Proceedings of the 5th International Cryptology and Information Security Conference 2016

31st May – 2nd June 2016
Kota Kinabalu, Malaysia

Published by
Institute for Mathematical Research (INSPEM)
Universiti Putra Malaysia
43400 UPM Serdang
Selangor Darul Ehsan

First Print 2016

Perpustakaan Negara Malaysia Cataloguing – in – Publication Data

Proceedings of the 5th International Cryptology and Information Security
(5th.: 2016: Kota Kinabalu)
CONFERENCE PROCEEDINGS CRYPTOLOGY 2016: Proceedings of The 5th International Cryptology and Information Security Conference 2016 31st May – 2nd June 2016, Kota Kinabalu, Sabah, MALAYSIA / Editors: Amir Hamzah Abd Ghafar, Goi Bok Min, Hailiza Kamarulhaili, Heng Swee Huay, Moesfa Soeheila Mohamad, Mohamad Rushdan Md. Said, Muhammad Rezal Kamel Ariffin.

ISBN 9789834406950
1. Cryptography- - Malaysia - - Congresses.
2. Computer security - - Malaysia- - Congresses.
I. Amir Hamzah Abd Ghafar II. Goi Bok Min. III. Hailiza Kamarulhaili. IV. Heng Swee Huay. V. Moesfa Soeheila Mohamad. VI. Mohamad Rushdan Md. Said. VII. Muhammad Rezal Kamel Ariffin.

Title.005.8072

OPENING REMARKS

First and foremost, I would like to thank the Malaysian Society for Cryptology Research (MSCR) in collaboration with CyberSecurity Malaysia together with universities: UMS, UPM, and USM, Sabah's Ministry of Resource Development & Information Technology and Sabah's State Computer Service Department for their continuous efforts and commitment to host this premier event, the International Cryptology and Information Security Conference for the fifth time. This biannual conference series which started in 2008 has been organized and hosted at several locations in Malaysia, beginning from Kuala Lumpur to Melaka, and then on to Langkawi and Putrajaya. This year Kota Kinabalu has been chosen as the venue.

Cryptology is an area of study and research which has numerous applications especially in the area of information and communication technology. The 5th International Cryptology and Information Security Conference 2016 (Cryptology2016) is one of the many steps towards enhancing and realizing research and applications of cryptology in Malaysia through research collaboration and discussion with research counterparts from the international arena. This is an open forum, provided for contributions which cover research on the theoretical foundations, applications and any related issues in cryptology, information security and other underlying technologies. Pertaining to these challenges MSCR was formed in 2007. Since then, collaborative efforts between Malaysian cryptography centres of excellence has been galvanised and are no longer working in silo. The ministry greatly appreciates these efforts, where research activities do not only generate significant impact towards the improvement of the body of cryptographic knowledge but also related to society needs, especially in the area of information security.

Once again MSCR, CyberSecurity Malaysia and the above mentioned local institutions of higher learning has put in a huge effort to bring together distinguished researchers from various countries as speakers and participants to share knowledge and novel ideas. We congratulate MSCR and its partners for tirelessly organizing beneficial and meaningful events that contribute to further strengthen Research and Development (R&D) in Malaysia. I would also like to take this opportunity to give special thanks to the conference corporate supporters namely Augmented Technology Sdn Bhd in supporting this event. Augmented Technology Sdn Bhd is one of the few companies in Malaysia that engage directly with the academia to enhance its competitiveness. This company has direct application of cryptography in their business products and services. I urge researchers attending this event to take this opportunity to enquire on various industrial related issues regarding cryptography. It is hoped that Augmented Technology will continue to support this conference in the future.

Our keynote and invited speakers are highly distinguished researchers and are from different parts of the world. They bring with them invaluable knowledge and experience. We hope that participants will take this opportunity to interact and benefit from this conference. We would like to express our sincere appreciation and gratitude to the speakers for accepting our invitation despite their busy schedule and contributing to this meaningful event.

I also note that, the participation of researchers from various cryptographic disciplines is impressive and this signifies the interdisciplinary nature of the topic of the conference. It is hoped that the network among researchers and institutions will grow and new research collaboration can be built. We hope that through these type of events and activities we can overcome the shortage of expertise in this area of research.

I am confident that this Conference would serve as a platform to further discuss the research collaboration and afford brilliant ideas in the cryptology and information security. To all our participants, I wish you a fruitful and productive Conference. To our international participants, I wish you would take some time to travel in this country, especially in Kota Kinabalu to experience and enjoy the rich Malaysian environment and warm hospitality.

Thank you.

DATUK SERI PANGLIMA MADIUS TANGAU
Minister of Science Technology and Innovation (MOSTI), MALAYSIA

WELCOMING NOTES

I am most pleased to welcome speakers from various countries across the world to the 5th International Cryptology and Information Security Conference 2016 (Cryptology2016). It is our hope that participants will grab this advantage and gain valuable experience either through formal dialogue or social interaction during this intellectual congregation in order to further understand cryptology and its applications.

Cryptography is an area of study and research that has numerous applications especially in the area of communication technology. In this respect, the conference will provide an avenue for participants to engage current topics related to research in cryptology. It is also aimed at promoting and encouraging exchange of ideas and at the same time identifying areas of collaborative research between local and foreign researchers.

Information security has never become as important in our daily lives as we are experiencing today. We are now on the brink of experiencing cryptography and its deployment in every corner of our day to day experiences. Thus, research in this area has become so important that without continuous research in the area one would not be able to ascertain the degree of security being deployed.

In this conference, we have organized 3 keynote speeches to be delivered by renowned researchers in their respective areas for the benefit of participants. Also, 24 papers are scheduled to be presented encompassing various areas of cryptology such as theoretical foundations, applications, network security and other underlying technologies in this interesting mathematical field. I hope this conference will be bringing Malaysia one more step closer towards realizing research in theory and applications of cryptology.

It goes without saying that a conference of this kind could not have been held without the committed efforts of various individuals. I would like to take this opportunity to congratulate and thank everyone involved for their excellent work and in particular to Universiti Malaysia Sabah (UMS), Universiti Putra Malaysia (UPM), Universiti Sains Malaysia (USM) and CyberSecurity Malaysia for taking up the challenge of organizing this conference. I would also like to thank our corporate partner Augmented Technology, who has helped to realize this event. I wish all participants to have an enjoyable and beneficial event. With that, I once again thank all presenters and participants in making this conference possible and a successful event.

Thank you.

DATO DR HJ KAMEL ARIFFIN BIN HJ MOHD ATAN
President,
Malaysian Society for Cryptology Research

EDITORIAL PREFACE

Since the time of Julius Caesar and possibly up until the Greek era, cryptography (a word that is derived from the Greek term cryptos) has been an integral tool for organizations (and indeed for individuals too) to ensure information that is intended only for authorized recipients remain confidential only to this set of people. Cryptography had far reaching implications for organizations in the event information leakage occurred. Often referred to as the last bastion of defence after all other mechanisms had been overcome by an adversary, encrypted information would still remain useless to the attacker (i.e. that is, under the usual security assumptions). Nevertheless, this simple fact has remained oblivious to the practitioners of information security omitting cryptographic mechanism for data being transferred and also during storage.

Fast forward to World War 2, the war between cryptographic and cryptanalytic techniques. While the Germans were efficiently transferring information via the Enigma encryption machine, the Allies in Bletchley Park, England were busy intercepting these ciphered information being transmitted via telegraph by the Germans. Leading mathematicians, linguists, engineers etc. were all working to cryptanalyze these ciphers in the most information way. It is here that the first electrical machine (i.e. the bomba) was born and revolutionized computing. Post World War 2 saw the emergence of the computer. Every organization that had to process data had to acquire a computer so as not to be left behind by their competitor. The banking sector advanced on a global scale due to the invention of the computer. Techniques to secure information among the headquarters of these banks had to be developed. Encryption procedures using the same key (i.e. symmetric encryption) played this role in the early days. Then came the unthinkable problem computers were being deployed almost everywhere. How is it possible to deploy cryptographic keys in secure manner so that symmetric encryption could take place? Thus, leading to the so-called key distribution problem. It was not until 1975, when Diffie and Hellman provided us with a secure key exchange method and in 1976 when Rivest, Shamir and Adleman with the asymmetric encryption scheme (i.e. to encrypt using key e and decrypt using key d, where $e \neq d$). Since then, cryptographic procedures evolved, not only playing the role of ensuring confidentiality of data, but also to ensure integrity and authenticity of data. It is also able to ensure that non-repudiating of data does not occur.

Mechanisms to transfer and store data has changed of the centuries and more so every 5 years (in this modern age). Cryptography that has long existed before mechanisms changed from manual telegraphic electrical electronic (WAN/LAN/internet) wired until wireless procedures, has to be properly deployed in order to maintain a high level of security confidence among the stakeholders of a certain organization. The concept of securing information via encryption procedures has to be properly understood in order to avoid a null intersection to occur between cryptography and computer security practitioners. This scenario would not be to the best interest for stakeholders. As a friendly reminder, this scenario could already been seen in other discipline of knowledge where the minuting (minute-ting) of knowledge has forced the original body of knowledge to look as though it is independent and disassociated. Ever since mass usage of computers became a reality, computer security issues have never been this complicated. However, as the human race advances so will ingenious ideas emerge to overcome challenges.

It is hoped that Cryptology2016 will not only provide a platform for every participant to exchange ideas in their respective fields, but also to exchange new ideas on a broader scale for the advancement of the field of cryptology and computer security. The organizing committee hopes every participant will have an enjoyable and beneficial conference.

Thank you.

Editorial Board,
Cryptology2016

Table of Contents

Board of Editors

	Zahari Mahad
Associate Editors	Amir Hamzah Abd Ghafar Nor Azlida Aminudin
Cover Design	Zahari Mahad

Watch your Constants: Malicious Streebog

Prof. Dr. Amr M. Youssef

Concordia University, Montreal,
Quebec, Canada

youssef@ciise.concordia.ca

ABSTRACT

The belief that governmental spy agencies work hard to incorporate backdoors in their primitives, which enable the efficient manipulation of certain security properties, has always been lurking in the cryptographic community. This belief was further strengthened after Edward Snowden exposed the existence of the NSAs Bullrun decryption project. In this talk, we provide an overview of malicious block ciphers and hash functions and review some of the recent results in this area. Then we present some of our results on the new Russian hash function standard GOST R 34.11-2012, also known as Streebog, in the context of malicious hashing where we are able to obtain a practical collision for a malicious version of the full hash function. We will also discuss some future research directions in this area. This is a joint work with my PhD student, Riham Altawy.

Keywords: backdoor mechanisms, malicious block ciphers, malicious hash, hash collision

Information Security for the Protection of National Sovereignty

Dr. Amirudin Abdul Wahab

CyberSecurity Malaysia,
Malaysia

ABSTRACT

Digital revolution has introduced new security challenges. Today, we are facing emerging threats of malware, organised cyber crime, cyber espionage, hactivism, disgruntled employees and other types of cyber threats. There are also alarming trends of Advanced Persistent Threats (APTs) committed at the levels of organization and nation which are targeting critical sectors that give impact on national security. Such threats need to be addressed adequately via a more adaptive approach encompassing predictive, detective, responsive and mitigative actions. This presentation highlights the current trends of information security landscape and the various cyber threats aimed at compromising confidentiality, integrity and availability of information. It also highlights Malaysias current initiatives in protecting her e-sovereignty. Among them is the National Cryptography Policy. This policy is a central strategy for the protection of information security in the aspects of confidentiality, integrity, authentication and disallowing repudiation

Keywords: digital revolution, Advanced Persistent Threats (APT), e-sovereignty

Post Quantum Cryptography

Prof. Dr. Abderrahmane Nitaj

Université de Caen Normandie,
France

abderrahmane.nitaj@unicaen.fr

ABSTRACT

A quantum computer with Shor's algorithm will solve the integer factorization and the discrete logarithm problems upon which most of the widely used cryptosystems such as RSA (Rivest, Shamir, Adleman) and ECC (elliptic curve cryptography) are based. Nevertheless, some cryptosystems running on conventional computers, such as NTRU, LWE and McEliece are still resisting to quantum computers. Such cryptosystems are good candidates the post quantum cryptography. In this talk, we will present the most promising post quantum cryptosystems and discuss their security.

Keywords: Shor's algorithm, quantum computers, post quantum cryptosystems

HXDTRU Cryptosystem Based On Hexadecnion Algebra

Hassan R.Yassein *[1] and **Nadia M. G. Al-Saidi**[2]

[1]*Department of Mathematics, College of Education, Al-Qadisiyah University, Iraq*
[2]*Department of Applied Sciences, University of Technology, Iraq*

E-mail: hassan.yaseen@qu.edu.iq
Corresponding author

ABSTRACT

In this paper, a new public key cryptosystem based on a new proposed algebra called hexadecnion algebra, which is a non associative, non-commutative and alternative is introduced; it is called HXDTRU. The HXDTRU is principally works as the NTRU, but its operations taken place in the hexadecnion algebra. The security of HXDTRU with N dimension equals the security of NTRU with the $16N$ dimension, and HXDTRU with N dimension is sixteen times faster than NTRU with the $16N$ dimension, which is a respectable improvement especially for large N.

Keywords: NTRU, HXDTRU, hexadecnion algebra, lattice hexadecnion algebra.

1 INTRODUCTION

The NTRU (number theory research unit) public key cryptosystem was founded in 1996 by three mathematicians Jeffery Hoffstein, Joseph H. Silverman and Jill Piper (Hoffstein et al., 1998), the basic collection of objects used by the NTRU public key cryptosystem take place in a truncated polynomial ring of degree $N-1$ with integer coefficients in $Z[x]/(x^N-1)$. It is the first public key cryptosystem that do not depend on factorization (as RSA cryptosystem) or discrete algorithmic problems (as EL Gamal cryptosystem and ECC cryptosystem)(Blömer and May, 2001). In comparison with RSA cryptosystem and Ecc cryptosystem, NTRU is faster and has significantly smaller keys. Many researchers have tried to improve the NTRU cryptosystem through choosing a different ring and applying a more efficient linear transformation. In 2002, P.Gaborit et al., introduced CTRU based on the ring of the polynomials in one variable over a finite field (Gaborit et al., 2002). In 2005, M.Coglianese and B.Goi, presented a new cryptosystem called MaTRU by using ring of $k \times k$ matrices of polynomials of order n (Coglianese and Goi, 2005). In 2009, Malekian et al.,introduced QTRU cryptosystem based on quaternion algebra (Malekian and Zakerolhosseini, 2010b, Malekian et al., 2009). They also introduced OTRU cryptosystem based on octonions algebra (Malekian and Zakerolhosseini, 2010a,b), also, N.Vats presented a new cryptosystem NNRU (Vats, 2009). In 2011, K.Jarvis presented ETRU based on Eisenstein integers (Jarvis, 2011, Jarvis and Nevins, 2015). In 2015, Majeed introduced CQTRU cryptosystem based on commutative quaternions algebra (Alsaidi et al., 2015, Majeed, 2015). In this paper, we presented a new multidimensional public key cryptosystem HXDTRU based on hexadecnion algebra.

This paper is organized as follows;the mathematical description of the hexadecnion algebra where the HXDTRU is defined is described in Section 2, as well as its algebraic structure and the proposed cryptosystem. The resistance of the proposed system against some known attacks is investigated in Section 3. Finally, we summarize our conclusions in Section 4.

2 THE PROPOSED HXDTRU CRYPTOSYSTEM

The parameters N, p and q are similar to the parameters in NTRU, the constant d_f, d_g, d_m and d_Φ are integers less than N. Let $K = Z[x]/(x^N - 1)$ be the truncated polynomials ring of degree N-1. We define a new algebra as follows:

2.1 HEXADECNION ALGEBRA (HD)

In this section, we define hexadecnion algebra and its properties. It is a vector space of sixteen dimension over the real number R defined as follows: HD= $\{w|w = r_0+\sum_{i=1}^{15} r_i x_i | r_0, r_1, \cdots, r_{15} \in R\}$ where $\beta = \{1, x_1, x_2, \cdots, x_{15}\}$ form the basis of the hexadecnion algebra and r_i s are scalars in a set of real number. Let w_1 , w_2 $\in$HD such that: $w_1 = r_0 + r_1x_1 + r_2x_2 + ... + r_{14}x_{14} + r_{15}x_{15}$, $w_2 = r'_0 + r'_1x_1 + r'_2x_2 + ... + r'_{14}x_{14} + r'_{15}x_{15}$

The addition of w_1 and w_2 is found by adding their corresponding coefficients such that; $w_1 + w_2 = (r_0 + r'_0) + (r_1 + r'_1)x_1 + (r_2 + r'_2)x_2 + \cdots + (r_{14} + r'_{14})x_{14} + (r_{15} + r'_{15})x_{15}$.

The multiplication table in Appendix is given to define the multiplication between w_i and w_j, where w_i and w_j $\in$HD, and $x_i^2 = -1, x_ix_j = -x_jx_i, i \neq j$, and $i, j = 1, 2, \cdots, 15$. The multiplication is non commutative and non associative but it is alternative.

For any scalar α, we have,

$$\alpha w = \alpha(r_0 + r_1x_1 + r_2x_2 + \cdots + r_{14}x_{14} + r_{15}x_{15})$$

$$= \alpha r_0 + \alpha r_1x_1 + \alpha r_2x_2 + \cdots + \alpha r_{14}x_{14} + \alpha r_{15}x_{15}$$

The conjugate of a hexadecnion $w = r_0 + \sum_{i=1}^{15} r_ix_i$ is defined as follows $\overline{w} = r_0 - \sum_{i=1}^{15} r_ix_i$ and the square norm is given by $N(w) = w\overline{w} = \sum_{i=1}^{15} {r_i}^2$.

The multiplicative inverse of any non zero element w in HD is given by

$$w^{-1} = N(w)^{-1}\overline{w}$$

where $gcd(N(w), 15) = 1$.

2.2 ALGEBRAIC STRUCTURE OF HXDTRU

Let K be any arbitrary finite ring of characteristic is not equal to 2, we define the hexadecnion algebra Ψ over K as follows:

$$\Psi = \{r_0 + \sum_{i=1}^{15} r_i x_i | r_0, r_1, \cdots, r_{15} \in K\}$$

Where the multiplication, the multiplicative inverse and the norm has the same properties as the real hexadecnion algebra HD. Note that Ψ is a non associative and since the usual multiplication of matrices is associative then it does not have any matrix representation. Now, consider the truncated polynomial rings $K_p(x) = (Z/pZ)[x]/(x^N - 1)$ and $K_q(x) = (Z/qZ)[x]/(x^N - 1)$. We define three hexadecnion algebras Ψ, Ψ_p and Ψ_q as follows:

$$\Psi = \{f_0 + \sum_{i=1}^{15} f_i(x)x_i | f_0, f_1, \ldots, f_{15} \in \Re\}$$
$$\Psi_p = \{f_0 + \sum_{i=1}^{15} f_i(x)x_i | f_0, f_1, \ldots, f_{15} \in K_p\}$$
$$\Psi_q = \{f_0 + \sum_{i=1}^{15} f_i(x)x_i | f_0, f_1, \ldots, f_{15} \in K_q\}$$

Now, let Φ_1 , $\Phi_2 \in \Psi_p$ or Ψ_q such that;

$$\Phi_1 = f_0(x) + f_1(x)x_1 + f_2(x)x_2 + \cdots + f_{14}(x)x_{14} + f_{15}(x)x_{15}$$
$$\Phi_2 = g_0(x) + g_1(x)x_1 + g_2(x)x_2 + \cdots + g_{14}(x)x_{14} + g_{15}(x)x_{15}$$

where f_i , $g_i \in K_p$ or K_q.
The addition of $\Phi_1$and Φ_2 is done by adding their corresponding coefficients including $16N$, $mod\ p$ or $mod\ q$

$$\Phi_1 + \Phi_2 = f_0(x) + g_0(x) + (f_1(x) + g_1(x))x_1 + (f_2(x) + g_2(x))x_2 + \cdots + \\ (f_{14}(x) + g_{14}(x))x_{14} + (f_{15}(x) + g_{15}(x))x_{15}$$

The multiplication of Φ_1 and Φ_2 is defined as follows:

$$\Phi_1 * \Phi_2 = (f_0 * g_0 - f_1 * g_1 - f_2 * g_2 - f_3 * g_3 - f_4 * g_4 - f_5 * g_5 - f_6 * g_6 - f_7 * g_7 - f_8 * g_8 \\ - f_9 * g_9 - f_{10} * g_{10} - f_{11} * g_{11} - f_{12} * g_{12} - f_{13} * g_{13} - f_{14} * g_{14} - f_{15} * g_{15}) \\ + (f_0 * g_1 + f_1 * g_0 + f_2 * g_3 + f_3 * g_5 + f_4 * g_5 f_5 * g_4 - f_6 * g_7 + f_7 * g_6 + f_8 * g_9 - f_9 * g_8 + f_{10} * g_{11} \\ - f_{11} * g_{10} + f_{12} * g_{13} - f_{13} * g_{12} + f_{14} * g_{15} f_{15} * g_{14})x_1 + \cdots + (f_0 * g_{15} + f_1 * g_{14} \\ - f_2 * g_1 3 + f_3 * g_{12} - f_4 * g_{11} - f_5 * g_{10} + f_6 * g_9 + f_7 * g_8 - f_8 * g_7 - f_9 * g_6 \\ + f_{10} * g_5 + f_{11} * g_4 - f_{12} * g_3 + f_{13} * g_2 - f_{14} * g_1 + f_{15} * g_0)x_{15}$$

such that $*$ is a convolution product.

2.3 The PROPOSED HXDTRU

The security of HXDTRU cryptosystem depended on the parameters N, p and q, where N is a prime, $gcd(p,q) = 1$ and q much larger than p. The subsets $\mathcal{L}_f$, $\mathcal{L}_g$, $\mathcal{L}_m$ and $\mathcal{L}_\Phi \subset \Psi$ are defined as follows:

$\mathcal{L}_f = \{f_0(x) + f_1(x)x_1 + f_2(x)x_2 + \ldots + f_{14}(x)x_{14} + f_{15}(x)x_{15} \in \Psi |\ f_i \in K$ has d_f coefficients equal to +1, $(d_f - 1)$ equal to -1,the rest are 0 $\}$,
$\mathcal{L}_g = \{g_0(x)+g_1(x)x_1+g_2(x)x_2+\ldots+g_{14}(x)x_{14}+g_{15}(x)x_{15} \in \Psi | g_i \in K$ has d_g coefficients equal to +1, d_g equal to -1,the rest are 0 $\}$,
$\mathcal{L}_m = \{m_0(x) + m_1(x)x_1 + m_2(x)x_2 + \ldots + m_{14}(x)x_{14} + m_{15}(x)x_{15} \in \Psi |$ coefficients of $m_i(x) \in \Psi$ are chosen modulo p, between $p/2$ and $p/2\}$ and
$\mathcal{L}_\Phi = \{\Phi_0(x) + \Phi_1(x)x_1 + \Phi_2(x)x_2 + \ldots + \Phi_{14}(x)x_{14} + \Phi_{15}(x)x_{15} \in \Psi | \Phi_i \in k$ has d_Φ coefficients equal to +1, d_Φ equal to -1,the rest are 0 $\}$

Also, d_f, d_g and d_Φ are constant parameters similar to those in NTRU. Then HXDTRU can be described through three phases:

a) KEY GENERATION
To generate the public key and the private key, two small norm F and $G \in \Psi$ are randomly generated, such that:

$$F = f_0(x) + f_1(x)x_1 + f_2(x)x_2 + \cdots + f_{14}(x)x_{14} + f_{15}(x)x_{15},$$
$$f_0, f_1, f_2, \cdots, f_{14}, f_{15} \in \mathcal{L}_f$$
$$G = g_0(x) + g_1(x)x_1 + g_2(x)x_2 + \cdots + g_{14}(x)x_{14} + g_{15}(x)x_{15},$$
$$g_0, g_1, g_2, \cdots, g_{14}, g_{15} \in \mathcal{L}_g$$

Here, F must have multiplication inverses over Ψ_p and Ψ_q. If F is not invertible (when the inverse of $\sum_{i=1}^{15} f_i^2(x)$, is not existed in K_p or K_q) then a new hexadecnion F should be chosen. The inverses of F are denoted by F_P and F_q over algebra Ψ_p and Ψ_q respectively. Now, the public key is calculated as follows:

$H = F_q \cdot G \in \Psi_q = h_0(x) + h_1(x)x_1 + h_2(x)x_2 + ... + h_{14}(x)x_{14} + h_{15}(x)x_{15}$. F, F_P and F_q must be kept secret in order to be used in decryption phase. When the same parameters N, p and q are used in NTRU and HXDTRU, the key generation phase of NTRU is faster than that of HXDTRU, but the computational time of this phase depends on the computation of the inverses which is greater than the time of this phase in traditional NTRU with the same parameters.

b) ENYCRYPTION

At the beginning of the encryption process, the message M should be converted to the form, $M = m_0(x) + m_1(x)x_1 + m_2(x)x_2 + \cdots + m_{14}(x)x_{14} + m_{15}(x)x_{15}$

where $m_i(x) \in \mathcal{L}_m$, $i = 0, 1, \cdots, 15$ and Φ is another small hexadecnion that is randomly chosen. It computes the encrypted message M as follows:

$$E = pH \circ \Phi + M \in \Psi_q,$$

the encryption in HXDTRU needs one hexadecnion multiplication including 256 convolution multiplication.

c) DECRYPTION

After receiving the encrypted message E, the receiver decrypt the message through the following steps: At the first, E is multiplied by the private key F on the left and then on right as follows:

$A = (F \circ E) \circ F \in \Psi_q$
$= (F \circ (pH \circ \Phi + M)) \circ F \circ \Psi_q$
$= p(F \circ (H \circ \Phi)) \circ F + (F \circ M) \circ F \in \Psi_q$
$= p(F \circ H) \circ (\Phi \circ F) + (F \circ M) \circ F \in \Psi_q$ (by moufang identity)
$= p(F \circ (F_q \circ G)) \circ (\Phi \circ F) + (F \circ M) \circ F \in \Psi_q$
$= pG \circ (\Phi \circ F) + (F \circ M) \circ F \in \Psi_q$

The coefficients of sixteen polynomial in $pG \circ (\Phi \circ F) + (F \circ M) \circ F$ must lie in the intervals $(-q/2, q/2]$ and the last reduction mod q does not required. When the term $(\Phi \circ F) + (F \circ M) \circ F$ is reduced mod p, the term $F \circ M$ (mod p) remains and $pG \circ (\Phi \circ F)$ vanishes.

Hence, $A = F \circ M \in \Psi_p$. Multiplying $A = F \circ M$ (mod p) by F_p, the message $M = F_p \circ A$ is constructed and its coefficients are adjusted to lying in the interval $[-p/2, p/2]$.

2.4 SUCCESSFUL DECRYPTION

If all hexadecnion coefficients of $pG \circ (\Phi \circ F) + (F \circ M) \circ F$ lies within the interval $(-q/2, q/2]$ then the probability of successful decryption is increased. Now, to compute this probability, let

$$A = pG \circ (\Phi \circ F) + (F \circ M) \circ F$$

which can be written in the form,

$$A = a_0(x) + a_1(x)x_1 + a_2(x)x_2 + \cdots + a_{14}(x)x_{14} + a_{15}(x)x_{15}.$$

The polynomial $a_0(x)$ represents the constant coefficients of A such that;

$a_0 = p(g_0\Phi_0 f_0 - g_0\Phi_1 f_1 - g_0\Phi_2 f_2 - g_0\Phi_3 f_3 - g_0\Phi_4 f_4 - g_0\Phi_5 f_5 - g_0\Phi_6 f_6 - g_0\Phi_7 f_7 - g_0\Phi_8 f_8 - g_0\Phi_9 f_9 - g_0\Phi_{10} f_{10} - g_0\Phi_{11} f_{11} - g_0\Phi_{12} f_{12}$
$- g_0\Phi_{13} f_{13} - g_0\Phi_{14} f_{14} - g_0\Phi_{15} f_{15}$
$+ \ldots + g_{15}\Phi_0 f_{15} + g_{15}\Phi_1 f_{14} - g_{15}\Phi_2 f_{13} + g_{15}\Phi_3 f_{12} - g_{15}\Phi_4 f_{11} - g_{15}\Phi_5 f_{10} +$

$g_{15}\Phi_6 f_9 + g_{15}\Phi_7 f_8 - g_{15}\Phi_8 f_7 - g_{15}\Phi_9 f_6 + g_{15}\Phi_{10} f_5 + g_{15}\Phi_{11} f_4 - g_{15}\Phi_{12} f_3 + g_{15}\Phi_{13} f_2 - g_{15}\Phi_{14} f_1 + g_{15}\Phi_{15} f_0) + \ldots + (f_0^2 m_0 + f_1^2 m_0 + f_2^2 m_0 + f_3^2 m_0 + f_4^2$
$m_0 + f_5^2 m_0 + f_6^2 m_0 + f_7^2 m_0 + f_8^2 m_0 + f_9^2 m_0 + f_{10}^2 m_0 + f_{11}^2 m_0 + f_{12}^2 m_0 + f_{13}^2 m_0 + f_{14}^2 m_0 + f_{15}^2 m_0$
$= [a_{0,0}, a_{0,1}, a_{0,2}, \ldots, a_{0,N-1}]$

Each polynomial of $a_1, a_2, \ldots, a_{15}$ is calculated in the similar method. Now, by the definition of $\mathcal{L}_f, \mathcal{L}_g, \mathcal{L}_m$ and $\mathcal{L}_\Phi$ we obtain,
$f_i = [f_{i,0}, f_{i,1}, f_{i,2}, \cdots, f_{i,N-1}] \quad i = 0, 1, 2, \cdots, 15$

$$g_i = [g_{i,0}, g_{i,1}, g_{i,2}, \ldots, g_{i,N-1}] \quad i = 0, 1, 2, \ldots, 15$$

$$\Phi_i = [\Phi_{i,0}, \Phi_{i,1}, \Phi_{i,2}, \ldots, \Phi_{i,N-1}] \quad i = 0, 1, 2, \ldots, 15$$

$$Pr(f_{i,j} = 1) = \frac{d_f}{N},\ Pr(f_{i,j} = -1) = \frac{d_f - 1}{N} \cong \frac{d_f}{N},\ Pr(f_{i,j} = 0) = 1 - \frac{2d_f}{N}$$

$$Pr(g_{i,j} = 1) = Pr(g_{i,j} = -1) = \frac{d_g}{N},\ Pr(g_{i,j} = 0) = 1 - \frac{2d_g}{N}$$

$$Pr(\Phi_{i,j} = 1) = Pr(\Phi_{i,j} = 1) = \frac{d_\Phi}{N},\ Pr(\Phi_{i,j} = 0) = 1 - \frac{2d_\Phi}{N}$$

$Pr(m_{i,j} = \gamma) = \frac{1}{p}$, where $\gamma \in [-p/2, p/2],\ i, j = 0, 1, 2, \ldots, 15$

Assume that all $f_{i,s}$, $g_{j,t}$ and $\Phi_{k,u}$ are pairwise independent random variables. Then we get, $E(f_{i.s}.g_{j,t}.\Phi_{k.u}) = 0$, and $E(f_{i.s}.f_{j.t}.m_{k.u}) = 0$

$Var(f_{i.s}.g_{j,t}.\Phi_{k.u}) = \frac{8d_f d_g d_\Phi}{N^3},\ Var(f_{i.s}.f_{j.t}.m_{k.u}) = \frac{d_f^2(P-1)(P+1)}{3N^2}$ and

$$Var(f_{i.s}{}^2 m_{k.u}) = \frac{d_f(p-1)(p+1)}{6N}$$

Assume that the covariance of $f_{i.s}$ and $f_{i.t}$ is negligible, the final result is obtained as follows:

$$Var((f_{i.s}.g_{j,t}.\Phi_{k.u})y) = Var(\Sigma\Sigma_{s+t+u=y(modN)} f_{i.s} g_{j.t} \Phi_{k.u}) = \frac{8d_f d_g d_\Phi}{N}$$

$$Var((f_i.f_j.m_k)y) = Var(\Sigma\Sigma_{s+t+u=y(modN)} f_{i.s} g_{j.t} m_{k.u}) = \frac{d_f^2(p-1)(p+1)}{3}$$

$$Var((f_i^2 m_k)y) = Var(\Sigma\Sigma_{s+t+u=y(modN)} f_{i.s} g_{j.t} m_{k.u}) \approx \frac{d_f^2(N-1)(P-1)(p+1)}{3N} + \frac{d_f(p-1)(p+1)}{6}$$

Therefore,

$$Var(a_0, k) \approx \frac{2048p^2 d_f d_g d_\Phi}{N} + 20d_f^2(p-1)(p+1) + \frac{16d_f^2(N-1)(P-1)(p+1)}{3N} + \frac{8d_f(p-1)(p+1)}{3}$$

By applying the same procedure, we obtain

$$Var(a_{0,k}) = Var(a_{1,k}) = Var(a_{2,k}) = \ldots = Var(a_{15,k})$$

$$\approx \frac{2048p^2 d_f d_g d_\Phi}{N} + 20d_f^2(p-1)(p+1) + \frac{16d_f^2(N-1)(P-1)(p+1)}{3N} + \frac{8d_f(p-1)(p+1)}{3}$$

If the probability of all coefficients a_{ik} lie within $[(-q+1)/2 \ldots (q+1)/2]$, then the successful decryption is acheived.

With the assumption that $a_{ik}s$ are independent random variable and have normal distribution $N(0, \sigma^2)$ we obtained,

$$Pr(\mid a_{i,k} \mid \leq \tfrac{q-1}{2}) = Pr(-\tfrac{q-1}{2} \leq a_{i,k} \leq \tfrac{q-1}{2}) = 2\mathcal{N}\tfrac{q-1}{2\sigma} \ ,$$

where $\sigma = \sqrt{\frac{2048p^2 d_f d_g d_\Phi}{N} + 20d_{f^2}(p-1)(p+1) + \frac{16d_{f^2}(N-1)(P-1)(p+1)}{3N} + \frac{8d_f(p-1)(p+1)}{3}}$

The probability for successful decryption in HXDTRU can be calculated by the following two observation;

i) The probability for any one of the messages $M_0, M_1, \ldots, M_{15}$ to be successfully decrypted is $(2\mathcal{N}(\frac{q-1}{2\sigma}) - 1)^N$

ii) The probability for all the messages $M_0, M_1, \ldots, M_{15}$ to be successfully decrypt $(2\mathcal{N}(\frac{q-1}{2\sigma}) - 1)^{16N}$

3 SECURITY ANALYSIS

In this section, some of the known attacks are discussed to show the security improvement of the proposed cryptosystem upon the classical NTRU.

3.1 BRUTE FORCE ATTACK

In HXDTRU an attacker who knows the public parameters, as well as, the public key $H = Fq \circ G$, must try all possible hexadecnion $F \in \mathcal{L}_f$ and check to see if $F \circ H$ turns into hexadecnion with small coefficients until find private key, the size of the subset $\mathcal{L}_f$ is calculated as follows: $\mid \mathcal{L}_f \mid = (\frac{N!}{(d_f!)^2(N-2d_f)!})^{16}$. Similarly, the attacker can search in the space $\mathcal{L}_\Phi$ to get the message original from the ciphertext and this search must be done in the order of the space $\mathcal{L}_\Phi$, where its size is calculated as follows: $\mid \mathcal{L}_\Phi \mid = (\frac{N!}{(d_\Phi!)^2(N-2d_\Phi)!})^{16}$

3.2 LATTICE BASED ATTACKS

The lattice attack against HXDTRU is more difficult because it is a non-commutative algebra and has dimension 16. When the attacker succeeds to obtain hexadecnion F satisfying $H = Fq \circ G \in \Psi_q$, the HXDTRU cryptosystem is broken. The only way for attacking the HXDTRU cryptosystem is by diffusion $H = F_q \circ G$ as follows:

$f_0 * h_0 - f_1 * h_1 - f_2 * h_2 - f_3 * h_3 - f_4 * h_4 - f_5 * h_5 - f_6 * h_6 - f_7 * h_7 - f_8 * h_8 - f_9 * h_9 - f_{10} * h_{10} - f_{11} * h_1 - f_{12} * h12 - f_{13} * h_{13} - f_{14} * h_{14} - f_{15} * h_{15} = g_0 + qv_0$

$f_0 * h_1 + f_1 * h_0 + f_2 * h_3 - f_3 * h_2 + f_4 * h_5 - f_5 * h_4 - f_6 * h_7 + f_7 * h_6 + f_8 * h_9 - f_9 * h_8 + f_{10} * h_{11} - f_{11} * h_{10} + f_{12} * h_{13} - f_{13} * h_{12} + f_{14} * h_{15} - f_{15} * h_{14} = g_1 + qv_1$

$\vdots$

$f_0 * h_{15} + f_1 * h_{14} - f_2 * h_{13} + f_3 * h_{12} - f_4 * h_{11} - f_5 * h_{10} + f_6 * h_9 + f_7 * h_8 - f_8 * h_7 - f_9 * h_6 + f_{10} * h_5 + f_{11} * h_4 - f_{12} * h_3 + f_{13} * h_2 - f_{14} * h_1 + f_{15} * h_0 = g_{15} + qv_{15}$

All polynomials $h_0, h_1, , h_{15}$ can be represented in their matrix isomorphic representation as follows:

$$(H_i)_{N \times N} = \begin{pmatrix} h_{i,0} & h_{i,1} & \cdots & h_{i,N-1} \\ h_{i,N-1} & h_{i,0} & \cdots & h_{i,N-2} \\ h_{i,N-2} & h_{i,N-1} & \cdots & h_{i,N-3} \\ \vdots & \vdots & \ddots & \vdots \\ h_{i,2} & h_{i,3} & \cdots & h_{i,1} \\ h_{i,1} & h_{i,2} & \cdots & h_{i,0} \end{pmatrix}$$

Under the above assumptions, we can describe the HXDTRU lattice of dimension $32N$ spanned by the rows of the matrix $\mathcal{M}_{32N \times 32N} = \begin{pmatrix} I_{16N \times 16N} & H_{16N \times 16N} \\ 0_{16N \times 16N} & qI_{16N \times 16N} \end{pmatrix}$ such that H is the fundamental matrix for the h_i's satisfied $F \circ H = G$, which is descriped in Appendix , where I denoted the identity matrix, qI denotes q times the identity matrix and 0 denoted zero matrix. The vector $(f_0, f_1, \ldots, f_{15}, g_0, g_1, \ldots, g_{15})_{1 \times 32N}$ belong to the HXDTRU lattice which is denoted by $\mathcal{L}_{HXDTRU}$. Using a lattice reduction algorithm, a short vector in HXDTRU lattice can be found. For simplicity assuming $d = d_f = d_g = d_\Phi \approx N/3$, since the determinant of $\mathcal{L}_{HXDTRU}$ is equal to the determinant of $\mathcal{M}_{32N \times 32N}$ which is an upper triangle matrix, so its determinant equal to q^{16N}, $\parallel (f_0, f_1, \ldots, f_{15}, g_0, g_1, \ldots, g_{15}) \parallel \approx \sqrt{64d} \approx 4.62\sqrt{N}$. The Gaussian heuristic expected that the length of the shortest nonzero vector in HXDTRU lattice is calculated as $\delta(\mathcal{L}_{HXDTRU}) = \sqrt{\frac{16N}{\pi e}}\sqrt{q} \approx 1.369\sqrt{Nq}$. Also $\frac{\|(f_0, f_1, \ldots, f_{15}, g_0, g_1, \ldots, g_{15})\|}{\delta} = \frac{4.62\sqrt{N}}{1.369\sqrt{Nq}} \approx \frac{3.37}{\sqrt{q}}$, hence the proposed vectors in $\mathcal{L}_{HXDTRU}$ are shorter than that expected by the Gaussian heuristic, also the dimension of $\mathcal{L}_{HXDTRU}$ is sixteen times larger than the dimension of $\mathcal{L}_{NTRU}$ with the same value of N. Therefore, the resistance of the HXDTRU against lattice attacks is much more than NTRU.

4 CONCLUSIONS

In this paper, the HXDTRU cryptosystem based on Hexadecnion algebra, which is a non-commutative, non-associative and alternative is proposed. The speed of HXDTRU is slower than NTRU with same parameter, but we can overtake this problem by taking small N. The HXDTRU is a multi dimension cryptosystem which has the ability to encrypt message of length $16N$ in one round (i.e. sixteen messages from a single source or sixteen independent messages from sixteen different sources). This property may be important in some applications such as electronic voting. When the coefficients of $x_1, x_2, \ldots, x_{15}$ are equal to zero, HXDTRU converts to NTRU. The security of HXDTRU with dimension N has been similar to that of NTRU with dimension $16N$.

5 APPENDIX

Table 1: The Multiplication Table

*	1	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
1	1	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
x_1	x_1	-1	x_3	$-x_2$	x_5	$-x_4$	$-x_7$	x_6	x_9	$-x_8$	x_{11}	$-x_{10}$	x_{13}	$-x_{12}$	x_{15}	$-x_{14}$
x_2	x_2	$-x_3$	-1	x_1	x_6	x_7	$-x_4$	$-x_5$	x_{10}	$-x_{11}$	$-x_8$	x_9	x_{14}	$-x_{15}$	$-x_{12}$	x_{11}
x_3	x_3	x_2	$-x_1$	-1	x_7	$-x_6$	x_5	$-x_4$	x_{11}	x_{10}	$-x_9$	$-x_8$	x_{15}	x_{14}	$-x_{13}$	$-x_{12}$
x_4	x_4	$-x5$	$-x_6$	$-x_7$	-1	x_1	x_2	x_3	x_{12}	$-x_{13}$	$-x_{14}$	$-x_{15}$	x_8	x_9	x_{10}	x_{11}
x_5	x_5	x_4	$-x_7$	x_6	$-x_1$	-1	$-x_3$	x_2	x_{13}	x_{12}	$-x_{15}$	x_{14}	$-x_9$	$-x_8$	x_{11}	$-x_{10}$
x_6	x_6	x_7	x_4	$-x_5$	$-x_2$	x_3	-1	$-x_1$	x_{14}	x_{15}	x_{12}	$-x_{13}$	$-x_{10}$	$-x_{11}$	$-x_8$	x_9
x_7	x_7	$-x_6$	x_5	x_4	$-x_3$	$-x_2$	x_1	-1	x_{15}	$-x_{14}$	x_{13}	x_{12}	$-x_{11}$	x_{10}	$-x_9$	$-x_8$
x_8	x_8	$-x_9$	$-x_{10}$	$-x_{11}$	$-x_{12}$	$-x_{13}$	$-x_{14}$	$-x_{15}$	-1	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_9	x_9	x_8	x_{11}	$-x_{10}$	x_{13}	$-x_{12}$	x_{15}	x_{14}	$-x_1$	-1	x_3	$-x_2$	x_5	$-x_4$	x_7	$-x_6$
x_{10}	x_{10}	$-x_{11}$	x_8	x_9	x_{14}	x_{15}	$-x_{12}$	$-x_{13}$	$-x_2$	$-x_3$	-1	x_1	x_6	$-x_7$	$-x_4$	x_5
x_{11}	x_{11}	x_{10}	$-x_9$	x_8	x_{15}	$-x_{14}$	x_{13}	$-x_{12}$	$-x_3$	x_2	$-x_1$	-1	x_7	x_6	$-x_5$	$-x_4$
x_{12}	x_{12}	$-x_{13}$	$-x_{14}$	$-x_{15}$	x_8	x_9	x_{10}	x_{11}	$-x_4$	$-x_5$	$-x_6$	$-x_7$	-1	x_1	x_2	x_3
x_{13}	x_{13}	x_{12}	x_{15}	$-x_{14}$	$-x_9$	x_8	x_{11}	$-x_{10}$	$-x_5$	x_4	x_7	$-x_6$	$-x_1$	-1	x_3	$-x_2$
x_{14}	x_{14}	$-x_{15}$	x_{12}	x_{13}	$-x_{10}$	$-x_{11}$	x_8	x_9	$-x_6$	$-x_7$	x_4	x_5	$-x_2$	$-x_3$	-1	x_1
x_{15}	x_{15}	x_{14}	x_{13}	x_{12}	$-x_{11}$	x_{10}	$-x_9$	x_8	$-x_7$	x_6	$-x_5$	x_4	x_3	x_2	$-x_1$	-1

$H_{16NX16N} =$

$$\begin{vmatrix}
H_0 & H_1 & H_2 & H_3 & H_4 & H_5 & H_6 & H_7 & H_8 & H_9 & H_{10} & H_{11} & H_{12} & H_{13} & H_{14} & H_{15} \\
-H_1 & H_0 & -H_3 & H_2 & -H_5 & H_4 & H_7 & -H_6 & -H_9 & H_8 & -H_{11} & H_{10} & -H_{13} & H_{12} & -H_{15} & H_{14} \\
-H_2 & -H_3 & H_0 & -H_1 & -H_6 & -H_7 & H_4 & H_5 & -H_{10} & H_{11} & H_8 & -H_9 & -H_{14} & H_{15} & H_{15} & -H13 \\
-H_3 & -H_2 & H_1 & H_0 & -H_7 & H_6 & -H_5 & H_4 & -H_{11} & -H_{10} & H_9 & H_8 & -H_{15} & -H_{14} & H_{13} & H_{12} \\
-H_4 & H_5 & H_6 & H_7 & H_0 & H_1 & -H_2 & -H_3 & -H_{12} & H_{13} & H_{14} & H_{15} & H_8 & -H_9 & -H_{10} & -H_{11} \\
-H_5 & -H_4 & H_7 & -H_6 & H_1 & H_0 & H_3 & -H_2 & -H_{13} & -H_{12} & H_{15} & -H_{14} & H_9 & H_8 & -H_{11} & H_{10} \\
-H_6 & -H_7 & -H_4 & H_5 & H_2 & -H_3 & H_0 & H_1 & -H_{14} & -H_{15} & -H_{12} & -H_{13} & H_{10} & H_{11} & H_8 & -H_9 \\
-H_7 & H_6 & -H_5 & -H_4 & -H_3 & H_2 & -H_1 & H_0 & -H_{15} & H_{14} & -H_{13} & -H_{12} & H_{11} & -H_{10} & H_9 & H_8 \\
-H_8 & H_9 & H_{10} & H_{11} & H_{12} & H_{13} & H_{14} & H_{15} & H_0 & -H_1 & -H_2 & -H_3 & -H_4 & -H_5 & -H_6 & -H_8 \\
-H_9 & -H_8 & -H_{11} & H_{10} & -H_{13} & H_{12} & H_{15} & -H_{14} & H_1 & H_0 & -H_3 & H_2 & -H_5 & H_4 & -H_7 & H_6 \\
-H_{10} & H_{11} & -H_8 & -H_9 & -H_{14} & -H_{15} & H_{12} & H_{13} & H_2 & H_3 & H_0 & -H_1 & -H_6 & H_7 & H_4 & -H_5 \\
-H_{11} & -H_{10} & H_9 & -H_8 & -H_{15} & H_{14} & -H_{13} & H_{12} & H_3 & -H_2 & H_1 & H_0 & -H_7 & -H_6 & H_5 & H_4 \\
-H_{12} & H_{13} & H_{14} & H_{15} & -H_8 & -H_9 & -H_{10} & -H_{11} & H_4 & H_5 & H_6 & H_7 & H_0 & -H_1 & -H_2 & -H_3 \\
-H_{13} & -H_{12} & -H_{15} & H_{14} & H_9 & -H_8 & -H_{11} & H_{10} & H_5 & -H_4 & H_7 & H_6 & H_1 & H_0 & -H_3 & H_2 \\
-H_{14} & H_{15} & -H_{12} & -H_{13} & H_{10} & H_{11} & -H_8 & -H_9 & H_6 & H_7 & H_4 & -H_5 & H_2 & H_3 & H_0 & -H_1 \\
-H_{15} & -H_{14} & -H_{13} & -H_{12} & H_{11} & -H_{10} & H_9 & -H_8 & H_7 & -H_6 & H_5 & H_4 & H_3 & -H_2 & H_1 & H_0
\end{vmatrix}$$

REFERENCES

Alsaidi, N., Saed, M., Sadiq, A., and Majeed, A. A. (2015). An improved ntru cryptosystem via commutative quaternions algebra. In *Proceedings of the International Conference on Security and Management (SAM)*, page 198. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

Blömer, J. and May, A. (2001). Low secret exponent RSA revisited. In *Cryptography and Lattices*, pages 4–19. Springer.

Coglianese, M. and Goi, B.-M. (2005). MaTRU: A new NTRU-based cryptosystem. In *Progress in Cryptology-INDOCRYPT 2005*, pages 232–243. Springer.

Gaborit, P., Ohler, J., and Solé, P. (2002). CTRU, a polynomial analogue of NTRU. *Rapport de recherche*.

Hoffstein, J., Pipher, J., and Silverman, J. H. (1998). NTRU: A ring-based public key cryptosystem. In *Algorithmic number theory*, pages 267–288. Springer.

Jarvis, K. (2011). *NTRU over the Eisenstein integers*. PhD thesis, Carleton University.

Jarvis, K. and Nevins, M. (2015). Etru: Ntru over the eisenstein integers. *Designs, Codes and Cryptography*, 74(1):219–242.

Majeed, A. A. (2015). *CQTRU Cryptosystem Pased on Commutative Rings of Quaternion*. PhD thesis, University of Technology,IBaghdad.

Malekian, E. and Zakerolhosseini, A. (2010a). NTRU-like public key cryptosystems beyond dedekind domain up to alternative algebra. In *Transactions on computational science X*, pages 25–41. Springer.

Malekian, E. and Zakerolhosseini, A. (2010b). OTRU: A non-associative and high speed public key cryptosystem. In *Computer Architecture and Digital Systems (CADS), 2010 15th CSI International Symposium on*, pages 83–90. IEEE.

Malekian, E., Zakerolhosseini, A., and Mashatan, A. (2009). QTRU: A Lattice Attack Resistant Version of NTRU PKCS Based on Quaternion Algebra. *Preprint, Available from the Cryptology ePrint Archive: http://eprint. iacr. org/2009/386. Pdf.*

Vats, N. (2009). NNRU, A noncommutative analogue of NTRU. *arXiv preprint arXiv:0902.1891*.

A Review on Four-Dimensional GLV method and ISD method for Scalar Multiplication on Elliptic Curves

Siti Noor Farwina Mohamad Anwar Antony[*1] and
Hailiza Kamarulhaili[2]

[1,2]*School of Mathematical Sciences, Universiti Sains Malaysia,
11800 USM Pulau Pinang, Malaysia*

E-mail: sitinoorfarwina@yahoo.com, hailiza@usm.my
[*]*Corresponding author*

ABSTRACT

In this paper, both four-dimensional GLV (Gallant, Lambert and Vanstone) method and the ISD (Integer sub-decomposition) method are reviewed and revisited. The comparison between these two methods are discussed. Both methods used shortest lattice method to compute the decomposed scalars. In order to compute the shortest vector, both methods implemented the Euclidean algorithm to obtain the required sequences. Both methods used efficiently computable endomorphism. In four-dimensional GLV method, two endomorphisms are being adopted, while in the ISD method, three fast endomorphisms are adopted.

Keywords: Elliptic curve cryptography, integer decomposition method, Efficient computable endomorphisms, Integer sub-decomposition (ISD) method.

1 INTRODUCTION

Supposed E be an elliptic curve where

$$E : y^2 = x^3 + ax + b \pmod p, . \tag{1}$$

which is defined over prime field F_p and let $P, R \in E\left(F_p\right)$. The important operation in elliptic curve cryptography are point multiplication kP and point multiexponentiation $kP + lR$ (Sica et al., 2002). GLV method is an approach to speed up the elliptic curve point multiplication which represented by kP (Gallant et al., 2001). In GLV, the scalar k is decomposed into two scalars where the bit length of the decomposed scalars has been reduced into half of its original length bits. Hence, this method able to accelerate the computation by 50%.

GLV proposed the decomposed scalar k_1 and k_2 to fall between $-\sqrt{n}$ and $\sqrt{n}$ but does not provided any proof. GLV method needs the GLV generator $\{v_1, v_2\}$ which in the kernel of the homomorphism T where $T : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}/n$. But since GLV is not efficient when dealing with larger field, a higher dimension GLV method has been proposed to further accelerate the original GLV method (Hankerson et al., 2004) such as a four-dimensional GLV method. The four-dimensional GLV method decomposed the scalar k into four scalars.

Four-dimensional GLV method has almost similar form as another method which has been proposed to solve the scalar multiplication if the decomposed values does not fall between $-\sqrt{n}$ and $\sqrt{n}$ where GLV is not applicable. This method is called as integer sub-decomposition (ISD) method. ISD helps to improve the computation of scalar multiplication on elliptic curve. In this study, we compare four dimensional GLV method and ISD method. Section 2 describes GLV decomposition method. Section 3 describes four-dimensional GLV method. Section 4 describes ISD method. Section 5 presents the comparison between four-dimensional GLV method and ISD method. And lastly, Section 6 provides the conclusions.

2 GLV METHOD

Supposed E be an elliptic curve which is defined over prime field F_p and let $P = (x, y)$ be a point lies on E with prime order n. And there exist only one subgroup of order n. Supposed there exist a public key $Q = kP$ which is an elliptic point multiplication, where k is a secret key (scalar) chosen from the interval $[1, n-1]$. In GLV method, the scalar k is decomposed into two scalars k_1 and k_2 where $\max\{\,|k_1|\,,\ |k_2|\,\} \leq n$. In order to find the decomposed scalar, first we need to find the generator vectors v_1 and v_2 using extended Euclidean Algorithm such that those vectors belong to kernel T, in other words the transformation $T(v_1) = 0$ and $T(v_2) = 0$. The transformation T is defined as $T : (x_1, x_2) \mapsto x_1 + x_2\lambda \ (mod\, n)$.

The extended Euclidean Algorithm applied to n and λ to produce a sequence of relations $s_i n + t_i \lambda = r_i$ and followed the Bezout's Identity where $\gcd(n, \lambda) = s_i n + t_i \lambda = 1$ since n, λ are relatively prime. From the algorithm, the shortest vector $v_1 = (r_{m+1}, -t_{m+1})$ is chosen where m is the largest integer for $r_m \geq \sqrt{n}$ and the second shortest vector v_2 is chosen between $(r_{m+2}, -t_{m+2})$ and $(r_m, -t_m)$ which ever has the smallest rectangle norm. Then, we need to find the vector $v = \mathbb{Z}v_1 + \mathbb{Z}v_2$ that is closed to $(k, 0) = \mathbb{Q}v_1 + \mathbb{Q}v_2$. And find the short vector u where $u = (k, 0) - v$ such that $T(u) = k$.

The general form for the decomposition of kP is given by

$$kP = k_1P + k_2\Phi(P) = k_1P + k_2\lambda P, \tag{2}$$

where $\max(|k_1|, |k_2|) \leq C\sqrt{n}$ for some explicit constant $C > 0$.

This GLV method used non-trivial endomorphism $\Phi : E \to E$ implying $\Phi(P) = \lambda P$ for some $\lambda \in [1, n-1]$ where λ is the root of the characteristic polynomial acting on Φ . But the original GLV fails to provide an explicit upper bound for k_1 and k_2 . The determination of bound on vector v_1 and v_2 is important in finding the vector u and later helps to improve the upper bound for k_1 and k_2. Some researchers were able to find the bounds on the decomposed scalars by relating it to the bound on v_1 and v_2 which can be obtained based on the characteristic polynomial $X^2 + rX + s$.

3 FOUR-DIMENSIONAL GLV METHOD

Since GLV method are only applicable to GLV curves with an endomorphism of small degree over F_p, Galbraith et al. (2009) has extended the study by exploiting Frobenius endomorphism to obtain more curve over F_{p^2} (GLS curve). In this four-dimensional GLV method, GLV and GLS curves are combined to extend the GLV method into higher dimension. The general form of four-dimensional GLV method is given by

$$kP = k_1P + k_2\Phi(P) + k_3\Psi(P) + k_4\Psi\Phi(P) \tag{3}$$

where the $\Psi = Frob_p$, the p-Frobenius endomorphism of E/F_p such that $\Psi^m(P) = P$ with characteristic polynomial $\Psi^2 + 1 = 0$ and $\Phi^2 + r\Phi + s = 0$ as the characteristic polynomial for Φ which has the same characteristic polynomial for GLV endomorphism , ϕ. Given that $\Psi = \psi Frob_p \psi^{-1}$ and $\Phi = \psi\phi\psi$ which are defined over F_{p^2} where $\psi : E \mapsto E'$ an isomorphism defined over F_{p^4} and E' is a twist of degree d of E over F_{p^d}. Then we have $\Phi(P) = \lambda P$ and $\Psi(P) = \mu P$.

Since in four-dimensional GLV method involved lattice of dimension four, there are two methods in order to obtain the generator vectors which are Lenstra-Lenstra-Lovasz algorithm, (LLL algorithm) [1] and Cornacchia algorithm [2]. In LLL algorithm, they need to have the basis of kernel T which is given by $w_1 = (n, 0, 0, 0), w_2 = (-\lambda, 1, 0, 0), w_3 = (-\mu, 0, 1, 0)$ and $w_4 = (\lambda\mu, -\mu, -\lambda, 1)$ where reduction map $T : \mathbb{Z}^4 \to \mathbb{Z}/n$ or $T : (x_1, x_2, x_3, x_4) \mapsto x_1 + x_2\lambda + x_3\mu + x_4\lambda\mu$. In other words, they have $T(w_1) = T(w_2) = T(w_3) = T(w_4) \equiv 0 \; mod \; n$. Then, LLL algorithm is applied in order to obtain the reduced basis v_1, v_2, v_3, v_4 which are the generator vectors and later being used to find the short vector u such that $u = (k, 0, 0, 0) - v$ which corresponds to the decomposed scalar k_1, k_2, k_3, k_4 respectively.

The LLL-basis can be written in matrix form as $L = \begin{pmatrix} n & 0 & 0 & 0 \\ -\lambda & 1 & 0 & 0 \\ -\mu & 0 & 1 & 0 \\ \lambda\mu & -\mu & -\lambda & 1 \end{pmatrix}$ where $d(L) = n$ and by following the theorem below, they obtained the upper bound for the generator vector in four-dimensional GLV method. We provided proofs of the following theorems and lemmas obtained from Birkner et al. (2012) and Longa and Sica (2014). We have filled the gaps in these proofs for better understanding.

Theorem 3.1. *Cohen (1996)*

Let $v_1, \ldots, v_n$ be an LLL-reduced of a lattice L. Then, $d(L) \leq \prod_{i=1}^{n} |v_i| \leq 2^{n(n-1)/4} d(L)$.

[1] LLL algorithm used to obtain the shortest vector that is closed to orthogonal for lattice that has dimension greater than two. It transform a lattice into a nice lattice by implementing Gram-Schmidt method.

[2] Cornacchia algorithm was introduced in 1908 to solve non-linear Diophantine equation in form of $x^2 + dy^2 = p$ where $d > 0, p$ is prime.

From the theorem above, we have $\prod_{i=1}^{n}|v_i| \leq 2^3 d(L) = 8\left[\mathbb{Z}^4 : kerT\right] = 8n$ given the kernel T has index $\left[\mathbb{Z}^4 : kerT\right] = n$ inside $\mathbb{Z}^4$. In order to sharpen this bound, they have come out with a lemma and a theorem correspond to the norm function of the four-dimensional GLV method.

Lemma 3.1. *Birkner et al. (2012), Longa and Sica (2014)*

Let $N : \mathbb{Z}^4 \to \mathbb{Z}$

$$(x_1, x_2, x_3, x_4) \mapsto \sum_{i_1,i_2,i_3,i_4} b_{i_1}b_{i_2}b_{i_3}b_{i_4}x_1^{i_1}x_2^{i_2}x_3^{i_3}x_4^{i_4}$$

be the norm of an element where b_i*'s lies in* $\mathbb{Z}$*. Then, for any non-zero* $v \in kerT$*, one has* $|v| \geq \frac{n^{1/4}}{\left(\sum_{i_1,i_2,i_3,i_4}|b_{i_1}b_{i_2}b_{i_3}b_{i_4}|\right)^{1/4}}$.

Proof. Ref.Birkner et al. (2012), Longa and Sica (2014)

Consider $v = x_1 + x_2\Phi + x_3\Psi + x_4\Phi\Psi$ with reduction map $T : v \mapsto x_1 + x_2\lambda + x_3\mu + x_4\lambda\mu$. The definition of norm in algebraic number field $Q[\Phi, \Psi]$ is given by $N(v) = v\bar{v}$. If $v \in kerT$, $T(v) = 0$ and since is a homomorphism $T(N(v)) = N(v) = T(v)T(v_1)T(v_2)T(v_3) \equiv 0$ $mod\ n$ where v_1, v_2, v_3 are the conjugates for v with being each of the form $v = x_1 + x_2\bar{\Phi} + x_3\bar{\Psi} + x_4\bar{\Phi}\bar{\Psi}$. $N(v) \equiv 0\ mod\ n$ if and only if $v = 0$. If $v \neq 0$, then $N(v) \geq 0$ and this in turn implies if $|v| < \frac{n^{1/4}}{\left(\sum_{i_1,i_2,i_3,i_4}|b_{i_1}b_{i_2}b_{i_3}b_{i_4}|\right)^{1/4}}$ and hence $N(v) < n$. This is a contradiction. □

Theorem 3.2. *Birkner et al. (2012), Longa and Sica (2014)*

Let E/F_p *be a GLV curve and* E'/F_{p^2} *a twist, together with the two efficient endomorphism* Ψ *and* Φ *. Supposed that the minimal polynomial of* Φ *is* $\Phi^2 + r\Phi + s = 0$ *. Let* $P \in E\left(F_{p^2}\right)$ *a generator of the large subgroup of prime order* n *. And there exists an efficient algorithm, which for any* $k \in [1, n-1]$ *, finds integers* k_1, k_2, k_3, k_4 *such that* $kP = k_1P + k_2\Phi(P) + k_3\Psi(P) + k_4\Psi\Phi(P)$ *with* $\max_i(|k_i|) \leq 16B^3n^{1/4}$ *where*
$B = \left(4 + 4s^2 + 8s + 8|r| + 8|r|s + 2\left(r^2 + 2s\right) + 2\left|r^2 - 2s\right|\right)^{1/4}$.

