strategies and technologies for on-line delivered content," *Journal of End User Computing*, vol. 12, no. 2, pp. 4–10, 2000.

[2] D. T. Clark, S. P. Goodwin, T. Samuelson and C. Coker, "A qualitative assessment of the Kindle E-book reader: Results from initial focus groups," *Perform Measure Metrics*, vol. 9, no. 2, pp. 118–129, 2008.

[3] N. Hoffelder, *AAP Reports eBooks, Audiobook Sales Up in First Part of 2014*, Oct. 28, 2014. (http://the-digital-reader.com/2014/10/28/aap-reports-ebook-audiobook-sales-first-part-2014/\#.VFn6KE8rjVg)

[4] C. C. Lina, W. C. Chiou, S. S. Huang, "The challenges facing E-book publishing industry in Taiwan," *Procedia Computer Science*, vol. 17, pp. 282–289, 2013.

[5] C. Loebbecke, P. Bartscher, T. Weiss, S. Weniger, "Consumers' attitudes to digital rights management (DRM) in the German trade ebook market," in *Ninth International Conference on Mobile Business and Ninth Global Mobility Roundtable (ICMB-GMR'10)*, 2010.

[6] A. Perrin, *Book Reading 2016*, Sept. 1, 2016. (http://www.pewinternet.org/2016/09/01/book-reading-2016/)

[7] The Publishers Association, *One Third of Publishers Revenues are Digital*, May 2, 2014. (https://www.publishers.org.uk/news/press-releases/2014/one-third-of-publishers-revenues-are-digital/)

[8] A. Savelyev, "Copyright in the blockchain era: Promises and challenges," *Computer Law & Security Review*, Available online 8 Dec. 2017.

[9] C. Shapiro and H. Varian, *Information Rules: A Strategic Guide to the Network Economy*, Boston: Harvard Business School Press, 1999.

[10] M. D. Smith and R. Telang, "Competing with free: The impact of movie broadcasts on DVD sales and internet piracy," *MIS Quarterly*, vol. 33, pp. 321–338, 2009.

[11] K. Zickuhr and L. Rainie, *E-Reading Rises as Device Ownership Jumps*, Jan. 16, 2014. (http://www.pewinternet.org/2014/01/16/e-reading-rises-as-device-ownership-jumps/)

[12] G. Zyskind, O. Nathan and A. Pentland, "Decentralizing privacy: Using blockchain to protect personal data," in *IEEE Security and Privacy Workshops*, San Jose, CA, pp. 180–184, 2015.

## Biography

**Cheng-Yi Tsai** received his B.S. degree from Department of Business Administration, Chaoyang University of Technology (CYUT), Taiwan in 2001 and M.S. degree from Computer Science & Information Engineering, Asia University, Taiwan in 2005 . He is currently pursuing the Ph.D. degree from Department of Computer Science and Information Engineering, Asia University, Taiwan. His research interests include blockchain, information security, and cloud computing.

**Cheng-Ying Yang** received the M.S. degree in Electronic Engineering from Monmouth University, New Jersey, in 1991, and Ph.D. degree from the University of Toledo, Ohio, in 1999. He is a member of IEEE Satellite & Space Communication Society. Currently, he is employed as an Associate Professor at Department of Computer Science, University of Taipei, Taiwan. His research interests are performance analysis of digital communication systems, error control coding, signal processing and computer security.

**Iuon-Chang Lin** received the B.S. in Computer and Information Sciences from Tung Hai University, Taichung, Taiwan, Republic of China, in 1998; the M.S. in Information Management from Chaoyang University of Technology, Taiwan, in 2000. He received his Ph.D. in Computer Science and Information Engineering in March 2004 from National Chung Cheng University, Chiayi, Taiwan. He is currently a professor of the Department of Management Information Systems, National Chung Hsing

University,and Department of Photonics and Communication Engineering, Asia University, Taichung, Taiwan, ROC. His current research interests include electronic commerce, information security, cryptography, and mobile communications.

**Min-Shiang Hwang** received the Ph.D. degree in computer and information science from the National Chiao Tung University, Taiwan in 1995. Dr. Hwang was the Chairman of the Department of Information Management, Chaoyang University of Technology (CYUT), Taiwan, during 1999-2003. He was a distinguished professor and Chairman of the Department of Management Information Systems, NCHU, during 2003-2011. He obtained the 1997, 1998, 1999, 2000, and 2001 Excellent Research Award of National Science Council (Taiwan). His current research interests include information security, cryptography, image compression, and mobile computing. Dr. Hwang has published over 200+ articles on the above research fields in international journals.

# Guide for Authors
## International Journal of Network Security

IJNS will be committed to the timely publication of very high-quality, peer-reviewed, original papers that advance the state-of-the art and applications of network security. Topics will include, but not be limited to, the following: Biometric Security, Communications and Networks Security, Cryptography, Database Security, Electronic Commerce Security, Multimedia Security, System Security, etc.

### 1. Submission Procedure

Authors are strongly encouraged to submit their papers electronically by using online manuscript submission at http://ijns.jalaxy.com.tw/.

### 2. General

Articles must be written in good English. Submission of an article implies that the work described has not been published previously, that it is not under consideration for publication elsewhere. It will not be published elsewhere in the same form, in English or in any other language, without the written consent of the Publisher.

### 2.1 Length Limitation:

All papers should be concisely written and be no longer than 30 double-spaced pages (12-point font, approximately 26 lines/page) including figures.

### 2.2 Title page

The title page should contain the article title, author(s) names and affiliations, address, an abstract not exceeding 100 words, and a list of three to five keywords.

### 2.3 Corresponding author

Clearly indicate who is willing to handle correspondence at all stages of refereeing and publication. Ensure that telephone and fax numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address.

### 2.4 References

References should be listed alphabetically, in the same way as follows:

For a paper in a journal: M. S. Hwang, C. C. Chang, and K. F. Hwang, ``An ElGamal-like cryptosystem for enciphering large messages," *IEEE Transactions on Knowledge and Data Engineering*, vol. 14, no. 2, pp. 445--446, 2002.

For a book: Dorothy E. R. Denning, *Cryptography and Data Security*. Massachusetts: Addison-Wesley, 1982.

For a paper in a proceeding: M. S. Hwang, C. C. Lee, and Y. L. Tang, ``Two simple batch verifying multiple digital signatures," in *The Third International Conference on Information and Communication Security (ICICS2001)*, pp. 13--16, Xian, China, 2001.

In text, references should be indicated by [number].

## Subscription Information

Individual subscriptions to IJNS are available at the annual rate of US$ 200.00 or NT 7,000 (Taiwan). The rate is US$1000.00 or NT 30,000 (Taiwan) for institutional subscriptions. Price includes surface postage, packing and handling charges worldwide. Please make your payment payable to "Jalaxy Technique Co., LTD." For detailed information, please refer to http://ijns.jalaxy.com.tw or Email to ijns.publishing@gmail.com.