Proof. Ref.Birkner et al. (2012), Longa and Sica (2014)

From Theorem 2 and Lemma 1, we have $|v| \leq 8n^{1/4}B^3$. And since we have $\Psi^2 + 1 = 0$ and $\Phi^2 + r\Phi + s = 0$ where $\Psi \equiv \imath\ mod\ n$ with $\imath^2 = -1$ and $\Phi = \frac{-r+\sqrt{r^2-4s}}{2} \Rightarrow \frac{-r+\left|\sqrt{r^2-4s}\right|\imath}{2}$ since $D = r^2 - 4s < 0$. So for any $v \in kerT$, we have $v = x_1 + x_2\Phi + x_3\Psi + x_4\Phi\Psi$ and

$N(v) = v\bar{v} = (x_1 + x_2\Phi + x_3\Psi + x_4\Phi\Psi)\left(x_1 + x_2\Phi + x_3\bar{\Psi} + x_4\Phi\bar{\Psi}\right)$
$\left(x_1 + x_2\bar{\Phi} + x_3\Psi + x_4\bar{\Phi}\Psi\right)\left(x_1 + x_2\bar{\Phi} + x_3\bar{\Psi} + x_4\bar{\Phi}\bar{\Psi}\right)$

$= x_1^4 + 2x_1^2x_2^2 + 2x_1^3x_2\Phi + 2x_1x_2x_3^2\Phi + 2x_1^2x_3x_4\Phi + 2x_3^3x_4\Phi + x_1^2x_2^2\Phi^2 + x_2^2x_3^2\Phi^2 + x_1^2x_4^2\Phi^2 + x_3^2x_4^2\Phi^2 + 2x_1^3x_2\bar{\Phi} + 2x_1x_2x_3^2\bar{\Phi} + 2x_1^2x_3x_4\bar{\Phi} + 2x_3^3x_4\bar{\Phi} + 4x_1^2x_2^2\Phi\bar{\Phi} + 8x_1x_2x_3x_4\Phi\bar{\Phi} +$

$$4x_3^2x_4^2\Phi\bar{\Phi} + 2x_1x_2^3\Phi^2\bar{\Phi} + 2x_2^2x_3x_4^2\Phi^2\bar{\Phi}$$
$$+2x_1x_2x_4^2\Phi^2\bar{\Phi}+2x_3x_4^3\Phi^2\bar{\Phi}+x_1^2x_2^2\bar{\Phi}^2+x_2^2x_3^2\bar{\Phi}^2+x_1^2x_4^2\bar{\Phi}^2+x_3^2x_4^2\bar{\Phi}^2+2x_1x_2^3\Phi\bar{\Phi}^2+2x_2^2x_3x_4^2\Phi\bar{\Phi}^2+$$
$$2x_1x_2x_4^2\Phi\bar{\Phi}^2 + 2x_3x_4^3\Phi\bar{\Phi}^2 + x_2^4\Phi^2\bar{\Phi}^2 + 2x_2^2x_4^2\Phi^2\bar{\Phi}^2$$
$$+ x_4^4\Phi^2\bar{\Phi}^2$$

Where the coefficient of terms in the norm function is represented as $\sum_{i_1,i_2,i_3,i_4} |b_{i_1}b_{i_2}b_{i_3}b_{i_4}|$ and we have

$$B = \left(\sum_{i_1,i_2,i_3,i_4} |b_{i_1}b_{i_2}b_{i_3}b_{i_4}|\right)^{1/4}$$
$$= \left(4 + 4s^2 + 8s + 8|r| + 8|r|s + 2\left(r^2 + 2s\right) + 2\left|r^2 - 2s\right|\right)^{1/4}. \qquad \square$$

But since LLL algorithm could not provide a sharper bound, they preferred to use Cornacchia algorithm which required Euclidean algorithm in it. In this four dimensional GLV method, they used Cornacchia algorithm twice where:

1. In order to obtain a Gaussian prime $\nu \in F_{p^2}$ where $\nu = a + b\imath$ dividing n and $\nu P = 0$ ($a \in \Re(\nu), b \in \Im(\nu), \imath^2 = -1$). Cornacchia algorithm need to solve $\mu^2 \equiv -1\ mod\ n$ where $n \equiv 1\ mod\ 4$ is a prime. The Euclidean algorithm in $\mathbb{Z}$ is being applied to n, μ to obtain ν from the sequence of the algorithm such that the initial data is given as $\begin{pmatrix} r_1 & s_1 & t_1 \\ r_0 & s_0 & t_0 \end{pmatrix} = \begin{pmatrix} \mu & 1 & 0 \\ n & 0 & 1 \end{pmatrix}$. The iteration stop at m where m is the largest integer for which $r_m \geq \sqrt{n}$ and $r_{m+1} < \sqrt{n}$. From property 6 given in Lemma 2 in (Birkner et al., 2012, Longa and Sica, 2014), we have $r_0s_j + r_1t_j = r_j$ which means $ns_{m+1} + \mu t_{m+1} = r_{m+1} \Rightarrow r_{m+1} - \mu t_{m+1} = ns_{m+1} \equiv 0\ mod\ n$. So we have $(r_{m+1} - \mu t_{m+1})(r_{m+1} + \mu t_{m+1}) = r_{m+1}^2 + t_{m+1}^2 \equiv 0\ mod\ n$ but it does not necessary to be $r_{m+1}^2 + t_{m+1}^2 = n$. Since $0 \leq r_{m+1}, t_{m+1} < \sqrt{n}$, we have $0 \leq r_{m+1}^2 + t_{m+1}^2 < n + n = 2n$, this implies $r_{m+1}^2 + t_{m+1}^2 = n$ which has same concept for Gaussian prime $|\nu| = a^2 + b^2 = n$. Hence, we have the shortest vector as $(a, b) = (r_{m+1}, -t_{m+1}) \Rightarrow \nu = a + b\imath = r_{m+1} - \mu t_{m+1}$.(Algorithm 1 Ref. Birkner et al. (2012), Longa and Sica (2014)).

2. In order to obtain generator vectors in compact form which contain real and imaginary parts of the vectors. The Euclidean algorithm in $\mathbb{Z}[\imath]$ is being applied the algorithm to such that the initial data is given as $\begin{pmatrix} r_1 & s_1 & t_1 \\ r_0 & s_0 & t_0 \end{pmatrix} = \begin{pmatrix} \nu & 1 & 0 \\ \lambda & 0 & 1 \end{pmatrix}$. where $\lambda^2 + r\lambda + s = 0$ to obtain two $\mathbb{Z}[\imath]$ -linearly independent vectors v_1, v_2 .The iteration stop at m where m is the largest integer for which $r_m \geq n^{1/4}$ and $r_{m+1} < n^{1/4}$. From Lemma 2 in (Birkner et al., 2012, Longa and Sica, 2014), we have $r_0s_j + r_1t_j = r_j$ which means $\lambda s_{m+1} + \nu t_{m+1} = r_{m+1} \Rightarrow r_{m+1} - \lambda s_{m+1} = \nu t_{m+1} \equiv 0\ mod\ \nu$. Hence, we have shortest vectors as $v_0 = (r_m, -s_m), v_1 = (r_{m+1}, -s_{m+1})$ .While the other two generator vectors are $v_2 = \imath v_0, v_3 = \imath v_1$. (Algorithm 2,3 Ref. Birkner et al. (2012), Longa and Sica (2014))

This algorithm gives uniform improvement when switching from 2-dimension to 4-dimensional GLV and since it used Euclidean algorithm, it gives stronger upper bound for the decomposed scalars. Later, they came out with the following theorem:

Theorem 3.3. *Birkner et al. (2012), Longa and Sica (2014)*

When performing an optimal lattice reduction on kernel T , it is possible to decompose k

into k_1, k_2, k_3, k_4 *such that* $kP = k_1P + k_2\Phi(P) + k_3\Psi(P) + k_4\Psi\Phi(P)$ *with* $\underset{j}{max}(|k_j|) \leq 103\sqrt{1 + |r| + s}n^{1/4}$.

4 ISD METHOD

In ISD method (Ajeena and Kamarulhaili, 2013, 2014a,b,c), the condition for the decomposed scalar k_1 and k_2 is $\max\{|k_1|, |k_2|\} > \sqrt{n}$, a complement to GLV method where the condition of the decomposed scalar k_1 and k_2 is $\max\{|k_1|, |k_2|\} \leq \sqrt{n}$. Basically, ISD needs six generator vectors. They used extended Euclidean Algorithm to obtain the generator vectors since it involves two-dimensional problem where the algorithm need to be applied in two stages in order to obtain $v_1, \ldots, v_6$.

In the first stage, ISD method is similar to GLV, where we need to find the shortest vector v_1 and v_2 such that it belong to kernel T. Then those vectors are crucial in finding short vector u which correspond to value of decomposed scalar k_1 and k_2. If the value do not exceed $\sqrt{n}$, then GLV method is applicable. But if it exceeds $\sqrt{n}$, then we need to proceed to second stage where k_1 and k_2 are sub decomposed into the scalars k_{11}, k_{12}, and k_{21}, k_{22} respectively. In this stage,the short vectors v_3, v_4 are obtained from the sequence of $s_in + t_i\lambda_1 = r_i \Rightarrow r_i - t_i\lambda_1 = s_in \equiv 0$ $mod\ n$ while v_5, v_6 are obtained from the sequence of $s_in + t_i\lambda_2 = r_i \Rightarrow r_i - t_i\lambda_2 = s_in \equiv 0$ $mod\ n$. Then, the vector u' and u'' are obtained where $u' = (k_1, 0) - v'$ and $u'' = (k_2, 0) - v''$ and the transformation acting on it are $T(u') = k_1$ and $T(u'') = k_2$ respectively. This later correspond to the decomposed scalar k_{11}, k_{12}, and k_{21}, k_{22}.

The general form of ISD method is given by

$$kP = k_{11}P + k_{12}\Phi_1(P) + k_{21}(P) + k_{22}\Phi_2(P). \tag{4}$$

ISD method used three GLV non-trivial endomorphism which are $\Phi \equiv \lambda\ (mod\ n)$,$\Phi_1 \equiv \lambda_1$ $(mod\ n)$ and $\Phi_2 \equiv \lambda_2\ (mod\ n)$. So, they have the equation as

$$kP = k_{11}P + k_{12}\lambda_1(P) + k_{21}(P) + k_{22}\lambda_2(P). \tag{5}$$

They have not found any explicit upper bound but rather general bounds for the decomposed scalars are given by Antony and Kamarulhaili (2015) in the following theorems.

Theorem 4.1. *Ref.(Antony and Kamarulhaili, 2015)*

Let $kP = k_1P + k_2\lambda(P)$. *Then, the upper bound for the sub-decomposed scalars is given by* $|k_1| \leq |n - 1 - (A)|$ *and* $|k_2| \leq |\sqrt{n}D|$.

Theorem 4.2. *Ref.(Antony and Kamarulhaili, 2015)*

Let $|k_1| \leq |n - 1 - A| > \sqrt{n}$ *and* $k_1P = k_{11}P + k_{12}\lambda_1(P)$. *Then, the upper bound for the sub-decomposed scalars is given by* $|k_{11}| \leq |n - 1 - (A + A')|$ *and* $|k_{12}| \leq |\sqrt{n}D'|$.

Theorem 4.3. *Ref.(Antony and Kamarulhaili, 2015)*

Let $|k_2| \leq |\sqrt{n}D\lambda|$ *and* $k_2P = k_{21}P + k_{22}\lambda_1(P)$. *Then, the upper bound for the sub-decomposed scalars is given by* $|k_{21}| \leq |\sqrt{n}D\lambda - A''|$ *and* $|k_{22}| \leq |\sqrt{n}D''|$.

The values for A, A', A'', D, D', D'' varies depending on the shortest vectors. Furthermore, ISD method gives more successful computation for kP when compared to GLV method (Ajeena and Kamarulhaili, 2014b)

5 COMPARISON BETWEEN FOUR-DIMENSIONAL GLV METHOD AND ISD METHOD

The similarities between four-dimensional GLV method and ISD method is both methods are the extension of the original GLV method which decomposed scalar k into four scalars which are given by

Four-dimensional GLV : $kP = k_1P + k_2\Phi\left(P\right) + k_3\Psi\left(P\right) + k_4\Psi\Phi\left(P\right)$

ISD : $kP = k_{11}P + k_{12}\Phi_1\left(P\right) + k_{21}\left(P\right) + k_{22}\Phi_2\left(P\right)$.

The concept being used in both methods mostly are the same where they used shortest vector problem. But the four-dimensional GLV method requires two types of fast endomorphism which are the p-Frobenius endomorphism and GLV endomorphism while ISD method requires three types of endomorphism. They used different algorithm in order to find the generator vectors. The four-dimensional GLV method used Cornacchia algorithm while ISD method used Extended Euclidean algorithm (EEA). Even these are two different methods but they form the same shortest vector since they are required to use the Euclidean algorithm which make them look much similar. The four-dimensional GLV method used Cornacchia algorithm in $\mathbb{Z}$ and $\mathbb{Z}\left[\imath\right]$ while ISD method used Euclidean algorithm in $\mathbb{Z}$ only. The four-dimensional GLV method managed to obtain the bound on its generator vectors which is $|v_i| \leq 8B^3n^{\frac{1}{4}}$ and when using Cornacchia algorithm to find the vectors, it gives a sharper bound which is $|v_i| \leq 51.5\left(\sqrt{1+|r|+s}\right)n^{1/4}$, while in ISD method the generator vectors are bounded by $|v_i| \leq \sqrt{n}$ as provided in GLV method . Since they used different approach, they have different upper bound of the decomposed scalars. In four-dimensional GLV method, the upper bound of the decomposed scalars is given by $\max_i\left(|k_i|\right) \leq 103\left(\sqrt{1+|r|+s}\right)n^{1/4}$ while ISD method have not found its explicit upper bound but the general upper bound is given by $\max_i\left(|k_i|\right) < C\sqrt{n}$ where the value for C varies depending on the generator vectors (Antony and Kamarulhaili, 2015) .Lastly, Cornacchia algorithm works well when moving a two-dimensional GLV over F_p into a four-dimensional GLV-GLS over F_{p^2} since it used Frobenius endomorphism which make the four-dimensional GLV method applicable on GLV-GLS curves. While ISD method only applicable for a two-dimensional GLV over F_p since it only used GLV endomorphism..

6 CONCLUSIONS

GLV method has improved point multiplications on elliptic curves since it speeds up the computation of scalar multiplication on E which defined over finite field. But when the fields become much larger such as E defined over quadratic extension field or complex field, the GLV method might takes some time to compute the point multiplication. Since then, a lot of methods being discovered which able to improve the original GLV method such as four-dimensional GLV method and ISD method. Both methods might look similar since it involved the decomposition of scalar k into four decomposed scalars. The four-dimensional GLV method helps to speed up the GLV computation and solved more curves since it is applicable on GLV-GLS curves which works on quadratic extension field. Meanwhile, ISD method helps to speed up the calculation of point multiplication on GLV curve where the original GLV method is not applicable, which happened when the decomposed scalar k_1 and k_2 falls outside the range $\pm\sqrt{n}$ but only works on finite field and it has not being applied to any GLS curves yet. A four-dimensional GLV method needs two types of fast endomorphism to accomplish while ISD method needs three GLV endomorphism to accomplish. Other than that, four-dimensional GLV method needs four generator vectors which can be obtained using Cornacchia's algorithm while ISD method needs six generator vectors which can be obtained from Extended Euclidean algorithm. But both methods still used shortest lattice method to find the decomposed scalars and they still implemented the Euclidean algorithm in order to obtain their generator vectors. From the observation, four-dimensional GLV method and ISD method might look similar but it is actually have different characteristics and it works for different type of curves.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to the Ministry of Higher Education and the Universiti Sains Malaysia for the Fundamental Research Grant Scheme (FRGS), account no. 203/PMATHS/6711320. And, the authors also would like to thank Francesco Sica for some explanation from previous works.

REFERENCES

Ajeena, R. and Kamarulhaili, H. (2013). Analysis on the elliptic scalar multiplication using integer sub decomposotion method. *International Journal of Pure and Applied Mathematics*, 87(1):95–114.

Ajeena, R. and Kamarulhaili, H. (2014a). Glv-isd method for scalar multiplication on elliptic curves. *Australian Journal of Basic and Applied Sciences*, 8(15):1–14.

Ajeena, R. and Kamarulhaili, H. (2014b). Point multiplication using integer sub decomposition for elliptic curve cryptography. *Journal of Applied Mathematics and Information Sciences*, 8(2):517–525.

Ajeena, R. and Kamarulhaili, H. (2014c). Two dimensional sub decomposition method for point multiplication on elliptic curves. *Journal of Mathematical Sciences: Advances and Applications*, 25:43–56.

Antony, S. and Kamarulhaili, H. (2015). On the upper bounds of the sub decomposition values of the scalar k for elliptic scalar multiplication. *Global Journal of Pure and Applied Mathematics*, 11(6):4035–4046.

Birkner, P., Longa, P., and Sica, F. (2012). Four dimensional gallant-lambert-vanstone scalar multiplication. *ASIACRYPT*.

Cohen, H. (1996). *A Course in Computational Algebraic Number Theory*.

Galbraith, S., Lin, X., and Scott, M. (2009). Endomorphisms for faster elliptic curve cryptography on a large class of curve. *EUROCRYPT*, pages 518–535.

Gallant, R., Lambert, R., and Vanstone, S. (2001). Faster point multiplication on elliptic curve with efficient endomorphism. *CRYPTO 2001,Advances in Cryptology*, pages 190–200.

Hankerson, D., Menezes, A., and Vanstone, S. (2004). *Guide to Elliptic Curve Cryptography*. Springer.

Longa, P. and Sica, F. (2014). Four dimensional gallant-lambert-vanstone scalar multiplication. *ASIACRYPT,Journal of Cryptology*, 27(2):248–283.

Sica, R., Ciet, M., and Quisquater, J.-J. (2002). Analysis of the gallant-lambert-vanstone method based on efficient endomorphisms: Elliptic and hyperelliptic curves. *SAC 2002,Selected Areas in Cryptography,9th Annual International Workshop*, 2595:21–36.

A Comparative *S*-Index in Factoring RSA Modulus via Lucas Sequences

Nur Azman Abu[*1], **Shekh Faisal Abdul-Latip**[1] and **Muhammad Rezal Kamel Ariffin**[2]

[1]*INSFORNET, Faculty of ICT, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, Durian Tunggal, 76100 Melaka, Malaysia*
[2]*AI-Kindi Cryptography Research Laboratory, Institute for Mathematical Research, Universiti Putra Malaysia,*
43400 Serdang, Selangor, Malaysia

E-mail: *{nura, shekhfaisal}@utem.edu.my, rezal@upm.edu.my*
**Corresponding author*

ABSTRACT

General Lucas sequences are practically useful in cryptography. In the past quarter century, factoring large RSA modulo into its primes is one of the most important and most challenging problems in computational number theory. A factoring technique on RSA modulo is mainly hindered by the strong prime properties. The success of factoring few large RSA modulo within the last few decades has been due to computing prowess overcoming one strong prime of RSA modulo. In this paper, some useful properties of Lucas sequences shall be explored in factoring RSA modulo. This paper introduces the *S*-index formation in solving quadratic equation modulo *N*. The *S*-index pattern is very useful in designing an algorithm to factor RSA modulo. At any instance in the factoring algorithm, the accumulative result stands independently. In effect, there is no clear direction to maneuver whether to go left or right. The S-index will add another comparative tool to better maneuver in a factoring process. On one hand, it shall remain a theoretical challenge to overcome the strong prime properties. On the other hand, it shall remain a computational challenge to achieve a running time within polynomial time to factor RSA modulo. This paper will propose an avenue to do both using general Lucas sequences.

Keywords: Lucas sequence, RSA factoring, RSA modulus

1 INTRODUCTION

General Lucas sequences have made significant contribution to the field of cryptography. Lucas sequence V has been proposed to be used for public key cryptosystem (Smith and Lennon, 1994), in a manner similar to the famous RSA (Rivest et. al., 1978), but using Lucas sequences modulo a composite number instead of exponentiation. It has stipulated to have the same security level as RSA for the same size key, but is about twice as slow. A special Lucas sequence has been used to directly factor pseudo prime numbers especially Carmichael numbers (Abu et. al., 2004).

An efficient computation of general Lucas sequences can be found in (Joye and Quisquater, 1996). Zhenxiang Zhang has shown on how to factor an RSA modulo into its primes near both multiples of group orders $P-1$ or $P+1$ and respectively $Q-1$ or $Q+1$ using Lucas sequences. An asymmetric key GM cryptosystem has been developed by Shafi Goldwasser and Silvio Micali in 1982. It is semantically secure based on intractability of the quadratic residue problem modulo $N = PQ$ where P and Q are large primes. The difficulties of decrypting the ciphertext without the key pair (P, Q) is solely based on a comparative interactive challenge on whether a given ciphertext c is a quadratic residue modulo N when the Jacobi symbol for c is +1.

The non-positional nature of Residue Number Systems (RNS) is very efficient in a single arithmetic computing without any hassle of carry propagations. Unlike in the common index number system, RNS has a drawback in comparison. There is no ease general method for magnitude comparison in RNS. This inability to compare two numbers whichever is larger makes it difficult to operate on large modulo efficiently especially in the field of cryptography. (Sousa, 2007). The magnitude comparison in RNS is equivalent to the Comparative *S*-Index in this paper.

2 CRITERIA OF STRONG RSA PRIMES

Let N be the product of two primes, P and Q. It may be desirable to use strong primes for P and Q. These are prime numbers with certain properties that make the product N difficult to factor by known factoring methods. The selection of P and Q as strong primes has been recommended, prior to the year 2000, as a way to safeguard the well-known classical factoring algorithm (Rivest and Silverman, 2001). However, these basic strong prime criteria are independently imposed on P or Q.

Among the properties of strong RSA modulo $N = PQ$ are as follows.

Criterion 1: $P-1$ and $P+1$ consists of a large prime factor.

Let $P-1 = P_0^- \cdot P_1^- \cdot \cdots \cdot P_{k^-}^-$ and $P+1 = P_0^+ \cdot P_1^+ \cdot \cdots \cdot P_{k^+}^+$. The largest prime factors $P_{k^-}^-$ and $P_{k^+}^+$ should be larger than 256-bit for 512-bit P.

Criterion 2: $Q-1$ and $Q+1$ consist of a large prime factor.

Let $Q-1 = Q_0^- \cdot Q_1^- \cdot \cdots \cdot Q_{k^-}^-$ and $Q+1 = Q_0^+ \cdot Q_1^+ \cdot \cdots \cdot Q_{k^+}^+$. Respectively, the largest prime factors $Q_{k^-}^-$ and $Q_{k^+}^+$ should be larger than 256-bit for 512-bit Q.

Criterion 3: Recursively, for each largest factor, $P_{k^-}^- - 1$and $P_{k^+}^+ - 1$must also consist of large enough prime factor, namely, $P_{k^{--}}^{--}$ and $P_{k^{+-}}^{+-}$ following the notation in (Rivest and Silverman, 2001).

Criterion 4: Each largest factor of the prime $Q_{k^-}^- - 1$and $Q_{k^+}^+ - 1$must also consist of large enough prime factor namely, $Q_{k^{--}}^{--}$ and $Q_{k^{+-}}^{+-}$ respectively.

Factoring the RSA modulo N is well known to be unfeasible. Recently, (Boudaoud, 2009) explores another practical approach to surmount this major difficulty by finding the factorization of an integer in a small neighborhood of N instead of N. (Bakhtiari and Maarof, 2012) pointed out that there are more than one set of decryption key (d, N) on a given set of RSA encryption key (e, N). However the distance between them is lcm($P-1$, $Q-1$) which is ruled by the basic strong prime criteria.

Let an elliptic curve be the set of points

$$E(a, b) = \{ (x, y, z) : y^2z \equiv x^3 + axz^2 + bz^3 \text{ (mod p)} \}$$

By the end of the century, it has been noted to be useless to concentrate on strong primes. It is unnecessary to protect against factoring attacks by building large prime factors into $P-1$ or $P+1$ since the adversary can instead attempt to overcome by finding an elliptic curve $E(a, b)$ whose size

$$P+1-2\sqrt{P} \le |E(a,b)| \le P+1+2\sqrt{P}$$

is smooth (Rivest and Silverman, 2001).

3 GENERAL LUCAS SEQUENCES

Given integer parameters $p>2$ and $q>0$, the general Lucas sequences give rise to two functions similar to exponentiation, namely, U_n and V_n.

$$U_0 = 0,\ U_1 = 1,\ U_n = p\cdot U_{n-1} - q\cdot U_{n-2}$$
$$V_0 = 2,\ V_1 = p,\ V_n = p\cdot V_{n-1} - q\cdot V_{n-2}$$

Calculating an element of a Lucas sequence can be done in a very similar pattern to exponentiation using a power modulo operation. It may be helpful to think of p as the base and the index n as the exponent. The closed forms of the general Lucas sequences are:

$$U_n = \frac{\alpha^n - \beta^n}{\alpha - \beta} \text{ and} V_n = \alpha^n + \beta^n .$$

where α and β are the two roots of the quadratic polynomial $x^2 - px + q$.

These classical Lucas sequences U_n and V_n are generated from second order recursions with integer variables (p, q) and discriminant $\delta = p^2 - 4q$. In the case of (p, q) = (1, 1), the Lucas sequence U_n is popularly known as Fibonacci numbers, and their companions V_n are the Lucas numbers. The requirement on P and Q, to be strong primes by making P±1 and Q±1 to have large prime factors, may no longer appear to be adequately substantiated in the view of the best factorisation algorithms known today.

Pollard Rho Method basically can achieve rapid factorization if P–1 consists of only small prime factors. On the other hand, similar result can be said also about P+1. This method of integer factorisation is originally described in (Williams, 1982). It can find a large factor P very quickly when P+1 is composed of only small factors. (Zhang, 2001) has also shown how the general Lucas Sequence can be employed to exploit any weak primes from both sides, the P–1 and P+1.

4 CRITERIA ON GENERAL LUCAS SEQUENCES

Let $N = PQ$. For a given parameters p and q, take $\delta = p^2 - 4q$. Let $\varepsilon_P = \left(\frac{\delta}{P}\right)$ and $\varepsilon_Q = \left(\frac{\delta}{Q}\right)$. The subscript to the epsilon ε is usually left out within the context of known prime P or Q and $\varepsilon_N = \left(\frac{\delta}{N}\right) = \left(\frac{\delta}{P}\right) \cdot \left(\frac{\delta}{Q}\right) = \varepsilon_P \cdot \varepsilon_Q$.

For instance,

$$\varepsilon_P = \left(\frac{\delta}{P}\right) = \begin{cases} +1, & \delta \text{ is quadratic residue mod } P \\ -1, & \delta \text{ is non-quadratic residue mod } P \end{cases}$$

Here the criteria of general Lucas sequences are being compactly summarised. They are very practical tools in factoring process.

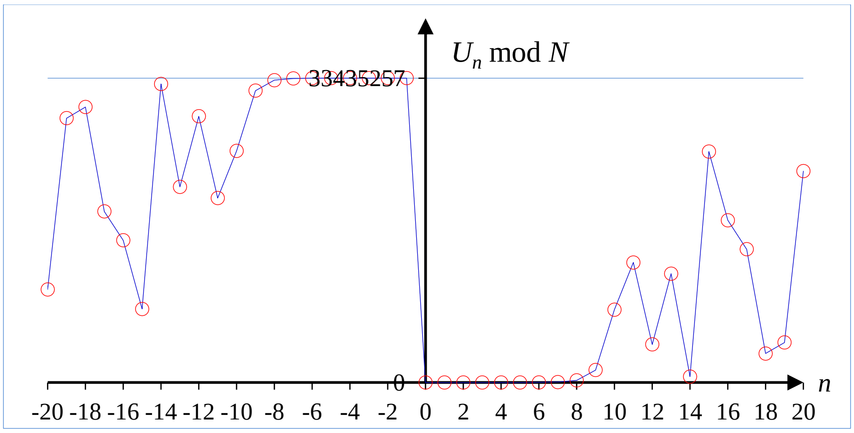


Figure 1: U_n mod N sequence is odd with respect to the center period C.

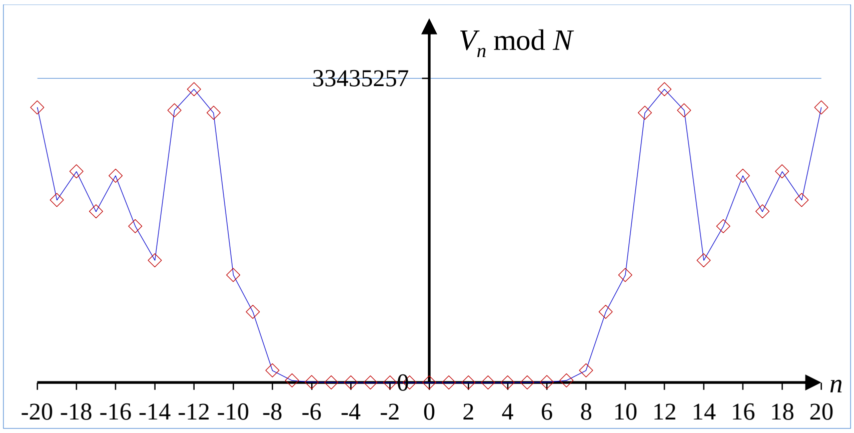


Figure 2: V_n mod N sequence is even with respect to the center period C.

Criterion 1: All the operations here are done modulo N. The maximum period of the general Lucas sequences U and V modulo N of parameters p and q is $C = \text{lcm}(P - \varepsilon_P)(Q - \varepsilon_Q)$. This criteria has been regarded as a generalisation of the Euler totient function for Lucas functions, the Lehmer totient function (Lehmer, 1930).

Table 1: The values of general Lucas sequences U_n *mod N* and V_n *mod N* near the center C.

n	U_n	V_n
-20	10216491	30209367
-19	29036528	20045158
-18	30261649	23191067
-17	18792338	18795473
-16	15621865	22711257
-15	8068338	17166298
-14	32788163	13416017
-13	21484355	29894547
-12	29247453	32210237
-11	20259335	29625847
-10	25438043	11803817
-9	32063152	7761798
-8	33199841	1331714
-7	33394866	228486
-6	33428327	39202
-5	33434068	6726
-4	33435053	1154
-3	33435222	198
-2	33435251	34
-1	33435256	6
0	0	2
1	1	6
2	6	34
3	35	198
4	204	1154
5	1189	6726

6	6930	39202
7	40391	228486
8	235416	1331714
9	1372105	7761798
10	7997214	11803817
11	13175922	29625847
12	4187804	32210237
13	11950902	29894547
14	647094	13416017
15	25366919	17166298
16	17813392	22711257
17	14642919	18795473
18	3173608	23191067
19	4398729	20045158
20	23218766	30209367

Criterion 2: The Lucas sequence U is odd while V is even with respect to the period as shown in the Figures 1 and 2 above, i.e. $U_{kC-n} = - U_{kC+n}$ and $V_{kC-n} = V_{kC+n}$ for any integer k and positive integer n from the center period C=0.

Let the parameters of general Lucas sequences be (p, q) = (6, 1). The values of both Lucas sequences have been listed in Table 1. The graphs in Figures 1 and 2 above show typical characteristics of an odd sequence U_n (mod N) and an even sequence V_n mod N for $N = PQ = 4073 \cdot 8209 = 33435257$. This criterion has made Lucas sequence V appear to be a better reference than U in the LUC public-key system.

Criterion 3: The center values of the general Lucas sequences U and V modulo RSA primes are as follows;

i. $U_{k(P-\varepsilon)} \equiv 0 \pmod{P}$ for any positive integer k.
ii. $V_{k(P-\varepsilon)} \equiv 2q^{\frac{k(1-\varepsilon)}{2}} \pmod{P}$ for any positive integer k.
iii. $U_{k(Q-\varepsilon)} \equiv 0 \pmod{Q}$ for any positive integer k.
iv. $V_{k(Q-\varepsilon)} \equiv 2q^{\frac{k(1-\varepsilon)}{2}} \pmod{Q}$ for any positive integer k.

Preferably the second parameter q is set to be one(1) so that the sequence V will always have consistent output 2 modulo N at a multiple instance of period C.

Criterion 4: These following characteristics have been observed based on the previous research on general Lucas sequences. Most researchers insist on Criterion 3 as a more practical form for factoring purposes. Nevertheless, these criteria are more flexible in factoring angles to choose from.

i. $U_{j(P-\varepsilon)+L} - U_{k(P-\varepsilon)+L} \equiv 0 \pmod{P}$ for some positive integers j and k.
ii. $V_{j(P-\varepsilon)\pm L} - V_{k(P-\varepsilon)\pm L} \equiv 0 \pmod{P}$ for some positive integers j and k.
iii. $U_{j(Q-\varepsilon)+L} - U_{k(Q-\varepsilon)+L} \equiv 0 \pmod{Q}$ for some positive integers j and k.
iv. $V_{j(Q-\varepsilon)\pm L} - V_{k(Q-\varepsilon)\pm L} \equiv 0 \pmod{Q}$ for some positive integers j and k.

It is a necessary condition that $j \neq k$ for integer $-R < L < R$ where R is typically referred to the absolute difference between the primes P and Q. This last criterion is the most useful but by far the most elusive characteristic of the general Lucas sequences in designing a factoring algorithm. It is also noted that Criterion 4 is useful for factoring algorithm if it does not happen simultaneously i.e. the sequence U or V is not equal to the ones modulo N.

Criterion 5: Alternatively, all the criteria above may be summarised in terms of primes P and Q as follows. There are integers $0 \leq a_j, b_k < Q$ and $0 \leq c_j, d_k < P$ such that

i. $U_{j(P-\varepsilon)+L} = a_j \cdot P + UL \pmod N$
ii. $V_{j(P-\varepsilon)\pm L} = b_k \cdot P + VL \pmod N$
iii. $U_{j(Q-\varepsilon)+L} = cj \cdot Q + UL \pmod N$
iv. $V_{j(Q-\varepsilon)\pm L} = dk \cdot Q + VL \pmod N$

for every integer L. Thus, an RSA prime can be extracted respectively by taking the greatest common divisor as follows;

i. $P = gcd(U_{j(P-\varepsilon)+L} - U_L, N)$
ii. $P = gcd(V_{j(P-\varepsilon)\pm L} - V_L, N)$
iii. $Q = gcd(U_{j(Q-\varepsilon)+L} - U_L, N)$
iv. $Q = gcd(V_{j(Q-\varepsilon)\pm L} - V_L, N)$

5 NEW PROPOSAL ON RSA FACTORING

On one hand, it shall remain a theoretical challenge to overcome the strong prime properties. On the other hand, it shall remain a computational challenge to keep the running time within polynomial time to factor RSA modulo.

According to the Proposition 3.3 in (Khadir, 2008) Let N be the product of two prime factors P and Q, $2 < P < Q$. If we can compute efficiently two odd integers r and s such that $s < P$ and $|sQ - rP| \leq 2^{\frac{K+5}{4}}$ where K is the bit-size of the integer rsN, then we can compute the factors P and Q.

In this paper, a more relaxed requirement shall be made.

Suppose $\varepsilon_N = \left(\frac{c}{N}\right) = \left(\frac{c}{P}\right) \cdot \left(\frac{c}{Q}\right) = \varepsilon_P \cdot \varepsilon_Q = (+1)(+1) = 1$. Let $R < P < Q$ such that $R = Q - P$.

$$\begin{aligned} N-1 &= (P-1)(Q-1) + (P-1)+(Q-1) \\ &= (P-1)(Q-1) + 2(P-1) + R = (P-1)(Q-1) + 2(Q-1) - R \end{aligned}$$

For a given odd w,

$$\begin{aligned} N-1 + w &= (P-1)(Q-1) + 2(P-1) + (R+w) \\ &= (P-1)(Q-1) + 2(Q-1) - (R-w) \end{aligned}$$

and

$$N-1 - w = (P-1)(Q-1) + 2(P-1) + (R-w)$$

$$= (P-1)(Q-1) + 2(Q-1) - (R+w)$$

Preferably, $w = 1$ is a good starting point.

Let V_n be the special Lucas sequence with parameters $(p, q) = (p, 1)$ so that $p^2 - 4$ is a quadratic residue of N. Then we need to set a special even Lucas sequence such that $V_0 = 2$, $V_1 = p$, $V_2 = p^2 - 2$ and $V_3 = p \cdot V_2 - V_1 = p \cdot (p^2 - 2) - p = p^3 - 3p$.

Let $N_0 = N-1$. Suppose an odd indexed sequence only is readily available. Nevertheless, it is sufficient to generate the values of V sequences along other large odd indexes. Since $N_0 - w$ and $N_0 + w$ are odd, V sequence modulo N can be computed using a special algorithm below. The running time of this textbook Algorithm 1 is still $O(n^3)$ compared to the running time of general Lucas sequences.

Algorithm 1: A textbook algorithm to compute an odd Lucas sequence V.

```
Function Vodd ( $p$, $K$, $N$)
Set $K=b_{n-1}b_{n-2}\ldots b_2b_1b_0$ be odd such that $b_{n-1}$=1 and $b_0$=1.
Left = $V_1$, Right = $V_3$.
for $i$=$n$–2 down to 1,
        if $b_i$ =0,
                Right = Left*Right – $p$ mod $N$,
                Left   = Left$^2$ – 2 mod $N$.
if $b_i$ =1,
                Left = Left*Right – $p$ mod $N$,
                Right = Right$^2$ – 2 mod $N$.
end(*for*)
return Left.
```

Following the Lucas sequence V criterion 5, there are integers a, b, c and d such that

$$V_{(N-1)-w} = aP + V_{R-w} = bQ + V_{R+w} \qquad (1)$$

$$V_{(N-1)+w} = cP + V_{R+w} = dQ + V_{R-w} \qquad (2)$$

Let us compute

$$S = V_{(N-1)-w} + V_{(N-1)+w} \equiv V_{R-w} + V_{R+w} \pmod N$$

$$T = V_{(N-1)-w} \cdot V_{(N-1)+w} \equiv V_{R-w} \cdot V_{R+w} \pmod N$$

Let us scan for a candidate of x of V_r and y of V_s. respectively the satisfy the conditions

$$x + y \equiv S \pmod N \qquad (3)$$

$$x \cdot y \equiv T \pmod N \qquad (4)$$

From (3), let $y = S - x$, equation (4) will become,

$$x \cdot y = x \cdot (S - x) \equiv T \pmod{N} \tag{5}$$

Consequently, the problem has been reduced down to solving the quadratic equation modulo *N*. We shall search for the root of the function

$$f(x) = x \cdot (S - x) - T \pmod{N}.$$

Let us take the (2*m*+1) terms at one time as the error function,

$$g(x) = \sum_{i=x-m}^{x+m} f(i)$$

A sample case for *N*= 4073·8209 = 33435257 is made here. Let the Lucas sequence parameters (*p*, *q*) = (6, 1), *m*=1 and *w*=3. From (1) and (2),

$$\begin{aligned} V_{(N-1)-3} &= 146 \cdot P + V_{R-3} = -146 \cdot Q + V_{R+3} \\ V_{(N-1)+3} &= 1561 \cdot P + V_{R+3} = -1561 \cdot Q + V_{R-3} \end{aligned}$$

The strategy is to locate the values of V_{R-3} and V_{R+3}. The error function has been plotted within the surrounding region of $V_{(N-1)+3}$ in the Figure 3. We would like to collect the points near zeros.

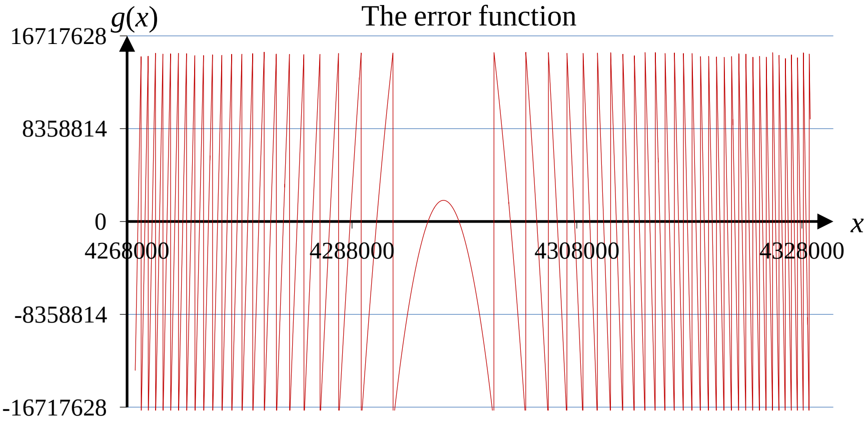


Figure 3: The error function near the zero value.

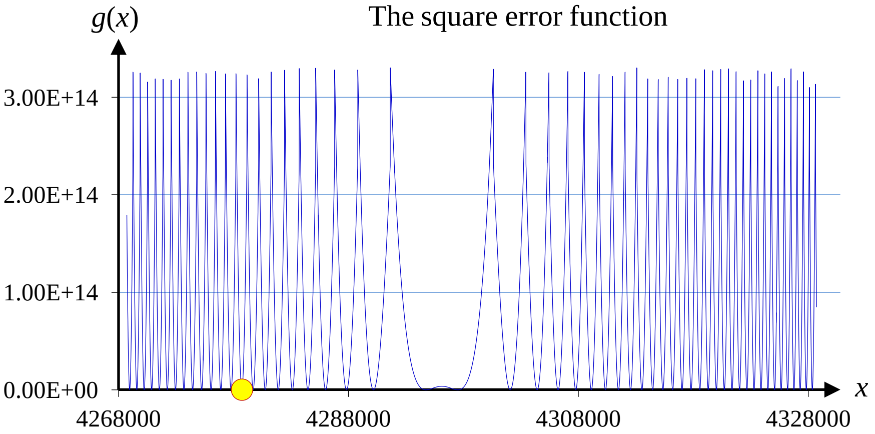


Figure 4: Taking the square on the error function.

Let us *take* the square of the error function so that we can see the error function value near zeros as depicted in Figure 4. The yellow dot is the target value for $V_{(N-1)+w}$. The touchdown points have been observed here as shown in Figure 5. The errors are probabilistically getting larger as the points are moving away from the center critical point. They are much easier to locate as the points of local minima as shown in Figure 4. The green dot is the target value for $V_{(N-1)+3}$.

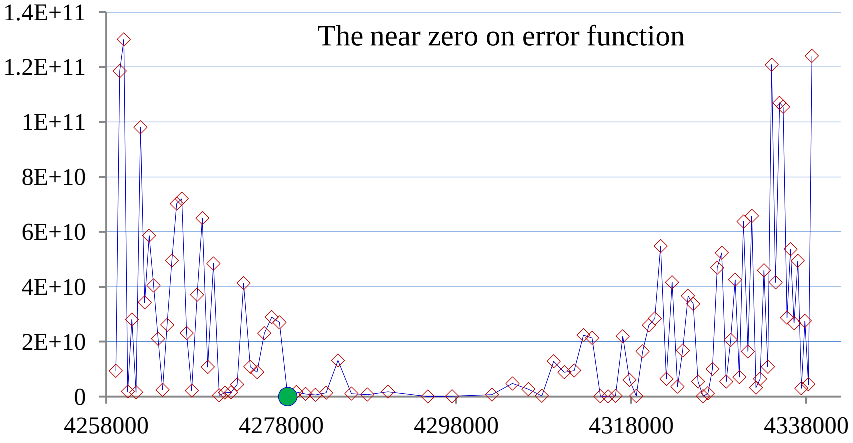


Figure 5: The point of local minima on the error function.

It has *also* been observed that the distances between the local minima is getting smaller as the points go further away from the center. The list of points x has been plotted in the Figure 6 which form the S pattern.

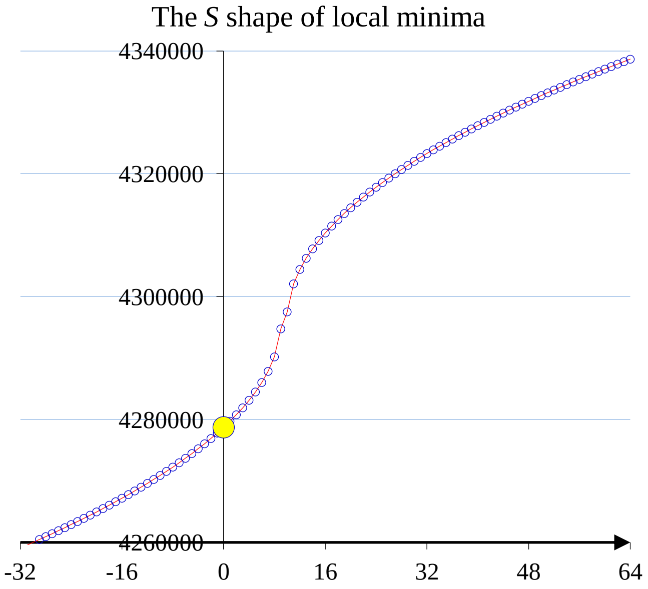


Figure 6: The point of local minima on the error function forms the S shape.

According to basic Calculus, a point x to the left of the critical inflection point z, is said to be concaved up and to the right of the critical inflection point z is concaved down respectively.

6 DISCUSSION

Checking on 3 consecutive ‘touch-down’ at any given point x, will give us a good estimate of the concavity of the surrounding region. A major hurdle in reducing the sub-exponential running time in breaking RSA down to super polynomial running time is the comparative mechanism. At any one time in the factoring algorithm, there has been no mechanism to compare the current position and where to go next. In effect, there is no direction to maneuver whether to go left or right. The S index pattern is very useful in designing an algorithm to factor RSA modulo. For instance, in order to determine the quadratic residue on ciphertext c of N, it suffices to predict whether the Lucas sequence V follow the S index pattern case 0 or case 1. The S-index pattern follows the similar behaviour on all root of the quadratic equation (5) at $V_{(N-1)-3}$, $V_{(N-1)+3}$, V_{R-3} and V_{R+3}. Rather than locating the periodic center of general Lucas sequences U and V as shown in Figures 1 and 2, it is much easier and we stand better chances in locating the S pattern on the quadratic equation (5) modulo N.

7 CONCLUSION

Factoring large integers into primes is one of the most important and most difficult problems in computational number theory. A factoring technique on RSA modulo has been previously hindered by the strong prime properties. Few algorithms have overcome the strong prime criteria of RSA modulo. Nevertheless, they are still subjected to the size of the primes. In this paper, some useful properties of general Lucas sequences have been explored in factoring RSA modulo. A major hurdle in reducing the sub-exponential running time in breaking RSA down to super polynomial running time is the comparative mechanism. At any instance in the factoring algorithm, the accumulative result stands independently. In effect, there is no clear direction to maneuver whether to go left or right. This paper has introduced the S index formation in solving quadratic equation modulo N. The S index pattern is very useful in designing an algorithm to factor RSA modulo. It shall remain a computational challenge to see whether the running time of factoring RSA modulo can be reduced down to a super polynomial time.

ACKNOWLEDGEMENT

We would like to thank anonymous reviewers of CRYPTOLOGY 2016 for giving valuable suggestions and corrections to this paper. This research paper is supported by Fundamental Research Grant Scheme FRGS/1/2015/ICT05/FTMK/02/F00293 funded by the Ministry of Higher Education, Malaysia.

REFERENCES

Abu, N. A., Suryana, N. and Sahib, S. 2004. Factoring Carmichael Numbers using General Lucas Sequences, *Jurnal Matematika*, 4(1):131–136.

Bakhtiari, M. and Maarof, M. A., 2012. Serious Security Weakness in RSA Cryptosystem, *International Journal of Computer Science Issues*, 9.1(3): 175−178.

Boudaoud, Abdelmadjid, 2009. Decomposition of Terms in Lucas Sequences, *Journal of Logic & Analysis*, 1(4):1–23.

Goldwasser, S. and Micali, S., 1984. Probabilistic Encryption, *Journal of Computer and System Sciences*, 28:270−299.

Joye, M. and Quisquater, J.-J., 1996. Efficient Computation of Full Lucas Sequences, *Electronics Letters*, 32(6):537–538.

Khadir, Omar, 2008. Algorithm for Factoring some RSA and Rabin Moduli. *J. Discrete Math. Sci. Cryptography*, 11(5):537−543.

Lehmer, D. H., 1930. An Extended Theory of Lucas' Functions, *Annals of Mathematics, Second Series*, 31(3): 419−448.

Smith, Peter J. and Lennon, Michael J. J., 1994. LUC: A New Public Key System, *Proceedings of the 9th IFIP International Symposium on Computer Security '93*, pp. 097−111, Elsevier Science Publications.

Rivest, R., Shamir, A. and Adleman L., 1978. A Method for Obtaining Digital Signatures and Public Key Cryptosystems, *Communications of the ACM*, 21(2):120−126.

Rivest, Ronald L. and Silverman, Robert D., 2001. Are Strong Primes Needed for RSA? *IACR Cryptology ePrint archive*, paper 2001/007, 30 January 2001.

Sousa, L., 2007. Efficient Method for Magnitude Comparison in RNS Based on Two Pairs of Conjugate Moduli, *18th IEEE Symposium on Computer Arithmetic (ARITH '07)*, 25-27 June 2007, pp. 240 – 250.

Williams, H. C., 1982. A P+1 Method of Factoring, *Mathematics of Computation*, 39(159): 225−234.

Zhang, Z., 2001. Using Lucas Sequences to Factor Large Integers near Group Orders, *Fibonacci Quarterly*, 39(3):228–237.

Families of Cyclotomic Cosets with Application to Secret Sharing Scheme

Denis C.K. Wong

Lee Kong Chian Faculty of Engineering and Science
Universiti Tunku Abdul Rahman, Sungai Long, Malaysia

E-mail: deniswong@utar.edu.my

ABSTRACT

In this paper, all 2-cyclotomic cosets modulo p^n are obtained, when 2 is a primitive root modulo p^n and when the order of 2 is $\frac{p-1}{2}$ modulo p. Furthermore, some results on all s-cyclotomic cosets modulo pq are obtained by considering three different possible orders of s modulo p and q, for distinct odd primes p, q. As an illustration, we use the 2-cyclotomic cosets to construct binary codes of length 49 and hence the access sets for the secret sharing scheme based on some of these families of binary codes are discussed in details.

Keywords: Cyclotomic cosets, minimum distance, secret sharing

1 INTRODUCTION

Let q be a prime and $gcd(q, n) = 1$. The q- cyclotomic coset modulo n containing i is defined by $C_i = \{iq^j (\text{mod}\, n) \in Z_n \mid j = 0, 1, 2, \dots\}$. A subset $\{i_1, \dots, i_t\}$ of Z_n is called a complete set of representatives of q-cyclotomic cosets modulo n if $C_{i_1}, C_{i_2}, \dots, C_{i_t}$ are distinct and $\bigcup_{j=1}^{t} C_{i_j} = Z_n$. It is well-known that two cyclotomic cosets are either equal or disjoint. Hence, the cyclotomic cosets partition Z_n. Dating back to 1948, the birth of coding theory was inspired by the paper called ”A Mathematical Theory of Communication” written by Shannon (Shannon, 1948). Coding theory is the study of the properties of error-correcting codes which are used for data compression, cryptography and network coding. A special type of linear code is cyclic code which was first studied by Prange in 1957. In recent years, many authors have used the cyclotomic cosets approach to construct various families of cyclic codes, see F.J. MacWilliams (1977). Construction of binary idempotents from the cyclotomic cosets is easy. However, there is not much information can be obtained from the generated codes. In years 1997 and 2003, respectively, Arora and Pruthi gave an explicit expression for all q-cyclotomic cosets modulo p^n when q is a primitive root modulo p^n (S.K. Arora, 1997) and when q has order $\frac{\phi(p^n)}{2}$ modulo p^n (S.K. Arora, 1999). Then, in Sharma et al. (2004), the authors obtained all q-cyclotomic cosets modulo p^n with a more subtle conditions. Later in year 2012, Sharma and Bakshi (A. Sharma, 2012) considered a more general type of q-cyclotomic cosets modulo p^m to compute the weight distribution of some irreducible cyclic codes. In K. Singh (2010), q-cyclotomic cosets modulo 2^n when q is quadratic residue modulo 2^n is obtained. More recently, l-cyclotomic cosets modulo the product of two distinct primes power are studied in S.K. Arora (2002) and A. Sahni (2012). In this papers, we construct 2-cyclotomic cosets modulo p^n when 2 is a primitive root modulo p^n and when the order of 2 modulo p is $\frac{p-1}{2}$. Furthermore, we investigate the structures of all s-cyclotomic cosets modulo pq when s is a primitive root modulo q and s is a primitive root modulo p, s has order $\frac{q-1}{2}$ modulo q and s has order $\frac{p-1}{2}$ modulo p, and s has order $\frac{q-1}{2}$ modulo q and s is a primitive root modulo p. For all these cases, p and q are distinct odd primes. Finally, we construct three binary cyclic codes of length 49 from 2-cyclotomic cosets modulo 49. Hence, we investigate the access sets for the secret sharing based on $[49, 3, 28]$-, $[49, 6, 14]$-, and $[49, 4, 21]$- binary cyclic codes.

2 CYCLOTOMIC COSETS

Let $n > 1$ and $gcd(a, n) = 1$. The order of a modulo n is the smallest integer k such that $a^k \equiv 1(\text{mod}\, n)$. When $k = \phi(n)$, then a is a primitive root of the integer n, where ϕ is the Euler-phi function. We need the following result from A. Sharma (2012).

Lemma 2.1. *Suppose α is a primitive root modulo p^n. Then, α is a primitive root modulo p^{n-j} also, for all j, $0 \leq j \leq n-1$.*

2.1 2-Cyclotomic Cosets Modulo p^n when 2 is a Primitive Root modulo p^n

Theorem 2.1. *Let p be an odd prime. Suppose 2 is a primitive root modulo p^2. Then, there are exactly two distinct nonzero 2-cyclotomic cosets modulo p^2.*

Proof. The 2-cyclotomic coset modulo p^2 with coset representative 1 is $C_1 = \{1, 2, 2^2, \ldots, 2^{t_1-1}\}$, where $2^{t_1} \equiv 1(\text{mod}\ p^2)$. Since 2 is a primitive root modulo p^2, then we have $t_1 = \phi(p^2) = p^2 - p = |C_1|$. Clearly, p being an odd prime cannot be in C_1. Thus, we construct the second 2-cyclotomic coset modulo p^2 with coset representative p, that is, $C_p = \{p, 2p, 2^2p, ..., 2^{t_p-1}p\}$, where $2^{t_p}p \equiv p(\text{mod}\ p^2)$. The condition $2^{t_p}p \equiv p(\text{mod}\ p^2)$ is equivalent to $p|(2^{t_p} - 1)$. By Lemma 2.1, we know that 2 is also a primitive root modulo p, then $t_p = \phi(p) = p - 1$, and so $|C_p| = p - 1$. Next, we show that the intersection of $C_1 = \{1, 2, 2^2, ..., 2^{p^2-p-1}\}$ and $C_p = \{p, 2p, 2^2p, ..., 2^{p-2}p\}$ is empty. Suppose $y \in C_1 \cap C_p$. Then, $y \in C_1$ and $y \in C_p$, and so $y = 2^j$ and $y = 2^k p$ for some integers j, k such that $j \geq k$. This implies $2^{j-k} = p$, which is a contradiction. Therefore, we conclude that $C_1 \cap C_p = \emptyset$. Next, for all $z \in \mathbb{Z}_{p^2} \subseteq C_1 \cup C_p \cup \{0\}$ and $|C_1 \cup C_p \cup \{0\}| = |\mathbb{Z}_{p^2}|$. Thus, we have $\mathbb{Z}_{p^2} = C_1 \cup C_p \cup \{0\}$ and the results follows directly. □

Theorem 2.2. *Let p be an odd prime. Suppose 2 is a primitive root modulo p^3. Then, there are exactly three distinct nonzero 2-cyclotomic cosets modulo p^3.*

Proof. Given 2 is a primitive root modulo p^3, then $2^{p^3-p^2} \equiv 1(\text{mod } p^3)$. By Lemma 2.1, we have $2^{p^2-p} \equiv 1(\text{mod } p^2)$ and $2^{p-1} \equiv 1(\text{mod } p)$. The 2-cyclotomic coset modulo p^3 with coset representative 1 is $C_1 = \{1, 2, 2^2, ...,$
$2^{t_1-1}\}$, where $2^{t_1} \equiv 1(\text{mod } p^3)$. Clearly, $t_1 = p^3-p^2$ and so $|C_1| = p^3-p^2$. Next, since $p \notin C_1$, we construct $C_p = \{p, 2p, 2^2p, ..., 2^{t_p-1}p\}$, where $2^{t_p}p \equiv p(\text{mod } p^3)$ which is equivalent to $2^{t_p} \equiv 1(\text{mod } p^2)$ and so $t_p = p^2 - p = |C_p|$. Now, we suppose that $p^2 \in C_p$. Then, $p^2 = 2^i p$ for some i implies that $p = 2^i$ which is a contradiction. Hence, $p^2 \notin C_p$. Finally, as $p^2 \notin C_1$ and $p^2 \notin C_p$, we consider $C_{p^2} = \{p^2, 2p^2, 2^2p^2, ..., 2^{t_{p^2}-1}p^2\}$, where $2^{t_{p^2}}p^2 \equiv p^2(\text{mod } p^3)$ which is equivalent to $2^{t_{p^2}} \equiv 1(\text{mod } p)$ and so $t_{p^2} = p - 1$. Thus, we have $|C_1| = p^3 - p^2, |C_p| = p^2 - p$ and $|C_{p^2}| = p - 1$. Hence, $|C_1| + |C_p| + |C_{p^2}| + |\{0\}| = |\mathbb{Z}_{p^3}|$. Suppose $x \in C_1 \cap C_p$. Then, $x \in C_1$ and $x \in C_p$ implies $x = 2^j$ and $x = 2^k p$ for some integers j, k such that $j \geq k$. It follows that $2^{j-k} = p$, which is a contradiction. Hence, $C_1 \cap C_p = \emptyset$. It can be shown in a similar way that $C_1 \cap C_{p^2} = \emptyset$ and $C_p \cap C_{p^2} = \emptyset$. Therefore, we conclude that $\mathbb{Z}_{p^3} = C_1 \cup C_p \cup C_{p^2} \cup \{0\}$ and $C_i \cap C_j = \emptyset$ for all $i,\ j \in \{1, p, p^2\}$. □

For a more general case, we have the following theorem.

Theorem 2.3. *Let p be an odd prime and $n \geq 2$. Suppose 2 is a primitive root modulo p^n. Then there are exactly n nonzero 2-cyclotomic cosets modulo p^n with $|C_1| = p^n - p^{n-1}, |C_p| = p^{n-1} - p^{n-2}, \ldots, |C_{p^{n-1}}| = p - 1$.*

Proof. We prove the theorem by using mathematical induction on the integer n. The case $n = 2$ is Theorem 2.1. Suppose there are exactly k nonzero cyclotomic cosets modulo p^k provided 2 is a primitive root modulo p^k for some $k \geq 2$. Hence, when $n = k + 1$, we suppose that 2 is a primitive root modulo p^{k+1}. Now, we have k nonzero cyclotomic cosets modulo p^k, that is $C_1, C_p, ...$ and $C_{p^{k-1}}$ and in total there are $p^k - 1$ elements. When we increase the modulus from p^k to p^{k+1}, the same elements remain in the same cyclotomic cosets but the same cosets can now be filled with more elements from $\mathbb{Z}_{p^{k+1}}$. Suppose there are exactly k nonzero cyclotomic cosets modulo p^{k+1}. Then, $C_1 \cup C_p \cup C_{p^2} \cup ... \cup C_{p^{k-1}} = \mathbb{Z}_{p^{k+1}} = \{1, 2, ..., p^k, p^k + 1, ..., p^{k+1} - 1\}$. Since $p^k + 1$ is not a multiple of p so it is not in $C_p, C_{p^2}, ...,$ and $C_{p^{k-1}}$ and it is forced to be in C_1. Same goes for other elements in $\mathbb{Z}_{p^{k+1}}$ that is not a multiple of p must be in C_1. Then, the remaining elements $..., p^k + p, .., p^k + p^2, ..., p^k + p^k, ...$ which are multiples of p modulo p^{k+1} will be in their respective cyclotomic cosets modulo p^{k+1}. For instance, $p^k + p^2 \in C_{p^2}$ but not in C_p as it is in the form of $2^j p^2$ instead of $2^i p$ for some integers i and j. However, p^k is not in any of these k cyclotomic cosets. Suppose we assume that p^k is in one of these cosets. Then, $p^k = p^l 2^t$ for some integers l and t where $0 \leq l < k$, implies $p^{k-l} = 2^t$ which is a contradiction. Hence, p^k must be the smallest integer which is in the new cyclotomic coset modulo p^{k+1} which is C_{p^k}. This contradicts the assumption that there are exactly k nonzero cyclotomic cosets modulo p^{k+1}. Hence, there are exactly $k + 1$ nonzero cyclotomic cosets modulo p^{k+1}. □

2.2 2-cyclotomic cosets modulo p^n when the order of 2 modulo p is $\frac{p-1}{2}$

Throughout this section, we let q be an odd prime such that $q < p$.

Theorem 2.4. *Suppose* 2 *has order* $\frac{p-1}{2}$ *modulo p and* 2 *has order* $\frac{p(p-1)}{2}$ *modulo* p^2*. Then there are exactly* 4 *distinct nonzero 2-cyclotomic cosets modulo* p^2*.*

Proof. The cyclotomic coset modulo p^2 with coset representative 1 is $C_1 = \{1, 2, 2^2, ..., 2^{t_1-1}\}$, where $2^{t_1} \equiv 1(\text{mod}\, p^2)$. Since 2 has order $\frac{p(p-1)}{2}$ modulo p^2, then we have $t_1 = \frac{p(p-1)}{2}$ and so $|C_1| = \frac{p(p-1)}{2}$. Next, we construct the second 2-cyclotomic coset modulo p^2 with coset representative q, that is, $C_q = \{q, 2q, 2^2q, ..., 2^{t_q-1}q\}$, where $2^{t_q}q \equiv q(\text{mod}\, p^2)$. The condition $2^{t_q}q \equiv q(\text{mod}\, p^2)$ can be reduced to $2^{t_q} \equiv 1(\text{mod}\, p^2)$, which gives us $t_q = \frac{p(p-1)}{2}$ and so $|C_q| = \frac{p(p-1)}{2}$. Now, since $p \notin C_1$ and $p \notin C_q$, we consider $C_p = \{p, 2p, ..., 2^{t_p-1}p\}$, where $2^{t_p}p \equiv p(\text{mod}\, p^2)$ which is equivalent to $2^{t_p} \equiv 1(\text{mod}\, p)$. Since 2 has order $\frac{p-1}{2}$ modulo p, then we have $t_p = \frac{p-1}{2}$ and so $|C_p| = \frac{p-1}{2}$. Finally, as $pq \notin C_1 \cup C_q \cup C_p$, we construct the last nonzero cyclotomic coset modulo p^2, that is, $C_{pq} = \{pq, 2pq, ..., 2^{t_{pq}-1}pq\}$, where $2^{t_{pq}}pq \equiv pq(\text{mod}\, p^2)$ which is equivalent to $2^{t_{pq}}q \equiv q(\text{mod}\, p)$. Similarly, it can be reduced to $2^{t_{pq}} \equiv 1(\text{mod}\, p)$. Then, we have $t_{pq} = \frac{p-1}{2}$ and so $|C_{pq}| = \frac{p-1}{2}$. Combining all above, we have $|C_1| = |C_q| = \frac{p(p-1)}{2}$ and $|C_p| = |C_{pq}| = \frac{p-1}{2}$. Clearly, $|C_1| + |C_q| + |C_p| + |C_{pq}| + |\{0\}| = |\mathbb{Z}_{p^2}|$. Hence, C_1, C_q, C_p and C_{pq} are the required 2-cyclotomic cosets modulo p^2. □

Theorem 2.5. *Suppose* 2 *has order* $\frac{p^2(p-1)}{2}$ *modulo* p^3*. Then there are exactly* 6 *distinct nonzero 2-cyclotomic cosets modulo* p^3*.*

Proof. Given 2 has order $\frac{p^2(p-1)}{2}$ modulo p^3, then $2^{\frac{p^2(p-1)}{2}} \equiv 1(mod\, p^3)$. Thus, by Euler's Theorem together with the definition of order, we have $2^{\frac{p(p-1)}{2}}$ $\equiv 1(mod\, p^2)$ and $2^{\frac{p-1}{2}} \equiv 1(mod\, p)$. The cyclotomic coset modulo p^3 with coset representative 1 is $C_1 = \{1, 2, 2^2, ..., 2^{t_1-1}\}$, where $2^{t_1} \equiv 1(\text{mod}\, p^3)$. From above, we have $t_1 = \frac{p^2(p-1)}{2}$ and so $|C_1| = \frac{p^2(p-1)}{2}$. Clearly, $q \notin C_1$, so we construct $C_q = \{q, 2q, 2^2q, ..., 2^{t_q-1}q\}$, where $2^{t_q}q \equiv q(\text{mod}\, p^3)$ which can be reduced to $2^{t_q} \equiv 1(\text{mod}\, p^3)$, which gives us $t_q = \frac{p^2(p-1)}{2}$ and so $|C_q| = \frac{p^2(p-1)}{2}$. Since $p \notin C_1 \cup C_q$, we consider the third cyclotomic coset modulo p^3 with coset representative p, that is, $C_p = \{p, 2p, 2^2p, ..., 2^{t_p-1}p\}$, where $2^{t_p}p \equiv p(\text{mod}\, p^3)$ which is equivalent to $2^{t_p} \equiv 1(\text{mod}\, p^2)$. Then, we have $t_p = \frac{p(p-1)}{2}$ and so $|C_p| = \frac{p(p-1)}{2}$. Similar to previous theorem, we consider pq which is not in all the previous cyclotomic cosets and construct $C_{pq} = \{pq, 2pq, ..., 2^{t_{pq}-1}pq\}$, where $2^{t_{pq}}pq \equiv pq(\text{mod}\, p^3)$ which is equivalent to $2^{t_{pq}}q \equiv q(\text{mod}\, p^2)$. It can be reduced to $2^{t_{pq}} \equiv 1(\text{mod}\, p^2)$. Then, we have that $t_{pq} = \frac{p(p-1)}{2}$ and so $|C_{pq}| = \frac{p(p-1)}{2}$. As $p^2 \notin C_1 \cup C_q \cup C_p \cup C_{pq}$, we consider $C_{p^2} = \{p^2, 2p^2, ..., 2^{t_{p^2}-1}p^2\}$, where $2^{t_{p^2}}p^2 \equiv p^2(\text{mod}\, p^3)$ which is equivalent to $2^{t_{p^2}} \equiv 1(\text{mod}\, p)$ and so $t_{p^2} = \frac{p-1}{2}$. Finally, we consider p^2q as it is also not in the other five cyclotomic cosets modulo p^3 and hereby construct the cyclotomic coset modulo p^3, that is, $C_{p^2q} = \{p^2q, 2p^2q, ..., 2^{t_{p^2q}-1}p^2q\}$, where $2^{t_{p^2q}}p^2q \equiv p^2q(\text{mod}\, p^3)$ which is equivalent to $2^{t_{p^2q}}q \equiv q(\text{mod}\, p)$. Later, it is reduced to $2^{t_{p^2q}} \equiv 1(\text{mod}\, p)$. Clearly, we see that $t_{p^2q} = \frac{p-1}{2} = |C_{p^2q}|$. The sum of all six cyclotomic cosets with $|\{0\}|$ gives the size of $\mathbb{Z}_{p^3}$. Hence, the required 2-cyclotomic cosets modulo p^3 are $C_1, C_q, C_p, C_{pq}, C_{p^2}$, and C_{p^2q}. □

2.3 $s-$ cyclotomic cosets modulo pq

Theorem 2.6. *Let p, q, s be distinct odd primes. Suppose s is a primitive root modulo q and s is a primitive root modulo p. Then there are $2 + h$ distinct nonzero s-cyclotomic cosets modulo pq, where $h = gcd(p-1, q-1)$. Furthermore, the h distinct nonzero s-cyclotomic cosets modulo pq have size $m = lcm(p-1, q-1)$.*

Proof. We fisrt construct the s-cyclotomic coset modulo pq which contains p. Note that $C_p = \{p, sp, \ldots, s^{t_p-1}p\}$, where $s^{t_p}p \equiv p(modpq)$. The condition $s^{t_p}p \equiv p(modpq)$ implies $q \mid s^{t_p} - 1$. Since s has order $q-1$ modulo q, then we have $t_p = q-1$ and so $|C_p| = q-1$. A similar argument shows that $|C_q| = p-1$. Next, we consider any $a \in \{1, 2, \ldots, pq\}$ with $gcd(a, pq) = 1$. Then, we have $gcd(a, p) = 1$ and $gcd(a, q) = 1$. The s-cyclotomic coset modulo n containing a is $C_a = \{a, as, \ldots, as^{t_a} - 1\}$, where $s^{t_a}a \equiv a(modpq)$. The choice of a ensures that $pq \mid s^{t_a} - 1$ which implies $p \mid s^{t_a} - 1$ and $q \mid s^{t_a} - 1$. Since s is a primitive root modulo p and modulo q, we obtain $t_a = lcm(p-1, q-1)$. Therefore, $|C_a| = lcm(p-1, q-1)$ for any $a \in \{1, 2, \ldots, pq\}$ with $gcd(a, pq) = 1$. Finally, we let h be the number of s-cyclotomic cosets modulo pq containing a and note that $|C_0| = |\{0\}| = 1$. Hence, $|C_0| + |C_p| + |C_q| + \sum_a |C_a| = pq$. We then have $1 + (q-1) + (p-1) + h.lcm(p-1, q-1) = pq$ and so $h = \frac{pq-q-p-1}{lcm(p-1,q-1)} = gcd(p-1, q-1)$. □

Theorem 2.7. *Let p, q, s be distinct odd primes. Suppose s has order $\frac{q-1}{2}$ modulo q and s has order $\frac{p-1}{2}$ modulo p. Then there are $4+h$ distinct nonzero s-cyclotomic cosets modulo pq, where $h = \frac{(p-1)(q-1)}{lcm(\frac{q-1}{2}, \frac{p-1}{2})}$. Furthermore, the h distinct nonzero s-cyclotomic cosets modulo pq have size $m = lcm(\frac{q-1}{2}, \frac{p-1}{2})$.*

Proof. Since s has order $\frac{q-1}{2}$ modulo q , then we have $|C_p| = \frac{q-1}{2}$. Also, as s has order $\frac{p-1}{2}$ modulo p, then we have $|C_q| = \frac{p-1}{2}$. Next, let k be a prime such that $q \nmid k$. The s-cyclotomic coset modulo pq containing kp is $C_{kp} = \{kp, skp, s^2kp, \ldots, s^{t_{kp}-1}k_p\}$, where $s^{t_{kp}}k_p \equiv kp(modpq)$ which implies $q \mid (s^{t_{kp}} - 1)$ and so $t_{kp} = \frac{q-1}{2}$. Thus, $|C_{kp}| = \frac{q-1}{2}$. Similarly, if $p \nmid j$, then $|C_{jq}| = \frac{p-1}{2}$. Next, we let h be the nonzero s-cyclotomic cosets modulo pq of size m, then we have $1 + 2.\frac{q-1}{2} + 2.\frac{p-1}{2} + h.m = pq$, that is, $hm = (p-1)(q-1)$. Finally, consider any $a \notin \{p, q, kp, jq\}$. The s-cyclotomic coset modulo pq containing a is $C_a = \{a, as, as^2, \ldots, as^{t_a-1}\}$, where $as^{t_a} \equiv a(modpq)$ which is equivalent to $pq|(s^{t_a} - 1)$. Thus, $t_a = m = lcm(\frac{q-1}{2}, \frac{p-1}{2})$ and so $h = \frac{(p-1)(q-1)}{m}$. □

The following theorem can be proved in the similar way.

Theorem 2.8. *Let p, q, s be distinct odd primes. Suppose s has order $\frac{q-1}{2}$ modulo q and s is a primitive root modulo p. Then there are $3 + h$ distinct nonzero s-cyclotomic cosets modulo pq, where $h = \frac{(p-1)(q-1)}{lcm(\frac{q-1}{2},p-1)}$. Furthermore, the h distinct nonzero s-cyclotomic cosets modulo pq have size $m = lcm(\frac{q-1}{2}, p-1)$.*

Remark 2.1. *Note that not much know results on 3-cyclotomic cosets modulo $2^n rs$, where r, s are distinct primes greater than 3, and $n \geq 1$. Suppose 3 is a primitive root modulo r and is also* *a primitive root modulo s. Furthermore, $gcd(\phi(r), \phi(s)) = 2$. Then, by the help of computer we obtain the four results in which we state the results without any proof; there are 9 3-cyclotomic cosets modulo $2rs$, there are 18 3-cyclotomic cosets modulo 2^2rs, there are 36 3-cyclotomic cosets modulo 2^3rs and there are 66 3-cyclotomic cosets modulo 2^4rs.*

3 CODES AND SECRET SHARING

In this section, we construct binary cyclic codes of length 49 by using the cyclotomic cosets obtained from the previous section. From Theorem 2.1, we consider the case $p = 7$. Then, all 2-cyclotomic cosets modulo 49 are

$$\begin{aligned}
&C_0 = \{0\},\\
&C_1 = \{1, 2, 4, 8, 16, 32, 15, 30, 11, 22, 44, 39, 29, 9, 18, 36, 23, 46, 43, 37, 25\},\\
&C_3 = \{3, 6, 12, 24, 48, 47, 45, 41, 33, 17, 34, 19, 38, 27, 5, 10, 20, 40, 31, 13, 26\},\\
&C_7 = \{7, 14, 28\} \text{ and}\\
&C_{21} = \{21, 42, 35\}.
\end{aligned}$$

In term of group ring $\mathbb{F}_2[\mathbb{Z}_{49}]$, we let $\Omega_1 = \sum_{s \in C_1} g^s$, $\Omega_2 = \sum_{r \in C_q} g^r$, $\Omega_3 = \sum_{t \in C_p} g^t$ and $\Omega_4 = \sum_{v \in C_{pq}} g^v$. Then, we have

$$\begin{aligned}
\Omega_0 =& g^0,\\
\Omega_1 =& g^1 + g^2 + g^4 + g^8 + g^{16} + g^{32} + g^{15} + g^{30} + g^{11} + g^{22} + g^{44} +\\
& g^{39} + g^{29} + g^9 + g^{18} + g^{36} + g^{23} + g^{46} + g^{43} + g^{37} + g^{25},\\
\Omega_2 =& g^3 + g^6 + g^{12} + g^{24} + g^{48} + g^{47} + g^{45} + g^{41} + g^{33} + g^{17} + g^{34} +\\
& g^{19} + g^{38} + g^{27} + g^5 + g^{10} + g^{20} + g^{40} + g^{31} + g^{13} + g^{26},\\
\Omega_3 =& g^7 + g^{14} + g^{28} \text{ and } \Omega_4 = g^{21} + g^{42} + g^{35}.
\end{aligned}$$

Secret sharing scheme is used to break down a secret into smaller portions call shares and later distributed to other participants by a dealer. The group of participants who hold the shares that can reconstruct the secret is called the access set. If a group of participants can recover the secret by combining their shares, then any group of participants containing this group can also recover the secret. The group of participants is known as the minimal access set if they can recover the secret with their shares, while any of its proper subgroups cannot do so. The following is a secret sharing scheme constructed based on balanced incomplete block design which gives a nice access structure. (X, A) is called a $(\nu, \kappa, \lambda) - BIBD$ if and only if $|X| = \nu$, for any nonempty subset $B \in A$ called block, $|B| = \kappa$, and any pair of distinct points occur in exactly λ blocks. Furthermore, we know that $|A| = \frac{\nu}{\kappa} \frac{\lambda(\nu-1)}{\kappa-1}$ and Every point occur in exactly $\frac{\lambda(\nu-1)}{\kappa-1}$ blocks. We observed that the secret sharing scheme based on a $(\nu, \kappa, \lambda) - BIBD$ is for sharing secrets among ν participants. There are $\frac{\nu}{\kappa} \frac{\lambda(\nu-1)}{\kappa-1}$ minimal access sets. Each minimal access set consists of κ participants, and each participants is a member of exactly $\frac{\lambda(\nu-1)}{\kappa-1}$ minimal access set.

Example 3.1. *Let $\Omega_1 = \sum_{s \in C_1} g^s \in \mathbb{F}_2[\mathbb{Z}_{49}]$. In term of group ring, the cyclotomic cosets of Ω_1 is $\Omega_1 = g^1 + g^2 + g^4 + g^8 + g^{16} + g^{32} + g^{15} + g^{30} + g^{11} + g^{22} + g^{44} + g^{39} + g^{29} + g^9 +$*

$g^{18} + g^{36} + g^{23} + g^{46} + g^{43} + g^{37} + g^{25}$. *The generator polynomial for* $< \Omega_1 >$ *is computed as follows:*

$$\begin{aligned} g(x) =& gcd(g^{49} - 1, \Omega_1) \\ =& 1 + g + g^3 + g^7 + g^8 + g^{10} + g^{14} + g^{15} + g^{17} + g^{21} + g^{22} + g^{24} + \\ & g^{28} + g^{29} + g^{31} + g^{35} + g^{36} + g^{38} + g^{42} + g^{43} + g^{45}. \end{aligned}$$

From above, we see that the dimension $k = 4$ *and the* 16 *codewords are as follows:*

0000000000000000000000000000000000000000000000000,
1111111111111111111111111111111111111111111111111,
1101000110100011010001101000110100011010001101000,
0110100011010001101000110100011010001101000110100,
0011010001101000110100011010001101000110100011010,
0001101000110100011010001101000110100011010001101,
1000110100011010001101000110100011010001101000110,
0100011010001101000110100011010001101000110100011,
1010001101000110100011010001101000110100011010001,
0101110010111001011100101110010111001011100101110,
0010111001011100101110010111001011100101110010111,
1001011100101110010111001011100101110010111001011,
1100101110010111001011100101110010111001011100101,
1110010111001011100101110010111001011100101110010,
0111001011100101110010111001011100101110010111001,
1011100101110010111001011100101110010111001011100.

From above, the minimum distance is 21. Hence, we have a $[49, 4, 21]$*-binary code. For the secret sharing scheme based on* $[49, 4, 21]$*-code, the* 7 *access sets are as follows:*

$$\begin{aligned} & \{1, 3, 7, 8, 10, 14, 15, 17, 21, 22, 24, 28, 29, 31, 35, 36, 38, 42, 43, 45\}, \\ & \{4, 5, 7, 11, 12, 14, 18, 19, 21, 25, 26, 28, 32, 33, 35, 39, 40, 42, 46, 47\}, \\ & \{2, 6, 7, 9, 13, 14, 16, 20, 21, 23, 27, 28, 30, 34, 35, 37, 41, 42, 44, 48\}, \\ & \{1, 4, 6, 7, 8, 11, 13, 14, 15, 18, 20, 21, 22, 25, 27, 28, 29, 32, 34, 35, 36, 39, \\ & 41, 42, 43, 46, 48\}, \{1, 2, 5, 7, 8, 9, 12, 14, 15, 16, 19, 21, 22, 23, 26, 28, 29, \\ & 30, 33, 35, 36, 37, 40, 42, 43, 44, 47\}, \{2, 3, 4, 7, 9, 10, 11, 14, 16, 17, 18, 21, \\ & 23, 24, 25, 28, 30, 31, 32, 35, 37, 38, 39, 42, 44, 45, 46\}, \{3, 5, 6, 7, 10, 12, 13, \\ & 14, 17, 19, 20, 21, 24, 26, 27, 28, 31, 33, 34, 35, 38, 40, 41, 42, 45, 47, 48\}. \end{aligned}$$

Participants $7, 14, 21, 28, 35, 42$ *appears in all access sets. Hence, any group who can determine the secret must include these* 6 *participants. The remaining participants must be in exactly* $k - 1 = 3$ *access sets.*

Example 3.2. *We use* $\langle \Omega_1 + \Omega_2 \rangle$ *which is a [*$49, 6, 14$*]- code to construct a secret sharing scheme. As the dimension is* 6*, there are* 64 *codewords. The dealer distributes the shares of the* *secret among* 48 *participants. The* $k = 6$ *access sets are as follows:*

$$\begin{gathered}\{1, 7, 8, 14, 15, 21, 22, 28, 29, 35, 36, 42, 43\},\\ \{2, 7, 9, 14, 16, 21, 23, 28, 30, 35, 37, 42, 44\},\\ \{3, 7, 10, 14, 17, 21, 24, 28, 31, 35, 38, 42, 45\},\\ \{4, 7, 11, 14, 18, 21, 25, 28, 32, 35, 39, 42, 46\},\\ \{5, 7, 12, 14, 19, 21, 26, 28, 33, 35, 40, 42, 47\},\\ \{6, 7, 13, 14, 20, 21, 27, 28, 34, 35, 41, 42, 48\}.\end{gathered}$$

Clearly, participants $7, 14, 21, 28, 35, 42$ *appears in all access sets. Hence, any group who can determine the secret must include these* 6 *participants. Each participant in the set* $\{1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 44, 45, 46, 47, 48\}$ *is in exactly* 1 *access set. The number of participants in each access set is* $d - 1 = 14 - 1 = 13$.

Example 3.3. *For secret sharing based on [*$49, 3, 28$*]-code, secret are distributed as shares to* 48 *participants by a dealer. The minimal codeword are listed below:*

$$\begin{gathered}\{0000000000000000000000000000000000000000000000000,\\ 0111010011101001110100111010011101001110100111010,\\ 0011101001110100111010011101001110100111010011101,\\ 0100111010011101001110100111010011101001110100111,\\ 1001110100111010011101001110100111010011101001110,\\ 1010011101001110100111010011101001110100111010011,\\ 1101001110100111010011101001110100111010011101001,\\ 1110100111010011101001110100111010011101001110100\}\end{gathered}$$

In the secret sharing constructed based on $[49, 3, 28]$*- cyclic code, the* 4 *access sets are as follows:*

$$\begin{gathered}\{3, 4, 5, 7, 10, 11, 12, 14, 17, 18, 19, 21, 24, 25, 26, 28, 31, 32, 33, 35, 38, 39, 40, 42, 45, 46, 47\},\\ \{2, 5, 6, 7, 8, 12, 13, 14, 16, 19, 20, 21, 23, 26, 27, 28, 30, 33, 34, 35, 37, 40, 41, 42, 44, 47, 48\},\\ \{1, 3, 6, 7, 8, 10, 13, 14, 15, 17, 20, 21, 22, 24, 27, 28, 29, 31, 34, 35, 36, 38, 41, 42, 43, 45, 48\},\\ \{1, 2, 4, 7, 8, 9, 11, 14, 15, 16, 18, 21, 22, 23, 25, 28, 29, 30, 32, 35, 36, 37, 39, 42, 43, 44, 46\}.\end{gathered}$$

Participants $7, 14, 21, 28, 35, 42$ *appears in all access sets. Hence, any group who can determine the secret must include these* 6 *participants. Each participant in the set* $\{1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 44, 45, 46, 47, 48\}$ *is in exactly* $k - 1 = 2$ *access sets. In each access set, there are* 27 *participants. Such a secret sharing scheme can be useful in big corporate where there are a few major shareholders to make a decision.*

4 CONCLUSIONS

In this paper, all 2-cyclotomic cosets modulo p^n are constructed, when 2 is a primitive root modulo p^n and when the order of 2 is $\frac{p-1}{2}$ modulo p. Also, some results on s-cyclotomic cosets modulo pq are obtained for three possible orders of s modulo p and q, respectively, for distinct odd primes p, q. We also used these families of cyclotomic cosets to construct some codes of length 49 with different minimum distance and dimension, and hence used these codes to define some secret sharing together with their corresponding access structures. The initial secret sharing results above are not comprehensive, further work in subsequent papers should provide a more substantiate outcome.

REFERENCES

A. Sahni, P. T. S. (2012). Minimal cyclic codes of length $p^n q$. *Finite fields and applications*, 18:1017–1036.

A. Sharma, G. B. (2012). The weight distribution of some irreducible cyclic codes. *Finite fields and applications*, 18:144–159.

F.J. MacWilliams, N. S. (1977). The theory of error-correcting codes. *North-Holland.*

K. Singh, S. A. (2010). The primitive idempotents in $fc_(2^n)$. *International Journal of algebra*, 4:1231–1241.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423.

Sharma, A., Bakshi, G. K., Dumir, V., and Raka, M. (2004). Cyclotomic numbers and primitive idempotents in the ring $GF(q)[x]/(xpn-1)$. *Finite Fields and Their Applications*, 10(4):653–673.

S.K. Arora, M. P. (1997). Minimal codes of prime-power length. *Finite fields and applications*, 3:99–113.

S.K. Arora, M. P. (1999). Minimal cyclic codes of length $2p^n$. *Finite fields and applications*, 5:177–187.

S.K. Arora, S. Batra, S. C. (2002). The primitive idempotents of a cyclic group algebra. *Southeast Asian Bulletin of Mathematics*, 26:549–557.

Analysis on the AA_β Cryptosystem

Muhammad Asyraf Asbullah^{*1,2} and **Muhammad Rezal Kamel Ariffin**^{1,2}

[1]*Al-Kindi Cryptography Research Laboratory, Institute for Mathematical Research, Universiti Putra Malaysia*
[2]*Department of Mathematics, Faculty of Science, Universiti Putra Malaysia*

E-mail: ma_asyraf@upm.edu.my, rezal@upm.edu.my
**Corresponding author*

ABSTRACT

In this paper, we present three types of algebraic analysis upon the AA_β cryptosystem. First, we give the congruence relation in order to solve the AA_β equation. For the second and third analysis, we bring in the continued fraction's method and the Coppersmith's theorems, which presents several potential ways to retrieve the prime factor of p and q from the AA_β public keys or the plaintext m from the AA_β ciphertext, respectively. Thus, based on such analysis, suggestions are offered as a counter measure on how to secure the AA_β cryptosystem during key generation and encryption process.

Keywords: AA_β cryptosystem, congruence relation, Legendre's theorem, Coppersmiths's method

1 INTRODUCTION

In 2013, a new public key cryptosystem namely the AA_β cryptosystem was proposed by Ariffin et al. (2013). The invention of the AA_β public key cryptosystem combines the concept of the Bivariate Function Hard Problem together with the integer factorization problem of the modulus of $N = p^2q$ (Ariffin et al., 2013). The AA_β cryptosystem was able to utilize the square root modulo problem and managed to overcome the decryption failure scenario exhibited by the Rabin cryptosystem (Rabin, 1979). It was proven analytically to be able to decrypt correctly without failure. This ability of the AA_β cryptosystem in 2013, concluded a long journey by various authors in trying to overcome the decryption failure scenario of the Rabin cryptosystem. Among them were Williams (1980), Kurosawa et al. (1988), Menezes et al. (1997) and Kurosawa et al. (2001).

The advantages for AA_β cryptosystem is exhibited by its encryption algorithm that does not involve complicated arithmetic operations, for instance, such as division, modular multiplication or exponentiation. Only basic multiplication and addition is required. Moreover, the decryption method is able to produce a unique solution without engaging with any padding or redundancies, while still occupying the Rabin primitive (Asbullah and Ariffin, 2014). In addition, AA_β cryptosystem acquired the quality to secure large data sets.

Our contributions. In this paper, we put forward rigorous mathematical analyses conducted upon the AA_β cryptosystem. First, we present the congruence relation of the AA_β equation and showing that to solve such congruence relation is currently infeasible. Secondly, we showed the algebraic analysis using the continued fraction's method by manipulating the AA_β public keys and recover its prime factor; p and q. The third analysis using the Coppersmith's theorems upon the AA_β ciphertext that present several possible ways to recover the plaintext m. Thus, several suggestions are provided on how to secure the AA_β cryptosystem during its key generation and encryption process.

Paper Organization. The remainder of the paper is structured as follows. In Section 2, we start with the description of the AA_β cryptosystem, followed by the Legendre's theorem and the Coppersmith's technique. In Section 3, we present the algebraic analysis, namely the congruence relation, the Legendre's theorem and the Coppersmith's method upon the AA_β cryptosystem. Finally, we conclude in Section 4.

2 PRELIMINARIES

In this section we start with the description of the AA_β cryptosystem. We then introduce the basic facts about the Legendre's theorem and the Coppersmith's method that are used in our analysis.

2.1 AA_β Cryptosystem

First of all, we will review the AA_β cryptosystem which is proposed earlier by Ariffin et al. (2013). We now describe the key generation, encryption and decryption procedure of AA_β cryptosystem as follows.

Algorithm 1 AA_β Key Generation Algorithm

```
Input: The size $k$ of the security parameter
Output: The public key $A_1, A_2$ and the private key $d, p$
    1. Choose two random and distinct primes $p$ and $q$ such that $2^k < p, q < 2^{k+1}$ satisfy $p, q \equiv 3 \pmod 4$
    2. Compute $A_2 = p^2 q$
    3. Compute a random integer $A_1$ such that $2^{3k+4} < A_1 < 2^{3k+6}$
    4. Compute an integer $d$ such that $A_1 d \equiv 1 \pmod{A_2}$
    5. Return the public key $A_1, A_2$ and the private key $d, p$
```

Algorithm 2 AA_β Encryption Algorithm

```
Input: The plaintext $m, t$ and the public key $A_1, A_2$
Output: A ciphertext $c$
  1. Choose a plaintext $2^{2k-2} < m < 2^{2k-1}$ such that $\gcd(m, A_2) = 1$
  2. Choose a plaintext $t$ such that $2^{4k} < t < 2^{4k+1}$
  3. Compute $c = A_1 m^2 + A_2 t$
  4. Return the ciphertext $c$
```

Algorithm 3 AA_β Decryption Algorithm (Asbullah and Ariffin, 2014)

```
Input: A ciphertext $c$ and the private key $d, p, q$
Output: The plaintext $m, t$
  1. Compute $w \equiv cd \pmod{A_2}$
  2. Compute $m_p \equiv w^{\frac{p+1}{4}} \pmod{p}$
  3. Compute $m_q \equiv w^{\frac{q+1}{4}} \pmod{q}$
  4. Proceed to solve $m_p \pmod{p}$ and $m_q \pmod{q}$ using Garner's algorithm to obtain the list $m_i$ for $i = 1, 2, 3, 4$
  5. Compute $t_i = \frac{c - A_1 m_i^2}{A_2}$ for $m_i < 2^{2k-1}$ for $i = 1, 2, 3, 4$
  6. Sort the pair $(m_i, t_i)$ for integer $t_i$, else reject
  7. Return the plaintext $m, t$
```

2.2 Legendre's theorem

In this section, we show the Legendre's theorem based on continued fractions as follows.

Theorem 2.1 (Legendre's Theorem). *(Hardy and Wright, 1965) Let R is a rational number. Let $x, y \in \mathbb{Z}, y \neq 0$ and $\gcd(x, y) = 1$. Suppose $\left|R - \frac{x}{y}\right| < \frac{1}{2y^2}$, then $\frac{x}{y}$ is a convergent of the continued fraction expansion of R.*

The theorem simply says that the unknown integers x and y can be retrieved from the list of continued fraction expansion of a rational number R satisfying the given inequality. We remark that the theory of continued fractions is one of the very important technique used in the analysis upon a public key cryptosystem. For instance, see Nitaj (2011) and Asbullah and Ariffin (2015).

2.3 Coppersmith's Theorem

In general, finding solutions to modular equations is easy if we know the factorization of the modulus. Else, it can be difficult. Consequently, a significantly powerful method for finding small roots of modular polynomial equations was invented by Coppersmith (1997). When working with modulo of a prime number, there is no reason to use the Coppersmith's theorem since there exist far better root-finding algorithm (for instance, Newton method), yet in cryptography we usually deal with a number of the product of primes (Galbraith, 2012). Moreover, this method has found many different applications in the area of cryptography and a vastly useful tool for cryptanalysis (Nitaj, 2013). We immediately provide the Coppersmith's theorem as follows.

Theorem 2.2. *(Coppersmith, 1997) Let N be an integer of unknown factorization. Let $f_N(x)$ be a univariate, a monic polynomial of degree δ. Then we can find all solutions x_0 for the equation $f_N(x) \equiv 0 \pmod{N}$ with $|x_0| < N^{\frac{1}{\delta}}$ in polynomial time.*

Theorem 2.3. *(May, 2003) Let N be an integer of unknown factorization, which has a divisor $b > N^{\beta}$. Furthermore, let $f_b(x)$ be a univariate, a monic polynomial of degree δ.Then we can find all solutions x_0 for the equation $f_b(x) \equiv 0 \pmod{b}$ with $|x_0| < \frac{1}{2}N^{\frac{\beta^2}{\delta}}$ in polynomial time.*

3 ANALYSIS AND DISCUSSION

In this section, we begin the analysis of the AA_β cryptosystem that focuses on the the congruence relation. Further analysis on the AA_β cryptosytem is then given in subsequence subsection. These include the analysis using the continued fraction's and the Coppersmith method upon the on the AA_β ciphertext equation.

3.1 Congruence Relation

Consider the Algorithm 1. Observe that the integer d is easily computed, nonetheless without the prime factors of A_2, we only ended up with the congruence relation of $cd \equiv m^2 \pmod{A_2}$ where $A_2 = p^2q$. Thus to solve the congruence $m^2 \pmod{A_2}$ reduces to solve the integer factorization problem, which is currently infeasible. Now, since the $\gcd(A_1, A_2) = 1$ then exist a unique integer d' such that $A_2d' \equiv 1 \pmod{A_1}$. In this section, we show that such integer d' can be use to solve the AA_β equation, but it is still far from feasible.

Theorem 3.1. *Suppose $c = A_1m^2 + A_2t$ be the AA_β equation. Let d' such that $A_2d' \equiv 1 \pmod{A_1}$. If at minimum 2^{k-4} is exponentially large, then it is infeasible to determine m^2 or t from its congruence relation.*

Proof. Let $c = A_1m^2 + A_2t$ be the AA_β equation. Since the $\gcd(A_1, A_2) = 1$ thus there exist the integer d' such that $A_2d' \equiv 1 \pmod{A_1}$. Suppose we take $cd' \equiv t \pmod{A_1}$. Set $a \equiv t \pmod{A_1}$. Then there exist for integer j such that

$$t = a + A_1j \tag{1}$$

Substitute (1) into $c = A_1m^2 + A_2t$, we obtain $c = A_1m^2 + A_2t = A_1m^2 + A_2(a + A_1j)$. Then we have $m^2 = \frac{c-A_2(a+A_1j)}{A_1} = \frac{c-A_2a}{A_1} - A_2j$. Note that $m^2 \in \mathbb{Z}$ also implies $\frac{c-A_2a}{A_1} \in \mathbb{Z}$. Hence setting $b = \frac{c-A_2a}{A_1}$, it follows that we have construct two parametric equation $t = a + A_1j$

and $m^2 = b - A_2j$ for $c = A_1m^2 + A_2t$. However, it is suffice only to find the integer j for $m^2 = b - A_2j$ such that $j = \frac{b-m^2}{A_2}$ satisfying $2^{4k-4} < m^2 < 2^{4k-2}$ and $\sqrt{b - A_2j} \in \mathbb{Z}$. We know that $2^{3k} < A_2 < 2^{3k+3}$. Hence we deduce that j should be in the range of

$$\frac{b - 2^{4k-2}}{2^{3k}} < j < \frac{b - 2^{4k-4}}{2^{3k}}$$

Therefore the difference between the upper and the lower bound of j is

$$\begin{aligned} \frac{b - 2^{4k-4}}{2^{3k}} - \frac{b - 2^{4k-2}}{2^{3k}} &= \frac{-2^{4k-4} + 2^{4k+1}}{2^{3k}} \\ &= \frac{(2^2 - 1) \cdot 2^{4k-4}}{2^{3k}} \\ &= 3 \cdot 2^{k-4} \\ &> 2^{k-4} \end{aligned}$$

The difference is very large and finding the correct j is need to sieve through approximately 2^{k-4} possible integer where 2^{k-4} is exponentially large. Hence finding the correct j using this approach is infeasible. □

3.2 Continued Fraction's Method

Suppose A_1 and A_2 are the public parameters from the AA_β cryptosystem. Based on the analysis in this section, we remark that it is important to carefully check for each parameter during the AA_β key generation process.

Theorem 3.2. *Let $A_1 = e_0 + apq$ for some integer e_0 and a. Suppose $\left|\frac{A_1}{A_2} - \frac{a}{p}\right| < \frac{1}{2p^2}$ then $\frac{a}{p}$ is a convergent of the continued fraction expansion of $\frac{A_1}{A_2}$.*

Proof. Consider the value $A_1 = e_0 + apq$ then it can be rewritten as $A_1 \equiv e_0 \pmod{pq}$. Suppose $e_0 \pmod{pq}$ with $e_0 < pq$. If we multiply $A_1 = e_0 + apq$ with p, then we have $A_1p = e_0p + ap^2q = e_0p + aA_1$. Hence

$$\begin{aligned} \left|\frac{A_1}{A_2} - \frac{a}{p}\right| &= \frac{|A_1p - aA_2|}{A_2p} \\ &= \frac{|e_0|}{A_2} \end{aligned}$$

If $\frac{|e_0|}{A_2} < \frac{1}{2p^2}$, that is if $e_0 < \frac{A_2}{2p^2} < \frac{q}{2}$, then by Theorem 2.1, $\frac{a}{p}$ is a convergent of the continued fraction expansion of $\frac{A_1}{A_2}$. This lead to finding p and then q. □

Remark 3.1. *Therefore, we put a remark that $A_1 \equiv e_0 \pmod{pq}$ should be chosen carefully (i.e. $e_0 > \frac{q}{2}$).*

Theorem 3.3. *Let $A_1 = e_1 + bp^2$ for some integer e_1 and b. Suppose $\left|\frac{A_1}{A_2} - \frac{b}{q}\right| < \frac{1}{2q^2}$, then $\frac{b}{q}$ is a convergent of the continued fraction expansion of $\frac{A_1}{A_2}$.*

Proof. Consider the value $A_1 = e_1 + bp^2$ then it can be rewritten as $A_1 \equiv e_1 \pmod{p^2}$. Suppose $e_1 \pmod{p^2}$ with $e_1 < p^2$. If we multiply $A_1 = e_1 + bp^2$ with q, then we have $A_1 q = e_1 q + bp^2 q = e_1 q + bA_1$ Hence

$$\begin{aligned} \left| \frac{A_1}{A_2} - \frac{b}{q} \right| &= \frac{|A_1 q - bA_2|}{A_2 q} \\ &= \frac{|e_1|}{A_2} \end{aligned}$$

If $\frac{|e_1|}{A_2} < \frac{1}{2q^2}$, that is if $e_1 < \frac{A_2}{2q^2} < \frac{p^2}{2q}$, then by Theorem 2.1, $\frac{b}{q}$ is a convergent of the continued fraction expansion of $\frac{A_1}{A_2}$. This lead to finding q and then p. □

Remark 3.2. *Therefore, we put a remark that $A_1 \equiv e_1 \pmod{p^2}$ should be chosen carefully (i.e. $e_1 > \frac{p2}{2q}$).*

3.3 Coppersmith's Method

We now analyze the AA_β cryptosystem based on the Coppersmith's method (i.e.Theorem 2.2 and Theorem 2.3) and obtain the following results.

Proposition 3.1. *Let $c = A_1 m^2 + A_2 t$ be the AA_β ciphertext. Let d such that $A_1 d \equiv 1 \pmod{A_2}$ where $A_2 = p^2 q$. If $m < A_2^{\frac{1}{2}}$, then it can be found in polynomial time.*

Proof. Since there exist an integer d such that $A_1 d \equiv 1 \pmod{A_2}$ where $A_2 = p^2 q$. Compute $w \equiv cd \equiv m^2 \pmod{A_2}$. Consider $f_{A_2}(x) \equiv x^2 - w \equiv 0 \pmod{A_2}$. Consider the Coppersmiths method (i.e. Theorem 2.2) hence $\delta = 2$, the root $x_0 = m$ can be recovered if $m < A_2^{\frac{1}{\delta}} = A_2^{\frac{1}{2}} \approx 2^{\frac{3k}{2}}$. □

Proposition 3.2. *Let $c = A_1 m^2 + A_2 t$ be the AA_β ciphertext. Let $w \equiv m^2 \pmod{p^2}$ such that p^2 is an unknown factor for A_2. If $m < A_2^{\frac{2}{9}}$, then m can be found in polynomial time.*

Proof. Suppose $w \equiv cd \equiv m^2 \pmod{p^2}$ such that p^2 is an unknown factor for A_2. Let $f_{p^2}(x) \equiv x^2 - w \equiv 0 \pmod{p^2}$ with $p^2 \approx 2^{2k} \approx A_2^{\frac{2}{3}}$. Consider the Theorem 2.3. We can find a solution $x_0 = m$ if $m < \frac{1}{2} A_2^{\frac{\beta^2}{\delta}} < A_2^{\frac{(\frac{2}{3})^2}{2}} = A_2^{\frac{2}{9}} \approx 2^{\frac{2k}{3}}$. □

Remark 3.3. *Therefore in order to avoid both attacks, we would set $m > 2^{\frac{3k}{2}}$ in the AA_β encryption algorithm.*

Proposition 3.3. *Let d_0 such that $A_1 d_0 \equiv 1 \pmod{p^2}$ where p^2 is an unknown factor for A_2. If $|d_0| < A_2^{\frac{4}{9}}$ then d_0 can be found in polynomial time.*

Proof. Let d_0 such that $A_1 d_0 \equiv 1 \pmod{p^2}$ where p^2 is an unknown factor for A_2. Consider $f_{p^2}(x) \equiv A_1 x - 1 \equiv 0 \pmod{p^2}$ with $p^2 \approx 2^{2k} \approx A_2^{\frac{2}{3}}$. Thus by applying Theorem 2.3, we can find solution $x_0 = d_0$ if $|d_0| < \frac{1}{2} A_2^{\frac{\beta^2}{\delta}} < A_2^{\frac{(\frac{2}{3})^2}{1}} = A_2^{\frac{4}{9}}$. then d_0 can be found in polynomial time. □

Corollary 3.1. *Let d_1 such that $A_1 d_1 \equiv 1 \pmod{pq}$ where pq is an unknown factor for A_2. If $|d_1| < A_2^{\frac{4}{9}}$ then d_1 can be found in polynomial time.*

Proof. Consider $f_{pq}(x) \equiv A_2 x - 1 \equiv 0 \pmod{pq}$ with $pq > A_2^{\frac{2}{3}}$. Then we reach the same conclusion as the Proposition 3.3. □

The significant of the result from Proposition 3.3 and Corollary 3.1 is that if one is able to compute either d_0 or d_1 then one is able to factor $A_2 = p^2 q$.

Proposition 3.4. *If $d_0 < A_2^{\frac{4}{9}}$ such that $A_1 d_0 \equiv 1 \pmod{p^2}$, then $A_2 = p^2 q$ can be factored in polynomial time.*

Proof. Consider the relation $A_1 d_0 \equiv 1 \pmod{p^2}$. Suppose the integer $d_0 < A_2^{\frac{4}{9}}$ could be computed using Proposition 3.3. Then we have the value $A_1 d_0 - 1 \equiv 0 \pmod{p^2}$ where $A_1 d_0 - 1$ is an integer multiple of p^2. Observe that if we take the $\gcd(A_1 d_0 - 1, A_2)$ resulting p^2, and then $\frac{A_2}{p^2} = q$. The same argument is applicable for $d_1 < A_2^{\frac{4}{9}}$ such that $A_1 d_1 \equiv 1 \pmod{pq}$. □

Corollary 3.2. *If $d_1 < A_2^{\frac{4}{9}}$ such that $A_1 d_1 \equiv 1 \pmod{pq}$, then $A_2 = p^2 q$ can be factored in polynomial time.*

Proof. Consider Proposition 3.4. The same argument is applicable for $d_1 < A_2^{\frac{4}{9}}$ such that $A_1 d_1 \equiv 1 \pmod{pq}$. □

Remark 3.4. *Consider Propostion 3.4 and Corollary 3.2. In order for the AA_β cryptosystem to be resistant against such methods, it is important to check for each $d_0, d_1 > A_2^{\frac{4}{9}}$ during the AA_β key generation process.*

4 SUMMARY

We now summarize the paper. In this paper, we put forward rigorous mathematical analyses conducted upon the AA_β cryptosystem. First, we present the congruence relation of the AA_β equation. We showed that to solve such congruence is infeasible. Secondly, we presented the analysis using the continued fraction's method which applies when $e_0 < \frac{q}{2}$ (or $e_1 < \frac{p2}{2q}$) satisfies an equation $A_1 = e_0 + apq$ (or $A_1 = e_1 + bp^2$), hence obtained the primes p and q. The third analysis is using the Coppersmith's theorems upon the AA_β ciphertext to recover the plaintext m. Finally, we provide a suggestion as a countermeasure during the AA_β key generation and encryption process, respectively.

REFERENCES

Ariffin, M. R. K., Asbullah, M. A., Abu, N. A., and Mahad, Z. (2013). A New Efficient Asymmetric Cryptosystem Based on the Integer Factorization Problem of $N = p^2q$. *Malaysian Journal of Mathematical Sciences*, 7(S):19–37.

Asbullah, M. A. and Ariffin, M. R. K. (2014). Comparative Analysis of Three Asymmetric Encryption Schemes Based Upon the Intractability of Square Roots Modulo $N = p^2q$. In *In the Proceeding of the 4^{th} International Cryptology and Information Security Conference 2014*, pages 86–99.

Asbullah, M. A. and Ariffin, M. R. K. (2015). New Attacks on RSA with Modulus $N = p^2q$ Using Continued Fractions. *Journal of Physics: Conference Series*, 622(1):012019.

Coppersmith, D. (1997). Small Solutions To Polynomial Equations, And Low Exponent RSA Vulnerabilities. *Journal Of Cryptology*, 10(4):233–260.

Galbraith, S. D. (2012). *Mathematics Of Public Key Cryptography*. Cambridge University Press.

Hardy, G. and Wright, E. (1965). *An Introduction to the Theory of Numbers*. Oxford University Press, London.

Kurosawa, K., Ito, T., and Takeuchi, M. (1988). Public Key Cryptosystem Using A Reciprocal Number With The Same Intractability As Factoring A Large Number. *Cryptologia*, 12(4):225–233.

Kurosawa, K., Ogata, W., Matsuo, T., and Makishima, S. (2001). IND-CCA Public Key Schemes Equivalent To Factoring $N = pq$. In *Public Key Cryptography*, pages 36–47. Springer.

May, A. (2003). *New RSA Vulnerabilities Using Lattice Reduction Methods*. PhD thesis, University Of Paderborn.

Menezes, A., Oorschot, P., and Vanstone, S. (1997). *Handbook Of Applied Cryptography*. CRC Press.

Nitaj, A. (2011). A New Vulnerable Class of Exponents in RSA. *JP Journal of Algebra, Number Theory and Applications*, 21(2):203–220.

Nitaj, A. (2013). Diophantine and Lattice Cryptanalysis of the RSA Cryptosystem. In *Artificial Intelligence, Evolutionary Computing and Metaheuristics*, pages 139–168. Springer.

Rabin, M. O. (1979). Digitalized Signatures and Public-Key Functions as Intractable as Factorization. *MIT Technical Report*, MIT/LCS/TR-212.

Williams, H. (1980). A Modification Of The RSA Public-Key Encryption Procedure. *Information Theory, IEEE Transactions On*, 26(6):726–729.

A New Efficient Rabin-like Signature Scheme

Zahari Mahad[*1] and **Muhammad Rezal Kamel Ariffin**[1,2]

[1]*Al-Kindi Cryptography Research Laboratory, Institute for Mathematical Research,*
Universiti Putra Malaysia
[2]*Department of Mathematics, Faculty of Science, Universiti Putra Malaysia*

E-mail: zaharimahad@upm.edu.my, rezal@upm.edu.my
[*]*Corresponding author*

ABSTRACT

In this work, we present a new Rabin-like signature scheme. This new scheme does not utilize Legendre and Jacobi symbols. It also does not use extra information either during signing or verification process. We name this scheme as the Rabin-RZ signature scheme.

Keywords: Rabin signature, Rabin cryptosystem

1 INTRODUCTION

The Rabin digital signature is similar to RSA digital signature but it has a number of advantages. One of the advantages is the use of a small exponent that makes the verification procedure faster than RSA. The Rabin digital signature was developed based on the square root modulo problem (Rabin (1979)). Its security is relies on the difficulty of finding the square roots of a quadratic residue modulus N and it has been proved to be as hard as factorization.

Previous works. Beginning in 1980, several several attempts have been done to develop the most efficient Rabin-like digital signature schemes. Williams (1980), proposed an implementation of the Rabin digital signature scheme by applying the Legendre and Jacobi symbols. However, this scheme has been cryptanalyzed by Sidorov (2015). Attempts by, Kurosawa et al. (1983), Kurosawa and Ogata (1999), Zheng et al. (2001) and Elia et al. (2011) all use the Legendre and Jacobi symbols in developing a Rabin-like digital signature scheme.

In this work, we revisit the Rabin cryptosystem and propose a new efficient and practical signing scheme without using the Legendre and Jacobi symbols. In this work, we generate a signing key s from the equation $s \equiv \frac{1}{2} \pmod{\frac{(p-1)(q-1)}{4}}$. Then for verification, we use the power of 8.

2 RABIN'S DIGITAL SIGNATURE SCHEME

The Rabin's digital signature scheme is defined as follows (Rabin (1979)). Let H be a hash function.

Algorithm 1 Key Generation Algorithm

```
Input: Two $n$-bits prime numbers, $p$ and $q$.
Output: The public key $(N, w)$ and the private key $(p, q)$.
Generate two random $n$-bit primes $p$ and $q$ where $p, q \equiv 3 \pmod 4$.
Set $N = pq$.
Pick an element $w \in \mathbb{Z}_N$ such that the Jacobi symbol of $w$ over $N$ is equal to $-1$. In other words, $w$ is a quadratic residue modulo exactly one of $p$ or $q$.
Output the public key $(N, w)$ and the private key $(p, q)$.
```

Algorithm 2 Signature Algorithm

```
Input: The public key $(N, w)$, the private key $(p, q)$ and the message $M$
Output: The signature $S$.
Compute $x = H(M) \in \mathbb{Z}_N$.
One can show that exactly one of $\pm x, \pm xw \in \mathbb{Z}_N$ must be a quadratic residue. Let $y \in \{\pm x, \pm xw\}$ be that value. To find $y$, find the unique element in $\{\pm x, \pm xw\}$ for which the Legendre symbol is equal to 1 over both $p$ and $q$.
Let $S \in \mathbb{Z}_N$ be the square root of $y$ in $\mathbb{Z}_N$. Output $S$ as the signature on $M$.
```

Algorithm 3 Verification Algorithm

```
Input: The message $M$, the public key $(N, w)$ and the signature $S$
Output: TRUE or FALSE
Compute $x^{'} = H(M) \in \mathbb{Z}_N$.
Check if $S^2 \in \{\pm x^{'}, \pm x^{'}w\}$. If so, accept the signature. Otherwise, reject.
```

3 PROPOSED DIGITAL SIGNATURE SCHEME

In this section, we propose a new Rabin-like digital signature scheme. Firstly, we assign the condition for the prime numbers $\frac{p-1}{2}$ and $\frac{q-1}{2}$ must be primes. We set $R = \frac{(p-1)(q-1)}{4}$ and GCD$(2, R) = 1$. Then we will have a multiplicative inverse of 2 modulo R and denoted as the signing key s to be used in signing process.

3.1 Proposed Digital Signature Scheme

In this section, we present a proposed Rabin-like digital signature scheme. The proposed scheme defined as follows.

Algorithm 4 Key Generation Algorithm

```
Input: Two $n$-bits prime numbers, $p$ and $q$.
Output: The public key $N$ and the signing key $s$.
1: Generate two random $n$-bit primes $p$ and $q$ where $\frac{p-1}{2}$ and $\frac{q-1}{2}$ are both primes.
2: Set $N = pq$.
3: Set $R = \frac{(p-1)(q-1)}{4}$.
4: Compute the signing key $2s \equiv 1 \pmod{R}$.
```

Remark 3.1. *The value of R is always odd.*

Algorithm 5 Signature Algorithm

```
Input: The public key $N$, the signing key $s$ and the message $M$.
Output: The signature $S$.
1: Compute $S = H(M)^s \pmod{N}$.
```

Remark 3.2. *Observe that in our proposed scheme, we did not use either the Legendre or Jacobi symbol and we did not need any extra information.*

Algorithm 6 Verification Algorithm

```
Input: The Signature $S$.
Output: TRUE or FALSE
1: Compute $V \equiv S^8 \pmod{N}$.
2: Verification is TRUE if $V \equiv H(M)^4 \pmod{N}$.
```

3.2 Proof of Correctness

Proposition 3.1. *The verification process is correct.*

Proof.

$$\begin{aligned} V &\equiv S^8 \pmod{N} \\ &\equiv (H(M)^s)^8 \pmod{N} \\ &\equiv H(M)^{4+\phi(N)k} \pmod{N} \\ &\equiv H(M)^4 \pmod{N} \end{aligned}$$

□

3.3 Security Analysis

3.3.1 Factoring Equivalence

Proposition 3.2. *Factoring $N = pq$ is reduced to obtaining the signing key.*

Proof. If $N = pq$ is factored, the signing key s can be computed via $2s \equiv 1 \pmod{\frac{\phi(N)}{4}}$. □

3.3.2 Forging A Signature On An Arbitrary Message

Proposition 3.3. *Assume the adversary wants to forge a signature on a specific message $M \in \mathbb{Z}_N$. The adversary will do the following:*

1. *Generate random $M_1 \in \mathbb{Z}_N$.*
2. *Compute $M_2 \equiv M_1^{-1} \cdot M^4 \pmod{N}$.*
3. *Requests signatures for M_1 and M_2 given by $A_1 \equiv M_1^d \pmod{N}$ and $A_2 \equiv M_2^d \pmod{N}$*
4. *Then, $A \equiv A_1 \cdot A_2 \pmod{N}$ is a valid signature of M^4.*

Proof. Let A_1 and A_2 are valid signature for M_1 and M_2 then,

$$A^8 \equiv (A_1 \cdot A_2)^8 \equiv (M_1^d \cdot M_2^d)^8 \equiv M_1^{8d} \cdot M_2^{8d} \equiv M_1 \cdot M_2 \equiv M^4 \pmod{N}$$

□

Remark 3.3. *Always hash the message first before signing.*

3.3.3 Low Signing Exponent Attack

Lemma 3.1. *Let $\psi(N) = \frac{\phi(N)}{4}$, then $s > \frac{\psi(N)}{2}$.*

Proof.

$$\begin{aligned} es &= 1 + \psi(N) \\ &> \psi(N) \end{aligned}$$

Since $e = 2$, then $s > \frac{\psi(N)}{2}$ □

Remark 3.4. *From the Lemma 3.1, the signing key that is $s > \frac{\psi(N)}{2}$ will not succumb to Wiener's attack since $\frac{\psi(N)}{2} > \frac{1}{3}N^{1/4}$ (Wiener, 1990) or the attack by Boneh and Durfee (Boneh and Durfee (1999)) since $\frac{\psi(N)}{2} > N^{0.292}$.*

Lemma 3.2. *The constant k in the key equation $es = 1 + \frac{\psi(N)}{4}k$ satisfies $k = 1$.*

Proof.

$$\begin{aligned} k &= \frac{es-1}{\psi(N)} \\ &< \frac{es}{\psi(N)}. \end{aligned}$$

Since $s \in \mathbb{Z}_{\psi(N)}$ and $e = 2$, we get $0 < k < 2$. So $k = 1$. □

3.4 Example

a) Key Generation Process
Let $n = 16$. Then Along will choose the prime $p = 38431$ and $q = 43319$. Along will compute $R = \frac{(p-1)(q-1)}{4} = \frac{1664710740}{4} = 416177685$ where $N = pq = 1664792489$. The public key will be N and the private key will d.

b) Signing Process
Along will compute the hash value of $M = 1461232481$ by using any hash function to get $H(M)$. To generate the signature Along will compute $S \equiv H(M)^d \pmod{N} = 1448161743$. Along will send (M, S) to Busu.

c) Verification Process
To verify Along's signature, Busu will compute $V \equiv S^8 \equiv 1448161743^8 \equiv 1489484796 \pmod{N}$. Then Busu will compute the hash value of message M send by Along denote as $H(M)^4$ and compare to V whether same or not.

4 CONCLUSION

We proposed a Rabin-like digital signature scheme without utilizing the Legendre or Jacobi symbol or added on extra information as in previous works. It has been proved that to obtain the signing key it is as hard as factoring $N = pq$.

REFERENCES

Boneh, D. and Durfee, G. (1999). *Advances in Cryptology — EUROCRYPT ’99: International Conference on the Theory and Application of Cryptographic Techniques Prague, Czech Republic, May 2–6, 1999 Proceedings*, chapter Cryptanalysis of RSA with Private Key d Less than N 0.292, pages 1–11. Springer Berlin Heidelberg, Berlin, Heidelberg.

Elia, M., Piva, M., and Schipani, D. (2011). The rabin cryptosystem revisited. *arXiv preprint arXiv:1108.5935*.

Kurosawa, K., Ito, T., and Takeuchi, M. (1983). Public key cryptosystem using a reciprocal number with same intractability as factoring a large number. *Cryptologia*, 7(4):225–233.

Kurosawa, K. and Ogata, W. (1999). Efficient rabin-type digital signature scheme. *Des. Codes Cryptography*, 16(1):53–64.

Rabin, M. O. (1979). Digital Signatures and Public Key Functions as Intractable as Factorization. Technical report lcs/tr212, Cambridge MA:MIT.

Sidorov, E. (2015). Breaking the rabin-williams digital signature system implementation in the crypto++ library. *IACR Cryptology ePrint Archive*, 2015:368.

Wiener, M. J. (1990). *Advances in Cryptology — EUROCRYPT ’89: Workshop on the Theory and Application of Cryptographic Techniques Houthalen, Belgium, April 10–13, 1989 Proceedings*, chapter Cryptanalysis of Short RSA Secret Exponents, pages 372–372. Springer Berlin Heidelberg, Berlin, Heidelberg.

Williams, H. C. (1980). A modification of the rsa public key encryption procedure. *IEEE Trans. Inf. Theory*, 26(6):726–729.

Zheng, R. K., Chen, S., and Qiu, W. (2001). New rabin-like signature scheme. In *Workshop Proceedings of the 7th International Conference on Distributed Multimedia Systems*, pages 185–188, Knowledge Systems Institute.

Fingerprint Minutiae Template Protection for Privacy Preserving

Zhe Jin*, **Wun-She Yap**, and **Bok-Min Goi**

Lee Kong Chian Faculty of Engineering and Science
Universiti Tunku Abdul Rahman, Sungai Long, Malaysia

E-mail: jin.zhe@gmail.com
**Corresponding author*

ABSTRACT

Due to the strong linkage between an individual and a claimed identity, biometric authentication is widely used to authenticate an individuals identity by matching an individuals biometric with a database of records. Yet, if the biometric template stored inside the database is compromised, invasion of user privacy is inevitable. For instance, a fingerprint image can be reconstructed with high accuracy using existing technologies if the corresponding set of fingerprint minutia is revealed to an adversary. In this paper, we first enhance the performance of original polar grid-based 3-tuple quantization by considering polar coordinate that only covers partial image of fingerprint. We then enhance the modified polar grid-based 3-tuple quantization with random bits toggling technique to protect fingerprint minutiae template. The promising experimental results on FVC2002 DB1 and DB2 justify the feasibility of the proposed method.

Keywords: Fingerprint, template protection, security, privacy

1 INTRODUCTION

Biometrics has been integrated in large-scale personal identification systems and the rapid proliferation of biometric recognition applications is a foreseeable trend in future. The foreseen pervasiveness of biometric authentication systems speeds up the growing biometric databases. However, if biometric databases breach is occurred, severe influences of biological nature of human are concerns. Particularly, the damage to persons privacy and security is permanent due to the irrevocability and irreplaceability nature of human traits. As a result,various biometric template protection techniques have been proposed to secure biometric templates.

Fingerprint is probably the most widely used biometric trait for the biometric-based authentication systems (Maltoni et al., 2009). Minutia represents the fingerprint ridge characteristics at local level. There are two most prominent ridge characteristics, called ridge bifurcations and ridge termination. Generally, fingerprint minutiae are stable and robust to fingerprint impression conditions (Maltoni et al., 2009). Each minutia can be associated with a number of attributes, including location coordinates, orientation, type (e.g. ridge termination or ridge bifurcation), a weight based on the quality of the fingerprint image in the neighborhood of each minutia, and so on. However, from the common practices, only two attributes are used to represent a minutia: x- and y-coordinates pertaining to the location of minutia in the fingerprint; the orientation of the ridge line to which the minutia is attached (Krivokuća, 2015). In this paper, fingerprint minutia is solely focused.

Related Works. Generally, biometric template protection refers a set of techniques that mitigate the aftermaths due to the compromise of biometric templates databases for the purpose of malicious use. Technically, biometric template protection is to design a protect function and apply it into unprotected template to generate protected template. The template protection methods proposed in literature can be broadly divided into two categories, namely, feature transformation approach (or cancellable biometrics) and biometric cryptosystem (or helper data methods) (Jain et al., 2008). Cancellable biometrics (Ratha et al., 2007) is truly meant designed for biometric template protection. It refers to the irreversible transform of the biometric template to ensure security and privacy of the actual biometric template. On the other hand, biometric cryptosystem serves the purpose of either securing the cryptographic key using biometric feature (i.e., key binding) or directly generating the cryptographic key from biometric feature (i.e., key generation) (Jain et al., 2008). In this paper, cancellable biometric is focused. Ratha et al. (2007) proposed three non-invertible transform functions, namely Cartesian, polar and surface-folding transformation. Although the three transformation functions were claimed to be non-invertible due to the many-to-one mapping property, a scheme by Feng et al. (2008) reveals that the surface-folding transform can be degenerated when the transformed template and parameters are revealed to the attacker. Wang and Hu (2012) proposed a cancellable fingerprint template based on a dense infinite-to-one mapping (DITOM). A complex vector is generated from the proposed method by applying a discrete Fourier transform and the final template is obtained by blending the complex vector with a randomly generated parametric matrix. In addition to DITOM, Wang and Hu (2014) proposed another cancellable fingerprint template based on curtailed circular convolution, which demonstrates an improvement on accuracy and security over DITOM. Jin et al. (2014) proposed a fingerprint template with strong non-invertibility, namely randomized graph-based hamming embedding (RGHE). This technique is able to protect the MVD features and preserve the recognition performance in the original feature space.

Contributions. Our contributions are two-fold:

1. Polar grid-based 3-tuple quantization is an alignment-free minutiae descriptor that utilizes variable-sized tessellated quantization in polar coordinate (Jin et al., 2012). In this approach, polar coordinate covers the entire image and produces a lengthy bit string, which is undesirable for practical applications due to larger storage of templates and higher equal error rate. We propose a modification on the original approach by considering partial image only to reduce the length of generated bit string and lower equal error rate.

2. To further improve the security and privacy of generated fingerprint template, we propose the use of random toggling technique on the modified fingerprint template such that it is more difficult for an attacker to recover the fingerprint from the template by introducing

noise into the template.

2 PROPOSED METHOD

Polar grid-based 3-tuple quantization (PGTQ) is an alignment-free minutiae descriptor that utilizes variable-sized tessellated quantization in polar coordinate (Jin et al., 2012). In this method, sectors near the reference minutia have smaller area and vice versa. This leads to a smaller (resp. larger) quantization step around (resp. further away from) the reference minutia to tolerate fingerprint elastic deformation. In the original PGTQ descriptor, polar coordinate covers the entire image and produces a lengthy bit string, which is undesirable for practical applications due to large storage of templates. As a solution, in this paper, we consider polar coordinate that only covers a part of the image limited by a circle with radius R. With this, the size of the resultant bit-string can be significantly reduced. The details of the modified PGTQ descriptor are described as follows:

1. Let $m_r = \{x_r, y_r, \theta_r\}$ be the reference minutiae. The neighboring minutiae within a circle with radius R in Euclidean distance is rotated and translated based on the reference minutiae using Eq. (1) and Eq. (2). The transformed minutiae are represented as $m^t = \{x_i^t, y_i^t, \theta_i^t | i = 1, \ldots, N_R - 1\}$, where N_R is the total number of minutiae within a predefined radius R.

$$\begin{bmatrix} x_i^t \\ y_i^t \end{bmatrix} = \begin{bmatrix} \cos\theta_r & -\sin\theta_r \\ \sin\theta_r & \cos\theta_r \end{bmatrix} \begin{bmatrix} x_i - x_r \\ -(y_i - y_r) \end{bmatrix} \tag{1}$$

$$\theta_i^t = \begin{Bmatrix} \theta_i - \theta_r; & \theta_i \geq \theta_r \\ 360^\circ + \theta_i - \theta_r; & \theta_i < \theta_r \end{Bmatrix} \tag{2}$$

2. The translated and rotated minutiae are then converted into polar coordinates using Eq. (3) and Eq. (4). ρ_i and α_i indicate the radial distance (in pixels) and the radial angle of the i-th minutia in Polar coordinates ($\alpha_i \in (0, 360^\circ]$), respectively.

$$\rho_i = \sqrt{(x_i^t)^2 + (y_i^t)^2} \tag{3}$$

$$\alpha_i = \arctan(\frac{y_i^t}{x_i^t}) \tag{4}$$

3. *3-Tuple-based Quantization.* The 3-tuple-based quantization is a sector-based quantization involving all minutiae in the neighborhood. Each quantized minutia can be represented as a vector $\omega = \{\rho^q, \alpha^q, \theta^q\}$, such that

$$\rho_i = \lfloor \rho_i / x \rfloor \tag{5}$$

$$\alpha_i = \lfloor \alpha_i / y \rfloor \tag{6}$$

$$\theta_i = \lfloor \theta_i / z \rfloor \tag{7}$$

where $/$ denotes quotient; x, y and z indicate the radius for each polar grid (in pixels), radial angle for tolerance ($y \in (0, 360°]$) and orientation angle to be tolerated $z \in (0, 360°]$, respectively. The quantization level is determined by x, y and z.

4. *Binarization*. The quantized minutiae ω are then binarized using the polar grids. We adopt a simple rule to map a polar grid to 1 if the polar grid contains more than one minutia, and otherwise, a polar grid is mapped to 0. By concatenating the individual output bits from the polar grids, we eventually obtain a binary vector with length equals to the number of polar grids $l = \lceil 360/x \rceil \lceil 360/y \rceil \lceil 360/z \rceil$, where $\lceil \cdot \rceil$ denotes the ceiling function. The above steps are repeated by changing the reference minutia with every remaining minutia to generate the full binary PGTQ descriptor. As the total minutiae number (N_m) extracted from each fingerprint image could be different, this template, denoted by $\Omega \in \{0, 1\}^{(N_m \times l)}$ is variable in size.

5. *Matching*. To evaluate the similarity between two sets of modified PGTQ descriptor, we adopt a typical two-stage matching strategy that is composed of local and global matching. The local descriptor matching searches for the intersections between two binary strings in which the PGTQ descriptor is represented. On the other hand, the global matching is to find the ratio of the matched descriptor pairs over all potential pairs as the final matching score.

Let $\Omega^e = [b_1^e; b_2^e, \ldots, b_{n^e}^e]$ and $\Omega^q = [b_1^q; b_2^q, \ldots, b_{n^q}^q]$ be the enrolled and query descriptor sets that consist of n^e and n^q l-bit binary strings, respectively. From this point onwards, we slightly abuse the notation of b, where $b_{i,k}$ represents the k-th bit for i-th binary string with $1 \leq k \leq l$ and $1 \leq i \leq n^e$ or n^q. To take into account the difference of minutiae quantity in the enrolled and query image, we normalize the similarity scores between two local descriptors Ω^e and Ω^q as follows:

$$S_{ij}^b = \frac{(N_j^q + N_i^e) \sum_{k=1}^{l} (b_{j,k}^q \cdot b_{i,k}^e)}{(N_j^q)^2 + (N_i^e)^2} \tag{8}$$

where S^b denotes the matching score between two binary strings, $\cdot$ represents a bitwise AND operator, $N_i^e = \sum_{k=1}^{l} (b_{i,k}^e)$ and $N_j^q = \sum_{k=1}^{l} (b_{j,k}^q)$ denote the total number of 1s of the enrolled and query bit-strings, respectively. The term $\sum_{k=1}^{l} (b_{j,k}^q \cdot b_{i,k}^e)$ in Eq. (8) counts the bit positions that have value '1' in both query and enrolled bit-strings. The scores in matrix $S^b \in \mathbb{R}^{n^q \times n^e}$ range from 0 to 1 where '1' indicates a perfect match.

Once the similarity score matrix S^b is calculated from the local descriptor matching; a global matching process is carried out. Given the score matrix $S^b = \{s_{ij}^b\}$, the final score can be calculated as:

$$S_{\text{PGTQ}} = \max\{\frac{1}{m} \sum_j s_j(\text{mbox}), \frac{1}{n} \sum_i s_i(\max)\} \tag{9}$$

where $s_{j(\text{max})} = \max_i\{s_{ij}^b\}$ and $s_{i(\text{max})} = \max_j\{s_{ij}^b\}$ represent the maximum score component of the i-th column and j-th row, respectively. The detailed matching process is illustrated in Algorithm 1.

Algorithm 1: Matching Two PGTQ-based Minutia Descriptors

```
Input: $\Omega^\theta, \Omega^q, n^\theta, n^q$
Function Prototype: $sim(\Omega^\theta, \Omega^q)$
$n^\theta \leftarrow \text{size}(\Omega^\theta)$
$n^q \leftarrow \text{size}(\Omega^q)$
for $i = 1 : n^\theta$ do
    $B_i^\theta = \Omega^\theta(i)$
    for $j = 1 : n^q$ do
        $B_j^1 = \Omega^q(j)$ Calculate similarity score $s_{ij}^b$ between $b_i^\theta$ and $b_j^q$ using Eq. (8)
    end
end
$S^b = \{s_{ij}^b\}$
$s_{j(\text{max})} = \max_i\{s_{ij}^b\}$
$s_{i(\text{max})} = \max_j\{s_{ij}^b\}$
$S_{\text{PGTQ}} = \max\{\frac{1}{m}\sum_j s_j(\text{mbox}), \frac{1}{n}\sum_i s_i(\text{max})\}$
Output: The matching score, $S_{\text{PGTQ}}$ between $\Omega^\theta$ and $\Omega^q$
```

The PGTQ descriptor as templates is required to be stored for verification. However, if the templates stored in database are compromised, the security and privacy of the system are vulnerable to template replay, spoof construction and targeted false accepts. To alleviate this problem, our treatment is to adopt a random bits-toggling process presented in Farooq et al. (2007). This process is to randomly select a fraction of bits and invert them. This process is a noise addition process that distorts the template data. Since PGTQ descriptor is a feature matrix; the bit-toggling process is applied in a row-wise basis. The overall authentication protocol consists of two stages: enrollment and verification. In the enrollment stage, PGTQ descriptor is extracted by presenting users genuine fingerprint. Thereafter, a random bits-toggling process is performed to distort the PGTQ descriptor and generate the protected template stored in database. While for the verification stage, the query PGTQ descriptor is extracted and matched against the protected template. Such protocol gives several observations:

1. PGTQ essentially is a many-to-one transformation that provides first layer of protection against feature (e.g. minutiae) inversion;

2. random bits-toggling process additionally offers second layer protection by introducing a significant portion of noise.

It thus can be expected that template replay, spoof construction and targeted false accepts can be prevented effectively.

3 EXPERIMENT ANALYSIS

To measure the feasibility of the proposed method, the experiments were conducted on six public fingerprint datasets, FVC2002 (DB1, DB2) (FVC, 2002). Each dataset consists of 100 users with 8 samples per user. In total, there are 800 (100×8) fingerprint images in each dataset. VeriFinger 6 SDK (ver) was used for minutia extraction. The performance of the proposed framework is evaluated using equal error rate (EER). For matching protocol, the first sample (gallery) of every identity is matched against the second samples (probe) of every identity for false rejection rate (FRR) calculation. On the other hand, the first sample of each identity is matched against the first sample of the remaining identities for false acceptance rate (FAR) calculation. This matching protocol yields 100 genuine scores and 4950 imposter scores for each dataset. Note the same setup has been employed by the existing methods for a fair comparison. Table 1 tabulates the parameters used in our experiments.

Symbols	Description	Value
R	Radius for polar coordinates (in pixel)	70
x	Radius for polar grid segment (in pixel)	10
y	Radius angle for polar grid segment (in degree)	20
z	Minutiae orientation angle (in degree)	30

Table 1: The parameters used in experiment.

As aforementioned discussion, the bit-toggling process is applied to distort the templates. However, Farooq et al. (2007) reveals that a large number of randomly toggled bits would deteriorate accuracy performance. This deterioration can be alleviated by carefully selecting a portion of bits for flipping. We show in Table 2 that, even though a significant portion of noise (50%) has been added, the accuracy would not degrade significantly. Then, a comparative study of performance is conducted between the proposed technique and the existing methods, and the corresponding EER performances are listed in Table 2. It is noticed that the proposed method achieves performance that is better than the existing methods (Wang and Hu, 2012, 2014). Further, the original PGTQ is discouraged because of the lengthy template and worse performance due to the maximum noise (e.g. spurious minutia) as the entire image is included.

Methods	DB1	DB2
Jin et al. (2012)	1.19	6.94
Wang and Hu (2012)	3.5	4.0
Wang and Hu (2014)	2.0	2.3
Modified PGTQ	1.08	2.03
Modified PGTQ with random toggling technique	1.04	2.02

Table 2: The EER performances for different methods using FVC2002

4 SECURITY ANALYSIS

In our experiment, a 5-bit toggling is considered; to correctly guess the 5-bits among 1512 bits-length template requires approximately 2^{45} attempts. Averagely, 40 minutiae point are extracted, subsequently, the size of binary feature is of 40×1512. Since 2^{45} attempts are required for single feature vector, the effort to recover the entire feature matrix requires 2^{1800} attempts. This is indeed computational hard in real time scenario. To make it harder, PGTQ descriptor generation essentially is a many-to-one transformation, i.e. multiple minutiae may map into an identical element in transformed domain. Information lost in this process is evitable. Therefore, privacy preserving is gained even the template is compromised.

5 CONCLUSION

In this paper, we modified the polar grid-based 3-tuple quantization to improve its EER performance and reduce the length of generated bit string. Subsequently, we proposed random bits togging based modified polar grid-based 3-tuple quantization to protect fingerprint minutia. The proposed method improves the security of PGTQ by preventing it from template replay, spoof construction and targeted false accepts. The experimental results vindicate that the proposed method could achieve satisfying recognition performance in comparison with several existing methods. In addition, with random bits togging procedure, a significant noise is salted to provide another layer of protection for fingerprint minutia without compromising recognition performance.

ACKNOWLEDGMENTS

Wun-She Yap and Bok-Min Goi would like to acknowledge MOSTI for financially funding their research through the MOSTI Science Fund numbers 01-02-11-SF0189 and 01-02-11-SF0201.

REFERENCES

VeriFinger SDK. *In Second International Fingerprint Verification Competition*. Accessed on 6tth January 2016. Sourced from `http://www.neurotechnology.com`.

(2002). FVC2002. `http:/bias.csr.unibo.it/fvc2002/`. *In Second International Fingerprint Verification Competition*. Accessed on 6tth January 2016.

Farooq, F., Bolle, R., Jea, T., and Ratha, N. (2007). Anonymous and revocable fingerprint recognition. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR2007)*, pages 1–7, Minneapolis, MN.

Jain, A. K., Nandakumar, K., and Nagar, A. (2008). Biometric template security. *EURASIP J. Adv. Signal Process.*, 579416:113.

Jin, Z., Lim, M. H., Teoh, A. B. J., and Goi, B.-M. (2014). A non-invertible randomized graph-based hamming embedding for generating cancelable fingerprint template. *Pattern Recog. Lett.*, 42(3):137–147.

Jin, Z., Teoh, A. B. J., Ong, T. S., and Tee, C. (2012). Fingerprint template protection with minutiae-based bit-string for security and privacy preserving. *Expert Syst. with Appl.*, 39(6):6157–6167.

Krivokuća, V. (2015). *Fingerprint Template Protection using Compact Minutiae Patterns*. PhD thesis, The University of Auckland, Auckland, NZ.

Maltoni, D., Maio, D., Jain, A., and Prabhakar, S. (2009). *Handbook of Fingerprint Recognition*. Springer-Verlag, London, UK, 2nd edition.

Ratha, N. K., Chikkerur, S., Connell, J. H., and Bolle, R. M. (2007). Generating cancelable fingerprint templates. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(4):561–572.

Wang, S. and Hu, J. (2012). Alignment-free cancelable fingerprint template design: A densely infinite-to-one mapping (ditom) approach. *Pattern Recogn.*, 45(12):4129–4137.

Wang, S. and Hu, J. (2014). Design of alignment-free cancelable fingerprint templates via curtailed circular convolution. *Pattern Recogn.*, 47(3):1321–1329.

Generating Non-Invertible Iris Template For Privacy Preserving

Yen-Lung Lai[*], **Zhe Jin**, **Bok-Min Goi**, and **Tong-Yuen Chai**

Lee Kong Chian Faculty of Engineering and Science
Universiti Tunku Abdul Rahman, Sungai Long, Malaysia

E-mail: yenlung@1utar.my, jin.zhe@1utar.my,
goibm@utar.edu.my ,chaity@utar.edu.my
[*]Corresponding author

ABSTRACT

Human iris has been widely used for personal identification and verification. The traditional iris recognition system stored the unprotected iris templates in a database for authentication. Since human iris is permanently associated with each individual, a compromised of the database implies a permanent loss in identity. One of the solutions for this is to incorporate a template protection scheme to protect the iris templates stored in a database. In this paper, a new iris template protection scheme is proposed based on "Winner Takes All" (WTA) hashing algorithm. The experiment results showed that our protected template size could be reduced to more than half of the original template without significant degrades of the recognition performance. Besides that, it is computationally hard for an adversary to regenerate the iris template from the protected template.

Keywords: Biometrics, Iris, Template protection, Security, Privacy

1 INTRODUCTION

Human iris is considered as a very reliable feature for personal identification and verification doe to its stability (Daugman (2004)). Daugman (2004, 2006) was the first one who came out with an *automated* iris recognition system using *iriscode*. His system which is now owned by the Iridian company has been tested under numerous studies and all reporting a zero failure rate (Masek (2003)).

Traditional iris recognition system stored the generated iris templates inside a database for authentication. Once the enrollment of individual iris is completed, only the right person can authenticate successfully through the database with the right iris information he/she is acquiring. However, iris template inside a database is potentially being attacked or compromised. Once this database is compromised, the attacker can use the stolen template to perform impersonation. Due to the fact that human iris is permanently associated with each individual, this implies a permanent loss of identity for each user. Moreover, information of human iris is very limited as every individual only possesses two iris for iris template generation. Thus, the compromised template cannot be revoked easily and further restricts the usage of the human iris in recognition purpose. The security and privacy issues become a major concern for a user and a solution is needed.

Biometric template protection is introduced to solve the existing security problem inside a biometric recognition system. The notion of cancellable biometric have been used by Jain et al. (2008) for:

1) Diversity: Different templates are used for different purposes. There is no cross-matching between the generated templates for different applications

2) Revocability : The generated template must be able to revoke when a database is compromised.

3) Non-invertible : Generated template need to be non-invertible to make sure it is computationally hard for the attacker to reconstruct the original template.

4) Performance : The newly generated template must preserve or at least come with insignificant degradation on the recognition performance.

2 LITERATURE REVIEW

Among various iris template protection schemes, protection in iris template can be classified into binary and non-binary.

For the protection of iris template in binary, it refers to the protection of the iris code. This iris code is generated by quantization of the iris Gabor-features which formed by convolution of the Gabor-filter with the extracted iris region. A well-known instance was reported by Rathgeb et al. (2013) proposed an alignment-free method for generating cancellable iris template using bloom filter. Bloom filter refers to a bit array which initialized with zeros. They first divided the iris code into several K equal blocks. Each single block consists of multiple columns of bits for a single bloom filter. The column-wise binary codewords constituted to an element of '1' in the corresponding index of the bloom filter. The final template is generated by concatenating all the bloom filters from different blocks. Since the angular resolution of the iris code is represented by the column bits, the template generated is alignment free. To achieve cancellability, they applied an application-specific secret key. When compromised case happens, different application-specific secret keys can be used to reissue a new template.

Ouda et al. (2010) proposed a tokenless template protection scheme for iris code. A random seed is used to generate different pseudo-random sequences. After this, consistence bits from several iris code of the same user are extracted through a training process. The many to one mapping between the consistence bits and the pseudo-random sequences is performed to output a Biocode which can discard safely into a database. Their algorithm does not require any user-specific token to be stored due to the fact that the consistence bits vectors are different for each individual. A new random sequence can be used to regenerate a new template.

For the protection of iris template in non-binary, it refers to the protection of iris real features. This iris real features can be described as the iris pixels texture, phase/amplitude values of the complex Gabor-features, etc depends on the proposed template protection schemes. Pillai et al. (2010) used a sectored random projection technique to achieve cancellability of iris template. The Gabor-features of each iris region is projected into a lower dimensions space using a same random Gaussian matrix. The recognition performance has maintained only when different Gaussian matrices are used for different users. Kong et al. (2006) stated the “stolen token case” need to take into consideration while evaluating the recognition performance. Because the random numbers, matrices, tokens or passwords are still vulnerable to be compromised.

Wang et al. (2008) have proposed a method for iris template protection using iris Gabor-features. They projected the Gabor-features using an optimal matrix. An adaptive non-uniform quantization method (ANUQ) is used to quantize the features vector. The quantized vector is hashed and stored inside the database. For cancellability, they suggested using different quantization templates. Their works require a large storage space for the optimal matrix, quantization templates, and the hashed output vector to be stored in the same database for matching.

Chong et al. (2006) generated cancellable iris templates using projection technique. The iris Gabor-features are projected by using a user specific-random matrix. Because of user-specific random matrices are used, the recognition performance showed improved after template protection scheme applied. However, the “stolen-token scenario” not included which omit the actual biometrics performance.

Hammerle-Uhl et al. (2009) used block remapping method to perform the non-invertible transform. The normalized iris texture is partitioned into several image blocks. The image blocks are permuted by using random permutation. After that, an image block remapping technique is used to generate cancellable iris template. Several image blocks can be reused and remapped causing an information loss for non-invertibility. A different permutation can be used for new template generation. The transformation setting maintained the recognition performance but leaving its security doubtable (Jenisch and Uhl (2011)).

3 THE “WINNER TAKES ALL” (WTA) HASHING

Yagnik et al. (2011) proposed a hashing technique based on rank correlation measures for fast similarity search namely “Winner Takes All” hashing. Their works have mentioned The benefits of using rank correlation measures due to its stability in perturbation, which gives rise to a good indication of inheriting similarity between items. “Winner Takes All” hashing computes ordinal embedding based on partial order statistics. Besides this, it gives rise to a resultant binary feature space where the hamming distance is highly correlated to the rank similarity measures. The hashing process involves several steps :

1. Perform random permutation
2. Select the first K items
3. Within the selected K items, choose the largest value
4. Record down the correspond index value using $log_2 K$

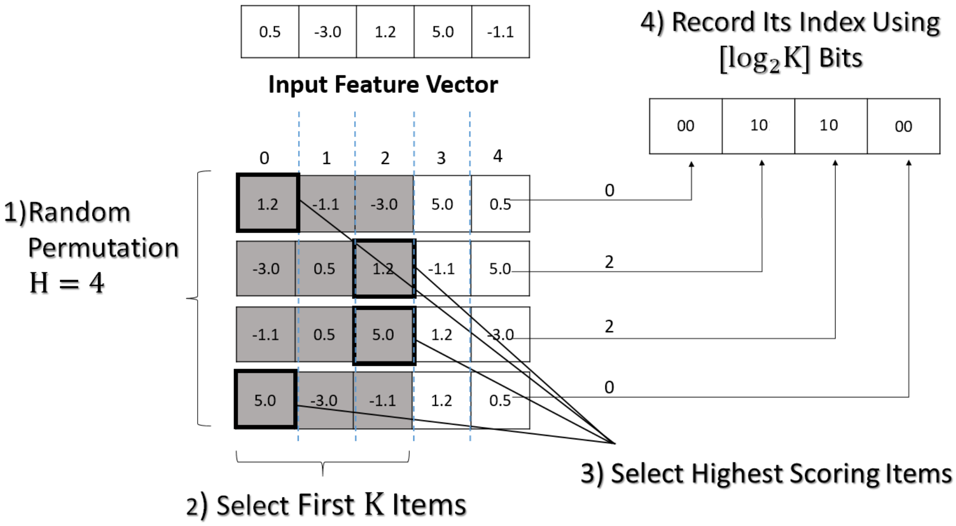


Figure 1: WTA hashing process

Since WTA hashing is using ordinal measures, the transformed feature space is not sensitive to the absolute value of the features dimension but rather than the implicit ordering. Thus, it is very resilience to noise and variation that do not affect the implicit order of the input features. When WTA hashing used for fast similarity search, the result showing WTA hashing outperform than other existing methods (Yagnik et al. (2011)). Besides that, it provides low computation cost and does not require any data-driven optimization (e.g training samples). Figure 1 showing the steps involved in WTA hashing.

4 PROPOSED METHOD

The traditional WTA hashing is not designed for biometric template protection. In our proposed method, a significant alteration of WTA hashing has been done. The modified WTA contributed to a new transformation which allowed us to generate non-invertible iris template from the iris code.

Our proposed transformation consists of certain benefits to generate non-invertible iris template which can be described as follow:

1) Stability: the stability of proposed method inherited from WTA hashing due to its resilient

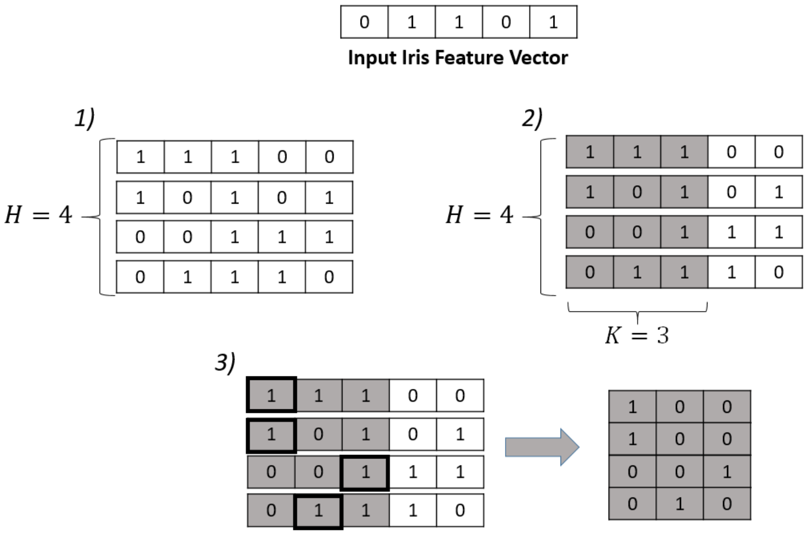


Figure 2: Proposed Method: 1)Generate H permutations for the original binary vector.; 2)Select K items,; 3)Remain first '1' value for every permutation

to noise

2) Simplicity: the proposed method is easy to be computed by tuning only two hyperparameters

The proposed transformation can be divided into three steps:

1. Perform random permutation.
2. Select the first K items.
3. Remain the first '1' bits value for every row.

In the last step, as iris code have several global indices with a value of '1', we only remain the first '1' bit value for every row within the binary array. By doing so, certain bits information will be lost resulted computationally difficult to reconstruct the original iris code. The proposed transformation function for iris code is described as H_t and shows in Figure 2.

Besides, conventional iris recognition system uses a noise mask to eliminate the noises caused by pupil dilation and in-concentric-pupil displacement (Daugman (2004)). This noise mask is generated together with the corresponding iris template during features encoding process. In order to integrate the use of noise mask in our proposed method, the noise masks have gone through a transformation with similar procedure in step 1 and step 2 as described in H_t. However, all the bits values have remained instead of only recording the first '1' bits for step 3. By doing so, the noise mask acts as control bits to confirm the validity of the template bits while matching. This transformation denoted as H_m is employed for every generated noise mask.

For matching, the dissimilarity between two templates can be computed by calculating the hamming distance (HD). While $0 \leq \text{HD} \leq 1$, higher HD indicates larger dissimilarity between

Table 1: Stolen Token Scenario

Proposed Method	G	K	EER%	
			With Mask	Without Mask
	240	1	4.01	8.06
	120	2	4.94	7.38
	80	3	5.25	9.60
	60	4	6.48	9.89
	48	5	7.77	9.53
Original Matching	-	-	3.56	7.11

the two templates. Eq.(1) refers to the equation to compute the dissimilarity between template A and template B

$$HD = \frac{|H_t(templateA) \otimes H_t(templateB) \cap H_m(MaskA \cap MaskB)|}{|H_m(MaskA \cap MaskB)|} \tag{1}$$

5 EXPERIMENT ANALYSIS

For the experiment, Masek (2003)'s algorithm has been adopted to generate the iris template. Every generated iris template with a corresponding noise mask both having a fixed dimension of $20 * 480$ with a total number of 9600 bits. CASIA database v1.0 is used in the experiment. This dataset contains 756 iris images from 108 unique eyes, with 7 different samples from each eye. For intra-class comparisons, each iris image is matched against the generated template extracted from other images of the same iris, leading to 7 genuine comparisons for each eye and a total of 2268 genuine comparisons. For inter-class comparisons, every iris image is matched against other iris images captured from different eyes, yielding 283122 different impostor comparisons.

During performance evaluation, two scenarios, "genuine-token" and "stolen-token scenario" are carried out. For the "genuine-token scenario", we assume that the secret token is securely stored, so user specific token is used here. In the" stolen-token scenario", it is assumed that a genuine user lost his/her token; therefore, the same permutation is used for all iris code. Table 1 shows the result of the "stolen-token scenario" after testing has been done by using different value of G and K.

Our result shows that the proposed method can almost preserve the recognition performance of the original recognition system in term of equal error rate (EER). With a higher value of K while keeping the value of G in constant, the recognition performance has degraded due to loss of information. This refers to the lesser number of '1' were used to calculate the hamming distance. Besides that, our experiment testified the improvement of the recognition performance when using iris noise mask in matching.

For the "genuine-token scenario", a user-specific token is used. Table 2 shows the result using a different value of G and K in the genuine token scenario.

Table 2: Genuine Token Scenario

Proposed Method	G	K	EER%	
			With Mask	Without Mask
	240	1	0.95	1.31
	120	2	1.60	4.00
	80	3	2.77	4.39
	60	4	5.16	4.93
	48	5	5.01	6.58
Original Matching	-	-	3.56	7.11

Table 3: Comparison between existing method with proposed method

Method used	Equal Error Rate(%) Without Template Protection	Equal Error Rate(%) With Template Protection
Wang et al Method	11.47	4.69
Osama et al Method	1.30	1.30
Proposed Method	3.56	4.01

Compare to Ouda' s method (Ouda et al. (2010)), our method does not involve any training process while generating the iris template. A new template can be regenerated by using different set of permutation without the requirement of multiple times re-enrollment. Our proposed method also achieved better recognition rate compare to Wang 's method (Wang et al. (2008)) as shown in Table 3.

To study the non-invertibility of our proposed method, we have evaluated two different scenarios by analyzing the computational hardness for an adversary to regenerate the original iris template. 1)If the token has been stolen or known by the adversary. 2)If the potential adversary has all the information including the protected template, token, parameters, and algorithm.

For the first scenario, because of the token have nothing associated with the original iris code, the adversaries are clueless and learn nothing from it. Furthermore, without knowing the K value, the adversary cannot guess the bits value using brute force due to dimensional differences between the original iris code and protected template.

For the second scenario, due to the reduced dimension in the protected template, It is hard to obtain more than 50% bits information of the original iris code by direct guessing. This can be shown by calculating the percentage of bits information left in the protected template. In our experiment, we set $G * K$=240, thus, the generated template size is reduced to $20 * 240$ with total 4800 bits. There are still leaving 4800 bits unknown. Taking into calculation of the bit loss due to the only first '1' bit value has remained, this lead to a minimum trial of 2^{4800} in order to regenerate the original iris code

6 CONCLUSION

In this paper, we have gone through several studies about different iris template protection schemes. A new method has been proposed to generate non-invertible iris template for template protection. The proposed method have shown improvement of the recognition performance in "genuine token" scenario. Even in a "stolen token" scenario, our protection scheme is able to preserve the recognition by getting a very close equal error rate (EER) compares to the original iris recognition system without template protection. It is also hard for an adversary to reconstruct our protected template by direct guessing. This proposed method is training free and easy to compute. A new template can be regenerated easily to replace the compromised template by using different permutation set. Thus, revocability is fulfilled.

ACKNOWLEDGMENTS

This works is under supports by MOSTI Science Fund numbers 01-02-11-SF0201.

REFERENCES

Chong, S. C., Teoh, A. B. J., and Ngo, D. C. L. (2006). High security iris verification system based on random secret integration. *Computer Vision and Image Understanding.*, 12(2):169–177.

Daugman, J. (2004). How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology.*, 14(1):21–30.

Daugman, J. (2006). Probing the uniqueness and randomness of iris codes: results from 200 billion iris pair comparisons. *Proceedings of the IEEE.*, 94(11):1927–1935.

Hammerle-Uhl, J., Pschernig, E., and Uhl, A. (2009). Cancelable iris biometrics using block re-mapping and image warping. In *Proceedings of the 12th International Conference on Information Security*. ISC 09.

Jain, A. K., Nandakumar, K., and Nagar, A. (2008). Biometric template security. *EURASIP J. Adv. Signal Process.*, 579416:113.

Jenisch, S. and Uhl, A. (2011). Security analysis of a cancelable iris recognition system based on block remapping. In *Proceedings of IEEE International Conference on Image Processing*, pages 3274–3277.

Kong, A., Cheunga, K.-H., Zhanga, D., Kamelb, M., and Youa., J. (2006). An analysis of biohashing and its variants. *Pattern Recogn.*, 9(7):1359–1368.

Masek, L. (2003). Recognition of human iris patterns for biometric identificationl. Master's thesis, University of Western Australia, Australia.

Ouda, O., Tsumura, N., and Nakaguchi, T. (2010). Tokenless cancelable biometrics scheme for protecting iris codes. *ICPR.*, pages 882–885.

Pillai, J. K., Patel, V. M., Chellappa, R., and Ratha, N. K. (2010). Sectored random projections for cancelable iris biometrics. In *Proceeding IEEE International Conference on Acoustics Speech and Signal Processing*, pages 1838–1841.

Rathgeb, C., Breitinger, F., and Busch, C. (2013). Alignment-free cancelable iris biometric templates based on adaptive bloom filters. In *Biometrics (ICB), 2013 International Conference*, pages 1–8,4–7.

Wang, Z. F., Han, Q., Niu, X., and Busch, C. (2008). A novel template protection algorithm for iris recognition. *Intelligent Systems Design and Applications, 2008. ISDA '08. Eighth International Conference.*, 2:340–345,26–28.

Yagnik, J., Strelow, D., Ross, D., and Lin, R.-S. (2011). The power of comparative reasoning. In *International Conference on Computer Vision*.

New Vulnerabilities of RSA Modulus Type $N = p^2 q$

Normahirah Nek Abd Rahman[*1] and **Muhammad Rezal Kamel Ariffin**[1,2]

[1]*Al-Kindi Cryptography Research Laboratory, Institute for Mathematical Research, Universiti Putra Malaysia*
[2]*Department of Mathematics, Faculty of Science, Universiti Putra Malaysia*

E-mail: mahirah_mayrah@yahoo.com, rezal@upm.edu.my
[*]*Corresponding author*

ABSTRACT

This paper proposes new attacks on RSA modulus of type $N = p^2 q$. Given k RSA moduli $N_i = p_i^2 q_i$ for $k \geq 2$ and $i = 1, ..., k$, the attack works when k RSA public keys (N_i, e_i) are such that there exist k relations of the shape $e_i x - N_i y_i = z_i - (ap_i^2 + bq_i^2)y_i$ or of the shape $e_i x_i - N_i y = z_i - (ap_i^2 + bq_i^2)y$ where the parameters x, x_i, y, y_i and z_i are suitably small in terms of the prime factors of the moduli. The proposed attacks utilizing the LLL algorithm enables one to factor the k RSA moduli N_i simultaneously.

Keywords: RSA, Factorization, LLL algorithm, Simultaneous diophantine approximations

1 INTRODUCTION

The RSA cryptosystem was developed by Rivest et al. (1978) is the well-known public key cryptosystem. The mathematical operations in RSA depend on three parameters, the modulus $N = pq$ which is the product of two large primes p and q, the public exponent e and the private exponent d, related by the congruence relation $ed \equiv 1 \pmod{\phi(N)}$ where $\phi(N) = (p-1)(q-1)$. Hence, the difficulty of breaking the RSA cryptosystem is based on three hard mathematical problems which is the integer factorization problem of $N = pq$, the e-th root problem from $C \equiv M^e \pmod{N}$ and to solve the diophantine key equation $ed + 1 = \phi(N)k$.

RSA is most commonly used for providing privacy and ensuring authenticity of digital data. Hence, many practical issues have been considered when implementing RSA in order to reduce the encryption or the execution decryption time. To reduce the encryption time, one may wish to use a small public exponent e. For discussion on security issues surrounding small encryption exponent see Boneh (1999). Logically, the RSA cryptosystem is likely to have faster decryption if the secret exponent d is relatively small. The knowledge of secret exponent d leads to factoring N in polynomial time. Thus, much research has been produced to determine the lower bound for d. Nevertheless, the use of short secret exponent will encounter serious security problems in various instance of RSA.

Based on the convergents of the continued fraction expansion of $\frac{e}{N}$, Wiener (1990) showed that the RSA cryptosystem is insecure when the secret exponent, $d < N^{1/4}$. Later, by using lattice basis reduction technique, Boneh and Durfee (1999) proposed an extension on Wiener's work. It was determined that the RSA cryptosystem is insecure when $d < N^{0.292}$. The work proposed by Blömer and May (2004) which combined lattice basis reduction techniques with continued fraction algorithm, showed that the RSA cryptosystem is insecure if there exist integers x, y and z satisfying the equation $ex - y\phi(N) = z$ with $x < \frac{1}{3}N^{1/4}$ and $|z| < exN^{-3/4}$. In cases where a single user generates many instances of RSA (N, e_i) with the same modulus and small private exponents, Howgrave-Graham and Seifert (1999) proved that the RSA cryptosystem is insecure in the presence of two decryption exponents (d_1, d_2) with $d_1, d_2 < N^{5/14}$. In the presence of three decryption exponents, they improved the bound to $N^{2/5}$ based on the lattice reduction method.

Then, Hinek (2007) showed that it is possible to factor k RSA moduli using equations $e_id - k_i\phi(N_i) = 1$ if $d < N^\delta$ with $\delta = \frac{k}{2(k+1)} - \varepsilon$ where ε is a small constant depending on the size of max $N_i = p_iq_i$. Later, Nitaj et al. (2014) proposed a new method to factor k RSA moduli N_i in the scenario that the RSA instances satisfy k equations of the shape $e_ix - y_i\phi(N_i) = z_i$ or of the shape $e_ix_i - y\phi(N_i) = z_i$ with suitably small parameters x_i, y_i, z_i, x, y where $\phi(N_i) = (p_i - 1)(q_i - 1)$. The analysis utilized the LLL algorithm.

As described in May (2004) the moduli of the form $N = p^2q$ is frequently used in cryptography and therefore they represent one of the most important cases. According to May, the modulus in the general form of $N = p^rq$ with $r \geq 2$ is more insecure than $N = pq$. Nevertheless, the modulus $N = p^2q$ is still tempting to be used. Examples of schemes are the RSA-Takagi Cryptosystem (1997), Okamoto-Uchiyama cryptosystem (1998), Pailier cryptosystem(1999), HIME(R) Cryptosystem (2002), Schmidt-Samoa Cryptosystem (2006) and AA_β Cryptosystem (2012). Differing from the modulus $N = pq$, research on the security of $N = p^2q$ is still scarce. Sarkar (2014) proved that the modulus $N = p^2q$ can be factored if $d < N^{0.395}$ using lattice reduction techniques. Later, Asbullah (2015) showed that one can factor $N = p^2q$ in polynomial time if e satisfies the equation $eX - (N - (ap^2 + bq^2)Y = Z$ where a, b are positive integer satisfying $\gcd(a, b) = 1$, $|ap^2 - bq^2| < N^{1/2}$, $|Z| < \frac{|ap^2-bq^2|}{3(ap^2+bq^2)}N^{1/3}Y$ and $1 \leq Y \leq X < \frac{N^{1/2}}{2(ap^2+bq^2)^{1/2}}$.

Our contribution. Hence, in this paper, we introduce new attacks to factor k RSA moduli of the form $N = p^2q$. The first attack is upon k-instances (N_i, e_i). The attack works when there exist an integer x, k integers y_i and k integers z_i satisfying $e_ix - N_iy_i = z_i - (ap_i^2 + bq_i^2)y_i$. We show that the k RSA moduli N_i can be factored in polynomial time

$$x < N^\delta, \;\; y_i < N^\delta, |z_i| < \frac{|ap_i^2 - bq_i^2|}{3(ap_i^2 + bq_i^2)}N^{1/3}y_i \;\text{ where }\; \delta = \frac{k}{3(1+k)},$$

with $N = \min_i N_i$.

The second attack works when there exist an integer y, k integers x_i and k integers z_i satisfying $e_i x_i - N_i y = z_i - (ap_i^2 + bq_i^2)y$. Similarly, we show that the k RSA moduli N_i can be factored in polynomial time

$$x_i < N^{\delta}, \;\; y < N^{\delta}, |z_i| < \frac{|ap_i^2 - bq_i^2|}{3(ap_i^2 + bq_i^2)} N^{1/3} y \;\text{ where }\; \delta = \frac{k(3\beta - 2)}{3(k+1)}$$

with $N = \min_i N_i$ and $\min_i e_i = N^{\beta}$.

For both attacks, we transform the equations into a simultaneous diophantine problem and apply lattice basis reduction techniques to find parameters (x, y_i) or (y, x_i). This leads to a suitable approximation of $ap^2 + bq^2$ which allow us to apply a theorem proposed by Asbullah (2015) in order to compute the prime factor p_i and q_i of each $N_i = p_i^2 q_i$ simultaneously.

The layout of the paper is as follows. In Section 2, we begin with a brief review on continued fractions expansion, lattice basic reduction, simultaneous diophantine approximation and also some useful results that will be used throughout the paper. In Section 3 and Section 4, we present our first and second attacks consecutively together with examples. Then, we conclude the paper in Section 5.

2 PRELIMINARIES

2.1 Lattice Basis Reductions

Let $u_1, ..., u_d$ be d linearly independent vectors of $\mathbb{R}^n$ with $d \leq n$. The set of all integer linear combinations of the vectors $u_1, ..., u_d$ is called a lattice and is in the form

$$\mathcal{L} = \left\{ \sum_{i=1}^{d} x_i u_i \;\mid\; x_i \in \mathbb{Z} \right\}.$$

The set $(u_i, ..., u_d)$ is called a basis of $\mathcal{L}$ and d is its dimension. The determinant of $\mathcal{L}$ is defined as $\det(\mathcal{L}) = \sqrt{\det(U^T U)}$ where U is the matrix of the u_i's in the canonical basis of $\mathbb{R}^n$. Define $\|v\|$ to be the Euclidean norm of a vector $v \in \mathcal{L}$. A central problem in lattice reduction is to find a short non-zero vector in $\mathcal{L}$. The LLL algorithm of Lenstra et al. (1982) produces a reduced basis and the following result fixes the sizes of the reduced basis vector (see May (2003)).

Theorem 2.1. *Let L be a lattice of dimension ω with a basis $\{v_1, ..., v_\omega\}$. The LLL algorithm produces a reduced basis $\{b_1, ..., b_\omega\}$ satisfying*

$$\|b_1\| \leq \|b_2\| \leq ... \leq \|b_i\| \leq 2^{\frac{\omega(\omega-1)}{4(\omega+1-i)}} \det(L)^{\frac{1}{\omega+1-i}},$$

for all 1 $\leq i \leq \omega$.

One of the important application of the LLL algorithm is it provides a solution to the simultaneous diophantine approximations problem which is defined as follows. Let $\alpha_1, ..., \alpha_n$ be n real numbers and ε a real number such that $0 < \varepsilon < 1$. A classical theorem of Dirichlet asserts that there exist integers $p_1, ..., p_n$ and a positive integer $q \leq \varepsilon^{-n}$ such that

$$|q\alpha_i - p_i| < \varepsilon \;\text{ for }\; 1 \leq i \leq n.$$

Lenstra et al. (1982) described a method to find simultaneous diophantine approximations to rational numbers which they consider a lattice with real entries. Hence, we state here a similar result for a lattice with integer entries.

Theorem 2.2. (Simultaneous Diophantine Approximations). *There is a polynomial time algorithm, for given rational numbers $\alpha_1, ..., \alpha_n$ and $0 < \varepsilon < 1$, to compute integers $p_1, ..., p_n$ and a positive integer q such that*

$$max_i|q\alpha_i - p_i| < \varepsilon \;\text{ and }\; q \leq 2^{n(n-3)/4} \cdot 3^n \cdot \varepsilon^{-n}.$$

Proof. See Appendix A (Nitaj et al. (2014)).

2.2 Approximation of The Prime in RSA

The following lemma shows that any approximation of $ap^2 + bq^2$ will lead to an approximation of q.

Lemma 2.1. (Asbullah, 2015). *Let $N = p^2q$ with $q < p < 2q$. Let a, b be suitable small integers with $gcd(a, b) = 1$. Let $|ap^2 - bq^2| < N^{1/2}$. Let S be an approximation of $ap^2 + bq^2$ such that $|ap^2 + bq^2 - S| < \frac{|ap^2-bq^2|}{3(ap^2+bq^2)}N^{1/3}$, then $abq = \left[\frac{S^2}{4N}\right]$.*

Proof. Set $S = ap^2 + bq^2 + r$ with $|r| < \frac{|ap^2-bq^2|}{3(ap^2+bq^2)}N^{1/3}$. Notice that

$$\begin{aligned}
(ap^2 - bq^2)^2 &= (ap^2)^2 - 2(abp^2q^2) + (bq^2)^2 \\
&= (ap^2)^2 - 2(abp^2q^2) + 2(abp^2q^2) - 2(abp^2q^2) + (bq^2)^2 \\
&= (ap^2 + bq^2)^2 - 4(abp^2q^2) \\
&= (ap^2 + bq^2)^2 - 4abqN
\end{aligned}$$

Hence we get

$$(ap^2 - bq^2)^2 = (ap^2 + bq^2)^2 - 4abqN \tag{1}$$

and consider

$$\begin{aligned}
S^2 - 4abqN &= (ap^2 + bq^2 + r)^2 - 4abqN \\
&= (ap^2 + bq^2)^2 + 2r(ap^2 + bq^2) + r^2 - 4abqN \\
&= (ap^2 + bq^2)^2 - 4abqN + 2r(ap^2 + bq^2) + r^2
\end{aligned}$$

By using (1) we can rewrite the equation as

$$S^2 - 4abqN = (ap^2 - bq^2)^2 + 2|r|(ap^2 + bq^2) + r^2 \tag{2}$$

Since $|ap^2 - bq^2| < N^{1/2}$ and $|r| < \frac{|ap^2-bq^2|}{3(ap^2+bq^2)}N^{1/3} < N^{1/3}$. Hence we have

$$\begin{aligned}\left|S^2 - 4abqN\right| &= (ap^2-bq^2)^2 + 2|r|(ap^2+bq^2) + r^2 \\ &< (N^{1/2})^2 + 2(ap^2+bq^2)\frac{|ap^2-bq^2|}{3(ap^2+bq^2)}N^{1/3} + (N^{1/3})^2 \\ &< N + \frac{2}{3}N^{1/2}N^{1/3} + N^{2/3} \\ &= N(1 + \frac{2}{3}N^{-1/6} + N^{-1/3}) \\ &< 2N\end{aligned}$$

Thus, we have $\left|S^2 - 4abqN\right| < 2N$. Divide by $4N$, we get

$$\left|\frac{S^2}{4N} - abq\right| < \frac{2N}{4N} = \frac{1}{2}$$

It follows that $abq = \left[\frac{S^2}{4N}\right]$. This terminates the proof. □

Lemma 2.2. (Asbullah, 2015). *Let $N = p^2q$ with $q < p < 2q$. Let a, b be suitable small integers with $gcd(a, b) = 1$. Let $eX - NY = Z - (ap^2 + bq^2)Y$ with $gcd(X, Y) = 1$. If $1 \le Y \le X < \frac{N^{1/2}}{2(ap^2+bq^2)^{1/2}}$ and $|Z| < \frac{|ap^2-bq^2|}{3(ap^2+bq^2)}N^{1/3}Y$, then $\frac{Y}{X}$ is a convergent of the continued fraction $\frac{e}{N}$.*

Proof. See (Asbullah (2015)).

Theorem 2.3. (Asbullah, 2015). *Let $N = p^2q$ with $q < p < 2q$. Let a, b be integers with $gcd(a, b) = 1$ such that $|ap^2 - bq^2| < N^{1/2}$. Let e be a public exponent satisfying the equation $eX - NY = Z - (ap^2 + bq^2)Y$ with $gcd(X, Y) = 1$. If $1 \le Y \le X < \frac{N^{1/2}}{2(ap^2+bq^2)^{1/2}}$ and $|Z| < \frac{|ap^2-bq^2|}{3(ap^2+bq^2)}N^{1/3}Y$, then N can be factored in polynomial time.*

Proof. Suppose e satisfies the equation $eX - NY = Z - (ap^2 + bq^2)Y$ with $\gcd(X, Y) = 1$. Let $1 \le Y \le X < \frac{N^{1/2}}{2(ap^2+bq^2)^{1/2}}$ and $|Z| < \frac{|ap^2-bq^2|}{3(ap^2+bq^2)}N^{1/3}Y$, hence satisfies Lemma 2.2. Therefore, $\frac{Y}{X}$ is a convergent of the continued fraction $\frac{e}{N}$. Rearrange the equation $eX - (N - (ap^2 + bq^2)Y = Z$ as $(ap^2 + bq^2)Y - NY + eX = Z$. Dividing by Y, we obtain

$$(ap^2 + bq^2) - N + \frac{eX}{Y} = \frac{Z}{Y}. \tag{3}$$

Set $S = N - \frac{eX}{Y}$. Then (3) can be written as $(ap^2 + bq^2) - S = \frac{Z}{Y}$ Hence, $|(ap^2 + bq^2) - S| = \frac{|Z|}{Y} < \frac{|ap^2-bq^2|}{3(ap^2+bq^2)}N^{1/3}$. By Lemma 2.1, this implies that $abq = \left[\frac{S^2}{4N}\right]$. It follows that obtaining $q = \gcd\left(\left[\frac{S^2}{4N}\right], N\right)$. □

The following thoerem is the estimation the number of the parameter $e < N$.

Theorem 2.4. (Asbullah, 2015). *Consider Theorem 2.3. The number of the parameter $e < N$ such that $e = \frac{Z+(N-(ap^2+bq^2))Y}{X}$ and $gcd(X,Y) = 1$ with*

$$1 \leq Y \leq X < \frac{N^{1/2}}{2(ap^2+bq^2)^{1/2}}, \ |Z| < \frac{|ap^2-bq^2|}{3(ap^2+bq^2)}N^{1/3}Y$$

is at least $N^{\frac{1}{6}-\epsilon}$ where $\epsilon > 0$ is arbitrarily small for suitably large N.

Proof. See (Asbullah (2015)).

3 THE FIRST ATTACK ON *k* RSA MODULI

In this section, suppose that we are given k RSA moduli $N_i = p_i^2q_i$ each with the same size N where $N = \min_i N_i$. We consider in this scenario, the RSA moduli satisfy k equations $e_ix - N_iy_i = z_i - (ap_i^2 + bq_i^2)y_i$. We show that it is possible to factor each RSA moduli $N_i = p_i^2q_i$ if the unknown parameters x, y_i and z_i are suitably small.

Theorem 3.1. *For $k \geq 2$, let $N_i = p_i^2q_i$, $1 \leq i \leq k$ be k RSA moduli each with the same size N where $N = \min_i N_i$. Let e_i, $i = 1, ..., k$ be k public exponents. Define $\delta = \frac{k}{3(1+k)}$. Let a, b be suitable small integers with gcd$(a,b) = 1$ such that $ap_i^2 + bq_i^2 < 2N^{2/3}$. If there exist an integer $x < N^\delta$ and k integers $y_i < N^\delta$ and $|z_i| < \frac{|ap_i^2-bq_i^2|}{3(ap_i^2+bq_i^2)}N^{1/3}y_i$ such that*

$$e_ix - N_iy_i = z_i - (ap_i^2 + bq_i^2)y_i$$

for $i = 1, ..., k$, then one can factor k RSA moduli $N_i = p_i^2q_i$ in polynomial time.

Proof. For $k \geq 2$ and $i = 1, ..., k$, the equation $e_ix - \Big(N_i - (ap_i^2 + bq_i^2)\Big)y_i = z_i$ can be written as $e_ix - N_iy_i = z_i - (ap_i^2 + bq_i^2)y_i$. Hence,

$$\left|\frac{e_i}{N_i}x - y_i\right| = \frac{|z_i - y_i(ap_i^2 + bq_i^2)|}{N_i} \tag{4}$$

Let $N = \min_i N_i$ and suppose that $y_i < N^\delta$ and $|z_i| < \frac{|ap_i^2-bq_i^2|}{3(ap_i^2+bq_i^2)}N^{1/3}y_i$. Then, $|z_i| < y_iN^{1/3} < N^{\delta+\frac{1}{3}}$. We set $ap_i^2 + bq_i^2 < 2N^{2/3}$, we will get

$$\begin{aligned} \frac{|z_i - y_i(ap_i^2 + bq_i^2)|}{N_i} &\leq \frac{|z_i| + y_i(ap_i^2 + bq_i^2)|}{N_i} \\ &< \frac{N^{\delta+1/3} + N^\delta(2N^{2/3})}{N} \\ &< \frac{2N^{\delta+\frac{2}{3}}}{N} \\ &= 2N^{\delta-\frac{1}{3}} \end{aligned} \tag{5}$$

Plugging (5) in (4), we obtain

$$\left|\frac{e_i}{N_i}x - y_i\right| = 2N^{\delta-\frac{1}{3}}$$

We now proceed to prove the existence of integer x. Let $\varepsilon = 2N^{\delta-\frac{1}{3}}$, $\delta = \frac{k}{3(1+k)}$. We have

$$N^{\delta} \cdot \varepsilon^k = 2^k N^{\delta + k\delta - \frac{k}{3}} = 2^k$$

Then, since $2^k < 2^{\frac{k(k-3)}{4}} \cdot 3^k$ for $k \geq 2$, we get $N^{\delta} \cdot \varepsilon^k < 2^{\frac{k(k-3)}{4}} \cdot 3^k$. It follows that if $x < N^{\delta}$, then $x < 2^{\frac{k(k-3)}{4}} \cdot 3^k \cdot \varepsilon^{-k}$. Summarizing for $i = 1, ..., k$, we have

$$\left|\frac{e_i}{N_i}x - y_i\right| < \varepsilon, \;\; x < 2^{\frac{k(k-3)}{4}} \cdot 3^k \cdot \varepsilon^{-k}$$

It follows the condition of Theorem 2.2 are fulfilled will find x and y_i for $i = 1, ..., k$. Next, using the equation $e_i x - \Big(N_i - (ap_i^2 + bq_i^2)\Big)y_i = z_i$, we get

$$(ap_i^2 + bq_i^2) - N_i + \frac{e_i x}{y_i} = \frac{z_i}{y_i}$$

Since $|z_i| < \frac{|ap_i^2 - bq_i^2|}{3(ap_i^2 + bq_i^2)} N^{1/3} y_i$ then $\frac{z_i}{y_i} < \frac{|ap_i^2 - bq_i^2|}{3(ap_i^2 + bq_i^2)} N^{1/3}$ and $S_i = N_i - \frac{e_i x}{y_i}$ is an approximation of $ap_i^2 + bq_i^2$. Hence, by using Lemma 2.1 and Theorem 2.3, this implies that $abq = \left[\frac{S^2}{4N}\right]$ for $S_i = N_i - \frac{e_i x}{y_i}$ for each $i = 1, ..., k$, we find $q_i = \gcd\left(\left[\frac{S_i^2}{4N_i}\right], N_i\right)$. This leads to the factorization of k RSA moduli $N_i, ..., N_k$. This terminates the proof. □

Example 3.1. As an illustration of the first attack, consider the following three RSA moduli and public exponents

$$\begin{aligned} N_1 &= 71405850537185027753153102246541344822641, \\ N_2 &= 83622086812913731772352818066844329216003, \\ N_3 &= 81753793974016274650553095189950555404543, \\ e_1 &= 21081727301452261853686566556023808926504, \\ e_2 &= 30528698360266832326319706490219542677843, \\ e_3 &= 56319280293205219671317298692421568858014. \end{aligned}$$

Then, $N = \max(N_1, N_2, N_3) = 83622086812913731772352818066844329216003$. Since $k = 3$, we get $\delta = \frac{k}{3(1+k)} = \frac{1}{4}$ and $\varepsilon = 2N^{\delta-\frac{1}{3}} \approx 0.000777$. Then, by using (11) in Nitaj et al. (2014) with $n = k = 3$, we find

$$C = \left[3^{n+1} \cdot 2^{\frac{(n+1)(n-4)}{4}} \cdot \varepsilon^{-n-1}\right] = 111114554156070.$$

Consider the lattice $\mathcal{L}$ spanned by the matrix

$$M = \begin{bmatrix} 1 & -[Ce_1/N_1] & -[Ce_2/N_2] & -[Ce_3/N_3] \\ 0 & C & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & C \end{bmatrix}.$$

Then, applying the LLL algorithm to $\mathcal{L}$, we get a reduced basis with the matrix

$$K = \begin{bmatrix} 315 & 1350 & 1500 & 2625 \\ -1277646367024 & -1242691890824 & 2734585090050 & -770203800216 \\ -7437619446777 & -3303197988912 & -3670219987680 & 4688570437167 \\ 3315401923149 & -11188746993606 & -85879531110 & 5405438528361 \end{bmatrix}.$$

Now, we obtain
$K \cdot M^{-1}$

$$= \begin{bmatrix} 315 & 93 & 115 & 217 \\ -1277646367024 & -377209879788 & -466442324469 & -880156386172 \\ -7437619446777 & -2195868598572 & -2715321385331 & -5123693396668 \\ 3315401923149 & 978832948739 & 1210384829086 & 2283943547058 \end{bmatrix}.$$

From the first row, we deduce $x = 315$, $y_1 = 93$, $y_2 = 115$ and $y_3 = 217$. By using x and y_i for $i = 1, 2, 3$, define $S_i = N_i - \frac{e_i x}{y_i}$ is an approximation of $ap_i^2 + bq_i^2$. Hence, by using Lemma 2.1 and Theorem 2.3, this implies that $abq = \left\lceil \frac{S^2}{4N} \right\rceil$ for $S_i = N_i - \frac{e_i x}{y_i}$. Then, we obtain $\left\lceil \frac{S_1^2}{4N_1} \right\rceil = 224732611461174$, $\left\lceil \frac{S_2^2}{4N_2} \right\rceil = 238418368587138$, $\left\lceil \frac{S_3^2}{4N_3} \right\rceil = 231335227128378$. For each $i = 1, 2, 3$, we find $q_i = \gcd\left(\left\lceil \frac{S_i^2}{4N_i} \right\rceil, N_i \right)$ and we obtain

$$q_1 = 37455435243529, \quad q_2 = 39736394764523, \quad q_3 = 38555871188063.$$

This leads us to the factorization of three RSA moduli N_1, N_2 and N_3 which $p_1 = 43662588028723$, $p_2 = 45873964046819$ and $p_3 = 46047780046031$.

4 THE SECOND ATTACK ON *k* RSA MODULI

In the section, we consider the second scenario when k RSA moduli $N_i = p_i^2 q_i$ satisfy k equations of the shape $e_i x_i - N_i y = z_i - (ap_i^2 + bq_i^2)y$ where the parameters x_i, y and z_i are suitably small unknown parameters.

Theorem 4.1. *For $k \geq 2$, let $N_i = p_i^2 q_i$, $1 \leq i \leq k$ be k RSA moduli each with the same size N where $N = \min_i N_i$. Let e_i, $i = 1, ..., k$ be k public exponents with $\min_i e_i = N^\beta$. Define $\delta = \frac{k(3\beta - 2)}{3(k+1)}$. Let a, b be suitable small integers with gcd$(a, b) = 1$ such that $ap_i^2 + bq_i^2 < 2N^{2/3}$. If there exist k integer $x_i < N^\delta$ and an integer $y < N^\delta$ and $|z_i| < \frac{|ap_i^2 - bq_i^2|}{3(ap_i^2 + bq_i^2)} N^{1/3} y$ such that*

$$e_i x_i - N_i y = z_i - (ap_i^2 + bq_i^2)y$$

for $i = 1, ..., k$, then one can factor k RSA moduli $N_i = p_i^2 q_i$ in polynomial time.

Proof. For $k \geq 2$ and $i = 1, ..., k$, the equation $e_i x_i - \Big(N_i - (ap_i^2 + bq_i^2)\Big)y = z_i$ can be written as $e_i x_i - N_i y = z_i - (ap_i^2 + bq_i^2)y$. Hence,

$$\left| \frac{N_i}{e_i} y - x_i \right| = \frac{|z_i - y(ap_i^2 + bq_i^2)|}{e_i} \tag{6}$$

Let $N = \max_i N_i$ and suppose that $y < N^{\delta}$ and $|z_i| < \frac{|ap_i^2 - bq_i^2|}{3(ap_i^2+bq_i^2)} N^{1/3} y$. Then, $|z_i| < yN^{1/3} < N^{\delta+\frac{1}{3}}$. Also, suppose that $\min_i e_i = N^{\beta}$. We set $ap_i^2 + bq_i^2 < 2N^{2/3}$, then we will get

$$\begin{aligned}
\frac{|z_i - y(ap_i^2 + bq_i^2)|}{N^{\beta}} &\leq \frac{|z_i| + y(ap_i^2 + bq_i^2)|}{N^{\beta}} \\
&< \frac{N^{\delta+1/3} + N^{\delta}(2N^{2/3})}{N^{\beta}} \\
&< \frac{2N^{\delta+\frac{2}{3}}}{N^{\beta}} \\
&= 2N^{\delta+\frac{2}{3}-\beta} \qquad (7)
\end{aligned}$$

Plugging (7) in (6), we obtain

$$\left|\frac{N_i}{e_i} x_i - y\right| = 2N^{\delta+\frac{2}{3}-\beta}.$$

We now proceed to prove the existence of integer y and the integers x_i. Let $\varepsilon = 2N^{\delta+\frac{2}{3}-\beta}$, $\delta = \frac{k(3\beta-2)}{3(k+1)}$. Then, we obtain

$$N^{\delta} \cdot \varepsilon^k = N^{\delta}(2N^{\delta+\frac{2}{3}-\beta})^k = 2^k(N^{\delta+\delta k+\frac{2}{3}k-\beta k}) = 2^k.$$

Then, since $2^k < 2^{\frac{k(k-3)}{4}} \cdot 3^k$ for $k \geq 2$, we get $N^{\delta} \cdot \varepsilon^k < 2^{\frac{k(k-3)}{4}} \cdot 3^k$. It follows that if $y < N^{\delta}$, then $y < 2^{\frac{k(k-3)}{4}} \cdot 3^k \cdot \varepsilon^{-k}$. Summarizing for $i = 1, ..., k$, we get

$$\left|\frac{N_i}{e_i} y - x_i\right| < \varepsilon, \quad y < 2^{\frac{k(k-3)}{4}} \cdot 3^k \cdot \varepsilon^{-k}, \text{ for } i = 1, ..., k,$$

It follows the condition of Theorem 2.2 are fulfilled will find y and x_i for $i = 1, ..., k$. Next, using the equation $e_i x_i - \Big(N_i - (ap_i^2 + bq_i^2)\Big) y = z_i$, we get

$$(ap_i^2 + bq_i^2) - N_i + \frac{e_i x_i}{y} = \frac{z_i}{y}$$

Since $|z_i| < \frac{|ap_i^2 - bq_i^2|}{3(ap_i^2+bq_i^2)} N^{1/3} y$, then $\frac{|z_i|}{y} < \frac{|ap_i^2 - bq_i^2|}{3(ap_i^2+bq_i^2)} N^{1/3}$ and $S_i = N_i - \frac{e_i x_i}{y}$ is an approximation of $ap_i^2 + bq_i^2$. Hence, by using Lemma 2.1 and Theorem 2.3, this implies that $abq = \left[\frac{S^2}{4N}\right]$ for $S_i = N_i - \frac{e_i x_i}{y}$ for each $i = 1, ..., k$, we find $q_i = \gcd\left(\left[\frac{S_i^2}{4N_i}\right], N_i\right)$. This leads to the factorization of k RSA moduli $N_i, ..., N_k$. This terminates the proof. □

Example 4.1. As an illustration of the second attack, consider the following three RSA moduli and public exponents

$$\begin{aligned}
N_1 &= 8969844924801174805316486803059616174 1473, \\
N_2 &= 85026897721391614213389394740488811877211, \\
N_3 &= 22195439947956831029781113035068543467503 7, \\
e_1 &= 41473476534022710186899569461496585030420, \\
e_2 &= 51495163408724921704345941632904239372728, \\
e_3 &= 15645965864952050680460908752975746984648 0.
\end{aligned}$$

Then, $N = \max(N_1, N_2, N_3) = 2219543994795683102978111303506854346750 37$. We also obtain $\min(e_1, e_2, e_3) = N^{\beta}$ with $\beta \approx 0.9791$. Since $k = 3$, we get $\delta = \frac{k(3\beta-2)}{3(k+1)} = 0.2367855052$ and $\varepsilon = 2N^{\delta+\frac{2}{3}-\beta} \approx 0.0010905142500234$. Then, by using (11) in Nitaj et al. (2014) with $n = k = 3$, we find

$$C = \left[3^{n+1} \cdot 2^{\frac{(n+1)(n-4)}{4}} \cdot \varepsilon^{-n-1}\right] = 28637140038889.$$

Consider the lattice $\mathcal{L}$ spanned by the matrix

$$M = \begin{bmatrix} 1 & -[CN_1/e_1] & -[CN_2/e_2] & -[CN_3/e_3] \\ 0 & C & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & C \end{bmatrix}.$$

Then, applying the LLL algorithm to $\mathcal{L}$, we get a reduced basis with the matrix

$$K = \begin{bmatrix} -43 & 268 & 169 & 124 \\ -2479239519408 & -1197495576067 & 420267865942 & 1155614885037 \\ 429520576715 & -1345051964694 & 3639724252239 & -1904597477943 \\ -1671575260344 & 1760449982945 & 575859503445 & -5169343448853 \end{bmatrix}.$$

Now, we obtain
$K \cdot M^{-1}$

$$= \begin{bmatrix} -43 & -93 & -71 & -61 \\ -2479239519408 & -5362076169883 & -4093628043674 & -3517060713579 \\ 429520576715 & 928963107779 & 709208394111 & 609319887898 \\ -1671575260344 & -3615267423535 & -2760042871731 & -2371304439093 \end{bmatrix}.$$

From the first row, we deduce $y = 43$, $x_1 = 93$, $x_2 = 71$ and $x_3 = 61$. By using y and x_i for $i = 1, 2, 3$, define $S_i = N_i - \frac{e_i x_i}{y}$ is an approximation of $ap_i^2 + bq_i^2$. Hence, by using Lemma 2.1 and Theorem 2.3, this implies that $abq = \left[\frac{S^2}{4N}\right]$ for $S_i = N_i - \frac{e_i x_i}{y}$. Then, we obtain $\left[\frac{S_1^2}{4N_1}\right] = 232951773063462$, $\left[\frac{S_2^2}{4N_2}\right] = 217660777111506$, $\left[\frac{S_3^2}{4N_3}\right] = 287031111167118$. For each $i = 1, 2, 3$, we find $q_i = \gcd\left(\left[\frac{S_i^2}{4N_i}\right], N_i\right)$ and we obtain

$$q_1 = 38825295510577, \quad q_2 = 36276796185251, \quad q_3 = 47838518527853.$$

This leads us to the factorization of three RSA moduli N_1, N_2 and N_3 which $p_1 = 48065679074407$, $p_2 = 48413190516781$ and $p_3 = 68115040311623$.

5 CONCLUSION

In conclusion, this paper presents two new attacks on k RSA moduli $N_i = p_i^2 q_i$. We focus on the system of generalized key equations of the form $e_i x - N_i y_i = z_i - (ap_i^2 + bq_i^2)y_i$ for the first attack and the form $e_i x_i - N_i y = z_i - (ap_i^2 + bq_i^2)y$ for the second attack. We show that both of the attacks are successful when the parameters x, x_i, y, y_i and z_i are suitably small. On top of that, we also prove that both of our attacks enables us to factor k RSA moduli of the form $N_i = p_i^2 q_i$ simultaneously based on LLL algorithm.

ACKNOWLEDGEMENT

The author gratefully acknowledged the MyPhD financial support from the Ministry of Higher Education Malaysia.

REFERENCES

Asbullah, M. A. (2015). *Cryptanalysis on the modulus $N = p^2q$ and design of Rabin-like Cryptosystem without decryption failure*. PhD thesis, Universiti Putra Malaysia.

Blömer, J. and May, A. (2004). A qeneralized Wiener attack on RSA. *Practice and Theory in Public Key Cryptography PKC 2004 LNCS Springer-Verlag*, 2947:1–13.

Boneh, D. (1999). Twenty years of attacks on the RSA cryptosystem. *Notices of the AMS*, 46(2):203–213.

Boneh, D. and Durfee, G. (1999). Cryptanalysis of RSA with private key d less than $N^{0.292}$. *Advance in Cryptology-Eurocrypt'99, Lecture Notes in Computer Science*, 1592:1–11.

Hinek, J. (2007). *On the security of some variants of RSA*. PhD thesis, Waterloo, Ontario, Canada.

Howgrave-Graham, N. and Seifert, J. (1999). Extending Wiener attack in the presence of many decrypting exponents. *In Secure Networking-CQRE (Secure)'99 LNCS 1740 Springer-Verlag*, 1740:153–166.

Lenstra, A. K., Lenstra, H. W., and Lovász, L. (1982). Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261:513–534.

May, A. (2003). *New RSA vulnerabilities using lattice reduction methods*. PhD thesis, University of Paderborn.

May, A. (2004). Secret exponent attacks on RSA-type scheme with moduli $N = p^r q$. *In PKC 2004 LNCS Springer-Verlag*, 2947:218–230.

Nitaj, A., Ariffin, M. R. K., Nassr, D. I., and Bahig, H. M. (2014). *New attacks on the RSA cryptosystem*, volume 8469 of *Lecture Notes in Computer Science*, pages 178–198. Springer-Verlag.

Rivest, R., Shamir, A., and Adleman, L. (1978). A method for obtaining digital signatures and public-key cryptosystems. *Communication of the ACM 21(2)*, 21(2):17–28.

Sarkar, S. (2014). Small secret exponent attack on rsa variant with modulus $N = p^2q$. *Designs, Codes and Cryptography*, 73(2):383–392.

Wiener, M. (1990). Cryptanalysis of short RSA secret exponents. *IEEE Transaction on Information Theory IT-36*, 36:553–558.

New Directions in Factoring the Prime Power RSA modulus $N = p^r q$

Sadiq Shehu*[1] and **Muhammad Rezal Kamel Ariffin**[1,2]

[1]*Al-Kindi Cryptography Research Laboratory, Institute for Mathematical Research, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.*
[2]*Department of Mathematics, Faculty of Science, Universiti Putra Malaysia , 43400 UPM Serdang, Selangor, Malaysia.*

*E-mail: *sadiqshehuzezi@gmail.com,rezal@upm.edu.my*
**Corresponding author*

ABSTRACT

Factoring large integers is a fundamental problem in algebraic number theory and modern cryptography, which many cryptosystem such as RSA are based on. This paper proposes three new attacks on the Prime Power RSA modulus $N = p^r q$. In the first attack we consider the class of public key exponents satisfying an equation $eX - NY - (ap^r + bq^r + Z) = 1$ where a, b are suitably small integers with $gcd(a, b) = 1$. Using the continued fraction algorithm, we show that such exponents yield the factorization of the RSA Prime Power modulus in polynomial time. Further, we show that the number of such weak keys is at least $N^{\frac{5}{6}-\varepsilon}$ where $\varepsilon > 0$ is arbitrarily small for large N. We furthered our analysis on k Prime Power RSA moduli $N_i = p_i^r q_i$ satisfying a variant of the above mentioned condition. We utilized the LLL algorithm on k Prime Power RSA public keys (N_i, e_i) with $N_i = p_i^r q_i$ and we were able to factorize the k Prime Power RSA moduli $N_i = p_i^r q_i$ simultaneously in polynomial time.

Keywords: Prime Power RSA, Cryptanalysis, Factorization, Continued fraction,

1 INTRODUCTION

The RSA cryptosystem is one of the most practical public key cryptosystems and is used throughout the world (See Rivest et al. (1978)).The mathematical operations in RSA depend on three parameters, the modulus $N = pq$ which is the product of two large primes p and q, the public exponent e and the private exponent d, related by the congruence $ed \equiv 1(mod(p-1)(q-1))$. The encryption and decryption in RSA require taking heavy exponential multiplications modulus the large integer $N = pq$. Many RSA variants have been proposed in order to ensure computational efficiency while maintaining the acceptable levels of security. One such important variant is the Prime Power RSA. In Prime Power RSA the modulus N is in the form $N = p^r q$ for $r \geq 2$.

An important attack on multi-power RSA $N = p^r q$ was showed by Takagi (1998) and Sarkar (2014), by extending the small private exponent attack of Boneh and Durfee on classical RSA. In particular, he showed that N can be factored efficiently for $r = 2$ with private exponent d satisfying $d < N^{0.395}$. As in the standard RSA cryptosystem, the security of the Prime Power RSA depends on the difficulty of factoring integers of the form $N = p^r q$. Recently for $N = p^r q$, Nitaj and Rachidi (2015) present three new attacks: In the first attack they consider a public exponent e satisfying an equation $ex - \phi(N)y = z$ where x and y are positive integers using a recent result of Lu et al. (2014), they show that one can factor N in polynomial time if $|xz| < N^{\frac{r(r-1)}{(r+1)^2}}$. Also recently for $N = p^2 q$, Asbullah and Ariffin (2015) in their paper considered a public exponent e satisfying an equation $eX - NY = Z - (ap^2 + bq^2)Y$ where $1 \leq Y \leq X < \frac{1}{2}N^{\frac{1}{6}-\frac{\alpha}{2}}$ and $Z < \frac{1}{3}N^{\frac{1}{3}+\alpha}Y$ using the result of Nitaj (2009), he we show that one can factor $N = p^2 q$ in polynomial time.

In Hinek (2007), showed that it is possible to factor k moduli N_i using k equations of the form $e_i d - k_i\phi(N_i) = 1$ if $d < N^\delta$ with $\delta = \frac{k}{2(k+1)} - \varepsilon$ where ε is a small constant depending on the size of max N_i. Very recently, with k RSA public keys (N_i, e_i), Nitaj et al. (2014), presented a method that factor the k RSA moduli N_i using k equations of the shape $e_i x - y_i\phi(N_i) = z_i$ or of the shape $e_i x_i - y\phi(N_i) = z_i$ where $N_i = p_i q_i$, $\phi(N_i) = (p_i - 1)(q_i - 1)$ and the parameters x, x_i, y, y_i, z_i are suitably small in terms of the prime factors of the moduli.

This paper combines the attack strategies of Nitaj et al. (2014) and M.A. Asbullah and Ariffin to mount a new attack on RSA prime power modulus $N = p^r q$, for $r \geq 2$. In the first attack we considered the class of public key exponents satisfying the equation $eX - NY - (ap^r + bq^r + Z) = 1$ where a, b are suitably positive integers with $gcd(X, Y) = 1$. If $X < \frac{N}{3(ap^r+bq^r)}$ and $|Z| < \frac{|ap^r-bq^r|}{3(ap^r+bq^r)}N^{\frac{1}{3}}$,then N can be factored in polynomial time. We also show that the number of the exponents e with $e < N$ for which this method works can be estimated as $N^{\frac{5}{6}-\varepsilon}$ where $\varepsilon > 0$ is arbitrarily small for large N.

The second attack works for $k \geq 2$, $r \geq 2$ moduli $N_i = p_i^r q_i$, $i = 1, ..., k$ when k instances (N_i, e_i) are such that there exist an integer x, k integers y_i, and k integers z_i satisfying $e_i x - N_i y_i - (ap_i^r + bq_i^r + z_i) = 1$. Hence we show that the k RSA moduli N_i can be factored in polynomial time if

$$x < N^\delta, \quad y_i < N^\delta, \quad |z_i| < \frac{ap_i^r - bq_i^r}{3(ap_i^r + bq_i^r)}N^{\frac{1}{3}} \quad \text{with} \quad \delta = \frac{k(1 - \alpha r - \alpha)}{(r + 1)}$$

where $N = min_i N_i$.

Third attack works when the k instance (N_i, e_i) of RSA are such that there exist an integer y, and k integers x_i and k integers z_i satisfying $e_i x_i - N_i y - (ap_i^r + bq_i^r + z_i) = 1$. Also we show that the k RSA moduli N_i can be factored in polynomial time if

$$x_i < N^\delta, \quad y < N^\delta, \quad |z_i| < \frac{ap_i^r - bq_i^r}{3(ap_i^r + bq_i^r)}N^{\frac{1}{3}} \quad \text{with} \quad \delta = \frac{kr(\beta - \alpha - 1) + k(\beta - \alpha)}{(r + 1)}$$

where $min_i N = min_i N_i$ and $e_i = N^{\beta}$. In both the second and third attacks we transformed the equations into simultaneous diophantine problem and apply lattice basis reduction techniques to find the parameters (x, y_i) or (y, x_i). This will result in the factorization of k RSA moduli N_i

The rest of the paper is organized as follows. In Section 2 we give a brief introduction to continued fractions, lattice basis reduction and simultaneous diophantine approximations with some useful lemmas required for the attacks. In Section 3 we present the first new attacks and estimation for the size of the public exponents for which our attacks work. In Section 4, and 5 we present the second and third attacks with numerical example respectively. Section 6 concludes the paper.

2 PRELIMINARIES

In this section, we give some definition and an important theorems concerning the continued fraction, lattice basis reduction and simultaneous diophantine approximations with some useful lemmas that we need to perform our attacks.

2.1 Continued fraction

The continued fraction of a real number R is an expression of the form

$$R = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + ...}}}$$

This expression is often used in the form $x = [a_0, a_1, a_2....]$. Any rational number $\frac{a}{b}$ can be expressed as a finite continued fraction $x = [a_0, a_1, a_2....a_m]$. For $i \geq 0$ $(0 \leq i \leq m$ in the finite case), we define the i^{th} convergent of the continued fraction $[a_0, a_1, a_2, ...]$ to be $[a_0, a_1, a_2, ..., a_i]$. Each convergent is a rational number.

Theorem 2.1. *Let $x = [a_0, a_1, a_2, ...a_m]$ be a continued fraction expansion of x. If X and Y are coprime integers such that*

$$\left| x - \frac{Y}{X} \right| < \frac{1}{2X^2}$$

Then $Y = p_n$ and $X = q_n$ for some convergent $\frac{p_n}{q_n}$ of x with $n \geq 0$.

Proposition 2.1. *Let $\frac{q}{p} = [a_0, a_1, ...a_m]$ be a continued fraction for $0 \leq i < m$, we have*

$$\left| \frac{q}{p} - \frac{p_i}{q_i} \right| < \frac{1}{q_i^2}$$

2.2 Lattice Basis Reductions

The most powerful attacks on RSA are based on techniques that use lattice basis reduction algorithms, such as the LLL algorithm. Lenstra et al. (1982), LLL is a polynomial time algorithm for lattice basis reduction with many applications in cryptography.

Definition 2.1. *Let $b_1, ..., b_n \in R^m$ be $n \leq m$ linearly independent vectors. The lattice generated by $b_1, ..., b_n$ is the set*

$$\mathcal{L} = \sum_{i=1}^{n} Zb_i = \left\{ \sum_{i=1}^{n} x_i b_i | x_i \in Z \right\}.$$

The set $B = \langle b_1, ..., b_n \rangle$ is called a lattice basis for $\mathcal{L}$. The lattice dimension is $dim(\mathcal{L}) = n$. If $n = m$ then $\mathcal{L}$ is said to be a full rank lattice.

A lattice $\mathcal{L}$ can be represented by a basis matrix. Given a basis B, a basis matrix M for the lattice generated by B is the $n \times m$ matrix defined by the rows of the set $b_1 ..., b_n$

$$M = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

It is often useful to represent the matrix M by B. A very important notion for the lattice $\mathcal{L}$ is the determinant.

Let $\mathcal{L}$ be a lattice generated by the basis $B = \langle b_1, ..., b_n \rangle$. The determinant of $\mathcal{L}$ is defined as

$$det(\mathcal{L}) = \sqrt{det(BB^T)}.$$

If $n = m$, we have

$$det(\mathcal{L}) = \sqrt{det(BB^T)} = |det(B)|.$$

Theorem 2.2. *Let L be a lattice of dimension ω with a basis $v_1, ..., v_\omega$. The LLL algorithm produces a reduced basis $b_1, ...b_\omega$ satisfying*

$$\|b_1\| \leq \|b_2\| \leq ... \leq \|b_i\| \leq 2^{\frac{\omega(\omega-1)}{4(\omega+1-i)}} det\mathcal{L}^{\frac{1}{\omega+1-i}}$$

for all $1 \leq i \leq \omega$
As an application of the LLL algorithm is that it provides a solution to the simultaneous diophantine approximations problem which is defined as follows. Let $\alpha_1, ..., \alpha_n$ be n real numbers and ε a real number such that $0 < \varepsilon < 1$. A classical theorem of Dirichlet asserts that there exist integers $p_1, ..., p_n$ and a positive integer $q \leq \varepsilon^{-n}$ such that

$$|q\alpha_i - p_i| < \varepsilon \quad for \quad 1 \leq i \leq n.$$

A method to find simultaneous diophantine approximations to rational numbers was described by Lenstra et al. (1982) In their work, they considered a lattice with real entries. Below a similar result for a lattice with integer entries.

Theorem 2.3. *(Simultaneous Diophantine Approximations). There is a polynomial time algorithm, for given rational numbers $\alpha_1, ..., \alpha_n$ and $0 < \varepsilon < 1$, to compute integers $p_1, ..., p_n$ and a positive integer q such that*

$$max_i \, |q\alpha_i - p_i| < \varepsilon \quad and \quad q \leq 2^{\frac{n(n-3)}{4}}.$$

Proof. See Nitaj et al. (2014) Appendix A.

2.3 Useful Lemma

Lemma 2.1. *Let $N = p^r q$ be an RSA modulus prime power with $q < p < 2q$. Then*

$$2^{-\frac{r}{r+1}} N^{\frac{1}{r+1}} < q < N^{\frac{1}{r+1}} < p < 2^{\frac{1}{r+1}} N^{\frac{1}{r+1}}$$

Proof. Suppose $N = p^r q$. Then multiplying by p^r we get $p^r q < p^r p < 2p^r q$ that is $N < p^{r+1} < 2N$ which implies $N^{\frac{1}{r+1}} < p < 2^{\frac{1}{r+1}} N^{\frac{1}{r+1}}$ Also since $N = p^r q$, then $q = \frac{N}{p^r}$ which in turn implies $2^{-\frac{r}{r+1}} N^{\frac{1}{r+1}} < q < N^{\frac{1}{r+1}}$ □

Lemma 2.2. *Let $N = p^r q$ be an RSA modulus prime power with $q < p < 2q$. Let a, b be a suitably small integers with $gcd(a, b) = 1$. Let $|ap^r - bq^r| < N^{\frac{1}{2}}$. Also let M be an approximation of $|ap^r + bq^r|$ such that $|ap^r + bq^r - M| < \frac{|ap^r - bq^r|}{3(ap^r + bq^r)} N^{\frac{1}{3}}$, then $abq^{r-1} = \left[\frac{M^2}{4N}\right]$.*

Proof. Set $M = ap^r + bq^r + x$ with $x < \frac{|ap^r - bq^r|}{3(ap^r + bq^r)} N^{\frac{1}{3}} < N^{\frac{1}{3}}$
Observe that

$$\begin{aligned}(ap^r - bq^r)^2 &= (ap^r - bq^r)(ap^r - bq^r) \\ &= a^2p^{2r} - abq^r p^r - abq^r p^r + b^2 q^{2r} \\ &= a^2p^{2r} - 2abq^r p^r + b^2 q^{2r} \\ &= a^2p^{2r} + 2abq^r p^r - 2abq^r p^r - 2abq^r p^r + b^2 q^{2r} \\ &= (ap^r + bq^r)^2 - 4abq^r p^r \\ &= (ap^r + bq^r)^2 - 4abNq^{r-1}\end{aligned}$$

Hence

$$(ap^r - bq^r)^2 = (ap^r + bq^r)^2 - 4abNq^{r-1} \tag{1}$$

Consider

$$\begin{aligned}&M^2 - 4abNq^{r-1} = (ap^r + bq^r + x)^2 - 4abNq^{r-1} \\ &= (ap^r + bq^r + x)(ap^r + bq^r + x) - 4abNq^{r-1} \\ &= a^2p^{2r} + 2abp^r q^r + b^2 q^{2r} + 2x(ap^r + bq^r) + x^2 - 4abNq^{r-1} \\ &= (ap^r + bq^r)^2 + 2x(ap^r + bq^r) + x^2 - 4abNq^{r-1} \\ &= (ap^r - bq^r)^2 + 2x(ap^r + bq^r) + x^2\end{aligned}$$

Therefore

$$M^2 - 4abNq^{r-1} = (ap^r - bq^r)^2 + 2x(ap^r + bq^r) + x^2 \tag{2}$$

Suppose that $|ap^r - bq^r| < N^{\frac{1}{2}}$ and $x < \frac{|ap^r - bq^r|}{3(ap^r+bq^r)} N^{\frac{1}{3}} < N^{\frac{1}{3}}$
Then from (2) we have

$$\begin{aligned}
&\left|M^2 - 4abNq^{r-1}\right| \\
&= \left|(ap^r - bq^r)^2 + 2x(ap^r + bq^r) + x^2\right| \\
&< \left(N^{\frac{1}{2}}\right)^2 + 2(ap^r + bq^r)\frac{|ap^r - bq^r|}{3\,(ap^r + bq^r)} N^{\frac{1}{3}} + (N^{\frac{1}{3}})^2 \\
&< N + \frac{2}{3}(N^{\frac{1}{2}})N^{\frac{1}{3}} + N^{\frac{2}{3}} \\
&< N + \frac{2}{3}N^{\frac{5}{6}} + N^{\frac{2}{3}} \\
&< 2N
\end{aligned}$$

Thus we have $\left|M^2 - 4abNq^{r-1}\right| < 2N$. Then dividing by 4N,

$$\begin{aligned}
\left|\frac{M^2}{4N} - abq^{r-1}\right| &= \frac{\left|M^2 - 4abNq^{r-1}\right|}{4N} \\
&< \frac{2N}{4N} = \frac{1}{2}
\end{aligned}$$

It follows that

$$abq^{r-1} = \left[\frac{M^2}{4N}\right]$$

□

Lemma 2.3. *Let $N = p^r q$ be an RSA modulus prime power with $q < p < 2q$. Let a, b be a suitably small integers with $gcd(a, b) = 1$. let M be an approximation of $|ap^r + bq^r|$ such that $|M - |ap^r + bq^r|| < \frac{|ap^r - bq^r|}{3(ap^r+bq^r)} N^{\frac{1}{3}}$ Then*

$$|H - |ap^r - bq^r|| < N^{\frac{1}{3}}$$

Where $H = \sqrt{M^2 - 4abNq^{r-1}}$

Proof. Let a and b be a suitably small integers and suppose that M satisfies $|M - |ap^r + bq^r|| < \frac{|ap^r - bq^r|}{3(ap^r+bq^r)} N^{\frac{1}{3}}$ Then

$$\begin{aligned}
&\left|H^2 - (ap^r - bq^r)^2\right| = \left|\left|M^2 - 4abNq^{r-1}\right| - (ap^r - bq^r)^2\right| \\
&\le \left|M^2 - 4abNq^{r-1} - (ap^r - bq^r)^2\right| \\
&= \left|M^2 - (ap^r + bq^r)^2\right| \\
&= \left|M^2 - M(ap^r + bq^r) + M(ap^r + bq^r) - (ap^r + bq^r)^2\right| \\
&= (M + |ap^r + bq^r|)\,|M - |ap^r + bq^r|| \\
&\le (M + |ap^r + bq^r|) \times \frac{|ap^r - bq^r|}{3\,(ap^r + bq^r)} N^{\frac{1}{3}}
\end{aligned}$$

Dividing by $|H + (ap^r - bq^r)|$, implies

$$|H - (ap^r - bq^r)| \leq \frac{(M + |ap^r + bq^r|)}{|H + (ap^r - bq^r)|} \times \frac{|ap^r - bq^r|}{3\,(ap^r + bq^r)} N^{\frac{1}{3}} \tag{3}$$

Using the approximation

$$|M - |ap^r + bq^r|| < \frac{|ap^r - bq^r|}{3\,(ap^r + bq^r)} N^{\frac{1}{3}}$$

we have

$$M < ap^r + bq^r + \frac{|ap^r - bq^r|}{3\,(ap^r + bq^r)} N^{\frac{1}{3}}$$

Let $|ap^r - bq^r| < |ap^r + bq^r|$ then

$$\begin{aligned} M + ap^r + bq^r &< \left(ap^r + bq^r + \frac{|ap^r - bq^r|}{3\,(ap^r + bq^r)} N^{\frac{1}{3}} \right) + ap^r + bq^r \\ &< 2(ap^r + bq^r) + \frac{|ap^r - bq^r|}{3\,(ap^r + bq^r)} N^{\frac{1}{3}} \\ &< 2(ap^r + bq^r) + \frac{(ap^r + bq^r)}{3\,(ap^r + bq^r)} N^{\frac{1}{3}} \\ &< 2(ap^r + bq^r) + \frac{1}{3} N^{\frac{1}{3}} \\ &< 2(ap^r + bq^r) + N^{\frac{1}{3}} \\ &< 3(ap^r + bq^r) \end{aligned}$$

Substituting the above in to (3) we obtain

$$|H - (ap^r - bq^r)| \leq \frac{3(ap^r + bq^r)}{|ap^r - bq^r|} \times \frac{|ap^r - bq^r|}{3(ap^r + bq^r)} N^{\frac{1}{3}} \quad < N^{\frac{1}{3}}$$

□

3 THE NEW ATTACK ON RSA PRIME POWER MODULUS

Lemma 3.1. *Let $N = p^r q$ be an RSA modulus prime power with $q < p < 2q$. Suppose that e is a public key exponent satisfying the equation $eX - NY - (ap^r + bq^r + Z) = 1$ with $gcd(X, Y) = 1$, if $X < \frac{N}{3(ap^r + bq^r)}$ and $|Z| < \frac{|ap^r - bq^r|}{3(ap^r + bq^r)} N^{\frac{1}{3}}$, then $\frac{Y}{X}$ is among the convergent of the continued fraction expansion of $\frac{e}{N}$.*

Proof. Suppose that $|Z| < \frac{|ap^r - bq^r|}{3(ap^r + bq^r)} N^{\frac{1}{3}}$ thus $Z < N^{\frac{1}{3}}$ also let $X < \frac{N}{3(ap^r + bq^r)}$. Consider the equation $eX - NY - (ap^r + bq^r + Z) = 1$ when dividing by NX we obtain

$$\begin{aligned}\left|\frac{e}{N} - \frac{Y}{X}\right| &= \frac{|eX - NY|}{NX} \\ &= \frac{|1 + ap^r + bq^r + Z|}{NX} \\ &< \frac{|ap^r + bq^r + Z|}{NX} \\ &\leq \frac{|Z|}{NX} + \frac{|ap^r + bq^r|}{NX} \\ &\leq \frac{N^{\frac{1}{3}} + |ap^r + bq^r|}{NX}\end{aligned}$$

If the condition $\frac{N^{\frac{1}{3}} + |ap^r + bq^r|}{NX} < \frac{1}{2X^2}$ hold, then $X < \frac{N}{2(ap^r + bq^r) + N^{\frac{1}{3}}}$
Therefore by Theorem 2.1, we conclude that $\frac{Y}{X}$ is among the convergent of the continued fraction expansion of $\frac{e}{N}$. Hence according to Lemma 3.3 such condition is satisfies if $X < \frac{N}{3(ap^r + bq^r)}$.
□

Theorem 3.1. *Let $N = p^r q$ be an RSA modulus prime power with $q < p < 2q$. let a, b be a suitably small integers with $gcd(a, b) = 1$ and suppose that e is a public key exponent satisfying the equation $eX - NY - (ap^r + bq^r + Z) = 1$ with $gcd(X, Y) = 1$, if $X < \frac{N}{3(ap^r + bq^r)}$ and $|Z| < \frac{|ap^r - bq^r|}{3(ap^r + bq^r)} N^{\frac{1}{3}}$. Then N can be factored in polynomial time.*

Proof. Suppose that e satisfies an equation $eX - NY - (ap^r + bq^r + Z) = 1$ with $gcd(X, Y) = 1$, let X and Z satisfy the condition of Lemma 4.1, then $\frac{Y}{X}$ is among the convergent of the continued fraction expansion of $\frac{e}{N}$. Also using X and Y, we define
$M = eX - NY$, $H = \sqrt{M^2 - 4abNq^{r-1}}$. Then since M is an approximation of $|ap^r + bq^r|$ and H is an approximation of $|ap^r - bq^r|$
Therefore combining M and H, we have

$$ap^r + bq^r + ap^r - bq^r = 2ap^r = M \pm H$$

Which implies that $ap^r = \frac{M \pm \sqrt{M^2 - 4\left[\frac{M^2}{4N}\right] N}}{2}$ Hence, we recover q using $q = gcd\,(ap^r, N)$. Then $p^r = \frac{N}{q}$ which lead to factorization of N in polynomial time. □

Algorithm 1

```
Input: A Prime Power RSA modulus $N = p^r q$,with $q < p < 2q$ and its public key exponent $e$.
Output: The prime factors $p$ and $q$.
1: Compute the continued fraction expansion of $\frac{e}{N}$.
2: For every convergent $\frac{Y}{X}$ of $\frac{e}{N}$, compute $M = eX - NY$.
3: Compute $K = \left[\frac{M^2}{4N}\right]$, and $H = \sqrt{|M^2 - 4KN|}$
4: Compute $q = gcd\left(N, \frac{M \pm H}{2}\right)$.
```

```
If $1 < q < N$ then
    Output the factors $q$ and $p^r = \left(\frac{N}{q}\right)$.
End if.
End for.
```

Example 3.1. *As an example to illustrate our attack with $r = 3$, let us take for N and e the numbers*

$$N = 35476872536008365493$$

$$e = 29091076177383811209$$

Suppose that N and e satisfy an equation of the form $eX - NY - (ap^r + bq^r + Z) = 1$ with $gcd(X, Y) = 1$, $X < \frac{N}{3(ap^r+bq^r)}$ and $|Z| < \frac{|ap^r-bq^r|}{3(ap^r+bq^r)}N^{\frac{1}{3}}$.
Following Lemma 3.1, $\frac{Y}{X}$ is one of the convergent of the continued fraction of $\frac{e}{N}$. The first convergents of the continued fraction expansion of $\frac{e}{N}$ are

$$\left[0, 1, \frac{4}{5}, \frac{5}{6}, \frac{9}{11}, \frac{32}{39}, \frac{41}{50}, \frac{14259}{17389}, \frac{14300}{17439}, \frac{142959}{174340}, \frac{157259}{191779}, \frac{300218}{366119}, \frac{757695}{924017}, \frac{1057913}{1290136}, \ldots\right]$$

Applying the factorization algorithm with the convergent $\frac{Y}{X} = \frac{41}{50}$, we obtain

$$M = eX - NY = 2034892847575237$$

$$K = \left[\frac{M^2}{4N}\right] = 29179495014$$

And

$$H = \sqrt{|M^2 - 4KN|} = 2403783919$$

$$\frac{M - H}{2} = 1017445221895659$$

Hence we compute $q = gcd\left(N, \frac{M-H}{2}\right) = (1017445221895659, 35476872536008365493) = 69737$. Finally for $q = 69737$ we compute $p = \sqrt[3]{\frac{N}{q}} = 79829$, which leads to the factorization of N.

3.1 The Number of Weak Exponents

Lemma 3.2. *Let $N = p^r q$ be an RSA modulus prime power with $q < p < 2q$. Let a, b be a suitably small integers with $gcd(a, b) = 1$. and Suppose that e is a public key exponent satisfying the two equation*

$$eX_1 - NY_1 - (ap^r + bq^r + Z_1) = 1$$

and

$$eX_2 - NY_2 - (ap^r + bq^r + Z_2) = 1$$

with $gcd(X_1, Y_1) = gcd(X_2, Y_2) = 1$, for $i = 1, 2$, $1 \leq Y_i \leq X_i < \frac{N}{3(ap^r+bq^r)}$ and $|Z_i| < \frac{|ap^r-bq^r|}{3(ap^r+bq^r)}N^{\frac{1}{3}}$, then $X_1 = X_2$, $Y_1 = Y_2$ and $Z_1 = Z_2$

Proof. Suppose that e is a public key exponent satisfying the two equation

$$eX_1 - NY_1 - (ap^r + bq^r + Z_1) = 1$$

$$eX_2 - NY_2 - (ap^r + bq^r + Z_2) = 1$$

with $gcd(X_1, Y_1) = gcd(X_2, Y_2) = 1$, for $i = 1, 2$, $1 \leq Y_i \leq X_i < \frac{N}{3(ap^r+bq^r)}$ and $|Z_i| < \frac{|ap^r-bq^r|}{3(ap^r+bq^r)}N^{\frac{1}{3}}$
Then

$$e = \frac{Z_1 + NY_1 + (ap^r + bq^r + 1)}{X_1}$$

$$e = \frac{Z_2 + NY_2 + (ap^r + bq^r + 1)}{X_2}$$

By equating e we get

$$X_2(Z_1 + NY_1 + ap^r + bq^r + 1) = X_1(Z_2 + NY_2 + ap^r + bq^r + 1)$$

$$Z_1X_2 + NY_1X_2 + (ap^r + bq^r + 1)X_2 = Z_2X_1 + NY_2X_1 + (ap^r + bq^r + 1)X_1$$

$$(ap^r + bq^r + 1)(X_2 - X_1) + Z_1X_2 - Z_2X_1 = N(Y_2X_1 - Y_1X_2) \tag{4}$$

let $|ap^r + bq^r + 1| > p^r > N^{\frac{r}{r+1}}$ and using the fact that $|ap^r - bq^r - 1| < |ap^r + bq^r + 1|$ in to (4) we get

$$\begin{aligned}
&|(ap^r + bq^r + 1)(X_2 - X_1) + Z_1X_2 - Z_2X_1| \\
&\leq |ap^r + bq^r + 1| \left(|X_2| + |X_1|\right) + |Z_1X_2| + |Z_2X_1| \\
&< \frac{2N(ap^r + bq^r + 1)}{3(ap^r + bq^r + 1)} + \frac{2(ap^r - bq^r - 1)}{9(ap^r + bq^r + 1)^2}N^{\frac{4}{3}} \\
&< \frac{2N}{3} + \frac{2}{3}N^{\frac{4(r+1)-3r}{3(r+1)}} < \frac{2N}{3} + \frac{2}{3}N^{\frac{4r+4-3r}{3(r+1)}} \\
&< \frac{2N}{3} + \frac{2}{3}N^{\frac{4+r}{3(r+1)}} \\
&< N
\end{aligned}$$

Therefore from the right hand side of (4) we get $Y_2X_1 - Y_1X_2 = 0.$ since the $gcd(X_1, Y_1) = 1 = gcd(X_2, Y_2)$ which shows that $X_1 = X_2$ and $Y_1 = Y_2$ and leads to $Z_1 = Z_2$. □

Theorem 3.2. *Let $N = p^rq$ be an RSA modulus prime power with $q < p < 2q$. Let a, b be suitably small integer such that $gcd(a, b) = 1$. Let the number of exponents $e < N$ satisfying an equation*

$$eX - NY - (ap^r + bq^r + Z) = 1$$

$$with \quad gcd(X, Y) = 1 \quad and \quad X < \frac{N}{3\,|ap^r + bq^r|}, \quad |Z| < \frac{|ap^r - bq^r|}{3\,|ap^r + bq^r|}N^{\frac{1}{3}}$$

is at least $N^{\frac{5}{6}-\varepsilon}$ where $\varepsilon > 0$ is arbitrarily small for suitably large N.

Proof. Suppose that the exponent e satisfying an equation $eX - NY - (ap^r + bq^r + Z) = 1$ with $gcd(X, Y) = 1$ and $X < \frac{N}{3|ap^r+bq^r|}$, $|Z| < \frac{|ap^r-bq^r|}{3|ap^r+bq^r|}N^{\frac{1}{3}}$
Then with $X < q$ and $gcd(X, N) = 1$. Hence we can express e as

$$e \equiv \frac{ap^r + bq^r + 1 + Z}{X} \pmod{N}$$

with the conditions given below

$$\xi = \sum_{Z=1}^{T_1} \sum_{\substack{X=1 \\ gcd(X,ap^r+bq^r+1+Z)=1}}^{T_2} 1 \tag{5}$$

where

$$T_1 = \left\lfloor \frac{|ap^r - bq^r|}{3(ap^r + bq^r)} N^{\frac{1}{3}} \right\rfloor \qquad \text{and} \qquad T_2 = \left\lfloor \frac{N}{3(ap^r + bq^r)} \right\rfloor$$

Using the given identity (See Nitaj (2011)), let m and n be positive integers then

$$\sum_{\substack{k=1 \\ gcd(k,n)=1}}^{m} 1 > \frac{cm}{(\log\log n)^2}$$

where c is a positive constant. Now from the above identity with $m = T_2$ and $n = ap^r + bq^r + 1 + Z$ we get

$$\sum_{\substack{X=1 \\ gcd(X,ap^r+bq^r+1+Z)=1}}^{T_2} 1 > \frac{cT_2}{(\log\log|ap^r + bq^r + 1 + Z|)^2} > \frac{cT_2}{(\log\log N)^2} = T_2 N^{-\varepsilon_1} \tag{6}$$

Where c is a positive constant and $\varepsilon_1 > 0$ is arbitrarily small for suitably large N. Substituting (6) into (5) we get

$$\begin{aligned}
\xi > \sum_{Z=1}^{T_1} T_2 N^{-\varepsilon_1} &= T_2 N^{-\varepsilon_1} \sum_{Z=1}^{T_1} 1 \\
&= 2T_2 T_1 N^{-\varepsilon_1} \\
&> 2 \left\lfloor \frac{|ap^r - bq^r|}{3\,|ap^r + bq^r|} N^{\frac{1}{3}} \right\rfloor \left\lfloor \frac{N}{3(ap^r + bq^r)} \right\rfloor N^{-\varepsilon_1} \\
&> \frac{2\,|ap^r - bq^r|}{9(|ap^r + bq^r|)^2} N^{\frac{4}{3}} \times N^{-\varepsilon_1} \\
&= N^{\frac{5}{6} - \varepsilon}
\end{aligned}$$

Where we used $|ap^r + bq^r| > N^{\frac{1}{2}}$ with $|ap^r - bq^r| < |ap^r + bq^r|$ and $\varepsilon > 0$ is arbitrarily small for suitably large N. □

4 THE SECOND ATTACK ON *k* PRIME POWER RSA MODULI

Theorem 4.1. *For $k \geq 2$, let $N_i = p_i^r q_i$, $1 \leq i \leq k$ be k RSA moduli with the same size N where $N = min_i\ N_i$. Let e_i, $i = 1, ..., k$, be k public exponents. Define $\delta = \frac{k(1-\alpha r-\alpha)}{(r+1)}$. Let a, b be suitably small integers with $gcd(a, b) = 1$ such that $ap_i^r + bq_i^r + 1 < N^{\frac{r}{r+1}+\alpha}$. If there exist an integer $x < N^\delta$ and k integers $y_i < N^\delta$ and $|z_i| < \frac{ap_i^r - bq_i^r}{3(ap_i^r + bq_i^r)} N^{\frac{1}{3}}$ such that $e_i x - N_i y_i - (ap_i^r + bq_i^r + z_i) = 1$ for $i = 1, ..., k$, then one can factor the k Prime Power RSA moduli $N_1, ...N_k$ in polynomial time.*

Proof. Rewrite the the equation $e_i x - N_i y_i - (ap_i^r + bq_i^r + z_i) = 1$ as

$$\left|\frac{e_i}{N_i} x - y_i\right| = \frac{|1 + ap_i^r + bq_i^r + z_i|}{N_i} \tag{7}$$

Let $N = min_i\ N_i$, and suppose that $y_i < N^\delta$, $|z_i| < \frac{1}{3} N^{\frac{1}{3}}$ and $|ap_i^r + bq_i^r + 1| < N^{\frac{r}{r+1}+\alpha}$ then we can have

$$\begin{aligned}\frac{|1 + ap_i^r + bq_i^r + z_i|}{N_i} &\leq \frac{|z_i + (ap_i^r + bq_i^r + 1)|}{N} \\ &< \frac{\frac{1}{3} N^{\frac{1}{3}} + N^{\frac{r}{r+1}+\alpha}}{N} \\ &< \frac{\frac{2}{3} N^{\frac{r}{r+1}+\alpha}}{N} \\ &< \frac{2}{3} N^{\frac{r}{r+1}+\alpha-1}\end{aligned}$$

Substitute in to (7), to get

$$\left|\frac{e_i}{N_i} x - y_i\right| < \frac{2}{3} N^{\frac{r}{r+1}+\alpha-1}$$

Next, to show the existence of the integer x, we let $\varepsilon = \frac{2}{3} N^{\frac{r}{r+1}+\alpha-1}$, with $\delta = \frac{k(1-\alpha r-\alpha)}{(r+1)}$. Then we have

$$N^\delta \varepsilon^k = \left(\frac{2}{3}\right)^k N^{\delta+\alpha k+\frac{kr}{r+1}-k} = \left(\frac{2}{3}\right)^k$$

Therefore since $\left(\frac{2}{3}\right)^k < 2^{\frac{k(k-3)}{4}} \cdot 3^k$ for $k \geq 2$, we get $N^\delta \varepsilon^k < 2^{\frac{k(k-3)}{4}} \cdot 3^k$. It follows that if $x < N^\delta$, then $x < 2^{\frac{k(k-3)}{4}} \cdot 3^k \cdot \varepsilon^{-k}$ Summarizing for $i = 1, ..., k$, we have

$$\left|\frac{e_i}{N_i} x - y_i\right| < \varepsilon, \qquad x < 2^{\frac{k(k-3)}{4}} \cdot 3^k \cdot \varepsilon^{-k}$$

Hence it satisfy the conditions of Theorem 2.3, and we can obtain x and y_i for $i = 1, ..., k$. Next from the equation $e_i x - N_i y_i - (ap_i^r + bq_i^r + z_i) = 1$ we get

$$(e_i x - N_i y_i) - (ap_i^r + bq_i^r + 1) = z_i$$

Since $|z_i| < \frac{1}{3} N^{\frac{1}{3}}$ and $S_i = e_i x - N_i y_i$ is an approximation of $ap_i^r + bq_i^r + 1$ with an error term of at most $N^{\frac{1}{3}}$. Hence using Lemma 2.2, implies that $abq_i^{r-1} = \left[\frac{S_i^2}{4N_i}\right]$ with $S_i = e_i x - N_i y_i$, for $i = 1, ..., k$, we compute $q_i = gcd\left(N_i, \left[\frac{S_i^2}{4N_i}\right]\right)$. Which leads to factorization of k Prime Power RSA moduli $N_i, ..., N_k$. □

Example 4.1. *Consider the following three Prime Power RSA moduli and its corresponding public exponents*

$$N_1 = 4229232763830021814528254027393024341671866665433042487$$

$$e_1 = 3902583256500382869836851951650840776799847451533540716$$

$$N_2 = 1352746686136293671601460604921536576239015034829144299$$

$$e_2 = 6965461729637445492078840892327549474576003556281543 92$$

$$N_3 = 3336812669259369997062738616040491080986136665280174101$$

$$e_3 = 1960882254438941326621946721392304235252988076448355264$$

Then $N = max(N_1, N_2, N_3) = 4229232763830021814528254027393024341671866665433042487$*. Then, since* $k = 3$ *and* $r = 3$ *with* $\alpha < \frac{1}{3}$*, we get* $\delta = \frac{k(1-\alpha r-\alpha)}{(r+1)} = 0.15$ *and* $\varepsilon = \frac{2}{3}N^{\frac{r}{r+1}+\alpha-1} = 0.001237643774$*. Using equation (11) of Theorem 2.3, with* $n = k = 3$*, we can obtain:*

$$C = [3^{n+1} \cdot 2^{\frac{(n+1)(n-4)}{4}} \cdot \varepsilon^{-n-1}] = 17261255360000$$

Consider the lattice $\mathcal{L}$ *spanned by the matrix*

$$M = \begin{bmatrix} 1 & -[Ce_1/N_1] & -[Ce_2/N_2] & -[Ce_3/N_3] \\ 0 & C & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & C \end{bmatrix}$$

Therefore applying the LLL algorithm to $\mathcal{L}$*, we obtain the reduced basis with following matrix*

$$K = \begin{bmatrix} -123153191 & -41719159 & -78136539 & -30149854 \\ 10363562265 & -4313737015 & -5793724315 & -21219818590 \\ 11690996060 & 11762614940 & -26319874260 & 3979643640 \\ 9302183256 & -41505455656 & 1786658424 & 15106524464 \end{bmatrix}$$

Next we compute

$$K \cdot M^{-1} = \begin{bmatrix} -123153191 & -113641317 & -63413117 & -72371131 \\ 10363562265 & 9563120980 & 5336327716 & 6090160687 \\ 11690996060 & 10788028946 & 6019839965 & 6870228863 \\ 9302183256 & 8583718762 & 4789810401 & 5466439948 \end{bmatrix}$$

Then from the first row we obtain $x = 123153191$*,* $y_1 = 113641317$*,* $y_2 = 63413117$*,* $y_3 = 72371131$*. Then using* x *and* y_i *for* $i = 1, 2, 3$*, define* $S_i = e_i x - N_i y_i$ *we get*

$$S_1 = 4254458568813966951710722613640591641893 77$$

$$S_2 = 184449236362338771616740629962420373094889$$

$$S_3 = 368917771225148087840868112672265896969193$$

And Lemma 2.2, implies that $abq_i^{r-1} = \left[\frac{S_i^2}{4N_i}\right]$ *for* $i = 1, 2, 3$*, which gives*

$$\left[\frac{S_1^2}{4N_1}\right] = 1069958709092351255593864061 4$$

$$\left[\frac{S_2^2}{4N_2}\right] = 628748921422642009828881744 6$$

$$\left[\frac{S_3^2}{4N_3}\right] = 1019688063249435916072889693 4$$

Therefore for $i = 1, 2, 3$ *we compute* $q_i = gcd\left(\left[\frac{S_i^2}{4N_i}\right], N_i\right)$*, that is*

$$q_1 = 42228716712137, q_2 = 32371513233671, q_3 = 41224751126183$$

And finally for $i = 1, 2, 3$ *we find* $p_i = \sqrt[3]{\frac{N_i}{q_i}}$*, hence*

$$p_1 = 46439182724351, p_2 = 34701732614789, p_3 = 43257152513563$$

Which leads to the factorization of the three Prime Power RSA moduli N_1*,*N_2*,and* N_3.

5 THE THIRD ATTACK ON k PRIME POWER RSA MODULI

The following section, For $k \geq 2$, and $r \geq 2$, we consider the scenario when the k RSA moduli satisfy k equations $e_i x_i - N_i y - (ap_i^r + bq_i^r + z_i) = 1$, with suitably small unknown parameters x_i,y and z_i.

Theorem 5.1. *For* $k \geq 2$*, and* $r \geq 2$ *let* $N_i = p_i^r q_i$*,* $1 \leq i \leq k$ *be* k *RSA moduli with the same size* N *where* $N = min_i\ N_i$*. Let* e_i*,* $i = 1, ...k$*, be* k *public exponents with* $min_i\ e_i = N^{\beta}$*. Let* $\delta = \frac{kr(\beta-\alpha-1)+k(\beta-\alpha)}{(r+1)}$*. If there exist an integer* $y < N^{\delta}$ *and* k *integers* $x_i < N^{\delta}$ *and* $|z_i| < \frac{ap_i^r - bq_i^r}{3(ap_i^r + bq_i^r)} N^{\frac{1}{3}}$ *such that* $e_i x_i - N_i y - (ap_i^r + bq_i^r + z_i) = 1$ *for* $i = 1, ..., k$*, then one can factor the* k *Prime Power RSA moduli* $N_1, ...N_k$ *in polynomial time.*

Proof. For $k \geq 2$, and $r \geq 2$, let $N_i = p_i^r q_i$, $1 \leq i \leq k$ be k RSA moduli. Then the equation $e_i x_i - N_i y - (ap_i^r + bq_i^r + z_i) = 1$ can be rewrite as

$$\left|\frac{N_i}{e_i} y - x_i\right| = \frac{|1 + ap_i^r + bq_i^r + z_i|}{e_i} \tag{8}$$

Let $N = max_i\, N_i$, and suppose that $y < N^{\delta}$, $|z_i| < \frac{1}{3}N^{\frac{1}{3}}\; min_i\, e_i = N^{\beta}$ and $|ap_i^r + bq_i^r + 1| < N^{\frac{r}{r+1}+\alpha}$. Then

$$\begin{aligned}\frac{|1 + ap_i^r + bq_i^r + z_i|}{e_i} &\leq \frac{|z_i + (ap_i^r + bq_i^r + 1)|}{N^{\beta}}\\ &< \frac{\frac{1}{3}N^{\frac{1}{3}} + N^{\frac{r}{r+1}+\alpha}}{N^{\beta}}\\ &< \frac{\frac{2}{3}N^{\frac{r}{r+1}+\alpha}}{N^{\beta}}\\ &< \frac{2}{3}N^{\frac{r}{r+1}+\alpha-\beta}\end{aligned}$$

Substitute in to (8), to get

$$\left|\frac{N_i}{e_i}y - x_i\right| < \frac{2}{3}N^{\frac{r}{r+1}+\alpha-\beta}$$

Hence to shows the existence of the integer y and integers x_i, we let $\varepsilon = \frac{2}{3}N^{\frac{r}{r+1}+\alpha-\beta}$, with $\delta = \frac{kr(\beta-\alpha-1)+k(\beta-\alpha)}{(r+1)}$. Then we have

$$N^{\delta}\varepsilon^k = \left(\frac{2}{3}\right)^k N^{\delta+\frac{kr}{r+1}+\alpha k-\beta k} = \left(\frac{2}{3}\right)^k$$

Therefore since $\left(\frac{2}{3}\right)^k < 2^{\frac{k(k-3)}{4}} \cdot 3^k$ for $k \geq 2$, we get $N^{\delta}\varepsilon^k < 2^{\frac{k(k-3)}{4}} \cdot 3^k$. It follows that if $y < N^{\delta}$, then $y < 2^{\frac{k(k-3)}{4}} \cdot 3^k \cdot \varepsilon^{-k}$ Summarizing for $i = 1, ..., k$,we have

$$\left|\frac{N_i}{e_i}y - x_i\right| < \varepsilon, \qquad y < 2^{\frac{k(k-3)}{4}} \cdot 3^k \cdot \varepsilon^{-k}$$

Hence it satisfy the conditions of Theorem 2.3, and we can obtain y and x_i for $i = 1, ..., k$. Next from the equation $e_i x_i - N_i y - (ap_i^r + bq_i^r + z_i) = 1$ we get

$$(e_i x_i - N_i y) - (ap_i^r + bq_i^r + 1) = z_i$$

Since $S_i = e_i x_i - N_i y$ is an approximation of $ap_i^r + bq_i^r + 1$ with an error term of at most $N^{\frac{1}{3}}$. Hence using Lemma 2.2, implies that $abq_i^{r-1} = \left[\frac{S_i^2}{4N_i}\right]$ with $S_i = e_i x_i - N_i y$ for $i = 1, ..., k$, we compute $q_i = gcd\left(N_i, \left[\frac{S_i^2}{4N_i}\right]\right)$. Which leads to factorization of k Prime Power RSA moduli $N_i, ..., N_k$.
□

Example 5.1. *Consider the following three Prime Power RSA moduli and its corresponding public exponents*

$$N_1 = 8316686807850485154364481602267030007234357923589115713$$

$$e_1 = 106043696928881338907210434052343577373210365263725083391$$

$$N_2 = 7238199574976743382622642035350480633266303235827761843$$

$$e_2 = 126798073597158975500581505905981226551910777660642734232$$

$$N_3 = 2811621101361791358541110559262991877781441285478967999$$

$$e_3 = 3981266763352114330670844984791661846869032483242345295 8$$

Then $N = max(N_1, N_2, N_3) = 831668680785048515436448160226703000$ 7234357923589115713*. Also* $min(e_1, e_2, e_3) = N^{\beta}$ *with* $\beta = 1.021275$*. Then since* $k = 3$ *and* $r = 3$ *with* $\alpha < \frac{1}{3}$ *we get* $\delta = \frac{kr(\beta-\alpha-1)+k(\beta-\alpha)}{r+1} = 0.213825$ *and* $\varepsilon = \frac{2}{3}N^{\frac{r}{r+1}+\alpha-\beta} =$ 0.00008118752087*. Using equation (11) of Theorem 2.3, with* $n = k = 3$*, we can obtain:*

$$C = [3^{n+1} \cdot 2^{\frac{(n+1)(n-4)}{4}} \cdot \varepsilon^{-n-1}] = 932175969300000000$$

Consider the lattice $\mathcal{L}$ *spanned by the matrix*

$$M = \begin{bmatrix} 1 & -[CN_1/e_1] & -[CN_2/e_2] & -[CN_3/e_3] \\ 0 & C & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & C \end{bmatrix}$$

Therefore applying the LLL algorithm to $\mathcal{L}$*, we obtain the reduced basis with following matrix*

$$K = \begin{bmatrix} 923423545819 & 524702578396 & 732777747416 & 897448085252 \\ 162400481624 & -32870838301984 & -1403515936064 & 20429922748192 \\ 47917650366395 & 5227702745180 & -81926411975720 & 14081061814660 \\ -96760812365818 & 67390041700088 & -18796402342352 & 74698387322056 \end{bmatrix}$$

Next we compute

$$K{\cdot}M^{-1} = \begin{bmatrix} 923423545819 & 72421319173 & 52713134571 & 65213342417 \\ 162400481624 & 12736579186 & 9270543816 & 11468928061 \\ 47917650366395 & 3758036566121 & 2735353201168 & 3384005265309 \\ -96760812365818 & -7588658213787 & -5523538734238 & -6833371336403 \end{bmatrix}$$

Then from the first row we obtain $y = 923423545819$*,* $x_1 = 72421319173$*,* $x_2 = 52713134571$*,* $x_3 = 65213342417$*. Then using* x *and* y_i *for* $i = 1, 2, 3$*, define* $S_i = e_i x_i - N_i y$ *we get*

$$S_1 = 7109034306874434647869117070886525038016 96$$

$$S_2 = 673658519698253852838292575819303073950055$$

$$S_3 = 329602723919139912751882442716306134273305$$

And Lemma 2.2, implies that $abq_i^{r-1} = \left[\frac{S_i^2}{4N_i}\right]$ *for* $i = 1, 2, 3$*, which gives*

$$\left[\frac{S_1^2}{4N_1}\right] = 151918576303162854284927182 14$$

$$\left[\frac{S_2^2}{4N_2}\right] = 15674332976771395603736133366$$

$$\left[\frac{S_3^2}{4N_3}\right] = 965972580394088675960736725 4$$

Therefore for $i = 1, 2, 3$ *we compute* $q_i = gcd\left(\left[\frac{S_i^2}{4N_i}\right], N_i\right)$*, that is*

$$q_1 = 50318746722463, q_2 = 51111533233331, q_3 = 40124235826453$$

And finally for $i = 1, 2, 3$ *we find* $p_i = \sqrt[3]{\frac{N_i}{q_i}}$*, hence*

$$p_1 = 54879082726751, p_2 = 52123937435137, p_3 = 41227152517627$$

Which leads to the factorization of three RSA moduli N_1*,*N_2*,and* N_3*.*

6 CONCLUSION

In this paper, we present three new attacks. The first attacks studied the class of exponents e satisfying an equation $eX - NY - (ap^r + bq^r + Z) = 1$ with $gcd(X, Y) = 1$, where a and b are suitably small integers. Using the continued fraction algorithm we showed that such exponents are vulnerable and lead to the factorization of the RSA prime power modulus $N = p^r q$ for $r \geq 2$ in polynomial time. We also showed that the new class of weak exponents is sufficiently large since the size of this class can be estimated as $N^{\frac{5}{6}-\varepsilon}$ where $\varepsilon > 0$ is arbitrarily small for large N. For $k \geq 2$, $r \geq 2$, we present second and third attacks on the Prime Power RSA with moduli $N_i = p_i^r q_i$ for $i = 1, ..., k$. The two attacks work when k RSA public keys (N_i, e_i) are with the condition such that there exist k relations of the form $e_i x - N_i y_i - (ap_i^r + bq_i^r + z_i) = 1$ or of the form $e_i x_i - N_i y - (ap_i^r + bq_i^r + z_i) = 1$ where the parameters x, x_i, y, y_i, z_i are suitably small in terms of the prime factors of the moduli. Based on LLL algorithm we show that our approach enable us to simultaneously factor the k Prime Power RSA moduli N_i.

REFERENCES

Asbullah, M. and Ariffin, M. (2015). New attacks on RSA with modulus $N = p^2 q$ using continued fractions. In *Journal of Physics: Conference Series*, volume 622, pages 191–199. IOP Publishing.

Hinek, M. J. (2007). *On the security of some variants of RSA. PhD Thesis, University of Waterloo.* PhD thesis, University of Waterloo.

Lenstra, A. K., Lenstra, H. W., and Lovász, L. (1982). Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261(4):515–534.

Lu, Y., Zhang, R., and Lin, D. (2014). New results on solving linear equations modulo unknown divisors and its applications. *IACR Cryptology ePrint Archive*, 2014:343.

Nitaj, A. (2009). Cryptanalysis of RSA using the ratio of the primes. In *Progress in Cryptology - AFRICACRYPT 2009*, pages 98 – 115. Springer.

Nitaj, A. (2011). A new vulnerable class of exponents in RSA. *Journal of Algebra, Number Theory and Applications* 23(2): 203 - 220.

Nitaj, A., Ariffin, M. R. K., Nassr, D. I., and Bahig, H. M. (2014). New attacks on the RSA cryptosystem. In *Progress in Cryptology - AFRICACRYPT 2014*, pages 178 –198. Springer.

Nitaj, A. and Rachidi, T. (2015). New attacks on RSA with moduli $N = p^r q$. In *Codes, Cryptology, and Information Security*, pages 352 – 360. Springer.

Rivest, R. L., Shamir, A., and Adleman, L. (1978). A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120 – 126.

Sarkar, S. (2014). Small secret exponent attack on RSA variant with modulus $N = p^r q$. *Designs, Codes and Cryptography*, 73(2):383 – 392.

Takagi, T. (1998). Fast RSA -type cryptosystem modulo $p^k q$. In *Advances in Cryptology - CRYPTO'98*, pages 318 – 326. Springer.

A New Improved Attack on RSA

Martin Bunder[1] and **Joseph Tonien**[*1,2]

[1]*School of Mathematics and Applied Statistics, University of Wollongong, NSW 2522, Australia*
[2]*School of Computing and Information Technology, University of Wollongong, NSW 2522, Australia*

E-mail: joseph_tonien@uow.edu.au
[*]*Corresponding author*

ABSTRACT

This paper presents a new improved attack on RSA based on Wiener's technique using continued fractions. In the RSA cryptosystem with public modulus $N = pq$, public key e and secret key d, if $d < \frac{1}{3}N^{\frac{1}{4}}$, Wiener's original attack recovers the secret key d using the convergents of the continued fraction of $\frac{e}{N}$. Our new method uses the convergents of the continued fraction of $\frac{e}{N'}$ instead, where N' is a number depending on N. We will show that our method can recover the secret key if $d^2e < 8N^{\frac{3}{2}}$, so if either d or e is relatively small the RSA encryption can be broken. For $e \approx N^t$, our method can recover the secret key if $d < 2\sqrt{2}\,N^{\frac{3}{4}-\frac{t}{2}}$ and certainly for $d < 2\sqrt{2}\,N^{\frac{1}{4}}$. Our experiments demonstrate that for a 1024-bit modulus RSA, our method works for values of d of up to 270 bits compared to 255 bits for Wiener.

Keywords: RSA, Wiener's attack, continued fractions.

1 INTRODUCTION

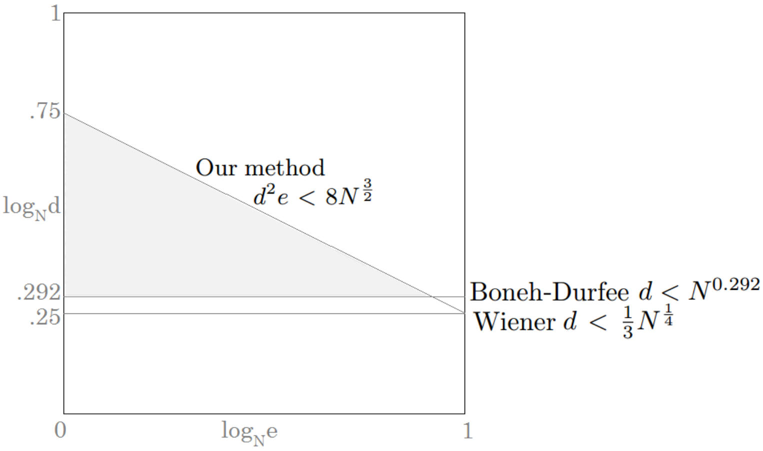
The RSA public-key cryptosystem is one of the most popular systems in use today. The key setup involves picking two large prime numbers p, q to form a product $N = pq$ and selecting two integers $e, d < \phi(N) = (p-1)(q-1)$ such that $ed = 1 \pmod{\phi(N)}$. Messages can be encrypted using the public key (N, e), whereas ciphertexts can be decrypted using the secret key (p, q, d). It is well known that RSA is not secure if the secret key d is relatively small.

Wiener (1990) showed that using continued fractions, one can efficiently recover the secret key d from the public information (N, e) as long as $d < \frac{1}{3}N^{\frac{1}{4}}$ (see also Boneh and Durfee (2000), Nassr et al. (2008)). Steinfeld et al. (2005) showed that, for linear attack, $N^{\frac{1}{4}}$ is the best bound in the sense that, for any fixed $\epsilon > 0$ and all sufficiently large modulus lengths, Wiener's attack succeeds with negligible probability over a random choice of $d < N^\delta$ as soon as $\delta > \frac{1}{4} + \epsilon$. Exploiting a non-linear equation satisfied by the secret key, Boneh and Durfee (2000) presented a lattice-based attack that succeeds in polynomial-time when $d < N^{0.292}$.

In this paper, we present a new improved attack on RSA based on Wiener's technique using continued fractions. As in Wiener's original attack, our method only uses the public information (N, e). The difference between our attack and Wiener's is that in Wiener's attack one is searching the convegents of the continued fraction of $\frac{e}{N}$ whereas in ours, one is searching the convegents of the continued fraction of $\frac{e}{N'}$ where N' is given by

$$N' = \left\lceil N - (1 + \frac{3}{2\sqrt{2}})N^{\frac{1}{2}} + 1 \right\rceil$$

We will show that our method can recover the secret key if $d^2 e < 8N^{\frac{3}{2}}$. So if $e \approx N^t$, then our method can recover the secret key if $d < 2\sqrt{2}\, N^{\frac{3}{4}-\frac{t}{2}}$ and certainly for $d < 2\sqrt{2}\, N^{\frac{1}{4}}$ – which is more than 8 times the Wiener's bound. In the following figure, the shaded part shows the area where our method is better than Wiener (1990) and Boneh and Durfee (2000).



There are other variants of Wiener's attack but these attacks need more than just the public information (N, e). For example, the de Weger (2002) attack exploited the small distance between the two RSA's secret primes: if $|p - q| = N^\beta$ and $d = N^\delta$ then d can be recovered if $2 - 4\beta < \delta < 1 - \sqrt{2\beta - \frac{1}{2}}$ or $\delta < \frac{1}{6}(4\beta + 5) - \frac{1}{3}\sqrt{(4\beta + 5)(4\beta - 1)}$. The Blömer and May (2004) attack assumed a linear relation between e and $\phi(N)$: $ex + y = 0 \mod \phi(N)$ with either $0 < x < \frac{1}{3}N^{\frac{1}{4}}$ and $y = \mathcal{O}(N^{-\frac{3}{4}}ex)$ (their Theorem 2) or $x < \frac{1}{3}\sqrt{\frac{\phi(N)}{e}\frac{N^{\frac{3}{4}}}{p-q}}$ and $|y| \leq \frac{p-q}{\phi(N)N^{\frac{1}{4}}}ex$ (their Theorem 4). These conditions are much more complex than ours: $d^2e < 8N^{\frac{3}{2}}$, particularly because they have in addition to p, q and d the unknown x and y. For the case $x = d$ and $y = -1$, used by Wiener and us, our result is better than Blömer–May's Theorem 2 result and also better than their Theorem 4 result if $\frac{9}{8} < \frac{p}{q} < 2$, and theirs is better if $1 < \frac{p}{q} < \frac{9}{8}$. The Nassr et al. (2008) attack required an approximation $p_o \geq \sqrt{N}$ of the prime p with $|p - p_0| \leq \frac{1}{8}n^\alpha$, $\alpha \leq \frac{1}{2}$, $\delta < \frac{1-\alpha}{2}$.

The Blömer and May (2001) attack is a variant of the Boneh and Durfee (2000) attack which works for $d < N^{0.29}$. Using an exhaustive search of about $8 + 2b$ bits, Verheul and van Tilborg (1997) improved Wiener's bound to $d < 2^b N^{\frac{1}{4}}$. Another exponential time attack similar to this is due to Dujella (2004).

The rest of the paper is organized as follows. In Section 2, we review some preliminary results on continued fractions and Wiener's attack. Section 3 presents our main result which says that the RSA encryption system is not secure if $e \approx N^t$ and $d < 2\sqrt{2}\,N^{\frac{3}{4}-\frac{t}{2}}$. As $t < 1$, this means that RSA encryption is not secure for $d < 2\sqrt{2}\,N^{\frac{1}{4}}$ compared to Wiener's result of $d < \frac{1}{3}N^{\frac{1}{4}}$. In Section 4, we show our experiment result with a 1024-bit modulus and 270-bit secret key. We show that our usage of continued fraction of $\frac{e}{N'}$ is essential because if we use the continued fraction expansion of $\frac{e}{N}$ as in Wiener's attack then the secret key cannot be found.

2 PRELIMINARIES

RSA is a public-key cryptosystem widely used for secure data transmission.

RSA Key Generation algorithm: Choose two distinct prime numbers p and q of similar bit-length. Compute $N = pq$, $\phi(N) = (p-1)(q-1)$. Choose e such that $(e, \phi(N)) = 1$. Determine $d = e^{-1} \pmod{\phi(N)}$. Keep p, q, d secret, publish N, e.

RSA Encryption-Decryption algorithm: For a message $m \in (1, N)$, the ciphertext c is $c = m^e \pmod N$. For a ciphertext $c \in (1, N)$, the message m is determined as $m = c^d \pmod N$.

The complexity of the decryption algorithm is based on the size of the decryption key d. In a cryptosystem with a limited resource such as a credit card, it is desirable to have a smaller value of d. Wiener's attack, uses the *continued fraction* method to expose the private key d when d is small ($d < \frac{1}{3}N^{\frac{1}{4}}$). A *continued fraction* is an expression of the form

$$x = [a_0, a_1, \ldots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}$$

If $k \le n$, the continued fraction $[a_0, a_1, \ldots, a_k]$ is called the k^{th} *convergent* of x. The following theorem gives us the *fundamental recursive formulas* to calculate the convergents.

Theorem 2.1. *The k^{th} convergent can be determined as*

$$[a_0, \ldots, a_k] = \frac{p_k}{q_k}$$

where the sequences $\{p_n\}$ and $\{q_n\}$ are specified as follows[1]*:*

$$\begin{aligned} p_{-2} = 0,\ \ p_{-1} = 1,\ \ p_n = a_n p_{n-1} + p_{n-2},\ \ \forall n \geq 0, \\ q_{-2} = 1,\ \ q_{-1} = 0,\ \ q_n = a_n q_{n-1} + q_{n-2},\ \ \forall n \geq 0. \end{aligned}$$

The following theorem is the basis for Wiener's attack.

Theorem 2.2. *(Hardy and Wright, 1965) Let p, q be positive integers such that*

$$0 < \left| x - \frac{p}{q} \right| < \frac{1}{2q^2}$$

then $\frac{p}{q}$ is a convergent of the continued fraction of x.

The following theorem summarises Wiener's attack (Boneh and Durfee, 2000, Wiener, 1990).

Theorem 2.3. *In a RSA algorithm, if the following conditions are satisfied*

- $q < p < 2q$ *(i.e. p and q are two primes of the same bit size)*
- $0 < e < \phi(N)$
- $ed - k\phi(N) = 1$
- $\boxed{d < \frac{1}{3}N^{\frac{1}{4}}}$

then $\frac{k}{d}$ is a convergent of $\frac{e}{N}$. Thus, the secret information p, q, d, k can be recovered from public information (e, N).

Since $\frac{e}{N}$ has $O(\log(N))$ number of convergents, Wiener's algorithm will succeed to factor N and output p, q, d, k in $O(\log(N))$ time complexity.

Example 1. In the following example, we have a 1024-bit modulus N, the upper bound $\frac{1}{3}N^{\frac{1}{4}}$ in Theorem 2.3 is 255-bit, d is 255-bit and we have found the convergent $c_{149} = \frac{p_{149}}{q_{149}} = \frac{k}{d}$ as asserted by Theorem 2.3.

[1] The convergents start with $\frac{p_0}{q_0}$, but it is a convention to extend the sequence index to -1 and -2 to allow the recursive formula to hold for $n = 0$ and $n = 1$

p	12137
	2429807756 5612551149 2629609691 9449141205 8680156593
	9661850265 4224438815 0519802020 4979508724 3102230079
	9409502534 6163494126 0471531617 7098769594 1320931493
		512 bits

q	9201
	0524322086 3900671386 8662660639 9738950237 2692456878
	2613825773 8431082681 6215281513 7070448098 3908271161
	4206768781 4447541784 7243525840 6453897707 3778553491
		512 bits

N	111675409
	0485730823 5978712392 1718417590 8091542898 6532382066
	5485087798 8534958587 2419428390 8818158158 7258671440
	7683378413 7900981405 8406611299 6495087782 9075022344
	5692173775 8022280271 1775885570 7370037539 5363272503
	0411307566 7128393688 9712399229 9533595050 1425299028
	6693467091 9270372721 8720248761 5489260235 4246992063
		1024 bits

$\phi(N)$	111675409
	0485730823 5978712392 1718417590 8091542898 6532382066
	5485087798 8534958587 2419428390 8818158158 7258671440
	7683378413 7900981405 8406611299 6495087782 9075001006
	2738043932 8509057735 0483615238 8181946096 3990659030
	8135631527 4472872192 2977315695 7483638227 4414797787
	3077195775 8659336811 1005191303 1936592933 9147507080
		1024 bits

Theorem 2.3 bound $\frac{1}{3}N^{\frac{1}{4}}$	3426637 2625316286 2968546235 7247145632 3454416288 1157194267 8892540948 5361638977	255 bits

e	45643085
	8324017120 3133152071 1529402253 9055348712 7592566099
	1853899212 7134329984 8723684744 2845550165 4714497720
	7173865355 1358820024 8341016147 1746464324 1362580067
	0745402653 2892481331 8307985083 2822164891 3129959216
	3726940854 8355291478 1683701096 4254131032 8949699809
	7582249761 4243019490 2375579169 7150271910 4226716997
		1023 bits

d	3426637 2625316286 2968546235 7247145632 3454416288 1157194267 8892540948 5361638973	255 bits
k	1400507 9544612205 2131699024 5626308122 5492430329 4046240953 0743691100 4314600526	253 bits
convergent of $\frac{e}{N}$	found $c_{149} = \frac{p_{149}}{q_{149}} = \frac{k}{d}$	

3 A NEW IMPROVED ATTACK BASED ON CONTINUED FRACTIONS

In this section, we present our main result. Instead of using the convergents of the continued fraction of $\frac{e}{N}$ as in the Wiener's original attack, we will use the convergents of the continued fraction of $\frac{e}{N'}$ where N' is given by

$$N' = \left[N - (1 + \frac{3}{2\sqrt{2}})N^{\frac{1}{2}} + 1\right]$$

We will show that for $e \approx N^t$, the secret key can be recovered if $d < 2\sqrt{2}\,N^{\frac{3}{4}-\frac{t}{2}}$.

This is our main theorem.

Theorem 3.1. *In a RSA algorithm, if the following conditions are satisfied*

- $q < p < 2q$

- $0 < e < \phi(N)$
- $ed - k\phi(N) = 1$
- $N > 2000000$
- $\boxed{d < 2\sqrt{2}\left(\frac{N}{e}\right)^{\frac{1}{2}} N^{\frac{1}{4}}}$

and

$$N' = \left[N - (1 + \frac{3}{2\sqrt{2}})N^{\frac{1}{2}} + 1\right]$$

then $\frac{k}{d}$ is a convergent of $\frac{e}{N'}$. Thus, the secret information p, q, d, k can be recovered from public information (e, N).

Proof. Let $\phi_1 = N + 1 - \frac{3}{\sqrt{2}}N^{\frac{1}{2}}$ and $\phi_2 = N + 1 - 2N^{\frac{1}{2}}$. It follows from $q < p < 2q$ that $1 < \sqrt{\frac{p}{q}} < \sqrt{2}$, so since the function $f(x) = x + \frac{1}{x}$ is increasing on $[1, +\infty)$,

$$\begin{aligned}
&2 < \frac{p+q}{N^{\frac{1}{2}}} = \sqrt{\frac{p}{q}} + \sqrt{\frac{q}{p}} < \sqrt{2} + \frac{1}{\sqrt{2}} = \frac{3}{\sqrt{2}} \\
&2N^{\frac{1}{2}} < p + q < \frac{3}{\sqrt{2}}N^{\frac{1}{2}} \\
&\phi_1 = N + 1 - \frac{3}{\sqrt{2}}N^{\frac{1}{2}} < \phi(N) < N + 1 - 2N^{\frac{1}{2}} = \phi_2
\end{aligned}$$

Let $\phi_{mid} = N - (1 + \frac{3}{2\sqrt{2}})N^{\frac{1}{2}} + 1$, then ϕ_{mid} is the midpoint of the interval $[\phi_1, \phi_2]$ and $N' = [\phi_{mid}]$. Since $\phi(N) \in (\phi_1, \phi_2)$,

$$|\phi(N) - N'| < |\phi(N) - \phi_{mid}| + |\phi_{mid} - N'| < \frac{1}{2}(\phi_2 - \phi_1) + 1 = \frac{1}{2}(\phi_2 - \phi_1 + 2)$$

We have

$$\begin{aligned}
\left|\frac{e}{N'} - \frac{k}{d}\right| &= \left|(\frac{e}{N'} - \frac{e}{\phi(N)}) + (\frac{e}{\phi(N)} - \frac{k}{d})\right| = \left|\frac{e(\phi(N) - N')}{N'\phi(N)} + \frac{1}{d\phi(N)}\right| \\
&= \left|\frac{e(\phi(N) - N')}{N'\phi(N)} + \frac{e}{\phi(N)(k\phi(N) + 1)}\right| \\
&< \frac{e|\phi(N) - N'|}{N'\phi(N)} + \frac{e}{\phi(N)(k\phi(N) + 1)} \\
&< \frac{e(\phi_2 - \phi_1 + 2)/2}{\phi_1^2} + \frac{e}{\phi_1^2} < \frac{e(\phi_2 - \phi_1 + 4)}{2(\phi_1 - 1)^2} = e\,\frac{(\frac{3}{\sqrt{2}} - 2)N^{\frac{1}{2}} + 4}{2(N - \frac{3}{\sqrt{2}}N^{\frac{1}{2}})^2}
\end{aligned}$$

For $N > 2000000$, it can be shown that [2]

$$\frac{(\frac{3}{\sqrt{2}} - 2)N^{\frac{1}{2}} + 4}{2(N - \frac{3}{\sqrt{2}}N^{\frac{1}{2}})^2} < \frac{1}{16N^{\frac{3}{2}}}. \tag{1}$$

Therefore,

$$\left|\frac{e}{N'} - \frac{k}{d}\right| < \frac{e}{16N^{\frac{3}{2}}} < \frac{1}{2d^2}. \ \blacksquare$$

The boxed condition in Theorem 3.1 amounts to $d^2e < 8N^{\frac{3}{2}}$, so if either d or e is relatively small then RSA encryption can be broken. When e is relatively small, the Wiener attack cannot be applied, whereas ours can.

This result is superficially like that of Blömer and May (2004)(Theorem 4), which is

Theorem 3.2. *(Blömer and May, 2004) Given an RSA public key tuple (N, e), where $N = pq$. Suppose that e satisfies an equation $ex + y = 0 \pmod{\phi(N)}$ with*

$$0 < x \leq \frac{1}{3}\sqrt{\frac{\phi(N)}{e}\frac{N^{\frac{3}{4}}}{p-q}} \text{ and } |y| \leq \frac{p-q}{\phi(N)\,N^{\frac{1}{4}}}ex$$

then N can be factored in time polynomial in $\log N$.

With $x = d$ and $y = -1$, these conditions amount to

$$ed^2 < \frac{\phi(N)\,N^{\frac{3}{2}}}{9(p-q)^2} \tag{2}$$

and

$$\phi(N)\,N^{\frac{1}{4}} < (p-q)ed, \tag{3}$$

whereas our only condition is $ed^2 < 8N^{\frac{3}{2}}$. Let R be the ratio between our bound and Blömer-May's bound (2)

$$R = \frac{8N^{\frac{3}{2}}}{\frac{\phi(N)\,N^{\frac{3}{2}}}{9(p-q)^2}} = \frac{72(p-q)^2}{\phi(N)}$$

then

$$R = \frac{N}{\phi(N)}\frac{72(p-q)^2}{pq} = \frac{N}{\phi(N)}\frac{72(\frac{p}{q}-1)^2}{\frac{p}{q}}$$

Since $q < p < 2q$, the quotient $\frac{p}{q}$ ranges in the interval $(1, 2)$. Consider the graph of the function $f(x) = \frac{72(x-1)^2}{x}$ for $x \in (1, 2)$, we can see that $f(x) < 1$ for $x \in (1, \frac{9}{8})$ and $f(x) > 1$ for

[2] $(1) \leftrightarrow 8N^{\frac{1}{2}}((\frac{3}{\sqrt{2}} - 2)N^{\frac{1}{2}} + 4) < (N^{\frac{1}{2}} - \frac{3}{\sqrt{2}})^2 \leftrightarrow (12\sqrt{2} - 16)N + 32N^{\frac{1}{2}} < N - 3\sqrt{2}N^{\frac{1}{2}} + \frac{9}{2} \leftrightarrow (32 + 3\sqrt{2})N^{\frac{1}{2}} < (17 - 12\sqrt{2})N + \frac{9}{2} \leftrightarrow \frac{32+3\sqrt{2}}{17-12\sqrt{2}} < N^{\frac{1}{2}} + \frac{9}{2(17-12\sqrt{2})N^{\frac{1}{2}}}$, this is true because $N > 200000 > \left(\frac{32+3\sqrt{2}}{17-12\sqrt{2}}\right)^2$.

$x \in (\frac{9}{8}, 2)$. Therefore, if $\frac{p}{q} \in (\frac{9}{8}, 2)$ then $R = \frac{N}{\phi(N)} f(\frac{p}{q}) > 1$ and our bound is better than Blömer-May's bound. Our experiment result in Section 4 also confirms this.

From Theorem 3.1, we have

Corollary 3.1. *In a RSA algorithm, if the following conditions are satisfied*

- $q < p < 2q$
- $0 < e < \phi(N)$
- $ed - k\phi(N) = 1$
- $N > 2000000$
- $\boxed{d < 2\sqrt{2}N^{\frac{1}{4}}}$

and

$$N' = \left\lceil N - (1 + \frac{3}{2\sqrt{2}})N^{\frac{1}{2}} + 1 \right\rceil$$

then $\frac{k}{d}$ *is a convergent of* $\frac{e}{N'}$*. Thus, the secret information* p, q, d, k *can be recovered from public information* (e, N)*.*

Note that Corollary 3.1 has $d < 2\sqrt{2}\,N^{\frac{1}{4}}$ while Wiener's result had $d < \frac{1}{3}N^{\frac{1}{4}}$.

4 EXPERIMENT RESULT

We will use the same 1024-bit modulus as in Example 1. With this 1024-bit modulus, the Wiener's upper bound $\frac{1}{3}N^{\frac{1}{4}}$ is 255-bit. Here, we show an example of a 270-bit secret key.

N	111675409	
	0485730823 5978712392 1718417590 8091542898 6532382066	
	5485087798 8534958587 2419428390 8818158158 7258671440	
	7683378413 7900981405 8406611299 6495087782 9075022344	
	5692173775 8022280271 1775885570 7370037539 5363272503	
	0411307566 7128393688 9712399229 9533595050 1425299028	
	6693467091 9270372721 8720248761 5489260235 4246992063	1024 bits
Theorem 3.1	111675409	
N'	0485730823 5978712392 1718417590 8091542898 6532382066	
	5485087798 8534958587 2419428390 8818158158 7258671440	
	7683378413 7900981405 8406611299 6495087782 9075000568	
	2159570564 0981693044 2093595665 5130899532 7328449321	
	6820552021 8559771355 1247634195 5201901221 0109431097	
	4104405733 7196789666 1898135689 1959781693 7504572404	1024 bits
e	9497738493 9533670765 7042840968 7659484313 7084252195	
	6357612333 8847198573 4448278894 7630928901 1796460405	
	3837337081 2904542700 5252696553 0732537894 7443876974	
	8735584808 1502373619 6458971201 9372820861 3917977593	
	0646731395 1290537294 6709829003 9830064227 6485488318	
	8298864198 1593551375 9303722339 5282843022 6076170323	997 bits
d	16 8426074727 9546104062 9984578341	
	1702121043 1469393463 8412655292 6172702449 5099104827	270 bits
k	1432 4253002139 3318566580	
	1576488907 6467402086 1953632340 7603167662 3143713764	244 bits
convergent of $\frac{e}{N}$	not found, $c_i \neq \frac{k}{d}, \forall i$	
convergent of $\frac{e}{N'}$	found $c_{146} = \frac{p_{146}}{q_{146}} = \frac{k}{d}$	

This experiment result shows that our usage of continued fractions of $\frac{e}{N'}$ is *essential*. If we use continued fractions of $\frac{e}{N}$ as in Wiener's original attack then no convergent c_i is found for which $c_i = \frac{k}{d}$.

For this example, the Blömer and May Theorems 2 and 4 results, with $x = d$ and $y = -1$, do not apply as neither of $d < \frac{1}{3}N^{\frac{1}{4}}$ and $d < \frac{1}{3}\sqrt{\frac{\phi(N)}{e}\frac{N^{\frac{3}{4}}}{p-q}}$ hold.

ACKNOWLEDGMENTS

The second author would like to thank the School of Mathematics & Applied Statistics, and the School of Computing & Information Technology, University of Wollongong, for financial support that enabled the author to attend the conference.

REFERENCES

Blömer, J. and May, A. (2001). Low secret exponent RSA revisited. In *Cryptography and Lattices, Proceedings of CALC 2001, Lecture Notes in Computer Science 2146*, pages 4–19.

Blömer, J. and May, A. (2004). A generalized Wiener attack on RSA. In *Practice and Theory in Public Key Cryptography, Proceedings of PKC 2004, Lecture Notes in Computer Science 2947*, pages 1–13.

Boneh, D. and Durfee, G. (2000). Cryptanalysis of RSA with private key d less than $N^{0.292}$. *IEEE Transactions on Information Theory*, 46:1339–1349.

de Weger, B. (2002). Cryptanalysis of RSA with small prime difference. *Applicable Algebra in Engineering, Communication and Computing*, 13:17–28.

Dujella, A. (2004). Continued fractions and RSA with small secret exponent. *Tatra Mountains Mathematical Publications*, 29:101–112.

Hardy, G. H. and Wright, E. M. (1965). *An Introduction to the Theory of Numbers*. Oxford University Press, London.

Nassr, D. I., Bahig, H. M., Bhery, A., and Daoud, S. S. (2008). A new RSA vulnerability using continued fractions. In *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications AICCSA 2008*, pages 694–701.

Steinfeld, R., Contini, S., Pieprzyk, J., and Wang, H. (2005). Converse results to the Wiener attack on RSA. In *Lecture Notes in Computer Science 3386*, pages 184–198.

Verheul, E. and van Tilborg, H. (1997). Cryptanalysis of 'less short' RSA secret exponents. *Applicable Algebra in Engineering, Communication and Computing*, 8:425–435.

Wiener, M. (1990). Cryptanalysis of short RSA secret exponents. *IEEE Transactions on Information Theory*, 36:553–558.

Securing Data on Outsourced Multistorage Repository

Moesfa Soeheila Mohamad[*1], **Geong Sen Poh**[1], and **Ji-Jian Chin**[2]

[1]*Information Security Lab, MIMOS Berhad*
[2]*Faculty of Engineering, Multimedia University*

E-mail: soeheila.mohamad@mimos.my
[*]*Corresponding author*

ABSTRACT

In utilising cloud storage services, users may not want to entrust complete control over all of their data to only one particular cloud provider. As such, this work proposes symmetric searchable encryption where no single server holds a complete document. This is achieved by partitioning the documents and scattering the encrypted blocks onto independent storage providers. Security model for SSE is adapted to fit this environment and a concrete scheme is presented and analysed.

Keywords: Searchable encryption, secure storage, cloud storage.

1 INTRODUCTION

As cloud technology is being widely accepted, the cloud owners offer storage service to the public. Among well-known cloud storage providers are Google Drive, Dropbox, Amazon S3 and Box. Such storage services can be utilized by users to optimize their data storage cost. However, users are aware of information disclosure by putting their documents on third party repositories. There are cloud storage providers with encryption feature such as Spider Oak. Using the client application or webpage, an encryption key is generated and user's files are encrypted before being uploaded to the storage server. This way the server is unable to read the files. However, the user can only identify files to use by the filename. but are unable to perform any additional operations such as searching on the content of the files.

To search contents of encrypted files, one could use symmetric searchable encryption (SSE). Considering users who subscribes to multiple cloud storage service, existing SSE schemes are only applicable in the trivial manner. In particular, the user can organize his documents into disjoint sets which he then stores at different storage providers via an SSE scheme. This is because currently, SSE is designed to be used with one storage. This setup has higher security than a single provider setup in the sense that each provider only has some and not all of the user's documents. Taking this concept further, we propose an SSE scheme in which documents are partitioned and the encrypted blocks are distributed over more than one storage providers.

2 RELATED WORKS

Song et al. (2000) introduced SSE by proposing that every word in the documents are encrypted, and hence search involves comparing the search token to every word. Next, Goh (2003) proposed index design with the IND-CKA security notion, and achieve search complexity linear to number of documents. The following proposals (Chang and Mitzenmacher, 2005, Curtmola et al., 2006) improved on index design to achieve index and search query security. More importantly, Curtmola et al. (2006) defined SSE security to encompass adaptive chosen keyword attacks. At this point, it is recognized that some information beyond document sizes and search result is leaked. Curtmola et al. (2006) defined search pattern as record of user's search queries whether it is the same keyword and access pattern as the list of document identifier or ciphertexts in the search results.

The concept of information leaked to server defined by Curtmola et al. (2006) is then parameterized by Chase and Kamara (2010) because type of information revealed may differ from scheme to scheme. The parameters are presented as leakage functions for two situations, setup leakage and search leakage. Besides the access pattern and search pattern (renamed as query pattern (QP)), Chase and Kamara (2010) introduced intersection pattern (IP) into the search leakage. For the schemes presented in (Chase and Kamara, 2010), the index hides the number of documents corresponding to each keyword by padding. As such, their structured schemes has less information in setup leakage compared to preceding schemes.

On the other hand, Cash et al. (2013) designed their index to hide the number of documents per keyword without padding. The index contains keyword-document pairs with only one document identifier per entry instead of a list of document identifiers as in Chase and Kamara (2010). The search complexity increased because the algorithm has to go through the index to find index keys corresponding to the search keyword. The number of documents resulting from the search is revealed after first search of a particular keyword.

Recently, some works focus on improving security based on the established security model. Naveed et al. (2014) designed a structure named blind storage to enable SSE strong against malicious servers, instead of honest-but-curious servers. He achieves this by storing the encrypted index on the server but only decrypted and read by the client. The cost of security improvement is the interactiveness of search. Search requires two rounds; first, the index is downloaded to obtain the document identifiers, and then the encrypted document blocks are extracted. Furthermore, this work is the first to partition the document or data into blocks. Consequently, a document size is hidden until the document is in a search result and downloaded by the user. Formally, this moves document sizes from being a setup leakage into a search leakage.

Our Contributions This work proposes the idea of utilizing more than one storage server for SSE. A multiserver SSE scheme is presented and analysed under security model fitting the new environment. By the design of index, the scheme delays the server's discovery of document sizes until the document is extracted at least twice.

3 PRELIMINARIES

3.1 Notation

We denote $\{1, 2, \ldots, n\}$ as $[n]$, $x \leftarrow A$ to mean x is an output of an algorithm A, $x \overset{R}{\leftarrow} X$ to mean random selection of a value x from a set X and $\|$ as concatenation. We use $negl(x)$ to indicate a negligible function, $poly(x)$ to indicate a polynomial function and $id(x) \in \{0,1\}^\lambda$ to denote an identifier that uniquely identify object x. The object may be a document, a block or a server.

This work assumes a user who owns n documents, $D = \{D_1, D_2, \ldots, D_n\}$ and utilizes s servers, $S = \{S_1, S_2, \ldots, S_s\}$ for storage. Every document D_f is divided into h_f equal-length blocks $d_{f,i} \in \{0,1\}^{len(\lambda)}$ and written as $D_f = \{d_{f,1}\|1, d_{f,2}\|2, \ldots, d_{f,h_f}\|h_f\}$. Denote max$(h_f) = m$. Further, denote the set of all blocks by $\mathbf{M} = \{M_1, \ldots, M_n\}$ where $M_f = \langle id(D_f), D_f \rangle$. The user also indicate the association of a keyword set to every document in an index, $\texttt{DB} = \{(id(D_f), W_f) \mid f = 1, \ldots, n\}$, where W_f denotes the set of keywords for document D_f. Indexes (or dictionaries) in our discussion is the data structure of the form $I[key] = value$. Given a key, the corresponding value is returned. For example, for the index DB, the key is the document identifier and the value is the associated set of keyword. Hence $\texttt{DB}[id(D_f)] = W_f$.

3.2 Symmetric encryption

A randomised symmetric encryption scheme $\mathcal{E} = (\texttt{Gen}, \texttt{Enc}, \texttt{Dec})$ consists of three PPT algorithms. Gen takes λ and outputs a secret key K; Enc takes a key K and a message $d \in \{0,1\}^{l(\lambda)}$ and outputs a ciphertext c. For all K from Gen and $d \in \{0,1\}^{l(\lambda)}$ we have $\texttt{Dec}(K, \texttt{Enc}(K, d)) = d$ with probability 1. We say $\mathcal{E}$ is indistinguishable under chosen-plaintext attack (CPA) if all PPT adversary has advantage $\mathbf{Adv}^{ind-cpa}_{\mathcal{E},\mathcal{A}}(\lambda)$,

$$\left|\Pr\left[\mathcal{A}(c){=}1 : c{\leftarrow}\texttt{Enc}(K,d), K{\leftarrow}\texttt{Gen}(1^\lambda)\right] - \Pr\left[\mathcal{A}(c){=}1 : c{\overset{R}{\leftarrow}}\{0,1\}^{l(\lambda)}\right]\right|$$

that is negligible, $\texttt{negl}(\lambda)$. Considering only IND-CPA symmetric encryption schemes, the adversary advantage in the following text is written as $\mathbf{Adv}_{\mathcal{E},\mathcal{A}}(\lambda)$.

3.3 Pseudorandom function

A function $F : \{0,1\}^\lambda \times \{0,1\}^* \to \{0,1\}^\lambda$ from $\mathcal{F}$ the set of all functions $\{0,1\}^* \to \{0,1\}^\lambda$ is pseudo-random if all PPT adversary $\mathcal{A}$ has the advantage $\mathbf{Adv}^{prf}_{F,\mathcal{A}}(\lambda)$,

$$\left|\Pr\left[\mathcal{A}(F_K) : K \overset{R}{\leftarrow} \{0,1\}^\lambda = 1\right] - \Pr\left[\mathcal{A}(g) : g \overset{R}{\leftarrow} \mathcal{F} = 1\right]\right|$$

that is negligible, $\texttt{negl}(\lambda)$. From this point, adversary advantage on pseudorandom function is written as $\mathbf{Adv}_{\mathcal{F},\mathcal{A}}(\lambda)$.

4 SSE WITH PARTITIONED DATA

The main objective of SSE with partitioned data (SSEwPD) is to disperse the power of storage provider on user's data by separating the data onto several independent storage in such away that none of the storage provider holds a complete document. SSEwPD consists of the same algorithms as SSE with similar correctness condition. The difference is in construction of index within the algorithms and in the adversary models.

4.1 Definition

The definitions of SSE in Chase and Kamara (2010), Curtmola et al. (2006) present five algorithms, `KeyGen`, `Encryption`, `Trapdoor`, `Search` and `Decryption`. By the description in this section, it is apparent that this work rename `Encryption` as `Setup`, and gather three algorithms, `Trapdoor`, `Search` and `Decryption` into one, `Search`.

Definition 4.1. *An SSEwPD scheme* Π *consists of:*

$k \leftarrow \mathtt{KGen}(1^\lambda)$*: A key generation algorithm, where on input the security parameter* λ*, it outputs a secret key* k*.*

$(\mathbf{K}, \mathbf{c}, I) \leftarrow \mathtt{Setup}(k, \mathtt{DB}, \mathbf{M}, S)$*: An index creation and encryption algorithm, where on input the secret key* k*, a database index* $\mathtt{DB}$*, the set of data blocks* $\mathbf{M}$ *and the list of storage providers* S*, it outputs a set of provider's keys* $\mathbf{K}$ *which is stored securely by the user and, a set of encrypted data blocks* $\mathbf{c}$ *and a set of (encrypted) indexes* I *for every server.*

$D^w := \mathtt{Search}(k, \mathbf{K}, w, I, S)$*: A search protocol, where on input the secret key* k*, the set of provider's keys* $\mathbf{K}$*, the set of (encrypted) indexes* I*, a query word* w *and the list of storage providers* S*, it outputs a set of decrypted documents* D^w *extracted from all the storage to the user.*

We say that Π *is correct if for any* λ*,* $\forall K \leftarrow \mathtt{KGen}(1^\lambda)$*,* $\forall \mathtt{DB}$*,* $\forall \mathbf{M}$*,* $\forall (K_S, \mathbf{c}, I) \leftarrow \mathtt{Setup}(K, \mathtt{DB}, \mathbf{M}, S)$*,* $\mathtt{Search}(K, K_S, w, I, S)$ *returns a set of encrypted data blocks,* $\mathbf{c}_f = (c_{f,j})_{j=1}^{h_f}$ *such that when the set of decrypted blocks* $D^w = (\mathcal{E}.\mathtt{Dec}(k_e, c_{f,j}))_{j=1}^{h_f}$*, where* $k_e \in K_S$*, are combined, the resulting documents have their identifiers in* $\mathtt{DB}(w)$*.*

Definition of an SSE scheme includes the leakage functions declaration of information allowed to be revealed to the storage provider during `Setup` and `Search`, $\mathcal{L}$=($\mathcal{L}^{setup}$,$\mathcal{L}^{query}$). The setup leakage, $\mathcal{L}^{setup}$ usually includes the total number of document and their sizes. Search leakage $\mathcal{L}^{query}$ includes access pattern (AP), query pattern (QP) and intersection pattern (IP). The AP records the identifiers in the search results and the encrypted data accessed during particular search instances. The QP records whether current search query is equal to a previous query. The IP records the identifiers in the search results and the encrypted data accessed during two or more different search queries.

4.2 Security

SSEwPD aims to achieve $\mathcal{L}$-security against adaptive chosen keyword attack by colluding storage providers which is defined by the following game:

Definition 4.2. *Given an SSE scheme,* $\Pi = (\texttt{KGen}, \texttt{Setup}, \texttt{Search})$ *with leakage function* $\mathcal{L} = (\mathcal{L}^{setup}, \mathcal{L}^{query})$*, semantic security is defined by two games.*

Real$_{\Pi,\mathcal{A}}(\lambda)$**:** *Adversary* $\mathcal{A}$ *selects a list of storage providers* S*, generates a set of documents* D*, a database index* $\texttt{DB} = \{(id(D_f), W_f) \mid f = 1, \ldots, n\}$ *and processes* D *as* $\mathbf{M} = (id_{D_f}, D_f)_{f=1}^{n}$*.* $\mathcal{A}$ *gives the challenger* $(\texttt{DB}, \mathbf{M}, S)$*. The challenger executes* $\texttt{KGen}(1^{\lambda})$ *to generate a secret key k and runs* $\texttt{Setup}(k, \texttt{DB}, \mathbf{M}, S)$ *resulting in* $(\mathbf{K}, \mathbf{c}, I)$*, in which* $\mathbf{K}$ *is stored secretly by the challenger while* $\mathbf{c}$ *and* $\mathcal{I} \in \{I, \perp\}$*, are given to* $\mathcal{A}$*, who then queries the* `Search` *protocols.*

For `Search`*,* $\mathcal{A}$ *chooses a sequence of query words* $w_1, \ldots, w_{p(\lambda)}$ *where* $p(.)$ *is a polynomial. The challenger executes* $\texttt{Search}(k,\mathbf{K},w_t, I, S)$ *and returns the corresponding set of encrypted blocks,* $E^{w_1}, \ldots, E^{w_{p(\lambda)}}$ *to* $\mathcal{A}$*.*

After receiving the challenger's replies, $\mathcal{A}$ *gives* $\mathcal{I}$*,* $(E^{w_1}, \ldots, E^{w_{p(\lambda)}})$*,* $(w_1, \ldots, w_{p(\lambda)})$*,* $((\texttt{op}_1, \texttt{in}_1), \ldots, (\texttt{op}_{p(\lambda)}, \texttt{in}_{p(\lambda)}))$ *to distinguisher* $\mathcal{D}$ *and obtain reply a bit b. Finally* $\mathcal{A}$ *outputs b.*

Ideal$_{\Pi,\mathcal{A},\mathcal{S}}(\lambda)$**:** *Adversary* $\mathcal{A}$ *selects a list of storage providers* S*, generates a database* $\texttt{DB}=\{(id(D_f), W_f) \mid f = 1, \ldots, n\}$ *and processes a set of documents* D *to form the set* $\mathbf{M} = \{(id(D_f), D_f) \mid f = 1, \ldots, n\}$*. The simulator* $\mathcal{S}$ *simulates* $(\mathbf{K}^*, \mathbf{c}^*, I*)$ *based on the leakage information from* $\mathcal{L}_{setup}$*, and gives* $\mathbf{c}^*$ *and* $\mathcal{I}^*$*, where* $\mathcal{I}^* \in \{I^*, \perp\}$*, to* $\mathcal{A}$*, who then queries the* `Search` *protocols. For* `Search`*,* $\mathcal{A}$ *chooses a sequence of query tokens* w_t *for* $t = 1, \ldots, p(\lambda)$ *where* $p(.)$ *is a polynomial. For every query* w_t $\mathcal{S}$ *is given* $\mathcal{L}_{query}$ *and returns to* $\mathcal{A}$ *the corresponding simulated encrypted data blocks* $E^{w_t^*}$*.*

After getting the replies from the challenger, $\mathcal{A}$ *gives* $(E^{w_1^*}, \ldots, E^{w^*_{p(\lambda)}})$*,* $\mathcal{I}^*$*,* $(w_1, \ldots, w_{p(\lambda)})$ *to distinguisher* $\mathcal{D}$ *which returns a bit b. Finally* $\mathcal{A}$ *outputs b.*

Π *is* $\mathcal{L}$*-secure against non-adaptive attacks if for all PPT adversaries* $\mathcal{A}$*, there exists a PPT simulator* $\mathcal{S}$ *such that*

$$\left|\Pr\left[\mathbf{Real}_{\Pi,\mathcal{A}}(\lambda) = 1\right] - \Pr\left[\mathbf{Ideal}^{\mathcal{S}}_{\Pi,\mathcal{A}}(\lambda) = 1\right]\right| \leq \texttt{negl}(\lambda).$$

5 A CONCRETE SCHEME

SSEwPD is designed similar to SSE, with encrypted index and encrypted data. However, instead of indexing and storing documents, the scheme works with document blocks. The document blocks are separated into many sets to be sent to different storage providers. Figure 1 shows the definition of an SSEwPD scheme, Π^S.

The `KGen` function generates a master key k. The `Setup` function uses k to derive a secret key K_j for every storage provider S_j using the identifiers of the server $id(S_j)$ as input. Following that, in `Setup` algorithm the master key k is used to derive keyword-associated keys K_x and K_v. Then for each keyword w, and given every $id(D_f)$ in $\mathtt{DB}(w)$, `Setup` generates a pseudo-random label l_j based on K_x and encrypts $id(D_f)$ with storage provider identifier $id(S_j)$ using K_v, so that each storage provider stores a unique encrypted $id(D_f)$. It iterates through all blocks of $id(D_f)$ by first selecting a storage provider S_j from all the storage providers. Temporary lists **t** are created to hold the encrypted $id(D_f)$ and the associated encrypted block identifiers for each storage provider. Then **t** is assigned to $\mathbf{QInx}_j$ for every key l_j. To search for documents with keyword w, the Π^S.`Search` protocol, as described in Figure 2, uses the master key k and w to derive K_x and K_v. K_x is broadcast to all storage providers and using this key each storage provider retrieves the encrypted document identifier $e_{f,j}$ and its associated encrypted block identifiers. These encrypted identifiers are sent back to the user, whom decrypts and uses

Preprocessing: Compile the list of storage providers, $S = \{S_1, S_2, \ldots, S_s\}$, database of document identifier/keyword pairs, $\mathtt{DB} = (id(D_f), W_f)_{f=1}^{n}$, and the document/data block pairs, $\mathbf{M} = (id(D_f), D_f)_{f=1}^{n}$, for $D_f = \{d_{f,1}\|1, d_{f,2}\|2, \ldots, d_{f,h_f}\|h_f\}$.

```
$k \leftarrow \mathtt{KGen}(1^\lambda)$:
Generate a random binary sequence, $k$, of length $\lambda$ and output to user.

$(\mathbf{K}, \mathbf{c}, I) \leftarrow \mathtt{Setup}(k, \mathtt{DB}, \mathbf{M}, S)$:
  1. For all $S_j \in S$, generate $K_j \leftarrow F(k, id(S_j))$ and set $\mathbf{K} = \{K_1, \ldots, K_s\}$.
  2. For each $w \in \mathbf{W}$:
     $K_x \leftarrow F(k, 1||w)$ and $K_v \leftarrow F(k, 2||w)$.
     Initialize a counter $z$, $z = 0$.
     For each $id(D_f) \in \mathtt{DB}(w)$:
         For all $S_j \in S$,
             $l_j \leftarrow F(K_x, z\|id(S_j))$
             $e_{f,j} \leftarrow \mathcal{E}.\mathtt{Enc}(K_v, id(D_f)\|id(S_j))$.
         Increment $z$.
         Initialise temporary lists $\mathbf{t}_1 = e_{f,1}, \ldots, \mathbf{t}_s = e_{f,s}$.
         For $i = 1$ to $h_f$:
             $j \xleftarrow{R} [s]$
             $c_{f,i} \leftarrow \mathcal{E}.\mathtt{Enc}(K_j, d_{f,i}||i)$
             $p_{f,i} \leftarrow \mathcal{E}.\mathtt{Enc}(K_v, id(d_{f,i}))$.
             Add $c_{f,i}$ to $\mathbf{c}_j$.
             Append $p_{f,i}$ to $\mathbf{t}_j$.
     For every $S_j$, set $\mathbf{QInx}_j[l_j] = \mathbf{t}_j$.
  3. Output to storage providers $(\mathbf{c}, I) = \{(\mathbf{c}_j, \mathbf{QInx}_j)$ for $S_j \mid j = 1, \ldots, s\}$,
     and output to user $\mathbf{K}$.
```

Figure 1: The Π^S.`KeyGen` and Π^S.`Setup` algorithms.

```
$D^w$ := Search($k$, $S$, $\mathbf{K}$, $w$):
    1. $K_x \leftarrow F(k, 1||w)$ and $K_v \leftarrow F(k, 2||w)$.
    2. Broadcast $K_x$ to all storage providers.
    3. For each storage provider $S_j$,
       Initialize a counter, $z = 0$.
       While matching entry on $z$ exists,
           $(e_{f,j}, \mathbf{p}) \leftarrow \mathbf{QInx}_j[F(K_v, z\|id(S_j))]$.
           Send $id(S_j)$, $(e_{f,j}$, $\mathbf{p})$ to the user.
           Increment $z$.
       Decrypt $e_{f,j}$: $id(D_f) \leftarrow \mathcal{E}.\mathtt{Dec}(K_v, e_{f,j})$.
       Decrypt every $p_{f,i} \in \mathbf{p}$: $id(d_{f,i}) \leftarrow \mathcal{E}.\mathtt{Dec}(K_v, p_{f,i})$
       Send every $id(d_{f,i})$ to $S_j$ to retrieve $c_{f,i}$.
       Decrypt $c_{f,i}$: $d_{f,i} \leftarrow \mathcal{E}.\mathtt{Dec}(K_j, c_{f,i})$
    4. Reconstruct $D_f$ by concatenation: $d_{f,1}\|_{f,2}\| \dots \|d_{f,h_f}$. Add $D_f$ to $D^w$.
    5. Output $D^w$ to user.
```

Figure 2: The `Search` algorithm for the construction with indexes on storage providers, Π^S.

them to retrieve the encrypted blocks from the storage providers.

After `Setup`, each $\mathbf{QInx}_j$ reveals its number of entries $N_j \leqslant N$, where $N = \Sigma_{w \in \mathbf{W}} |\mathtt{DB}(w)|$, and the number of encrypted block identifiers per entry $\left|\mathbf{p}_{l_j}\right|$. For search query w, let T_j^w be the set of resulting encrypted document identifiers from storage provider S_j. Then, the access pattern on storage provider S_j, AP$_j$ consists of the number of encrypted document identifier $\left|T_j^y\right|$ and their number of blocks $t_j = \{|\mathbf{p}_{f,j}| \mid e_{f,j} \in T_{S_j}\}$, and the set R_j^w which is ... From tokens queried, S_j can also build the query pattern QP$_j(w)$ and IP$_j(w)$. In summary, under scheme π^S, leakage function for storage provider S_j is $\mathcal{L}_j = (\mathcal{L}_j^{setup}, \mathcal{L}_j^{query})$ where

$$\mathcal{L}_j^{setup} = (|\mathbf{c}_j|, |c|, N_j, |\mathbf{p}_{l_j}| \text{ for } l_j \text{ in } \mathbf{QInx}_j)$$
$$\mathcal{L}_j^{query} = (\text{AP}_j = (|T_j|, t_j, R_j, \text{QP}_j(w), \text{IP}_j$$

Theorem 5.1. *The SSE scheme Π^S is $\mathcal{L}$-secure against non-adaptive chosen keyword attacks by colluding storage providers assuming F is a secure PRF and $\mathcal{E}$ is an IND-CPA symmetric encryption scheme.*

Proof. Based on the leakage functions defined above, we define the simulator $\mathcal{S}$ for non-colluding storage provider. Then we extend the simulator to be for colluding storage provider.

Let S_j be the storage provider being simulated while other storage providers perform Π^S. Simulator $\mathcal{S}$ prepares a simulated index $\mathbf{QInx}_j^*$ and block ciphertexts $\mathbf{c}_j^*$ from $\mathcal{L}_j^{setup} = (|\mathbf{c}_j|, |c|,$

$N_{S_j}, |\mathbf{p}_{l_j}|$ for l_j in $\mathbf{QInx}_j$) as follows. Randomly generate N_{S_j} binary strings x_i of length λ as the keys of the index, and randomly generate N_{S_j} binary strings of length λ as the encrypted document identifiers a_i. For each a_i, generate $|\mathbf{p}_{l_j}|$ random binary strings $\mathbf{b}_i$ to form $a_i \| \mathbf{b}_i$. Set $\mathbf{QInx}^*_{S_j} = \{(x_i, a_i \| \mathbf{b}_i) | i = 1, \ldots, N_{S_j}\}$. Finally, generate $|\mathbf{c}_j|$ random binary strings, each with length $|c|$ to form $\mathbf{c}^*_{S_j}$. Return $\mathbf{c}^*_j$, $\mathbf{QInx}^*_j$ to adversary.

For search queries the simulator $\mathcal{S}$ produces the transcript of interaction from $\mathcal{L}^{query}_j = (\text{AP}_j(w) = (|T_j|, t_j, |R_j(w)|), \text{QP}_j(w), \text{IP}_j(w))$. For query w, the simulator perform the following:

1. Choose $|T_j|$ entries from $\mathbf{QInx}^*_j$ such that their values part matches the numbers in t_j to form T^*_j.
   If QP(w) indicates this query has been made before, return the entries returned previously.
   Otherwise, select unused entries according to t.
2. Choose encrypted blocks to form D^*_w.
   If QP(w) indicates this query has been made before, return the previously returned D^*_w.
   If IP(w) indicates there are block identifiers accessed by this query are the same as those accessed by previous queries, choose the previously returned encrypted blocks.
   Otherwise, choose $|R_j|$ unused strings in $\mathbf{c}^*_j$.
3. Return to adversary the transcript (T^*_j,D^*_j).

Given that $\mathcal{E}$ is IND-CPA and F is a secure PRF, then the adversary advantage in distinguishing the random entries in $\mathbf{QInx}^*_j$ from the entries in the real $\mathbf{QInx}_j$ is at most the advantage of the adversaries for $\mathcal{E}$ and F. Similarly, the real search results T_j and R_j contain ciphertexts which are indistinguishable from random strings in (T^*_j,D^*_j) for IND-CPA $\mathcal{E}$. Hence,

$$\left|\Pr\left[\mathbf{Real}_{\Pi,\mathcal{A}_j}(\lambda) = 1\right] - \Pr\left[\mathbf{Ideal}^{\mathcal{S}}_{\Pi,\mathcal{A}_j}(\lambda) = 1\right]\right| \leqslant \mathbf{Adv}_F(\lambda) + \mathbf{Adv}_{\mathcal{E}}(\lambda)$$

which is negligible.

Now, consider the content of index $\mathbf{QInx}_j$. Every entry consists of ciphertexts dependent on storage provider identifier. As a result, $\mathbf{QInx}_j$ for all j is mutually exclusive. It follows that the output from each storage provider for `Search` are independent. So, the simulator for colluding storage providers, can prepare replies based on leakage of every storage provider $\mathcal{L}^{setup}_j = (|\mathbf{c}_j|$,
$|c|, N_{S_j}, |\mathbf{p}_{l_j}|$ for l_j in $\mathbf{QInx}_j$) and $\mathcal{L}^{query}_j = (\text{AP}_j = (|T_j|, t_j, R_j), \text{QP}_j, \text{IP}_j))$ independently for each j as for the non-colluding storage provider attack. So, we have that the advantage of the adversary for colluding storage providers $\mathcal{A}_B(\lambda)$ is the sum of advantage of the adversaries controlling single storage provider $\mathcal{A}_g(\lambda)$ for $S_j \in B$,

$$\left|\Pr\left[\mathbf{Real}_{\Pi,\mathcal{A}_B}(\lambda) = 1\right] - \Pr\left[\mathbf{Ideal}^{\mathcal{S}}_{\Pi,\mathcal{A}_B}(\lambda) = 1\right]\right| \leqslant s\big(\mathbf{Adv}_F(\lambda) + \mathbf{Adv}_{\mathcal{E}}(\lambda)\big)$$

which is also negligible. □ The performance of Π^S can be measured in terms of indexes creation during `Setup` and communications during `Search`.

The main setup cost is the storage of $\mathbf{QInx}_j$ for every storage provider S_j. The size of $\mathbf{QInx}_j$ is $O(Nm/s)$ given $N = \Sigma_{w \in \mathbf{W}} |\mathtt{DB}(w)|$ and since the overall index is partitioned into s storage providers. For a document, it stores maximally m encrypted block identifiers.

The search cost is $O(rz)$ for $r = |\mathtt{DB}(w)|$ and $z = \Sigma_{j=1}^{s}|\mathbf{QInx}_j[l_j]|$ since for a query on w, the total number of matches is r with z the size of retrievals from all storage providers, whereby each $\mathbf{QInx}_j[l_j]$ returns the list of encrypted block identifiers associated to the matching documents. The communication cost is $O(s)$, linear to the number of participating storage providers since the query is broadcast to all storage providers.

6 CONCLUSION

This work proposes extending searchable symmetric encryption schemes to ensure a storage provider would not be able to obtain user's complete document. The SSE security model is adapted to consider adversary who controls more than one service providers. One concrete SSEwPD scheme is presented and proven to achieve $\mathcal{L}$-security under chosen keyword attack against colluding storage providers.

REFERENCES

Cash, D., Jarecki, S., Jutla, C. S., Krawczyk, H., Rosu, M.-C., and Steiner, M. (2013). Highly-Scalable Searchable Symmetric Encryption with Support for Boolean Queries. In Canetti, R. and Garay, J. A., editors, *Advances in Cryptology - CRYPTO 2013 and IACR eprint archive*, volume 8042 of *LNCS*, pages 353–373. Springer.

Chang, Y. and Mitzenmacher, M. (2005). Privacy Preserving Keyword Searches on Remote Encrypted Data. In Ioannidis, J., Keromytis, A. D., and Yung, M., editors, *ACNS 2005*, volume 3531 of *LNCS*, pages 442–455. Springer.

Chase, M. and Kamara, S. (2010). Structured Encryption and Controlled Disclosure. In Abe, M., editor, *ASIACRYPT 2010*, volume 6477 of *LNCS*, pages 577–594. Springer.

Curtmola, R., Garay, J. A., Kamara, S., and Ostrovsky, R. (2006). Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions. In Juels, A., Wright, R. N., and di Vimercati, S. D. C., editors, *ACM Conference on Computer and Communications Security, CCS 2006*, pages 79–88. ACM.

Goh, E.-J. (2003). Secure indexes. Cryptology ePrint Archive, Report 2003/216. `http://eprint.iacr.org/2003/216/`.

Naveed, M., Prabhakaran, M., and Gunter, C. A. (2014). Dynamic Searchable Encryption via Blind Storage. In *2014 IEEE Symposium on Security and Privacy, SP 2014*, pages 639–654. IEEE Computer Society.

Song, D. X., Wagner, D., and Perrig, A. (2000). Practical Techniques for Searches on Encrypted Data. In *SP '00: Proceedings of the 2000 IEEE Symposium on Security and Privacy*, page 44. IEEE Computer Society.

S-box Optimisation using Heuristic Methods

Herman Isa[*1,2], **Norziana Jamil**[1], and **Muhammad Reza Z'aba**[2]

[1]*College of Information Technology, Universiti Tenaga Nasional (UNITEN), Malaysia*
[2]*Network Security Lab, MIMOS Berhad, Kuala Lumpur, Malaysia*

E-mail: herman.isa@mimos.my
[*]*Corresponding author*

ABSTRACT

Isa et al. (2013, 2016) proposed two heuristic algorithms (redundancy removal and bee waggle dance) to construct cryptographically strong substitution boxes (S-boxes). The resulting S-boxes produced by these algorithms are suitable for cryptographic use. Inspired by their work, this paper explores a new method to optimise an S-box by integrating these two algorithms. Our experiments show that at least three cryptographically strong S-boxes can be produced by the new method. The results also improves upon a previous construction by Mamadolimov et al. (2013) which utilises the redundancy removal algorithm.

Keywords: S-box, Nonlinearity, Heuristic, Redundancy Removal Algorithm, Bee Waggle Dance Algorithm

1 INTRODUCTION

A substitution box (S-box) is cores of nonlinear operation in symmetric cryptosystem especially block ciphers. An S-box typically used to obscure the relationship between key and ciphertext, such that fulfils Shannon's property of confusion (Shannon, 1949).

In general, there are three generic methods in the construction of an S-box, which are random searching approach, heuristic or evolutionary (i.e. heuristic) approach and mathematical functions or algebraic (i.e. mathematical) approach. Each approach has its advantages and weaknesses. As an example, the advantage of each approach is random searching being the simplest method; heuristic approach has better implementation in both software and hardware; and lastly known best cryptography properties (National Institute of Standards and Technology, 2001) achieved by mathematical approach. Yet, the weakness of each approach is low cryptographic properties exhibited by random and heuristic approaches, and extremely hard to find a mathematical function that give a complete set of cryptographically strong S-boxes.

However, in recent years, the uses of heuristic approach in S-box construction are gaining the attention of researchers. This can be seen through the increasing number of S-box constructions proposed in literature such as using evolution of theorem of permutation polynomials (Yang et al., 2011), gradient descent (Kazymyrov et al., 2013), redundancy removal algorithm (Isa et al., 2013), chaotic map-based technique (Alkhaldi et al., 2015), reversed genetic algorithm (Ivanov et al., 2016) and latest is the S-box construction inspired by bee waggle dance (Isa et al., 2016).

In this paper, we optimise the construction of S-box by combining two algorithms proposed by Isa et al. (2013, 2016), which are redundancy removal algorithm (RRA) and bee waggle dance (BWD) algorithm. Our objective is to construct a permutation S-box from a non-permutation initial S-box that performs the RRA and then followed by the BWD algorithm.

The rest of the paper is organised as follows. In the second section, the main cryptographic properties of an S-box are discussed. Then, we share our S-box optimisation together with the involved algorithms in the third section. The paper is concluded in the last section.

2 S-BOX PROPERTIES

In this paper, our focused result is on bijective S-boxes over finite field $\mathbb{F}_{2^8}$. Therefore, a cryptographically strong S-box should at least exhibits the optimal values on the following three properties: (1) high nonlinearity (NL), (2) low differential uniformity (DU), and (3) high algebraic degree (AD).

Let $\mathbb{F}_2$ and $\mathbb{F}_{2^n}$ be a finite field with 2 and 2^n elements, respectively. An $n \times n$ S-box is a Boolean map:

$$F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n} = (f_1(x_1, \dots, x_n), \dots, f_n(x_1, \dots, x_n)) \tag{1}$$

2.1 Nonlinearity

Let $c \cdot F = c_1 f_1 + c_2 f_2 + \dots + c_n f_n$ be a linear combination of the coordinate Boolean functions $(f_1, f_2, \dots, f_n)$ of F where $c = (c_1, c_2, \dots, c_n)$ be a nonzero elements in $\mathbb{F}_{2^n}$. The nonlinearity (NL) of an S-box F, is the Hamming distance between the set of all non-constant linear combination of component functions of F and the set of all affine functions over $\mathbb{F}_{2^n}$ as defined below:

$$\mathrm{NL}(F) = \min_{c \in \mathbb{F}_{2^n}, c \neq 0} \mathrm{NL}(c \cdot F) \tag{2}$$

Carlet (2011) suggested that the value of NL should be as close as to the best known NL (i.e. NL > 100) to thwart linear cryptanalysis (Matsui, 1994).

2.2 Differential Uniformity

The largest value present in difference distribution table, after omitting the trivial entry case (i.e. $a = b = 0$), determine the value of differential uniformity (DU). The value of DU is defined as:

$$\mathrm{DU}(F) = \max_{a,b \in \mathbb{F}_{2^n}, a \neq 0} |\{x \in \mathbb{F}_{2^n} : F(x+a) + F(x) = b\}| \tag{3}$$

Smaller value of DU is more preferable (i.e. $2 \leq \mathrm{DU} \leq 6$) (Carlet, 2011) to resist differential cryptanalysis (Biham and Shamir, 1991).

2.3 Algebraic Degree

The number of variables in the largest monomial for component function f of an S-box is denoted as $deg(f)$. Therefore, the algebraic degree (AD) of the S-box is determined by the maximum degree of all component functions:

$$\mathrm{AD}(F) = \max\{deg(f_1), deg(f_2), \ldots, deg(f_n)\} \tag{4}$$

Carlet (2011) suggested that AD ≥ 4 in order to resist higher order differential cryptanalysis (Knudsen, 1995).

3 S-BOX OPTIMISATION

As mentioned above, in this study, we analyse the combination of two different algorithms proposed by Isa et al. (2013, 2016) in constructing a cryptographically strong S-box. The referred algorithms are *Redundancy Removal Algorithm* and *Bee Waggle Dance* algorithm. The following subsections will describe the said algorithms in brief.

3.1 Redundancy Removal Algorithm

Isa et al. (2013) proposed an S-box construction from non-permutation power functions. One of the algorithms included in their construction is called *Redundancy Removal Algorithm* (RRA). In principle, this RRA is an improvement of the algorithm proposed by (Mamadolimov et al., 2013). As the name suggest, the RRA was meant to remove or replace the redundant elements in an initial S-box with the non-existent elements such that a bijective S-box is generated.

Figure 1 illustrate the process flow in RRA. The algorithm start with a non-permutation initial S-box as an input. From the input, the information about redundant elements and non-existent elements were extracted. Then, a table called as Distance Matrix (DM) is generated. This table contains the Hamming distances which were calculated based on bit error rates between the redundant elements and the non-existent elements in the initial S-box. Then, the smallest Hamming distance is selected and its corresponding redundant element will be replaced by

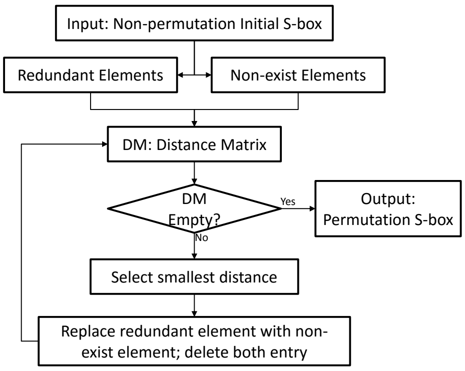

Input: Non-permutation Initial S-box
Redundant Elements
Non-exist Elements
DM: Distance Matrix
DM Empty?
Yes
No
Output: Permutation S-box
Select smallest distance
Replace redundant element with non-exist element; delete both entry

Figure 1: Redundancy Removal Algorithm (RRA)

non-existent element in the initial S-box. This process is repeated until there is none DM generated. As a result, a permutation S-box is constructed.

3.2 Bee Waggle Dance Algorithm

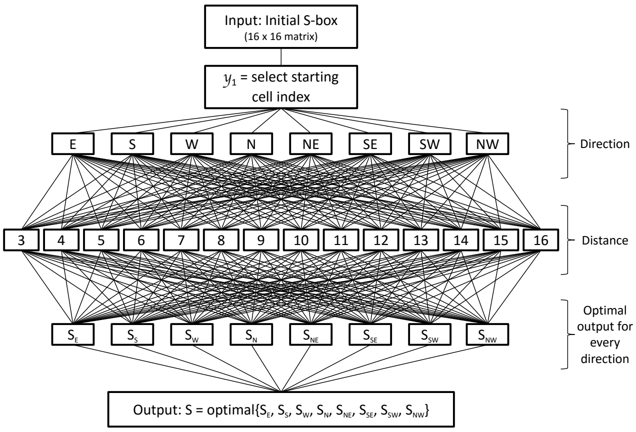

Input: Initial S-box (16 x 16 matrix)
y_1 = select starting cell index
E
S
W
N
NE
SE
SW
NW
Direction
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Distance
S_E
S_S
S_W
S_N
S_{NE}
S_{SE}
S_{SW}
S_{NW}
Optimal output for every direction
Output: S = optimal{S_E, S_S, S_W, S_N, S_{NE}, S_{SE}, S_{SW}, S_{NW}}

Figure 2: A loop of BWD Algorithm with static starting point

Recently, Isa et al. (2016) introduced a new algorithm inspired by bees behaviour to construct an S-box. They name the algorithm as *Bee Waggle Dance* (BWD)adopt algorithm. As illustrated in Figure 2, BWD algorithm requires a starting point (y_1), eight different directions (r), 14 distinct distances (d) and at least a loop (l) to be completed. The BWD function is defined as the following:

$$\texttt{BWD}(r, d, l, y_1) = y_1 \to y_2 \to \ldots \to y_j \to y_1 \tag{5}$$

where the right hand side of equation 5 denotes the movement of bee from cell indexed by y_1 (i.e. starting point) to cell indexed by y_2. This movement will continue so on and so forth, until it returns to cell indexed by y_1. Thus, one complete loop (l) is counted.

The third parameter of BWD algorithm (i.e. eight different directions (r)) is borrowed from the *points of the compass*. These points of the compass are composed of four cardinal directions (i.e. east (E), south (S), west (W) and north (N)) and four intercardinal directions (i.e. northeast (NE), southeast (SE), southwest (SW) and northwest (NW)).

The last parameter in BWD algorithm varies from the minimum distance (i.e. $d = 3$) to maximum distance (i.e. $d = 16$), thus make it 14 distinct distances as a whole. This distance is counted based on the height of the cell traversed on the dance floor. Basically, the dance floor is the initial S-box which was arranged in the form of a 16×16 matrix.

The final S-box in BWD algorithm is selected based on the most optimal cryptographic properties exhibited by the generated S-boxes.

3.3 Our Construction

Just like the construction proposed in Isa et al. (2013) and Isa et al. (2016), we also apply an initial S-box to be optimised by the RRA and BWD algorithms. However, in executing the optimisation, we have two options to perform the construction which are either to execute RRA first or BWD first on our initial S-box. Figure 3 illustrate these two options by representation of *Opt-1* and *Opt-2*.

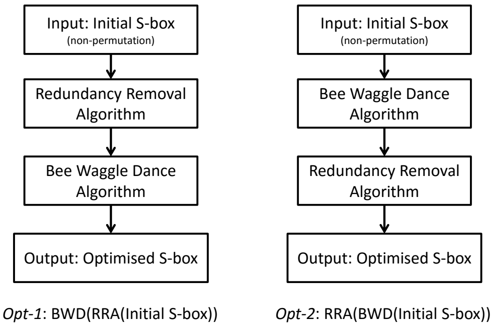

Input: Initial S-box
(non-permutation)
Redundancy Removal Algorithm
Bee Waggle Dance Algorithm
Output: Optimised S-box
Opt-1: BWD(RRA(Initial S-box))
Input: Initial S-box
(non-permutation)
Bee Waggle Dance Algorithm
Redundancy Removal Algorithm
Output: Optimised S-box
Opt-2: RRA(BWD(Initial S-box))

Figure 3: Option to Perform S-box Optimisation

In *Opt-1*, the initial S-box, which is non-permutation will be executed in RRA to generate a permutation S-box. Then, this generated S-box will perform the BWD algorithm until the most optimal S-box is identified. While in *Opt-2* construction, the initial S-box will perform the BWD algorithm first. Here, all the optimal outputs in every direction of BWD algorithm will be taken as candidates to perform RRA. This is because we cannot certain which candidate will exhibit the most optimal cryptographic properties after RRA is applied. Nevertheless, due to more time consuming and more detailed attention required in executing the *Opt-2* construction, therefore, in this paper we focus our attention on the construction of *Opt-1*.

In our construction, the initial S-box must be from a non-permutation function. Therefore, our first candidates to fulfil this requirement are the set of non-permutation power functions as used by Isa et al. (2013). However, preliminary investigation shows that almost no candidates from this set can retain its cryptographic properties once performing the RRA. It might because of the large number of non-existent elements in the function. Therefore, this set of candidates was discarded.

As a solution, we adopt the S-box proposed by Isa et al. (2014, 2015) called as S-Box2. This S-Box2 is generated using trinomial power functions over finite field $\mathbb{F}_{2^8}$ with the following function:

$$F_{S-Box2} = x^{29} + x^{89} + x^{164} \tag{6}$$

which exhibits (108, 6, 4) for its (NL, DU, AD), respectively. To fulfil our requirement for the initial S-box, at least three elements must be removed from S-Box2, such that a non-permutation initial S-box is obtained. Now, the current cryptographic properties of the initial S-box are changed to (106, 8, 4). Despite that, we manage to retain the cryptographic properties of (106, 6, 4) after the RRA is performed on the initial S-box. Afterward, we conduct the BWD algorithm.

3.3.1 Result

Table 1 shows the optimal results obtained from the initial S-box for every distance and direction performed using the BWD algorithm. The entries in Table 1 represents the cryptographic properties of nonlinearity, differential uniformity and algebraic degree of each result and the starting point, y_1 used in the BWD algorithm. We denote these entries as (NL, DU, AD, y_1). Note that, as discussed in Section 2, an S-box is considered as cryptographically strong if it satisfies the following requirements: i) NL > 100, ii) 2 $\leq$ DU $\leq$ 6 and iii) AD ≥ 4. From Table 1, there are three S-boxes that meet these requirements which are highligted in the table.

Direction	EAST	SOUTH	WEST	NORTH	NORTHEAST	SOUTHEAST	SOUTHWEST	NORTHWEST
Distance								
3	(102, 6, 7, 235)	(102, 8, 7, 194)	(104, 8, 7, 180)	(102, 8, 7, 207)	(104, 8, 7, 8)	(102, 8, 7, 3)	(102, 6, 7, 179)	(104, 8, 7, 181)
4	(102, 8, 7, 172)	(102, 8, 7, 184)	(102, 8, 7, 61)	(102, 8, 7, 175)	(102, 8, 7, 10)	(102, 8, 7, 1)	(102, 8, 7, 242)	(102, 6, 7, 94)
5	(102, 8, 7, 69)	(100, 8, 7, 162)	(100, 8, 7, 108)	(100, 8, 7, 172)	(102, 8, 7, 14)	(102, 8, 7, 5)	(102, 8, 7, 185)	(102, 8, 7, 254)
6	(100, 8, 7, 25)	(98, 8, 7, 155)	(100, 8, 7, 234)	(100, 8, 7, 103)	(102, 8, 7, 22)	(102, 8, 7, 65)	(102, 8, 7, 248)	(100, 8, 7, 253)
7	(98, 8, 7, 106)	(98, 8, 7, 134)	(100, 8, 7, 185)	(100, 8, 7, 154)	(100, 8, 7, 7)	(100, 8, 7, 23)	(100, 8, 7, 217)	(100, 8, 7, 216)
8	(100, 8, 7, 73)	(100, 8, 7, 98)	(98, 8, 7, 202)	(98, 8, 7, 137)	(98, 8, 7, 27)	(100, 8, 7, 132)	(100, 8, 7, 163)	(98, 8, 7, 250)
9	(94, 8, 7, 104)	(98, 8, 7, 118)	(94, 8, 6, 185)	(98, 8, 7, 126)	(98, 8, 7, 29)	(98, 8, 7, 39)	(98, 8, 7, 216)	(96, 8, 7, 169)
10	(98, 8, 7, 72)	(98, 10, 7, 114)	(98, 10, 7, 201)	(98, 8, 7, 126)	(96, 8, 7, 110)	(96, 8, 7, 1)	(100, 10, 7, 145)	(98, 8, 7, 237)
11	(98, 8, 7, 56)	(98, 10, 7, 114)	(98, 10, 7, 233)	(98, 10, 7, 123)	(98, 8, 7, 15)	(94, 8, 7, 34)	(96, 8, 7, 246)	(98, 8, 7, 238)
12	(98, 8, 7, 56)	(96, 10, 7, 114)	(98, 10, 7, 201)	(98, 10, 7, 123)	(98, 10, 7, 13)	(98, 8, 7, 35)	(92, 8, 7, 226)	(94, 8, 7, 206)
13	(98, 10, 7, 56)	(98, 10, 7, 115)	(98, 10, 7, 201)	(98, 10, 7, 124)	(98, 10, 7, 13)	(96, 10, 7, 1)	(98, 10, 7, 210)	(96, 10, 7, 221)
14	(96, 10, 7, 24)	(94, 10, 7, 116)	(96, 10, 7, 233)	(96, 10, 7, 125)	(96, 10, 7, 16)	(98, 10, 7, 3)	(96, 10, 7, 243)	(96, 10, 7, 239)
15	(96, 10, 7, 40)	(94, 12, 7, 115)	(96, 12, 7, 233)	(98, 10, 7, 126)	(96, 10, 7, 31)	(94, 10, 7, 2)	(96, 10, 7, 241)	(92, 10, 7, 239)
16	(94, 10, 7, 24)	(90, 12, 7, 114)	(94, 12, 7, 233)	(92, 10, 7, 127)	(94, 10, 7, 16)	(98, 10, 7, 1)	(94, 10, 7, 241)	(94, 12, 7, 256)

Table 1: Experiment Results on Initial S-box

The S-box obtained at direction northwest and distance 4 is represented in hexadecimal in Table 2. The first column in Table 2 denotes the first four bits of the input while the first row denotes the remaining four bits of the 8-bit input to the S-box. For instance, the input `4A` gives the output `74`, (i.e. F(`4A`) = `74`). The 13 highlighted elements in Table 2 represent the changed elements of the final S-box from the initial S-box, S-Box2.

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	00	01	17	6C	08	E9	35	CB	39	4A	DD	98	4B	E4	F1	D1
1	40	EB	A5	E0	78	9B	9A	3C	AA	76	14	F2	EF	CD	0E	61
2	43	EC	32	1F	6A	12	BE	80	A6	9D	7B	7D	8E	81	8C	6F
3	E1	47	B5	57	73	5F	37	55	4C	F5	5D	15	8F	B4	10	36
4	28	89	2D	DE	92	4E	E5	D3	B8	54	74	B0	FF	1D	45	93
5	88	BA	24	50	1A	C2	F8	BC	B3	0D	58	3B	A8	23	AE	D9
6	2E	13	41	77	FD	19	67	91	D8	AB	5A	B2	C5	6B	BD	4D
7	6D	48	E8	72	C8	FA	42	EA	85	A1	9F	FE	C1	7C	05	F6
8	7A	71	3F	CA	33	82	C7	29	0C	2C	63	69	C0	3A	D4	79
9	7F	51	44	4F	90	27	06	FB	0B	DC	B9	07	E2	46	1B	65
A	59	A0	A4	09	B6	8D	A7	96	5C	E6	95	AF	0F	AC	8A	75
B	C9	EE	03	F9	9E	F3	62	1C	18	20	04	AD	B1	CC	49	B7
C	6E	DA	3D	D5	F7	2A	26	DB	97	25	60	86	52	34	A2	30
D	53	3E	C3	A9	64	D0	D7	D6	C6	BB	38	D2	99	0A	ED	87
E	5B	E3	CE	21	02	66	1E	F0	FC	CF	E7	8B	9C	2B	BF	56
F	11	A3	68	84	DF	F4	16	70	22	83	5E	31	7E	94	2F	C4

Table 2: Optimal S-box using Proposed Technique

3.4 Discussion

By adopting and combining the *Redundancy Removal Algorithm* (Isa et al., 2013) and *Bee Waggle Dance* algorithm (Isa et al., 2016), we manage to obtain cryptographically strong S-boxes that compare well with the original proposed constructions. Table 3 summarised the main cryptographic properties (i.e. NL, DU, and AD) exhibited by each proposed S-box construction involved in this study which are Mamadolimov et al. (2013), Isa et al. (2013), Isa et al. (2016) and our result.

Construction	**NL**	**DU**	**AD**	**Technique**
Isa et al. (2016)	108	6	7	Bee Waggle Dance Algorithm
Isa et al. (2013)	104	6	7	Redundancy Removal Algorithm*
This paper	**102**	**6**	**7**	*Opt-1*: BWD (RRA (Initial S-box))
Mamadolimov et al. (2013)	102	8	7	Redundancy Removal Algorithm

*This technique is not utilised in the said paper. See discussion below.

Table 3: Comparison Result

In summary, the S-box generated using BWD algorithm (Isa et al., 2016) is ranked first since it exhibits NL = 108, DU = 6 and AD = 7. This might due to their selected initial S-box was generated from trinomial power function that is equivalent to inverse function (National Institute of Standards and Technology, 2001).

Isa et al. (2013) is ranked second since their proposed S-box exhibits (104, 6, 7) for its (NL, DU, AD), respectively. In brief, they took a non-permutation power function with highest NL (i.e. NL = 112) and lowest DU (i.e. DU = 2) as a base-function and add it with another power function over $\mathbb{F}_{2^8}$ to generate an initial S-box. If the initial S-box is found not bijective, then RRA is applied. Otherwise, at least any two elements in initial S-box were swapped to generate the final S-box.

Our S-box construction is ranked third. In this construction, we adopt the S-box generated in Isa et al. (2014, 2015) with slight modification which lastly exhibits (106, 8, 4) for its (NL, DU, AD) as an initial S-box. Then we perform the S-box optimisation through the RRA and followed by BWD algorithm to obtain the final bijective S-box.

The proposed S-box by Mamadolimov et al. (2013) is ranked last. This proposal was the first introducing the RRA in the construction of an S-box. However, since the value of their DU = 8, thus make the S-box fail to fulfil one of the pre-condition required to be considered as cryptographically strong (i.e. $2 \leq \text{DU} \leq 6$).

Our best result produced so far, however, fail to outperform the original proposals of Isa et al. (2013, 2016). The key factor of our lower result might due to the selection of initial S-box with lower NL (instead of NL = 112 as used in Isa et al. (2016)). Besides that, we notice that the exemplary construction in Isa et al. (2013) was not invoking the RRA. This is because, after an intermediate processing stage, their initial S-box is already bijective. Thus, their final S-box was generated from swapping the elements in the initial S-box. Nevertheless, as an alternative, our proposed method manage to produce several crytographically strong S-boxes that fulfills our pre-condition requirements as discussed in Section 2.

4 CONCLUSION

In this paper, we manage to combine two different algorithms proposed by Isa et al. (2013) and Isa et al. (2016) named as *Redundancy Removal Algorithm* and *Bee Waggle Dance* algorithm, respectively to construct an S-box. The construction performed RRA first and followed by BWD algorithm on the initial S-box. With the right selection of initial S-box, this combination managed to produce several cryptographically strong S-boxes that fulfils our pre-condition requirements of i) $\text{NL} > 100$, ii) $2 \leq \text{DU} \leq 6$ and iii) $\text{AD} \geq 4$.

ACKNOWLEDGMENTS

The authors would like to express their deepest gratitude to Ministry of Higher Education (MOHE), Malaysia for the financial support in this research project under the Exploratory Scheme Research Grant 2012/13.

REFERENCES

Alkhaldi, A. H., Hussain, I., and Gondal, M. A. (2015). A Novel Design for the Construction of Safe S-boxes based On TDERC Sequence. *Alexandria Engineering Journal*, 54(1):65–69.

Biham, E. and Shamir, A. (1991). Differential Cryptanalysis of DES-like Cryptosystems. In Menezes, A. J. and Vanstone, S. A., editors, *Advances in Cryptology - CRYPT0 '90*, volume 537 of *Lecture Notes in Computer Science*, pages 2–21. Springer Berlin Heidelberg.

Carlet, C. (2011). On Known and New Differentially Uniform Functions. In Parampalli, U. and Hawkes, P., editors, *Information Security and Privacy*, volume 6812 of *Lecture Notes in Computer Science*, pages 1–15. Springer Berlin Heidelberg.

Isa, H., Jamil, N., and Z'aba, M. R. (2013). S-box Construction from Non-Permutation Power Functions. In *Proceedings of the* 6^{th} *International Conference on Security of Information and Networks*, SIN '13, pages 46–53, New York, NY, USA. ACM.

Isa, H., Jamil, N., and Z'aba, M. R. (2014). Improved S-box Construction from Binomial Power Functions. In *Proceedings of the* 4^{th} *International Cryptology and Information Security Conference 2014 (CRYPTOLOGY2014)*, CRYPTOLOGY2014, pages 131–139, UPM Serdang, Selangor, Malaysia. Institute for Mathematical Research (INSPEM).

Isa, H., Jamil, N., and Z'aba, M. R. (2015). Improved S-box Construction from Binomial Power Functions. *Malaysian Journal of Mathematical Sciences*, 9(S)(1):21–35.

Isa, H., Jamil, N., and Z'aba, M. R. (2016). Construction of Cryptographically Strong S-Boxes Inspired by Bee Waggle Dance. *New Generation Computing*. (to be appeared).

Ivanov, G., Nikolov, N., and Nikova, S. (2016). Reversed genetic algorithms for generation of bijective s-boxes with good cryptographic properties. *Cryptography and Communications*, pages 1–30.

Kazymyrov, O., Kazymyrova, V., and Oliynykov, R. (2013). A Method for Generation of High-Nonlinear S-Boxes based on Gradient Descent. Cryptology ePrint Archive, Report 2013/578.

Knudsen, L. R. (1995). Truncated and Higher Order Differentials. In Preneel, B., editor, *Fast Software Encryption*, volume 1008 of *Lecture Notes in Computer Science*, pages 196–211. Springer Berlin Heidelberg.

Mamadolimov, A., Isa, H., and Mohamad, M. S. (2013). Practical Bijective S-box Design. *CoRR*, abs/1301.4723.

Matsui, M. (1994). Linear Cryptanalysis Method for DES Cipher. In Helleseth, T., editor, *EUROCRYPT '93*, volume 765 of *Lecture Notes in Computer Science*, pages 386–397. Springer Berlin Heidelberg.

National Institute of Standards and Technology (2001). Advanced Encryption Standard. Federal Information Processing Standard (FIPS) 197.

Shannon, C. E. (1949). Communication Theory of Secrecy Systems. *Bell System Technical Journal*, 28(7):656–715.

Yang, M., Wang, Z., Meng, Q., and Han, L. (2011). Evolutionary Design of S-box with Cryptographic Properties. In *Ninth IEEE International Symposium on Parallel and Distributed Processing with Applications Workshops (ISPAW), 2011*, ISPAW '11, pages 12–15. IEEE Computer Society.

CSM S-Box Evaluation Tool (CSET); Tool to Evaluate the Strength of an S-Box

Nik Azura Nik Abdullah[*1], **Abdul Alif Zakaria**[2],
Wan Zariman Omar[3], **Nor Azeala Mohd Yusof**[4], and
Hazlin Abdul Rani[5]

[1,2,3,4,5]*Cryptography Development Department*
CyberSecurity Malaysia
Kuala Lumpur, Malaysia

E-mail: [1]*azura@cybersecurity.my*, [2]*alif@cybersecurity.my,*[3]*wanzariman@cybersecurity.my,*
[4]*azeala@cybersecurity.my,* [5]*hazlin@cybersecurity.my*
**Corresponding author*

ABSTRACT

In this research paper, we introduce a cryptographic evaluation tool to evaluate the strength of an S-Box. This tool evaluates the Nonlinearity, Algebraic Degree and Differential Uniformity properties of an S-Box. It also analyses the Avalanche Effect, Completeness and Strong S-Box criteria of the tested S-Box. This tool will finally generates a complete report which provides a conclusion to indicate whether the tested S-Box passes or fails each test. It takes about seven seconds to produce a complete analysis report compared to an hour and a half when using conventional method. With the presence of this tool, process of evaluating an S-box will be simplified and automated which resulted in reducing the evaluation timeframe period significantly.

Keywords*: S-Box, S-Box Design Properties, Strict Avalanche Criterion*

1 INTRODUCTION

S-Box is one of the most important components in development of cryptography algorithm. It is a basic component of symmetric key algorithms which performs substitution. In block ciphers, they are typically used to obscure the relationship between the key and the ciphertext which applies Shannon's property of confusion (Bart, 2006). A good S-Box influences the strength of an algorithm. S-Box has been implemented in most of the frequently used cryptography algorithm, such as AES, DES, Blowfish, IDEA, and Skipjack. Rijndael S-Box, which the AES cryptographic algorithm was based on, is considered as one of the best S-Box among these algorithms (Selcuk and Melek, 2001).

In general, an S-Box takes some number of input bits, *m*, and transforms them into some number of output bits, *n*, where *n* is not necessarily equal to *m.* An $m\times n$ S-box can be implemented as a lookup table with 2*m* words of *n* bits each.

Two properties must be taken into account when designing an S-Box, which are the S-Box Design Properties and Strict Avalanche Criterion. S-Box Design Properties are closely related to the most significant attacks on S-Box which are linear attack and differential attack (Howard*,* 2011). These two attacks determine the value of Nonlinearity (NL), Algebraic Degree (AD), and Differential Uniformity (DU) which are the properties of S-Box. Meanwhile, Strict Avalanche Criterion is also crucial in determining the strength of an S-Box (Phyu and Khin*,* 2008). Three criteria that should be emphasized are Avalanche Effect, Completeness, and Strong S-Box. All of these criteria differentiate between a strong S-Box and a weak S-Box.

To evaluate S-Boxes that are used in any cryptography algorithm, we have developed an evaluation tool named CSM S-Box Evaluation Tool (CSET) which will examine six properties of S-Box as explained above. This application will assist cryptography analyst to evaluate the strength of S-Box and facilitate them to produce the evaluation report. Report templates for each test module are prepared to help the cryptography analyst or user in producing a complete analysis report. The report templates contain each tests module process flowchart, definition, result, conclusion, and reference. Cryptography analysts simply need to provide an S-Box file, name of the S-Box algorithm and length of the S-Box. Within a short period of time, the S-Box evaluation analysis report is ready to be presented.

With the presence of this tool, process of evaluating an S-Box will be simplified and automated which resulted in reducing the evaluation timeframe. All of the testing, analyzing and reporting will be managed simultaneously and automated. This will results in reducing the evaluation timeframe.

2 S-BOX DESIGN PROPERTIES

There are three properties in evaluating S-Box that need to be emphasized namely (a) Nonlinearity, (b) Algebraic Degree, and (c) Differential Uniformity. The said properties will determine the strength and weakness of an S-Box.

(a) Nonlinearity (NL)

The nonlinearity of a Boolean function can be defined as the distance between the function and the set of all affine functions. In other words we can say that non-linearity is the number of bits which must be changed in the truth table of a Boolean function to reach the closest affine function. The upper bound of nonlinearity is: $NL(f) = 2^{n-1} - \left(2^{-1} * 2^{n/2}\right)$, for S-Box in $GF(2^n)$. Note that *n*, is denoted as degree. For an S-Box with $GF(2^8)$, the optimal value of NL is 120 (Iqtadar *et. al.,* 2011). Smallest value of nonlinearity parameter (NL) must be between 100 < NL ≤ 120 otherwise the S-Box is susceptible to linear cryptanalysis.

Nonlinearity feature makes an S-Box designed to be resistant to linear cryptanalysis (Howard, 2011). This was done by minimizing the correlation between linear transformations of input/output bits, and at the same time minimizing the difference propagation probability. After applying all possible input values and examining the corresponding output values, the number of cases that holds true is finally observed.

(b) Algebraic Degree (AD)

High algebraic degree is a property of good S-Box. Consider a function $f\{0,1\}^r \rightarrow \{0,1\}^r$, where *r* denotes the degree. If *f* is a permutation, then the algebraic degree (*AD*) of any output bit as a function of the input bits is at most *r*-1. The higher the algebraic degree value is, the better the S-Box would be. Preferable measurement of $AD \geq 4$ is suggested in (Selcuk and Melek, 2001) in order to resist higher order differential cryptanalysis.

(c) Differential Uniformity (DU)

Difference uniformity table of an S-Box gives information about the security of the block cipher against differential cryptanalysis (Howard, 2011). The essence of a differential attack is that it exploits particular high-valued entries in the table. A necessary condition for an S-Box to be immune to differential cryptanalysis is, it does not have large values entries.

This test examines the difference pairs of input and output of an S-Box. A complete data for an S-Box is compiled in an XOR table (difference uniformity table). Each element of the table represents the number of occurrences of output difference value corresponding to input difference value which *DU* indicates the highest value. A good S-Box would have *DU* value in range of $2 \leq DU \leq 6$ otherwise the S-Box is susceptible to differential cryptanalysis (Selcuk and Melek, 2001).

3 STRICT AVALANCHE CRITERION

Other testing that has been implemented towards S-Box are based on new techniques for Strict Avalanche Criterion (SAC) that was proposed by Phyu Phyu Mar and Khin Maung Latt (Phyu and Khin, 2008). The process highlighted three main properties which are Avalanche Effect, Completeness, and Strong Function. Definitions of each property are described as follows:-

(a) Avalanche Effect

A function exhibits the avalanche effect if and only if an average of one half of the output bits changes whenever a single input bit is complemented.

Frequency Analysis of Various Hamming Weight is used to determine if the S-Box meets the property of Avalanche Effect (Phyu and Khin, 2008). The objective of this process is to observe total number of bit changes in each output. Output values of the algorithm which correspond to two inputs were chosen. Apply XOR function to compute the differential value of these two outputs and find the hamming weight in the differential value. For necessary count of testing, repeat above steps. The frequency of various differential values was analyzed by counting the number of '1's in each output. If the frequency of testing result graph shows normal distribution shape (bell shape), the S-Box satisfies the avalanche effect property.

(b) Completeness

A function is complete if and only if each output bit depends on all of the input bits. Thus, if it is possible to find the simplest Boolean expression for each output bit in terms of the input bits, each of these expressions would have to contain all of the input bits if the function is complete.

Frequency Analysis of Various Differential Value is used to determine if the S-Box meets the property of Completeness (Phyu and Khin, 2008). The objective of this process is to observe the value of each output. First, two inputs with their corresponding output values of the algorithm were chosen. Next, the differential value of these two outputs was computed by applying XOR function. Then, the hamming weight in the differential value of the outputs was determined. Above steps were repeated for necessary count of testing. The frequencies of various differential values were analyzed by counting the total number of occurrence for each output values. If the frequency of testing result graph shows uniform distribution shape, the S-Box satisfies the completeness property.

(c) Strong S-Box

An S-Box is considered strong if and only if each of its output bits should change with a probability of one half whenever a single input bit is complemented.

Analysis of Various Hamming Weight According to The Bit Position is used to determine if the S-Box meets the property of Strong S-Box (Phyu and Khin, 2008). The objective of this process is to observe total number of bit changes in each bit position. Two inputs and their corresponding output value of the algorithm were chosen. The differential values of these two outputs were computed by applying XOR function. The above steps were repeated for necessary count of testing. Hamming weight was analyzed according to the bit position of resulting differential values by counting total number of '1's in each bit position. If the frequency of testing result graph shows uniform distribution shape, the S-Box satisfies the strong S-Box property.

4 ALGORITHM OF CSET

CSET is a web-based application which has been developed using uBuntu 12.04 LTS; 23.5 Gb RAM; Intel Xeon CPU 5580 @3.20GHz Qual Core as the platform, Apache as the server and Ext-JS as the framework. The code of this application is written in C++ and Perl programming language.

Process flowchart of the CSET is presented in Figure 1. To run the test module, users are required to provide three inputs; 1) name of algorithm, 2) S-Box file, and 3) length of S-Box tested. There will be two properties options. The first option which is the S-Box Design Properties option will perform three tests; Nonlinearity, Algebraic Degree and Differential Uniformity, whereas the second option which is the Strict Avalanche Criterion option will perform another three tests; Avalanche Effect, Completeness and Strong S-Box. To execute this application, users are able to select either one of the two properties options or users can also select both to run at once. Results from these tests will be referred to the parameters as shown in Table 1 to determine the strength of the tested S-Box.

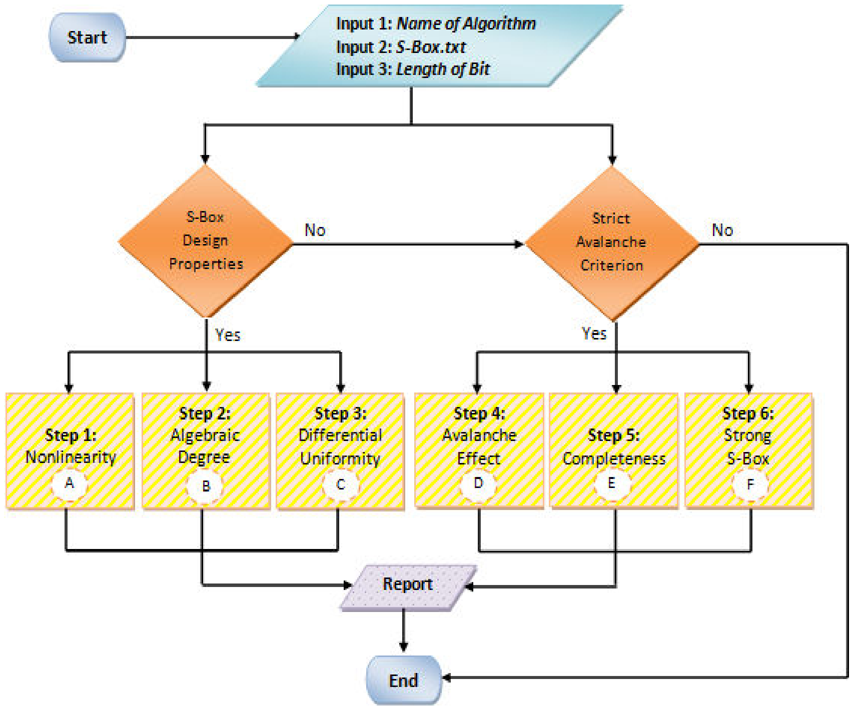
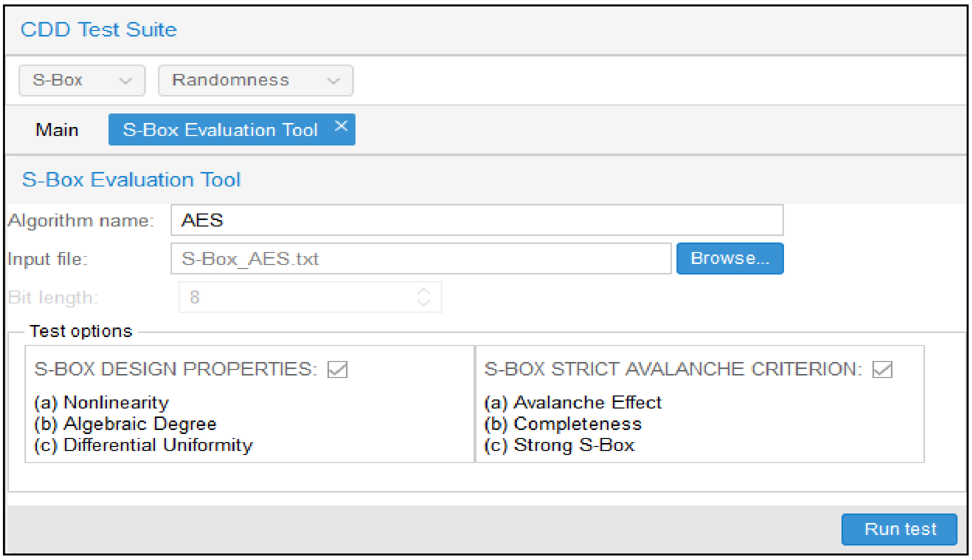


Figure 1: Process Flowchart of CSET

Table 1: S-Box Tests Description Table

Properties	**Tests Module**	**Description**	**Parameters**
S-Box Design Properties	A. Nonlinearity	To resist linear cryptanalysis	$100 \leq NL \leq 120$
	B. Algebraic Degree	To resist higher order differential cryptanalysis	$AD \geq 4$
	C. Differential Uniformity	To resist differential cryptanalysis	$2 \leq DU \leq 6$
Strict Avalanche Criterion	D. Avalanche Effect	Average of 1/2 of the output bits change whenever a single input bit is complemented	Normal Distribution
	E. Completeness	Each output bit depends on all of the input bits	Uniform Distribution
	F. Strong S-Box	Average of 1/2 of the output bits change whenever a single input bit is complemented	Uniform Distribution

Figure 2 shows the S-Box Evaluation Tool page, which contains all six tests as mention above. To conduct evaluation, user need to key in algorithm name and browse the S-Box file to be tested. The S-Box file must be in text file format which contains 256 bytes in hexadecimal representation separated by comma as shown in Figure 3. Next, user may choose either one or both test option and click **"Run Test"** to execute the tests.



Algorithm name: AES

Input file: S-Box_AES.txt

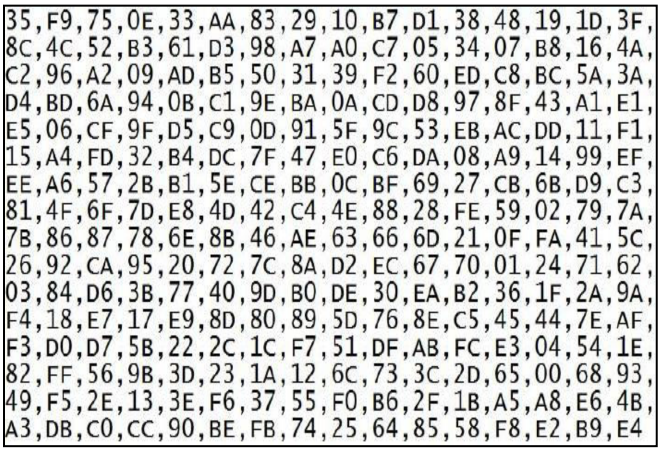
Bit length: 8

Test options

S-BOX DESIGN PROPERTIES: ☑
(a) Nonlinearity
(b) Algebraic Degree
(c) Differential Uniformity

S-BOX STRICT AVALANCHE CRITERION: ☑
(a) Avalanche Effect
(b) Completeness
(c) Strong S-Box

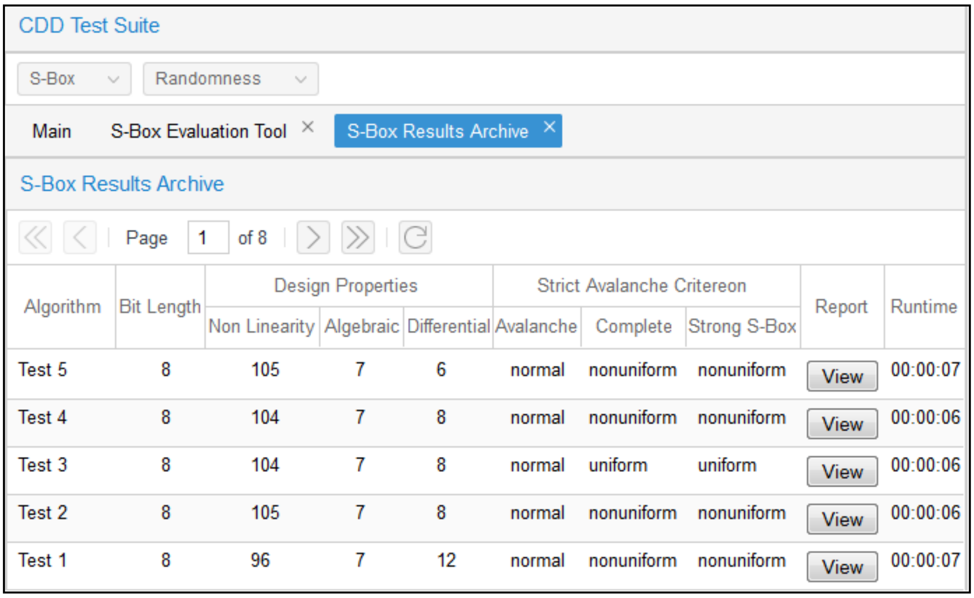
Figure 2: S-Box Evaluation Tool Page



```
35,F9,75,0E,33,AA,83,29,10,B7,D1,38,48,19,1D,3F,
8C,4C,52,B3,61,D3,98,A7,A0,C7,05,34,07,B8,16,4A,
C2,96,A2,09,AD,B5,50,31,39,F2,60,ED,C8,BC,5A,3A,
D4,BD,6A,94,0B,C1,9E,BA,0A,CD,D8,97,8F,43,A1,E1,
E5,06,CF,9F,D5,C9,0D,91,5F,9C,53,EB,AC,DD,11,F1,
15,A4,FD,32,B4,DC,7F,47,E0,C6,DA,08,A9,14,99,EF,
EE,A6,57,2B,B1,5E,CE,BB,0C,BF,69,27,CB,6B,D9,C3,
81,4F,6F,7D,E8,4D,42,C4,4E,88,28,FE,59,02,79,7A,
7B,86,87,78,6E,8B,46,AE,63,66,6D,21,0F,FA,41,5C,
26,92,CA,95,20,72,7C,8A,D2,EC,67,70,01,24,71,62,
03,84,D6,3B,77,40,9D,B0,DE,30,EA,B2,36,1F,2A,9A,
F4,18,E7,17,E9,8D,80,89,5D,76,8E,C5,45,44,7E,AF,
F3,D0,D7,5B,22,2C,1C,F7,51,DF,AB,FC,E3,04,54,1E,
82,FF,56,9B,3D,23,1A,12,6C,73,3C,2D,65,00,68,93,
49,F5,2E,13,3E,F6,37,55,F0,B6,2F,1B,A5,A8,E6,4B,
A3,DB,C0,CC,90,BE,FB,74,25,64,85,58,F8,E2,B9,E4
```

Figure 3: S-Box File

Upon completion, S-Box Results Archive page will appear as shown in Figure 4. This page displays all tested S-Boxes with their results that have been tested. To view the complete report, select the desired result file by clicking the view button.



Algorithm	Bit Length	Design Properties			Strict Avalanche Critereon			Report	Runtime
		Non Linearity	Algebraic	Differential	Avalanche	Complete	Strong S-Box		
Test 5	8	105	7	6	normal	nonuniform	nonuniform	View	00:00:07
Test 4	8	104	7	8	normal	nonuniform	nonuniform	View	00:00:06
Test 3	8	104	7	8	normal	uniform	uniform	View	00:00:06
Test 2	8	105	7	8	normal	nonuniform	nonuniform	View	00:00:06
Test 1	8	96	7	12	normal	nonuniform	nonuniform	View	00:00:07

Figure 4: S-Box Results Archive Page

A complete S-Box evaluation tests report is ready for user to view as shown in Figure 5(a) and Figure 5(b). The complete report generated from this application provides the definitions and process figures of each test modules selected during execution. Analysis graphs are presented for Nonlinearity, Avalanche Effect, Completeness and Strong S- Box tests. An analysis table is presented for the Differential Uniformity test. A conclusion which indicates whether the tested S-Box passes or fails the selected test is presented at the end of each test section.

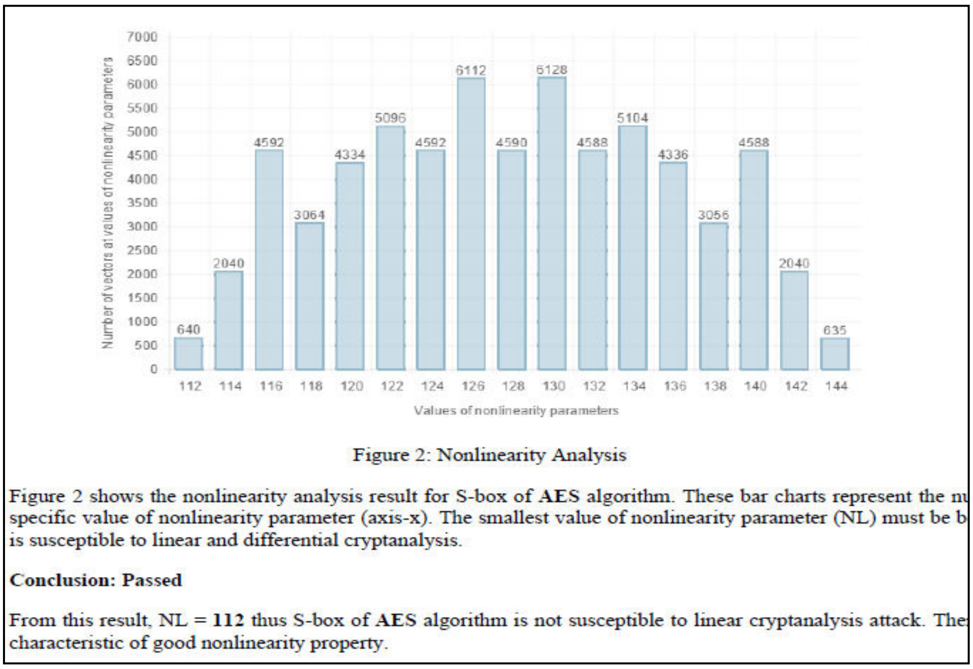


Figure 2: Nonlinearity Analysis

Figure 2 shows the nonlinearity analysis result for S-box of **AES** algorithm. These bar charts represent the nu specific value of nonlinearity parameter (axis-x). The smallest value of nonlinearity parameter (NL) must be b is susceptible to linear and differential cryptanalysis.

Conclusion: Passed

From this result, NL = **112** thus S-box of **AES** algorithm is not susceptible to linear cryptanalysis attack. The characteristic of good nonlinearity property.

Figure 5(a): S-Box Evaluation Report Page

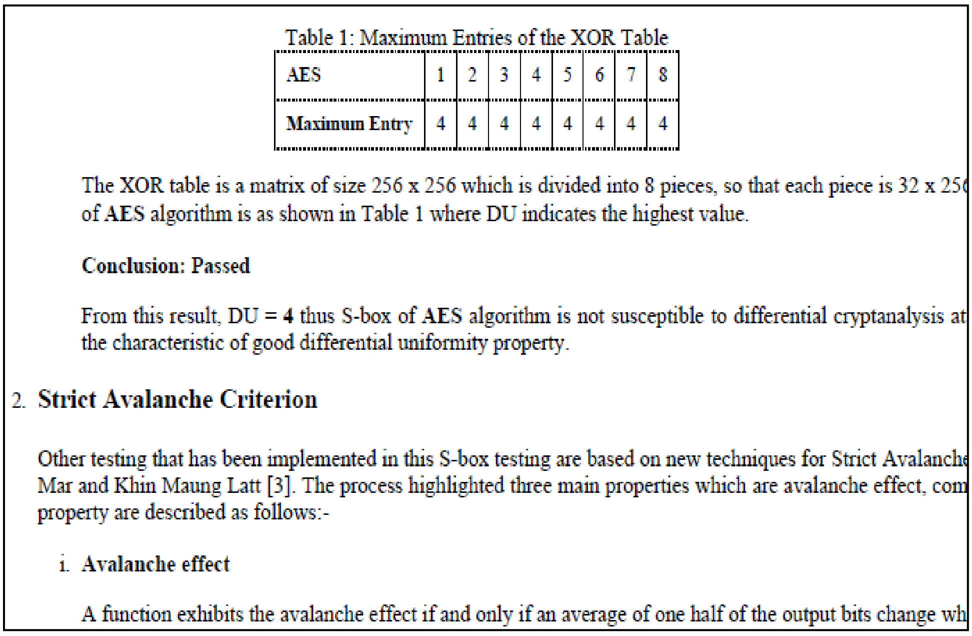


Table 1: Maximum Entries of the XOR Table

AES	1	2	3	4	5	6	7	8
Maximum Entry	4	4	4	4	4	4	4	4

The XOR table is a matrix of size 256 x 256 which is divided into 8 pieces, so that each piece is 32 x 25
of AES algorithm is as shown in Table 1 where DU indicates the highest value.

Conclusion: Passed

From this result, DU = **4** thus S-box of AES algorithm is not susceptible to differential cryptanalysis at
the characteristic of good differential uniformity property.

2. **Strict Avalanche Criterion**

Other testing that has been implemented in this S-box testing are based on new techniques for Strict Avalanch
Mar and Khin Maung Latt [3]. The process highlighted three main properties which are avalanche effect, com
property are described as follows:-

i. **Avalanche effect**

A function exhibits the avalanche effect if and only if an average of one half of the output bits change wh

Figure 5(b): S-Box Evaluation Report Page

Table 2 shows the comparison of execution time for evaluating S-Boxes with using CSET and without using CSET. It takes about 7 seconds to produce a complete analysis report compared to 5423 seconds (90.38minutes) without using the tool. In general, to manually produce a complete analysis report takes over an hour and a half. With the presence of this tool, process of evaluating an S-box will be simplified and automated which resulted in reducing the evaluation timeframe period significantly.

Table 2 : Comparison of Execution Time with and without CSET (in second)

Tests Module	Run Test		Analyse Result		Report Writing		Total	
	Without Tool	With Tool	Without Tool	With Tool	Without Tool	With Tool	Without Tool	With Tool
1.Nonlinearity	5	3	1	2	900	2	906	7
2.Algebraic Degree	1		1		900		902	
3.Differential Uniformity	5		1		900		906	
4.Avalance Effect	2		1		900		903	
5.Completeness	2		1		900		903	
6.Strong S-Box	2		1		900		903	
							5423	**7**

5 S-BOXES EVALUATION USING CSET

For testing purposes, eight different S-Boxes have been evaluated using CSET. These S-Boxes are used in AES, AES-RC4 (Abd-ElGhafar *et al.,* 2009), SKIPJACK (Hussain, 2011), Whirlpool (Stallings, 2006), and Camellia (Aoki, 2000) algorithms. Results of each evaluated S-Boxes using CSET are summarized in Table 3. Some of the results have been compared with results presented in the research papers mentioned above.

Table 3: Example of S-Box Evaluation Tests Using CSET

Algorithm Name	**S-Box Design Properties**			**Strict Avalanche Criterion**		
	Nonlinearity	**Algebraic Degree**	**Differential Uniformity**	**Avalanche Effect**	**Completeness**	**Strong S-Box**
Requirement of strong S-Box	$100 \leq NL \leq 120$	$AD \geq 4$	$2 \leq DU \leq 6$	Normal	Uniform	Uniform
AES	112	7	4	Normal	Uniform	Uniform
AES-RC4	89	7	10	Normal	Non-uniform	Non-uniform
SKIPJACK	98	7	10	Normal	Non-uniform	Non-uniform
Whirlpool	95	7	10	Normal	Non-uniform	Non-uniform
Camellia_1	111	7	4	Normal	Non-uniform	Non-uniform
Camellia_2	112	7	4	Normal	Uniform	Uniform
Camellia_3	111	7	4	Normal	Non-uniform	Non-uniform
Camelia_4	110	7	6	Normal	Non-uniform	Non-uniform

6 CONCLUSION

Normally, new cryptographic algorithm developers will consider well researched and widely known S-Boxes in their designs. However, sometimes there is a need to use new unused S-Boxes to cater for other S-Box sizes. In the process of designing new S-Box, this CSM S-Box Evaluation Tool will be very useful for cryptographic researches to ensure the strength of the new S-Box is within the desired requirement. It tests whether an S-Box has meets the characteristics of S-Box Design Properties and Strict Avalanche Criterion Properties that determine the strength of an S-Box. Furthermore, S-Box evaluation timeframe is reduced as this application will run all tests simultaneously. Analysis of tests results with a complete report and conclusion will also be generated automatically.

For further development of CSET, we are planning to add more relevant tests which will cover a wider aspect of cryptographic properties. We are also considering enhancing this tool by including variety of S-Box sizes to be tested.

REFERENCES

Abd-ElGhafar, I. *et al.* 2009. Generation of AES Key Dependent S-Boxes using RC4 Algorithm. *13th International Conference on Aerospace Sciences & Aviation Technology, ASAT-13.*

Aoki, K. *et al.* 2000. Generation of AES Key Dependent S-Boxes using RC4 Algorithm. *NTT and Mitsubishi Electric Corporation.*

Bart, P. 2006. Cryptographic Properties Of Boolean Functions And S-Boxes. Sourced from https://www.cosic.esat.kuleuven.be/publications/thesis-129.pdf

Howard, M. H. 2011. A Tutorial on Linear and Differential Cryptanalysis. Sourced from http://www.engr.mun.ca/~howard/PAPERS/ldc_tutorial.pdf

Hussain, *et al.* 2011. Analyses of SKIPJACK S-Box. *World Applied Sciences Journal.* 13(11):2385-2388.

Iqtadar, H., Tariq, S., Muhammad, A. G. and Waqar, A. K. 2011. Construction of Cryptographically Strong 8x8 S-boxes. *World Applied Sciences Journal.* **13**(11): 2389-2395.

Phyu, P. M. and Khin, M. L. 2008. New Analysis Methods on Strict Avalanche Criterion of S-Boxes. *World Academy of Science, Engineering and Technology*. **24**: 150-154.

Selcuk, K. and Melek, D. Y. 2001. On Some Cryptographic Properties of Rijndael. I*nformation Assurance in Computer Networks, LNCS*. **2052:** 300-311.

Stallings, W. 2006. The Whirlpool Secure Hash Function. *Cryptologia.* 30:55-67.

High Capacity Data Embedding Method with LSB and PVD Shift

Mehdi Hussain^{*1,2}, **Ainuddin Wahid Abdul Wahab**[1], **Yamani Idna Bin Idris**[1], **Noman Javed**[3], and **Rosli Salleh**[1]

[1]*Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia*
[2]*School of Electrical Engineering and Computer Science, National University of Science and Technology, 44000 Islamabad, Pakistan*
[3]*Department of Computer Science, Namal College Mianwali- 42250, (An associate college of University of Bradford UK), Pakistan*

E-mail: mehdi141@hotmail.com, ainuddin@um.edu.my
**Corresponding author*

ABSTRACT

In this paper, we proposed a high capacity data embedding method to improve the Yang et al.s method. The proposed method is based on the least significant bit (LSB), pixel value difference (PVD) and PVD shift methods. In this method, the secret data can be embedded in the smooth area by LSB embedding and on edge area by PVD embedding. Furthermore, PVD shift process is applied to increase the embedding capacity without sacrificing the visual quality of stego-images. The experimental results show that the propose method enhanced the embedding capacity up to $15,000$ bits while retained good visual quality (above $38dB$) and proved its security against RS steganalysis detection attacks.

Keywords: Data hiding, Steganography, LSB, PVD, PVD shift, Hybrid Steganography

1 INTRODUCTION

In the modern computer world, digital data can be easily copied, destroyed and altered by an unauthorized user. Therefore, the information hiding has a vital and significant role in data protection. Steganography is one of the branches of information hiding; which deals to conceal the secret information in the innocent carrier to provide data security. This innocent carrier can be an image, video and audio (Subhedar and Mankar, 2014). Image steganography algorithms gained a lot of attention in the last decade due to ease in camera devices connected to the Internet. Image steganography can be categorized into the spatial domain and frequency domain methods (Hussain and Hussain, 2013, Subhedar and Mankar, 2014). An image steganography method is evaluated by its three parameters, high payload (hidden capacity), imperceptibility (visual quality) and security (undetectability).

In literature, there are many image steganography methods proposed to improve the hidden capacity, visual quality and security (Balasubramanian et al., 2014, Chan and Cheng, 2004, Hussain and Hussain, 2010, 2011a,b, Jung, 2010, Jung et al., 2008, Jung and Yoo, 2015, Karim and Wong, 2015, Muhammad et al., 2015, Shen et al., 2015, Tsai et al., 2014, Wu and Tsai, 2003, Wu et al., 2005, Yang et al., 2010). LSB is one of the most common and fundamental image steganographic method (Chan and Cheng, 2004). Secret data bits are embedded into the least bits of pixels. More secret bits for embedding needs more least bits modification in a pixel. The LSB based methods can easily be exposed by statistical analysis (Fridrich et al., 2001). PVD method introduced by Wu and Tsai (2003), it hides secret data inside the pixel differences. It readjusts the difference between two consecutive pixels for secret data embedding. PVD based method improves the visual quality but suffers from low payload against LSB based methods and can be detected by histogram analysis (Hussain et al., 2015).

To improve the hidden capacity, the combination of PVD with other existing methods was explored in (Jung et al., 2008, Shen et al., 2015, Tsai et al., 2014, Wu et al., 2005). In (Wu et al., 2005) combined the LSB with PVD method, it divided the cover image into two levels, i.e. smooth (lower) level, and edgy (higher) level. LSB and PVD were applied on smooth and edgy levels respectively. This method improved the hidden capacity against Wu and Tsai (2003) method and exposed by RS steganalysis (Fridrich et al., 2001) method. In Yang et al. (2010) improved the Wu et al. (2005) method by introducing the lower level strategy. This lower level strategy improved the visual quality and undetectability against RS steganalysis while this method was unable to improve the hidden capacity.

There is still room to improve the embedding capacity without distortion. In this paper, a new data hiding method is proposed to improve Yang et al. (2010) scheme, where a high embedding capacity is a motive with the similar visual quality for both the human visual system and peak signal to noise ratio (PSNR). The proposed method utilized the lower and higher level concept to embed the secret data using LSB and PVD shift methods. Furthermore, an iterative PVD shift process applied to PVD based stego-pixels for embedding one extra secret bit between pixels pair block.

The rest of this paper is structured as follows. In Section 2, Yang et al. (2010) method is reviewed and the proposed method is described in Section 3. The experimental results are to be presented and discussed in Section 4. Finally, the conclusions are outlined in Section 5.

2 RELATED WORK

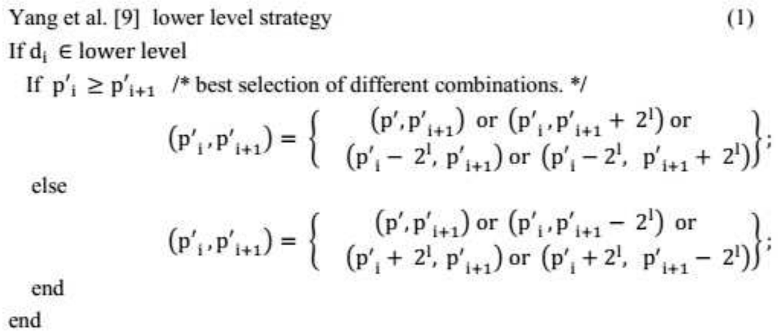
2.1 Yang et al's. Method

The Yang et al. (2010) extends the Wu et al. Wu et al. (2005) data embedding method by introducing the lower level readjustment strategy to improve the visual quality of the stego-

	Lower Levels		Higher Levels			
Lower-Upper bound	R1 ϵ [0,7]	R2 ϵ [8,15]	R3 ϵ [16,31]	R4 ϵ [32,63]	R5 ϵ [64,127]	R6 ϵ [128,255]

Table 1: Range table division for higher and lower level pixels blocks.

image. The lower level readjustment strategy is described in Eq. (1). The Yang et al. (2010) method combined LSB with PVD methods and applied the similar data embedding procedure of Wu et al. (2005) method. In this data embedding process, the cover image is divided into two non-overlapped pixels blocks. These pixels block categorized by lower and higher levels based on pixels pair differences (see range Table 1).



```
Yang et al. [9]  lower level strategy                                                   (1)
If $d_i \in$ lower level
  If $p'_i \geq p'_{i+1}$  /* best selection of different combinations. */
      $(p'_i, p'_{i+1}) = \left\{ \begin{array}{c} (p', p'_{i+1}) \text{ or } (p'_i, p'_{i+1} + 2^l) \text{ or} \\ (p'_i - 2^l, p'_{i+1}) \text{ or } (p'_i - 2^l, p'_{i+1} + 2^l) \end{array} \right\};$
   else
      $(p'_i, p'_{i+1}) = \left\{ \begin{array}{c} (p', p'_{i+1}) \text{ or } (p'_i, p'_{i+1} - 2^l) \text{ or} \\ (p'_i + 2^l, p'_{i+1}) \text{ or } (p'_i + 2^l, p'_{i+1} - 2^l) \end{array} \right\};$
   end
end
```

Let $pi, p_i + 1$ are the pixels of the ith block and its difference is calculated as $d_i = p_i + 1 - p_i$. If the difference d_i exists in the lower level of Table 1 then 3-bit LSB embedding is applied. The output of 3-bit LSB stego-pixels difference must remain into the lower level range of Table 1, otherwise, lower level readjustment strategy of Eq.(1) is applied to retain the range consistency of stego-pixels. For higher level difference d_i, Wu and Tsai (2003) PVD method is employed to hide secret data in the pixels block. First locate the R_k which d_i belongs to and further estimate the number of t secret bits using upper and lower bounds i.e. for $R_k = 4$, the upper and lower bounds are 63 and 32, the $t = log2(63 - 31 + 1) = 5$. Further, t secret bits read and converted to decimal as b. For calculating new difference $d`_i$ and stego-pixels $p`_i, p`_i + 1$ Eq.(2) and Eq.(3) are applied respectively. Repeat the embedding procedure for all lower and higher level pixels blocks of the cover image.

$$d_i{}' = \begin{cases} l_k + b & for\ d_i \geq 0 \\ -(l_k + b) & for\ d_i < 0 \end{cases} \tag{2}$$

$$(p'_i, p'_{i+1}) = \begin{cases} \left(p_i - \left\lceil \frac{d_i{}' - d_i}{2} \right\rceil,\ p_{i+1} + \left\lfloor \frac{d_i{}' - d_i}{2} \right\rfloor\right) & d\ \in odd \\ \left(p_i - \left\lfloor \frac{d_i{}' - d_i}{2} \right\rfloor,\ p_{i+1} + \left\lceil \frac{d_i{}' - d_i}{2} \right\rceil\right) & d\ \in even \end{cases} \tag{3}$$

In the data extraction process, a similar embedding procedure is employed for calculating the difference between each block pixels pair. Furthermore, the selection of lower and higher levels ranges computed by above pixels block difference. For lower level 3-bit LSB extraction method is applied while for higher level PVD extraction Wu and Tsai (2003) method is employed to recover the secret bits using Eq.(4). Further, convert the b' into the binary stream for recovering the secret bits.

$$b' = d_i' - l_k \tag{4}$$

3 PROPOSED METHOD

In this section, the proposed data hiding method is presented. The LSB and PVD shift methods are used for data embedding in the pixels blocks. Embedding and extracting methods will be illustrated in details (3.1 and 3.2) respectively. The proposed method can increase the embedding capacity by using two pixels for each block and can also keep a good image visual quality with undetectability against RS steganalysis.

3.1 Embedding Algorithm

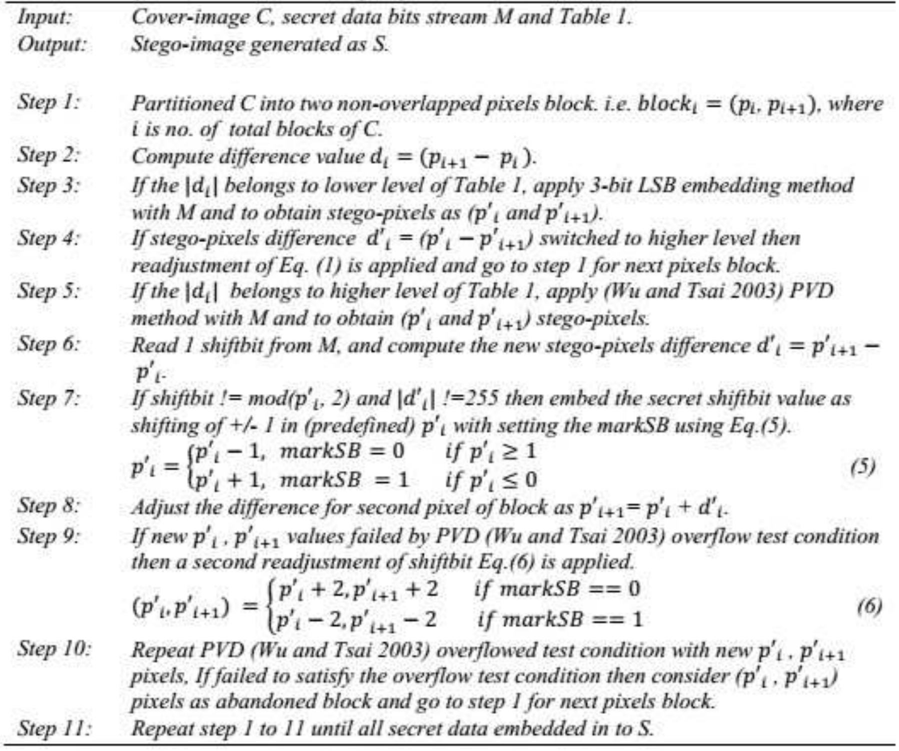
In the data embedding algorithm, the cover image is divided into two non-overlapping pair of pixels block. The difference between pixels pair is categorized into the lower level and higher level. This lower level and higher level (Table 1) indicate the smooth and edge area of the cover image. The 3-bit LSB is applied on the lower level and PVD is employed by higher level pixel block. Furthermore, an iterative PVD shift method is applied by +/- 1 on predefined selected pixel to increase the embedding capacity. The detailed embedding algorithm for secret data is described in Table 2 and block diagram is shown in Figure 1.

An example of embedding process is shown in Figure 2. Given the block with two pixels, $(p_i, p_i + 1) = (148, 184)$ for higher level. The difference value $d_i = 36$ is obtained. The number of secret bits are calculated, $log2(upper - lower) + 1 = 5 = log2(63 - 31 + 1)$. Transform the 5 secret bits into decimal $(01100)2 = (12)10$ and adjust the difference by Wu and Tsai (2003) PVD embedding method to obtain new stego-pixel $(144, 188)$. The $p_i = p_i + 1, 145 = 144 + 1$ can be calculated since a $shiftbit = 1$ is given. Furthermore, adjusted the difference for $p_i + 1 = 145 + (188 - 144) = 189$. Finally, the step 9 condition satisfied to retain $(p`_i, p`_i + 1) = (145, 189)$.

3.2 Extracting Algorithm

The proposed method can fully extract the secret data from the stego-image. The extraction method is similar to embedding process. First partitioned the cover image into two non-overlapping pixels blocks and calculates the difference between pixels pair. The difference value decides the extraction method, lower level difference value block targeted by LSB extraction method. On the other side, higher level difference value stego-pixels blocks recovered its secret data by PVD extraction method. Furthermore, after PVD extraction, extra shift bit can be extracted by predefined selected pixel from stego-pixels block. The detailed extracting steps are stated in Table 3.

Table 2: Proposed method embedding steps



Input:	*Cover-image C, secret data bits stream M and Table 1.*
Output:	*Stego-image generated as S.*
Step 1:	*Partitioned C into two non-overlapped pixels block. i.e.* $\mathrm{block}_i = (p_i, p_{i+1})$, *where* i *is no. of total blocks of C.*
Step 2:	*Compute difference value* $d_i = (p_{i+1} - p_i)$.
Step 3:	*If the* $\lvert d_i \rvert$ *belongs to lower level of Table 1, apply 3-bit LSB embedding method with M and to obtain stego-pixels as (*p'_i *and* p'_{i+1}*).*
Step 4:	*If stego-pixels difference* $d'_i = (p'_i - p'_{i+1})$ *switched to higher level then readjustment of Eq. (1) is applied and go to step 1 for next pixels block.*
Step 5:	*If the* $\lvert d_i \rvert$ *belongs to higher level of Table 1, apply (Wu and Tsai 2003) PVD method with M and to obtain (*p'_i *and* p'_{i+1}*) stego-pixels.*
Step 6:	*Read 1 shiftbit from M, and compute the new stego-pixels difference* $d'_i = p'_{i+1} - p'_i$.
Step 7:	*If shiftbit != mod(*p'_i*, 2) and* $\lvert d'_i \rvert$ *!=255 then embed the secret shiftbit value as shifting of +/- 1 in (predefined)* p'_i *with setting the markSB using Eq.(5).* $p'_i = \begin{cases} p'_i - 1, & markSB = 0 \quad if\ p'_i \geq 1 \\ p'_i + 1, & markSB = 1 \quad if\ p'_i \leq 0 \end{cases} \qquad (5)$
Step 8:	*Adjust the difference for second pixel of block as* $p'_{i+1} = p'_i + d'_i$.
Step 9:	*If new* p'_i , p'_{i+1} *values failed by PVD (Wu and Tsai 2003) overflow test condition then a second readjustment of shiftbit Eq.(6) is applied.* $(p'_i, p'_{i+1}) = \begin{cases} p'_i + 2, p'_{i+1} + 2 & if\ markSB == 0 \\ p'_i - 2, p'_{i+1} - 2 & if\ markSB == 1 \end{cases} \qquad (6)$
Step 10:	*Repeat PVD (Wu and Tsai 2003) overflowed test condition with new* p'_i , p'_{i+1} *pixels, If failed to satisfy the overflow test condition then consider (*p'_i , p'_{i+1}*) pixels as abandoned block and go to step 1 for next pixels block.*
Step 11:	*Repeat step 1 to 11 until all secret data embedded in to S.*

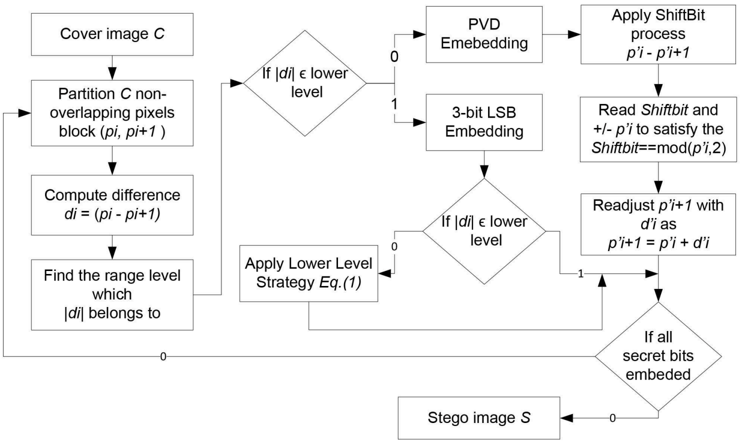

Cover image C
Partition C non-overlapping pixels block (pi, pi+1)
Compute difference di = (pi - pi+1)
Find the range level which |di| belongs to
If |di| ϵ lower level
0
1
PVD Emebedding
3-bit LSB Embedding
Apply ShiftBit process p'i - p'i+1
Read Shiftbit and +/- p'i to satisfy the Shiftbit==mod(p'i,2)
Readjust p'i+1 with d'i as p'i+1 = p'i + d'i
If |di| ϵ lower level
0
1
Apply Lower Level Strategy Eq.(1)
If all secret bits embeded
0
0
Stego image S

Figure 1: Proposed embedding block diagram.

4 EXPERIMENTAL RESULTS

In this section, the experimental results of the proposed method are evaluated on 512x512 grayscale images as shown in Figure 3. The standard images of Lena, Baboon, Pepper and etc. are used as cover images and the pseudo-random numbers generator is used to generate secret data for embedding.

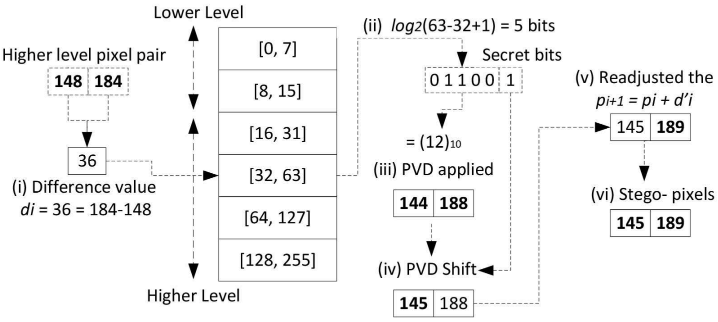


Figure 2: An example of the proposed embedding procedure.

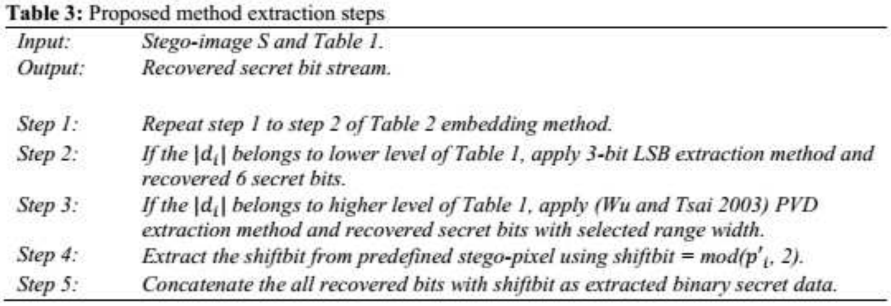


Table 3: Proposed method extraction steps

Input:	*Stego-image S and Table 1.*
Output:	*Recovered secret bit stream.*
Step 1:	*Repeat step 1 to step 2 of Table 2 embedding method.*
Step 2:	*If the $\lvert d_i \rvert$ belongs to lower level of Table 1, apply 3-bit LSB extraction method and recovered 6 secret bits.*
Step 3:	*If the $\lvert d_i \rvert$ belongs to higher level of Table 1, apply (Wu and Tsai 2003) PVD extraction method and recovered secret bits with selected range width.*
Step 4:	*Extract the shiftbit from predefined stego-pixel using shiftbit = mod(p'_i, 2).*
Step 5:	*Concatenate the all recovered bits with shiftbit as extracted binary secret data.*

(a) Lena (b) Baboon (c) Peppers (d) Jet
(e) Tank (f) Airplane (g) Truck (h) Elaine
(i) Couple (j) Boat (k) Tiffany (l) Lake

Figure 3: Cover images for experiments.

4.1 Embedding Capacity and Visual Quality Analysis

The embedding capacity of proposed method is calculated by two aspects. First, $embeddingcapacity(bits)$, the total number of secret bits hidden inside the stego-image. Second, the bit rate, $bitsperpixel(bpp)$ using Eq. (7), where W and H are the number of columns and rows of the cover image. For measuring the visual quality ratio, the PSNR is used as in Eq. (8). The MSE is the mean square error that is defined as Eq. (9), where $p`_i - p_i$ are the pixel value of cover and stego-pixels.

$$bpp = \frac{Embedding\ Capacity}{W \times H} \tag{7}$$

$$PSNR = 10\log_{10} \frac{255^2}{MSE} \tag{8}$$

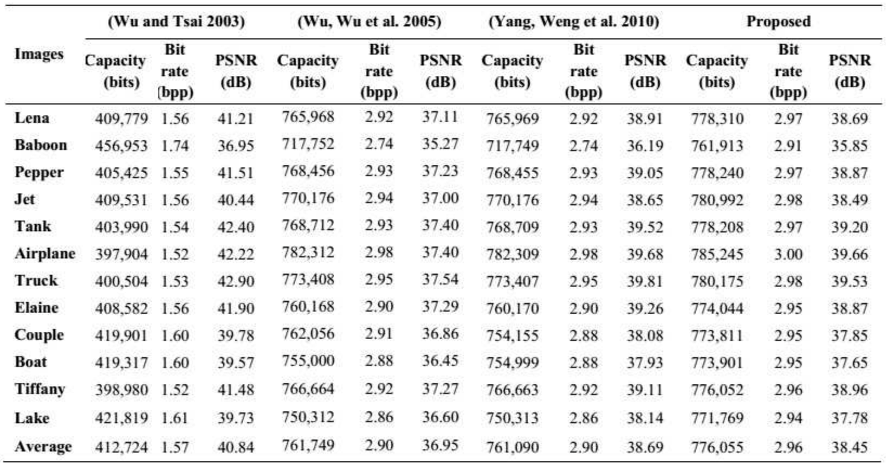
$$MSE = \sum_{i=1}^{W x H} \frac{\left(p'_i - p_i\right)^2}{W \times H} \tag{9}$$

The experimental results of the proposed method are shown in Table 4, where the embedding capacity is higher than existing methods. It has $761,913$ to $785,245$ bits and the PSNR has $35.85dB$ to $39.66dB$. Hidden capacity of Wu and Tsai (2003) method is $412,724$ bits on average. The average embedding capacity of Wu et al. (2005) and Yang et al. (2010) methods are same $761,749$ bits, where Yang et al. (2010) had better PSNR against Wu et al. (2005) PSNR. Furthermore, the average hidden capacity of the proposed method is higher around $776,055$ bits.

The proposed method average PSNR is similar to Yang et al. (2010) method but $2dB$ less than Wu and Tsai (2003) method. Generally, if the PSNR is higher than $30dB$ (the standard measurement), it is imperceptible to the human visual system, while the proposed method has average $38.45dB$ PSNR. However, the visual distortion of $2dB$ in PSNR is still acceptable, because the proposed method achieved more than $363,331$ hidden bits, if it directly compared to Wu and Tsai (2003) method. The results show that the proposed method can hide $14,304$ bits more than Wu et al. (2005) and Yang et al. (2010) method.

The improvements of our proposed method have a larger embedding capacity with a similar image quality over Wu et al. (2005) and Yang et al. (2010) methods. As the result, the comparison of hidden capacity of each image can be seen in Figure 4, and the visual quality of PSNR for all stego-images is shown in Figure 5.

Table 4: Experimental results of the proposed method with existing LSB+PVD based methods.



Images	(Wu and Tsai 2003)			(Wu, Wu et al. 2005)			(Yang, Weng et al. 2010)			Proposed		
	Capacity (bits)	Bit rate (bpp)	PSNR (dB)	Capacity (bits)	Bit rate (bpp)	PSNR (dB)	Capacity (bits)	Bit rate (bpp)	PSNR (dB)	Capacity (bits)	Bit rate (bpp)	PSNR (dB)
Lena	409,779	1.56	41.21	765,968	2.92	37.11	765,969	2.92	38.91	778,310	2.97	38.69
Baboon	456,953	1.74	36.95	717,752	2.74	35.27	717,749	2.74	36.19	761,913	2.91	35.85
Pepper	405,425	1.55	41.51	768,456	2.93	37.23	768,455	2.93	39.05	778,240	2.97	38.87
Jet	409,531	1.56	40.44	770,176	2.94	37.00	770,176	2.94	38.65	780,992	2.98	38.49
Tank	403,990	1.54	42.40	768,712	2.93	37.40	768,709	2.93	39.52	778,208	2.97	39.20
Airplane	397,904	1.52	42.22	782,312	2.98	37.40	782,309	2.98	39.68	785,245	3.00	39.66
Truck	400,504	1.53	42.90	773,408	2.95	37.54	773,407	2.95	39.81	780,175	2.98	39.53
Elaine	408,582	1.56	41.90	760,168	2.90	37.29	760,170	2.90	39.26	774,044	2.95	38.87
Couple	419,901	1.60	39.78	762,056	2.91	36.86	754,155	2.88	38.08	773,811	2.95	37.85
Boat	419,317	1.60	39.57	755,000	2.88	36.45	754,999	2.88	37.93	773,901	2.95	37.65
Tiffany	398,980	1.52	41.48	766,664	2.92	37.27	766,663	2.92	39.11	776,052	2.96	38.96
Lake	421,819	1.61	39.73	750,312	2.86	36.60	750,313	2.86	38.14	771,769	2.94	37.78
Average	412,724	1.57	40.84	761,749	2.90	36.95	761,090	2.90	38.69	776,055	2.96	38.45

4.2 Proposed method Security under RS-Steganalysis

In this section, the security of proposed method is measured by well-known RS steganalysis method. RS steganalysis is consist of discrimination function (DF) with M and $-M$ as flipping

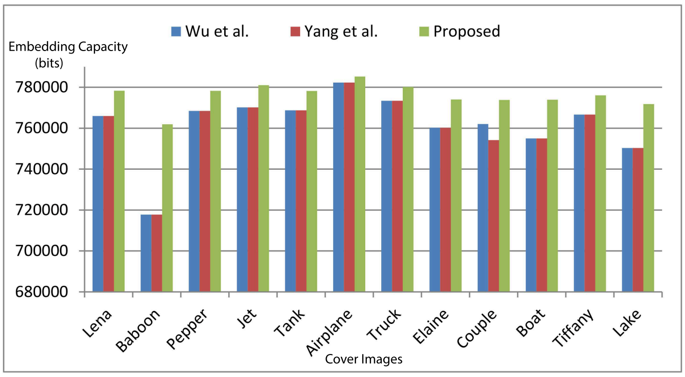


Figure 4: The comparison of embedding capacity.

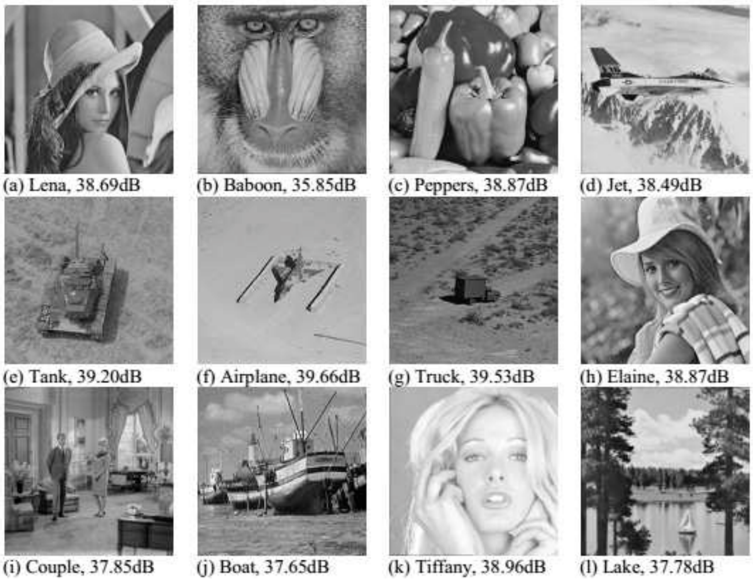
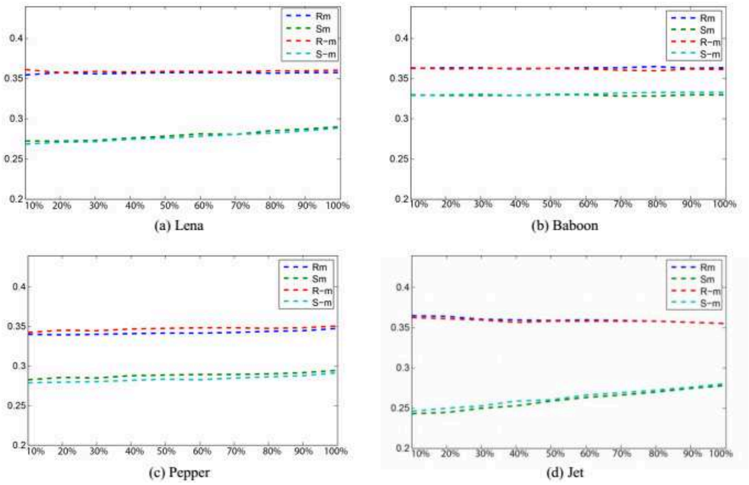

(a) Lena, 38.69dB (b) Baboon, 35.85dB (c) Peppers, 38.87dB (d) Jet, 38.49dB (e) Tank, 39.20dB (f) Airplane, 39.66dB (g) Truck, 39.53dB (h) Elaine, 38.87dB (i) Couple, 37.85dB (j) Boat, 37.65dB (k) Tiffany, 38.96dB (l) Lake, 37.78dB

Figure 5: Proposed method Stego-images.

masks. The singular and regular parameters i.e. R_M, R_{-M}, S_M and S_{-M} are used to find the magnitude of each pixel block using DF function. If both singular and regular parameters satisfy the $R_M \approx R_{-M} > S_M \approx S_{-M}$ conditions it indicates that the image is free of hidden data otherwise image has secret data. For experiments, the mask $M = [0110]$ and $M[0-1-10]$ used to apply RS analysis.

The RS analysis graph for first four Lena, Baboon, Pepper and Jet stego-images are shown in Figure 6. The x-axis showed the increasing embedding capacity from 10% to 100%. Similarly y-axis indicated the percentages of regular (R_M, R_{-M}) and singular (S_M, S_{-M}) parameters. As we can observe, in Lena images the differences between regular and singular groups are close to each other and it showed that RS steganalysis method was unable to detect the hidden data inside the stego-image. Similarly, in all other graphs i.e. Baboon, Pepper and Jet showed the closest singular and regular curves and was unable to detect hidden data by RS-steganalysis. This indicates that proposed method is secure under RS steganalysis detection attacks.



(a) Lena

(b) Baboon

(c) Pepper

(d) Jet

Figure 6: The RS steganalysis graph of proposed method.

5 CONCLUSIONS

In this paper, a high capacity data hiding method based on LSB, PVD and PVD shift has been proposed to improve the hidden capacity without degrading the visual quality of stego-image. The proposed method divided the cover image into two lower and higher levels, respectively to embed secret data by LSB and PVD embedding methods. Furthermore, PVD shift process applied to improve the hidden capacity. The experimental results demonstrated that the proposed method embedding capacity and the image quality was better than LSB and PVD based methods. The proposed method could keep the embedding capacity $776,055$ bits, $2.96bpp$, and the PSNR $38.45dB$ on average. Moreover, proposed method also had undetectability against RS steganalysis detection attacks. In the future work, proposed method could be integrated with reversible embedding approaches to further enhance the hidden capacity.

6 ACKNOWLEDGMENT

The work was supported by the National University of Science and Technology Islamabad Pakistan under faculty development program abroad [grant number 0972/F008
/HRD/FDP-14/] and the Ministry of Education, Malaysia under the University of Malaya [High Impact Research Grant UM.C/625/1/HIR/MoE/FCSIT/17].

REFERENCES

Balasubramanian, C., Selvakumar, S., and Geetha, S. (2014). High payload image steganography with reduced distortion using octonary pixel pairing scheme. *Multimedia Tools and Applications*, 73(3):2223–2245.

Chan, C.-K. and Cheng, L.-M. (2004). Hiding data in images by simple LSB substitution. *Pattern recognition*, 37(3):469–474.

Fridrich, J., Goljan, M., and Du, R. (2001). Reliable detection of lsb steganography in color and grayscale images. In *Proceedings of the 2001 workshop on Multimedia and security: new challenges*, pages 27–30. ACM.

Hussain, M. and Hussain, M. (2010). Pixel intensity based high capacity data embedding method. In *Information and Emerging Technologies (ICIET), 2010 International Conference on*, pages 1–5. IEEE.

Hussain, M. and Hussain, M. (2011a). Embedding data in edge boundaries with high PSNR. In *Emerging Technologies (ICET), 2011 7th International Conference on*, pages 1–6. IEEE.

Hussain, M. and Hussain, M. (2011b). Information hiding using edge boundaries of objects. *International Journal of Security and Its Applications*, 5(3):1–10.

Hussain, M. and Hussain, M. (2013). A survey of image steganography techniques. *International Journal of Advanced Science and Technology*, 54.

Hussain, M., Wahab, A. W. A., Anuar, N. B., Salleh, R., and Noor, R. M. (2015). Pixel value differencing steganography techniques: Analysis and open challenge. In *Consumer Electronics-Taiwan (ICCE-TW), 2015 IEEE International Conference on*, pages 21–22. IEEE.

Jung, K.-H. (2010). High-capacity steganographic method based on pixel-value differencing and LSB replacement methods. *The Imaging Science Journal*, 58(4):213–221.

Jung, K.-H., Ha, K.-J., and Yoo, K.-Y. (2008). Image data hiding method based on multi-pixel differencing and LSB substitution methods. In *Convergence and Hybrid Information Technology, 2008. ICHIT'08. International Conference on*, pages 355–358. IEEE.

Jung, K.-H. and Yoo, K.-Y. (2015). High-capacity index based data hiding method. *Multimedia Tools and Applications*, 74(6):2179–2193.

Karim, M. S. A. and Wong, K. (2015). Data embedding in random domain. *Signal Processing*, 108:56–68.

Muhammad, K., Ahmad, J., Farman, H., Jan, Z., Sajjad, M., and Baik, S. W. (2015). A Secure Method for Color Image Steganography using Gray-Level Modification and Multi-level Encryption. *KSII Transactions on Internet and Information Systems (TIIS)*, 9:1938–1962.

Shen, S., Huang, L., and Tian, Q. (2015). A novel data hiding for color images based on pixel value difference and modulus function. *Multimedia Tools and Applications*, 74(3):707–728.

Subhedar, M. S. and Mankar, V. H. (2014). Current status and key issues in image steganography: A survey. *Computer Science Review*, 13:95–113.

Tsai, Y.-Y., Chen, J.-T., and Chan, C.-S. (2014). Exploring LSB Substitution and Pixel-value Differencing for Block-based Adaptive Data Hiding. *IJ Network Security*, 16(5):363–368.

Wu, D.-C. and Tsai, W.-H. (2003). A steganographic method for images by pixel-value differencing. *Pattern Recognition Letters*, 24(9):1613–1626.

Wu, H.-C., Wu, N.-I., Tsai, C.-S., and Hwang, M.-S. (2005). Image steganographic scheme based on pixel-value differencing and LSB replacement methods. In *Vision, Image and Signal Processing, IEE Proceedings-*, volume 152, pages 611–615. IET.

Yang, C.-H., Weng, C.-Y., Wang, S.-J., and Sun, H.-M. (2010). Varied PVD+ LSB evading detection programs to spatial domain in data embedding systems. *Journal of Systems and Software*, 83(10):1635–1643.

Critical Analysis on Steganography Technique in Text Domain

Sunariya Utama, **Roshidi Din**[*], and **M. Mahmuddin**

School of Computing, UUM College of Arts and Sciences,
Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia

E-mail: sunariya_utama@ahgs.uum.edu.my, roshidi@.uum.edu.my [*],
massudi@uum.edu.my
[*]*Corresponding author*

ABSTRACT

This papers presents several steganography method on text domain based on the perspective of researchers effort in last decade. It has been analyzed the categories of method steganography in medium of text; text steganography and linguistic steganography. The following aim on this paper is identifying the typical these two methods in order to recognize the comparison technique used in previous study. Especially, the explication techniques of text steganography which consist of word-rule based and feature-based technique is critical concern in this paper. Finally, the advantage characteristic and drawback on these techniques in generally also presented in this paper.

Keywords: Steganography, text steganography, linguistic steganography, word-rule based technique, feature-based technique.

1 INTRODUCTION

Steganography as known as associated knowledge of hiding the messages via medium of data to become invisible and undetectable for human sense. Secure privacy information is critical point of steganography in applying performance as a part of information hiding. The implementation of steganography itself, it divides the methods into two categories; steganography in medium of image, audio, video and other digitally invisible code named technical steganography. Therefore, this paper is specifically focusing in steganography method in domain of text.

Generally, the process of steganography in text domain analogically can be illustrated using Prisoner's Problem. The analogy is represented in Figure 1, Ani is sending an original text (M) along with a cover message (C) in order to process embedding known as stego text (S) containing a stego key (K). Firstly, apply the invertible function $f(e) : \{M, C\} \rightarrow S$. Ani can plan an original text (M) using a stego key (K) through $e(M, C) = S$. Hence, S as stego object and it is invertible function, Widya will not discover it suspicious thing. Then, Budi will figure out $e - l(s) = \{M, C\}$ in order to retrieve original text M and cover message C with a stego key K for decoding the process use function. The process of embedding and extracting the information of stego text S will be known by Widya that clarify the process steganography is successful. Based on that idea, the general objective is signifying the development of steganography technique about technique steganography in medium of text that obtain from past researchers effort. The

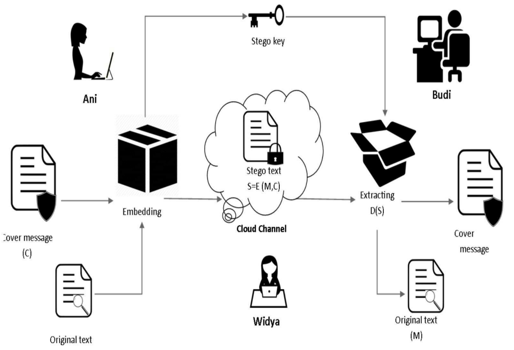


Figure 1: A typical steganography processes on text domain.

implementation of steganography in text domain can classify into two categories; text steganography and linguistic steganography. Therefore, the main concentration of this study is about text steganography and also discuss about linguistic steganography for comparison, with objective to signify the classification of technique. For text steganography is covering messages which manipulated the component in text such as, feature in word, space, line and any others character in sentence of text. Whereas, linguistic steganography is covering messages which modified the information that encoded the massage based on order of linguistically (Agarwal, 2013).

2 REVIEW OF TEXT STEGANOGRAPHY TECHNIQUES IN LAST DECADE

Text steganography consists of two kind of technique word-rule based and feature-based. Word-rules based known as technique embedded in the hidden message based on word pattern by shifting in the text. Meanwhile, feature-based can be define as technique that altered unique feature characteristic in text based on code word. It perhaps slightly moves up and down or code word decrease or increase its length to embed bits from hidden message that can be hidden in text data (Nasab and Shafiei, 2011).

2.1 Word-rule based

The implementation of word-rule based technique is divided in to be two kinds of technique to hide the message. The first technique, line-shift coding can be embedded vertically hidden message to concealed the message in the text. Meanwhile second technique, word-shift coding can embedded in horizontally the hidden message to concealed the message in the text. Word-rule based technique can be develop with other additions technique as show as in Figure 2. These techniques have their own process of embedding and also have advantage and disadvantage. There are several advantages of word-rule based technique that can divide three categories; firstly, the technique has high performance in hiding the hidden message like line-shift coding, technique using shifting second line and unique shaped technique(Roy and Manasmita, 2011). Then, in word-shift coding also has some technique has good performance like distribution white technique (Singh et al., 2009) and integrated inter character technique has decent performance in the way hiding the hidden message (Yang and Kot, 2004). Second advantage is the techniques can hide a lot of hidden message in text that will embed that the word-shift coding technique (Nasab and Shafiei, 2011).

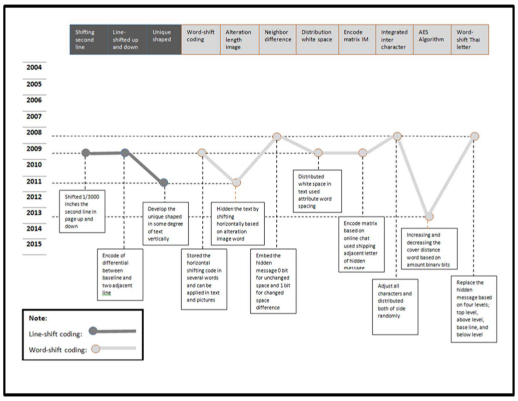

Shifting second line
Line-shifted up and down
Unique shaped
Word-shift coding
Alteration length image
Neighbor difference
Distribution white space
Encode matrix IM
Integrated inter character
AES Algorithm
Word-shift Thai letter
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
Shifted 1/3000 inches the second line in page up and down
Encode of differential between baseline and two adjacent line
Develop the unique shaped in some degree of text vertically
Stored the horizontal shifting code in several words and can be applied in text and pictures
Hidden the text by shifting horizontally based on alteration image word
Embed the hidden message 0 bit for unchanged space and 1 bit for changed space difference
Distributed white space in text used attribute word spacing
Encode matrix based on online chat used shipping adjacent letter of hidden message
Adjust all characters and distributed both of side randomly
Increasing and decreasing the cover distance word based on amount binary bits
Replace the hidden message based on four levels; top level, above level, base line, and below level
Note:
Line-shift coding:
Word-shift coding:

Figure 2: Recapitulation word-rule based technique in text steganography on last decade.

Thirdly , the advantages of those techniques useful in specific particular(Singh et al., 2009), useful in image alteration the length image word technique (Roy and Manasmita, 2011), useful for PDF document in technique (Li et al., 2008), encode matrix for IM that can use in internet based (Liu et al., 2009) and word-shift coding in Thai letter that can use technique steganography in Thai text (Samphaiboon and Dailey, 2008).

Meanwhile, some disadvantages are; firstly, several of techniques in word-rule based still have low security. The challenges of the technique steganography on how to make the hidden message cannot be detected , but several techniques still had low security that can be recognized by other person (Singh et al., 2009), word-shift coding(Nasab and Shafiei, 2011), and also in the alteration length in image technique (Roy and Manasmita, 2011). Then, vulnerable means the existences of hidden message will be lost if changes is done in the text. This disadvantage existed in technique second shift line and in unique shape technique (Roy and Manasmita, 2011). Secondly, the performance of algorithm is still complex in implemented the technique (Altigani and Barry, 2013) and technique neighbour difference (Li et al., 2008). Third disadvantage is can only embed few hidden message in distributing white space technique (Singh et al., 2009) and encode matrix (Liu et al., 2009). Finally, take a long time in developing of the system and process embedding/extracting hidden message in alteration image word technique (Roy and Manasmita, 2011).

2.2 Feature-based

Generally, feature-based technique alters the letter of feature that manipulated the shape, size, position that has relation with feature and structure of font in the text. It makes the reader cannot recognize the secret information in this text (Roy and Manasmita, 2011). Feature-based technique consist of two categories; language based that can use only in certain language and letter-based that can use in any language that use A-Z letters. Several category of feature-based technique commonly can conclude slightly similar with category of word-rule based that presented in Figure 3.

In feature-based, three advantages in word-rule based also occurred in feature-based technique. Firstly, high performance implementation; in English based technique using ECR technique has fast process embedding (Kataria et al., 2013), High invisibility in technique using Reversed Fatah technique (Memon et al., 2008) has standard model transition in Hindi (Changder et al., 2009) and in letter based using Back end interface web page (Mahato et al., 2013). Secondly, feature-based also can embed hidden text in large capacity likes in English based using Deoxyribonucleic Acid (DNA)technique (Reddy et al., 2014) and in Arabic based with reversed Fattah (Memon et al., 2008), vertical displacement technique (Odeh

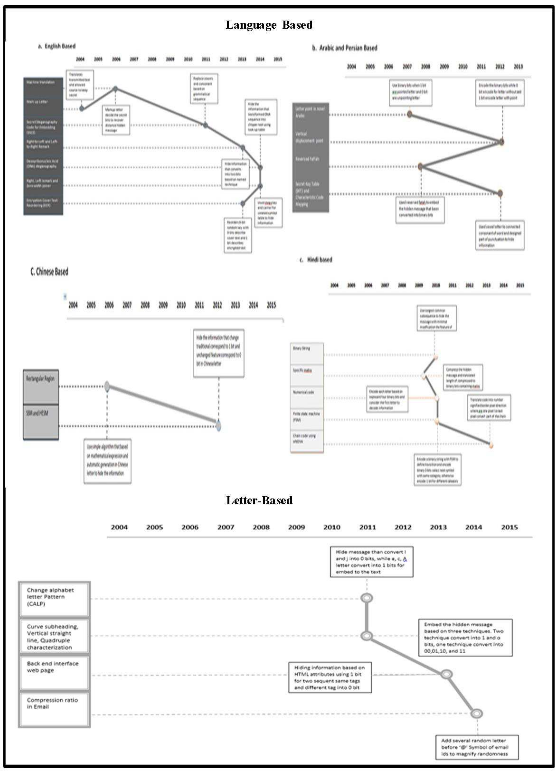

Language Based
a. English Based
b. Arabic and Persian Based
C. Chinese Based
c. Hindi based
Rectangular Region
SSM and HESM
Use simple algorithm that based on mathematical expression and automatic generation in Chinese letter to hide the information
Numerical code
Finite state machine (FSM)
Chain code using ANOVA
Letter-Based
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
Hide message than convert I and j into 0 bits, while a, c, & letter convert into 1 bits for embed to the text
Change alphabet letter Pattern (CALP)
Curve subheading, Vertical straight line, Quadruple characterization
Embed the hidden message based on three techniques. Two technique convert into 1 and o bits, one technique convert into 00,01,10, and 11
Back end interface web page
Hiding information based on HTML attributes using 1 bit for two sequent same tags and different tag into 0 bit
Compression ratio in Email
Add several random letter before '@' Symbol of email ids to magnify randomness

Figure 3: Recapitulation feature-based technique in text steganography on last decade.

et al., 2012). In Hindi based technique of specific matra (Changder et al., 2010)) and chain code technique. (Alam and Naser, 2014). Thirdly, feature-based technique more useful for letter English based using hypertext in markup letter technique (Sui and Luo, 2014) and for SSCE technique all of letter English font be the medium of hidden message (Banerjee et al., 2011). Fourthly, feature-based technique has implemented protection in embed the hidden message such as (Zhang et al., 2006) numerical code technique in Hindi based (Pathak, 2010) then, change alphabet letter pattern technique is recommended step to prevent steganalysis (Bhattacharyya et al., 2011).

Furthermore, some disadvantages feature-based are; firstly, similar condition word-rule based that is easy to detect the existence technique such as; Curve subheading, vertical straight line, and quadruple characterization by Dulera et al. (2011) that is easy to detect. However, easy to attack that can make the removed hidden message also existed in several technique likes retyping in Arabic based using reversed fatah technique (Memon et al., 2008) and vertical displacement of the point (Odeh et al., 2012), specific matra technique (Changder et al., 2009). Secondly, problem in implementation the algorithm in English based using machine translation has often error encode algorithm (Stutsman et al., 2006), in FSM technique in Hindi based (Changder et al., 2010), Right-to-Left Remark and Left-to-Right remark (Odeh et al., 2013). Some performance of technique also dependable ASCII in English based such as technique in remark joiner (Odeh et al., 2014), dependable with vowel and consonant word in technique SSCE (Bhattacharyya et al., 2011) and in Arabic based technique using letter point is dependable with extension character (Gutub and Fattani, 2007). The last disadvantages, Time consuming in process embedding/extracting hidden message in the text was also disadvantage in feature-based technique likes numerical code technique in Hindi based (Pathak, 2010).

3 REVIEW OF LINGUISTIC STEGANOGAPHY TECHNIQUES IN LAST DECADE

Linguistic steganography able cover the hidden message concern to language of word and order modification linguistically. The implementation of this technique based previous researchers in last decade as shown in Figure 4.

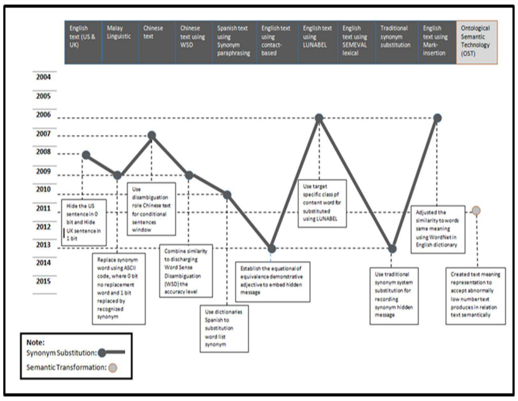

English text (US & UK)
Malay Linguistic
Chinese text
Chinese text using WSD
Spanish text using Synonym paraphrasing
English text using contact-based
English text using LUNABEL
English text using SEMEVAL lexical
Traditional synonym substitution
English text using Mark-insertion
Ontological Semantic Technology (OST)
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
Hide the US sentence in 0 bit and Hide UK sentence in 1 bit
Replace synonym word using ASCII code, where 0 bit no replacement word and 1 bit replaced by recognized synonym
Use disambiguation role Chinese text for conditional sentences window
Combine similarity to discharging Word Sense Disambiguation (WSD) the accuracy level
Use dictionaries Spanish to substitution word list synonym
Establish the equational of equivalence demonstrative adjective to embed hidden message
Use target specific class pf content word for substituted using LUNABEL
Use traditional synonym system substitution for recording synonym hidden message
Adjusted the similarity to words same meaning using WordNet in English dictionary
Created text meaning representation to accept abnormally low number text produces in relation text semantically
Note:
Synonym Substitution:
Semantic Transformation:

Figure 4: Recapitulation linguistic steganography on last decade.

Based on Figure 4, there are some advantages in linguistic steganography. Firstly, unlike text steganography, linguistic steganography especially in synonym substitution technique has own implementation performance; high invisibility in Malay linguistic technique (Muhammad et al., 2009) and English text using LUNABEL function (Chand and Orgun, 2006), simple variant in Chinese text synonym (Yuling et al., 2007) minimalized creating syntax error in English text using context-based (Wang et al., 2013) and in English text using mark-insertion can achieve maximum cumulative distortion (Topkara et al., 2006). Secondly, one of the techniques is traditional synonym substitution able to hide the hidden message in large capacity (Qi et al., 2013). Thirdly, linguistic steganography also has advantage in specific condition linguistic steganography which can useful in printing text using synonym substitution (Shirali-Shahreza and Shirali-Shahreza, 2008). Other technique has this advantage is synonym paraphrasing that very useful in Spanish language (Munoz et al., 2010).

Meanwhile, they also identified the disadvantages in this technique. Firstly, low security can also be the issue in linguistic steganography Shirali-Shahreza and Shirali-Shahreza (2008) and technique synonym paraphrasing in Spanish language is easy to attack by intruders (Munoz et al., 2010). Secondly, the limitation implementing of linguistic steganography is only capable in own language because this technique is based on linguistic. This technique has complex algorithm in English text using LUNABEL function (Chand and Orgun, 2006) and using traditional synonym substitution (Qi et al., 2013). In English text, using context-based has incomplete vocabulary (Wang et al., 2013). Then, semantic transformation technique possible generates a lot of semantic spam. Finally, in Malay linguistic consuming much in process embedding/extracting hidden message(Muhammad et al., 2009).

4 CONCLUSION

This paper is presented and explored several techniques steganography methods; text steganography and linguistic steganography in order to observe the development of these methods in last decade. In text steganography, the techniques consist of word-rule based technique and feature-based technique. Based on comparison in Figure 2 until Figure 4, this can conclude that text steganography is most widely researched by previous researcher especially in feature-based technique as shown in Figure 5 as following.

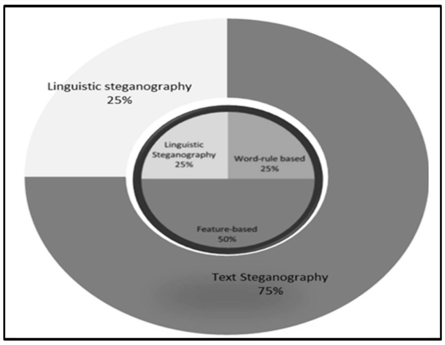


Figure 5: Amount of research on text and linguistic steganography in last decade,

There is a lot of efforts which have been proposed by previous researcher. The percentage of the research about the text steganography is much higher than linguistic steganography based on Figure 3. The text steganography has percentage of 75% that consist of word-rule based technique 25% and feature-based 50%, whereas linguistic steganography has similar percentage of only 25%.

Moreover, text and linguistic steganography almost has similar advantage in development, performance, and implement for hiding hidden message. Some techniques have high performance, recommended secure protection, or can embed large amount the hidden message. However, both of techniques also have certain issues which seems to be the limitation of those techniques; low security, complex algorithm performance, or time consuming. In future work, more investigation about on the techniques will be considering in the form of mathematical formulation as the substantiations of implementation.

5 ACKNOWLEDGMENTS

We would like thank to Assoc. Prof. Dr. Huda Ibrahim, Dean of School of Computing, Universiti Utara Malaysia (SoC CAS UUM) and Director of Awang Had Salleh Graduate School (AHSGS) for their moral support for the realization of this work. This research was financially supported by the Postgraduate Incentive Grant (Grant NO. 15817/2015) under AHSGS Grant, Universiti Utara Malaysia (UUM).

REFERENCES

Agarwal, M. (2013). Text steganographic approaches: A comparison. *International Journal of Network Security & Its Applications*, 5(1):91–106.

Alam, M. and Naser, M. (2014). Re-evaluating chain-code as features for bangla script. In *Electrical Information and Communication Technology (EICT), 2013 International Conference on*, pages 1–5.

Altigani, A. and Barry, B. (2013). A hybrid approach to secure transmitted messages using advanced encryption standard (aes) and word shift coding protocol. In *Computing, Electrical and Electronics Engineering (ICCEEE), 2013 International Conference on*, pages 134–139.

Banerjee, I., Bhattacharyya, S., and Sanyal, G. (2011). Novel text steganography through special code generation. In *Proceedings of International Conference on Systemics, Cybernetics and Informatics (ICSCI-2011)*, pages 298–303.

Bhattacharyya, S., Indu, P., Dutta, S., Biswas, A., and Sanyal, G. (2011). Hiding data in text through changing in alphabet letter patterns (CALP). *Journal of Global Research in Computer Science*, 2(3):33–39.

Chand, V. and Orgun, C. O. (2006). Exploiting linguistic features in lexical steganography: Design and proof-of-concept implementation. In *Proceedings of the Annual Hawaii International Conference on System Sciences*, volume 6, pages 1–10.

Changder, S., Das, S., and Ghosh, D. (2010). Text steganography through indian languages using feature coding method. In *ICCTD 2010 - 2010 2nd International Conference on Computer Technology and Development, Proceedings*, pages 501–505.

Changder, S., Debnath, N., and Ghosh, D. (2009). A new approach to hindi text steganography by shifting matra. In *Advances in Recent Technologies in Communication and Computing, 2009. ARTCom '09. International Conference on*, pages 199–202, New Delhi, India.

Dulera, S., Jinwala, D., and Dasgupta, A. (2011). Experimenting with the novel approaches in text steganography. *International Journal of Network Security & Its Applications*, 3(6):213–225.

Gutub, A. A.-A. and Fattani, M. M. (2007). A novel arabic text steganography method using letter points and extensions. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 1(3):483–486.

Kataria, S., Singh, K., Kumar, T., and Nehra, M. (2013). Ecr (encryption with cover text and reordering) based text steganography. In *Image Information Processing (ICIIP), 2013 IEEE Second International Conference on*, pages 612–616.

Li, L., Huang, L., Zhao, X., Yang, W., and Chen, Z. (2008). A statistical attack on a kind of word-shift text-steganography. In *Intelligent Information Hiding and Multimedia Signal Processing, 2008. IIHMSP '08 International Conference on*, pages 1503–1507.

Liu, M., Guo, Y., and Zhou, L. (2009). Text steganography based on online chat. In *Intelligent Information Hiding and Multimedia Signal Processing, 2009. IIH–MSP '09. Fifth International Conference on*, pages 807–810.

Mahato, S., Yadav, D. K., and Khan, D. A. (2013). A modified approach to text steganography using hypertext markup language. In *2013 Third International Conference on Advanced Computing and Communication Technologies (ACCT)*, pages 40–44.

Memon, J. A., Khowaja, K., and Kazi, H. (2008). Evaluation of steganography for urdu/arabic. *Journal of Theoretical & Applied Information Technology*, 4(3):232–237.

Muhammad, H. Z., Rahman, S. M. S. A., and Shakil, A. (2009). Synonym based malay linguistic text steganography. In *2009 Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA 2009)*, pages 423–427.

Munoz, A., Carracedo, J., and Alvarez, I. A. (2010). Measuring the security of linguistic steganography in spanish based on synonymous paraphrasing with wsd. In *Proceedings - 10th IEEE International Conference on Computer and Information Technology, CIT-2010, 7th IEEE International Conference on Embedded Software and Systems, ICESS-2010, ScalCom-2010*, pages 965–970.

Nasab, M. V. and Shafiei, B. M. (2011). Steganography in programming. *Australian Journal of Basic & Applied Sciences*, 3(12):1496–1499.

Odeh, A., Alzubi, A., Hani, Q., and Elleithy, K. (2012). Steganography by multipoint arabic letters. In *Systems, Applications and Technology Conference (LISAT), 2012 IEEE Long Island*, pages 1–7.

Odeh, A., Elleithy, K., and Faezipur, M. (2014). Steganography in text by using ms word symbols. In *Proceeding of zone 1 conference of the American Society Engineering Education*, pages 1–5.

Odeh, A., Elleithy, K., and Feazipour, M. (2013). Text steganography using language remarks. In *The American Society of Engineering Education, ASEE 2013*, pages 1–7.

Pathak, M. (2010). A new approach for text steganography using hindi numerical code. *International Journal of Computer Applications*, 1(8):56–59.

Qi, C., Xingming, S., and Lingyun, X. (2013). A secure text steganography based on synonym substitution. In *Conference Anthology, IEEE*, pages 1–3.

Reddy, R. P. K., Nagaraju, C., and Subramanyam, N. (2014). Text encryption through level based privacy using dna steganography. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, 3(3):168–172.

Roy, S. and Manasmita, M. (2011). A novel approach to format based text steganography. In *Proceedings of the 2011 International Conference on Communication, Computing &#38; Security*, pages 511–516, New York, NY, USA. ACM.

Samphaiboon, N. and Dailey, M. (2008). Steganography in thai text. In *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2008. ECTI-CON 2008. 5th International Conference on*, volume 1, pages 133–136.

Shirali-Shahreza, M. H. and Shirali-Shahreza, M. (2008). A new synonym text steganography. In *Proceedings - 2008 4th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIH–MSP 2008*.

Singh, H., Singh, P., and Saroha, K. (2009). A survey on text based steganography. In *Proceedings of the 3rd National Conference; INDIACom–2009Computing For Nation Development*, pages 26–27, New Delhi, India.

Stutsman, R., Atallah, M., Grothoff, C., and Grothoff, K. (2006). Lost in just the translation. In *Proceedings of the 2006 ACM Symposium on Applied Computing*, pages 338–345. ACM.

Sui, X.-G. and Luo, H. (2014). A new steganography method based on hypertext. In *Radio Science Conference, 2004. Proceedings. 2004 Asia-Pacific*, pages 181–184.

Topkara, U., Topkara, M., and Attalah, M. (2006). The hiding virtues of ambiguity: Quantifiably resilient watermarking of natural language text through synonym substitutions. In *Proceedings of the 8th workshop on Multimedia and security*, pages 163–174.

Wang, F., Huang, L.and Chen, Z., Yang, W., and Miao, H. (2013). A novel text steganography by context-based equivalent substitution. In *Signal Processing, Communication and Computing (ICSPCC), 2013 IEEE International Conference*, pages 1–6.

Yang, H. and Kot, A. (2004). Text document authentication by integrating inter character and word spaces watermarking. In *Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference on*, volume 2, pages 955–958.

Yuling, L., Xingming, S., Can, G., and Hong, W. (2007). An efficient linguistic steganography for chinese text. In *Proceedings of Multimedia and Expo, IEEE International Conference*, pages 2094–2097.

Zhang, W., Zeng, Z., Pu, G., and Zhu, H. (2006). Chinese text watermarking based on occlusive components. In *Information and Communication Technologies, 2006. ICTTA '06. 2nd*, volume 1, pages 1850–1854.

Oblivious Memory Trace in Graphics Hardware with Constant Blowup

Xiaoqi Yu[*1], **Nairen Cao**[1], **Linru Zhang**[1], **Gongxian Zeng**[1], and **Siu Ming Yiu**[*1]

[1]*Department of Computer Science, Hong Kong University*

E-mail: xqyu@cs.hku.hk, caonr@pku.edu.cn, zhanglr3@mail2.sysu.edu.cn,
naksi@connect.hku.hk, smyiu@cs.hku.hk
[*]*Corresponding author*

ABSTRACT

Attacks by access pattern can be devastating harms for cloud service applications and financial computations, while efficient solutions for building oblivious memory trace program is nontrivial. In this paper, we propose an efficient scheme that applys parallel techniques based on graphic hardware to achieve oblivious memory trace program with only constant blowup.

Keywords: Oblivious RAM, Oblivious Memory Trace, GPU

1 INTRODUCTION

Since cloud services have gained more popularity nowadays, there are more entities delegating their private data and computation to cloud service providers. However, security and privacy is still a big barrier in furthering this type of applications. Basic security requirements is to encrypt the sensitive data. Yet further research works show that even the contents of the data are encrypted, some side channels, e.g. memory access pattern, can also leak significant information.

In Chen et al. (2010), attackers will learn the investment structures from the data request access pattern. Specifically, the number of financial products that clients invest in can be leaked from client-server data request patterns in the Online Invest system. Furthermore, on the Online Tax website, eavesdroppers can also dig out some personal informations from the network request patterns. What's worse, investment bank will also leak the sensitive information on the corporates they are investing in, which can lead to devastating economic loss. Therefore, even the real data is encrypted, there are important information that can be inferred from the side channels of the access pattern. To design a program that can hide these types of information is thus urgent in the widespread of the new applications of cloud storage and computation.

In the domain of financial market, we can also notice the development of quantitive method and computation power (Bouchaud et al., 2001, Krawiecki et al., 2002), e.g, high-frequency trading, automatic trading, among which the basic idea is to train the financial models by applying the historical trading data to output some indicators and warnings, or generate strategy rules. However, the method would face the overfitting problem when the training data size is not large enough, which means that the results are too fitted to the historical data making it insignificant. On the other hand, large volume data will lead to heavy latency and computation pressure. Therefore parallel computation power is applied in the financial market (Preis et al., 2009). In Preis et al. (2009), they applied three main models to graphic hardware(GPU) to evaluate the fluctuations of the market. In this fold, they share some common computation patterns applying the financial data computation, and the memory configurations. Recently, some researchers (e.g. Pinkas and Reinman (2010)) also gave the examples on the possibility of revealing trading transactions if data access pattern is not perfectly protected.

From the memory access pattern analysis, we can easily infer the operations executed on the outsourced data, and financial activities run on the machine, which will lead to economic loss to some extends. Since in many financial firms, they will rely on the strategies to generate valuable rules to earn money. It will leak information of the investment firm-wide. Therefore, we aim to build a scheme that is oblivious memory access that almost leak no sensitive information even when the adversary can tab the memory bank in host machine.

From the aforementioned analysis, memory access pattern of a program can leak sensitive information that affect the privacy or lead to economic loss for clients. To solve these problems, previous researchers proposed a scheme called oblivious RAM(ORAM). ORAM was first proposed in Goldreich (1987), Goldreich and Ostrovsky (1996) and R.Ostrovsky (1990) by Goldreich and Ostrovsky *et. al.* to protect software privacy against reverse-engineering attacks. Other improvements includes applying techniques like cuckoo hash functions (Arbitman et al. (2009)) and randomised shell sort. Even so, the overhead of ORAM primitives is still a barrier to the practicality of such schemes. A breakthrough, which significantly simplifies ORAM structure, was proposed in Stefanov et al. (2013) by using a binary tree layout, which is followed by other improved tree-based results Chung and Pass (2013), Wang et al. (2015). However, we can not simply apply ORAM directly in our program since the best results until recently is still O($\log^2 N/\log\log N$) based on the input size N, whilst N is always big in the applications for cloud environments and financial computations.

Through observations, we notice that not all memory access pattern should be kept secret. In (Mitchell and Zimmerman, 2014, Wang et al., 2014), they compile the program with some data structures that exhibit some degree of predictability for access patterns, which can achieve some degree of efficiency improvements. Furthermore, it can also build into some sensitive partition of the memory obliviously that can be generally applied in many programs.

As the applications of graphic hardware(GPU) has gained more popularity in financial computation recently, and parallelism can indeed achieve obliviousness almost freely. Thus we have got some hints from the GPU techniques to build a parallel oblivious memory trace scheme that can achieve both efficiency and obliviousness simultaneously, which will be called GORAM in the paper.

In the following sections of this paper, we will first give out some preliminaries and related notations of the problems in Section 2. The details of GORAM and performance analysis will be introduced in Section 3 and Section 4 respectively. Last but not least, we will end with discussions and conclusions in Section 5.

2 PRELIMINARIES

2.1 Oblivious RAM

To begin with, we first introduce the definitions of access pattern for ORAM, which is the general oblivious requirements and the most restrict property of the oblivious RAM.

Definition 2.1 (Access pattern). *Let $A(\mathtt{y})$ denote the sequence of access to the remote server storage given the data request sequence* y *from client. Specifically,* $\mathtt{y} = (\mathtt{op_{L'}}, \mathtt{a_{L'}}, \mathtt{data_{L'}}), \ldots, (\mathtt{op_1}, \mathtt{a_1}, \mathtt{data_1})$*($L' = |\mathtt{y}|$), and op_i denotes a operation of either $read(a_i)$ or $write(a_i, data_i)$. In addition, a_i is the logical identifiers of the block.*

Definition 2.2 (Security definition). *An* ORAM *construction is said to be secure if (1) For any two data request* y *and* z *of the same length, their access pattern $A(\mathtt{y})$ and $A(\mathtt{z})$ are computationally indistinguishable by anyone but the client, (2) the* ORAM *construction is correct in the sense that it returns correct results for any input* y *with probability $\leq 1 - negl(|\mathtt{y}|)$.*

2.2 Oblivious Memory Trace

We also introduce the definition of oblivious memory trace given in the followings.

Definition 2.3 (Oblivious Memory Trace). *A program S satisfy memory trace obliviousness if for two memories $M_1 \sim_L M_2$ Liu et al. (2013). if M_1' is the memory after executing S with initial memory M_1 and memory traces t_1, and M_2' is the memory after executing S with initial memory M_2 and memory traces t_1. Then $t_1 \equiv t_2$, and $M_1' \sim_L M_2'$ Liu et al. (2013).*

3 GPU-BASED OBLIVIOUS MEMORY TRACE

3.1 A Toy Example

To clarify how access pattern can leak secret information through the executions of a program, we first present a toy example of a sample program. From Algorithm 1, we notice that the $while$ loop leak the information of the secret array $\mathcal{S}$. Supposed that the secret array is the action indicators in the trading activities, while positive and negative values means opposite actions that described in program block $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. Then the attacker can tell from memory trace by loading $\mathcal{S}_1$ or $\mathcal{S}_2$ to decide whether the corresponding element of $\mathcal{S}$ is positive or negative and infer more information of the strategies.

3.2 Existing solutions to transform to oblivious program

Previous solution Liu et al. (2013) is to add one $else$ branch with the program block $\mathcal{S}_2$ in Algorithm 1. However, it will double the size of the program running time. In addition, from the memory address of $\mathcal{S}_1$ and $\mathcal{S}_2$, the attacker can still tell whether the private value is positive or negative. To solve the problem, they should store $\mathcal{S}_1$ and $\mathcal{S}_2$ in an ORAM primitive. In this way, to access programs $\mathcal{S}_1$ or $\mathcal{S}_2$ is oblivious in the attacker's eyes. On the other hand, as we show in the previous sections, current ORAM constructions still have heavy overhead especially when the size of $\mathcal{S}_1$ or $\mathcal{S}_2$ is large. In this case, it is not a good solution in practical situations.

```
Input: Secret $\mathcal{S}[N]$
   Secret count
   While ($i \leq N$)
      if($\mathcal{S}[i] \leq 0$)
         $\mathcal{S}_1$
      else:
         $\mathcal{S}_2$
```

Algorithm 1: An example of transforming a program to oblivious memory trace in existing solutions

3.3 Our GPU-Based Solutions

The framework CUDA Fernando and Kilgard (2003),Kirk et al. (2007), which is developed for Graphics Processing Unit(GPU), has gained more popularity, we can apply it to build an efficient memory trace oblivious program. Via CUDA techniques, the intuitive idea is to firstly load the secret programs into every multi-processor(MP). Then load the corresponding data into each MP, after which is to reduce Preis et al. (2009) the intermediate results to calculate the final output. To elaborate the detailed scheme, we will demonstrate it with the previous example in Algorithm 1.

- **Allocating Blocks:** Assumed that parallelism parameter is P, then we can allocate $P/512+1$ Fernando and Kilgard (2003) Kirk et al. (2007) blocks with 512 threads in each block.
- **Load Program:** In this example, the secret programs whose trace should be hidden are in the set of statements of $\mathcal{S}_1$ and $\mathcal{S}_2$. Therefore, we firstly load the program set $\mathcal{S}_1$ and $\mathcal{S}_2$ into each processor, as demonstrated in Figure 1.
- **Load Input Data:** Load array segments $\mathcal{S}[m_{i-1}+1, m_i]$ to the $i\ th$ processor s.t the total input is divided into $P/512+1$ segments. As shown in Figure 2, the cross characters stand for the dummy data that we inserted into the array.
- **Reduce:** We can apply a reduce phase after the processing in each block including synchronising within each block, and this reduce process can be realized by a binary tree structure, as shown in Figure 3. After this reduction, the final results will be found in the first block.

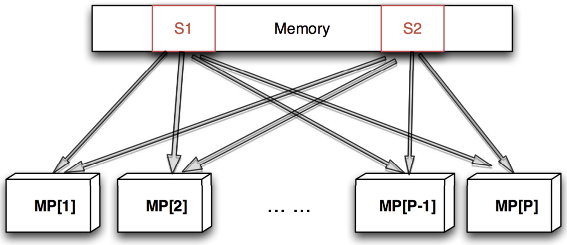


Figure 1: An example of transforming a program to oblivious memory trace with existing solutions

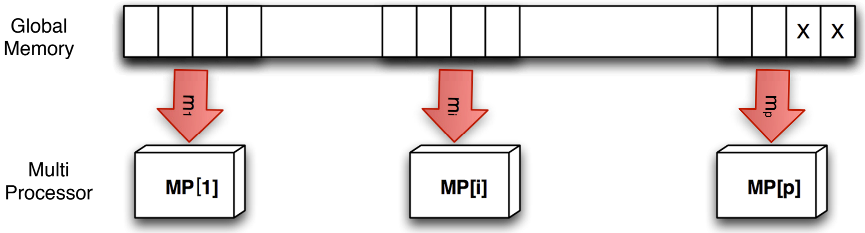


Figure 2: An example of loading data from global memory to multi-processor

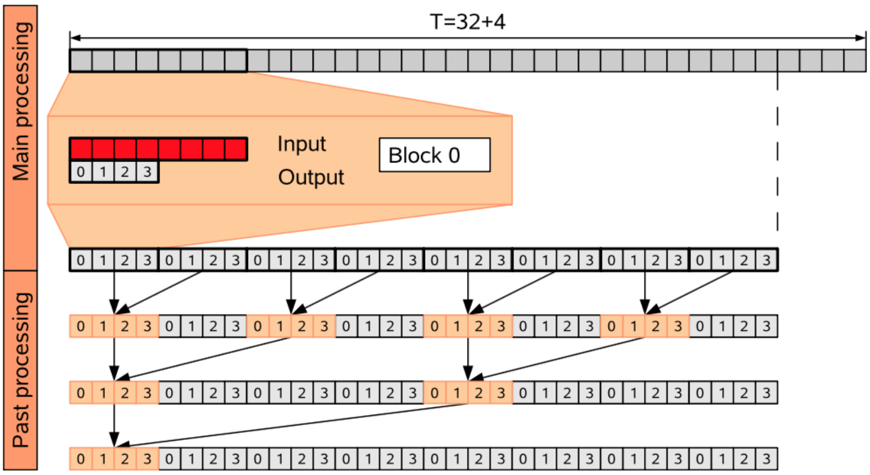


Figure 3: An example of reducing the intermediate results of each block to a final output

4 PERFORMANCE ANALYSIS FOR GORAM

4.1 Oblivious of the scheme

In this section, we outline the lemmas and theorem that complete our oblivious analysis of the scheme in Section 3, and give out the intuitive idea of them.

Lemma 4.1. ***Oblivious in Load Program*** *For any secret element in $\mathcal{S}[i]$, the program access pattern is identical, and no side information of $\mathcal{S}[i]$ is leaked to the attacker from the conditional logic statements in the execution of the program.*

The discussion of extraction of the secret program set is out of the scope of the paper, therefore supposed that we are given a set $\mathcal{S}$ of secret program statements. For example, in the toy example of Algorithm 1, $\mathcal{S} = \{\mathcal{S}_1\} + \{\mathcal{S}_2\}$. In the first step, we load the program set $\mathcal{S}$ into all the P threads in the framework. The memory access the attacker can learn is only the memory tab of the given set $\mathcal{S}$ once, which can not infer any meaningful information of the programs.

Lemma 4.2. ***Oblivious in Reduce*** *Supposed that computation in each multi-processor is private, then the memory trace in the reduce phase is oblivious in the attacker's view.*

In the reduce phase, the initial maximum step is decided by the allocating of the multi processors. After that, it will be reduced from the partial results in each blocks to one final output that is stored in the first block, which takes $\log N$ steps in total. This cost is insignificant compared to the $O(N)$ complexity to execute the program in sequence. Dummy steps may be inserted to achieve the same reduction steps regarding of the initial input data. Therefore, no extra information will be leaked to the adversary supposed that the calculation inside each processor is private.

Theorem 4.1. *Given a secret program statement set $\mathcal{S}$, a program that apply our* GPU*-based scheme is a oblivious memory trace program.*

Apparently, this theorem can be concluded directly from Lemmas 4.1 and 4.2.

4.2 Efficiency analysis for GORAM

Metric To clarify the performance evaluation of the scheme, we first introduce the metric in the paper. Specifically, we use the time complexity blowup to evaluate the results, and the blowup is defined as the time for oblivious program over their non-oblivious counterparts. Through this metric, we can give out the following theorem.

Theorem 4.2. *Blowup of the* GORAM *is $O(1)$, with the reasonable settings of parallelism parameter P.*

Supposed that P is the parallelism parameter, and size of the secret program block S_1 is s. Then the previous solution of applying partial ORAM will at least cause the blowup $\log^2 s/\log\log s$ if we only consider the extra cost in the step of loading the program.

When the input size is limited, the overhead of applying an partial ORAM is acceptable, and we can also build the program into a complete big ORAM. However, as demonstrated in section 1, we know that the key issue of the cases is that we always require big volume of input size. Therefore, parallel techniques can be applied and with slight modifications to achieve efficiency and obliviousness. Compared to its counterpart, our parallel oblivious memory trace scheme will only cause constant blowup. Specifically, all threads in each block will load the secret program to be executed in the memory at the same time, which need the time of normal request for $\mathcal{S}_1$ and $\mathcal{S}_2$, with only limited extra of syncrynization. Indeed it is not an concern since it is built into the CUDA framework within the same block.

5 DISCUSSIONS AND CONCLUSIONS

In this paper, we demonstrate the potential attacks by memory access pattern, and related existing solutions in cloud applications and financial markets. Furthermore, we apply graphic hardware framework to build a parallel oblivious memory trace scheme which is trace-free with constant blowup in terms of time complexity. Our model will achieve obvious improvement for the big data situations, which are prevalent in both cloud applications and financial markets.

We mainly adopt the techniques in one simplified example, and the CUDA source codes will be included in our full version. Additionally, it is supposed to be applied in more complete programs with regarding of the following considerations, which is the method to automatically build parse the program and extract the sensitive program set. Since this is out of the scope of this paper, more details can be found in paper Liu et al. (2013).

REFERENCES

Arbitman, Y., Naor, M., and Segev, G. (2009). Backyard Cuckoo Hashing: Constant Worst-Case Operations with a Succinct Representation. *CoRR*, abs/0912.5424.

Bouchaud, J.-P., Matacz, A., and Potters, M. (2001). Leverage effect in financial markets: The retarded volatility model. *Physical Review Letters*, 87(22):228701.

Chen, S., Wang, R., Wang, X., and Zhang, K. (2010). Side-channel leaks in web applications: A reality today, a challenge tomorrow. In *Security and Privacy (SP), 2010 IEEE Symposium on*, pages 191–206. IEEE.

Chung, K.-M. and Pass, R. (2013). A simple oram. Technical report, DTIC Document.

Fernando, R. and Kilgard, M. J. (2003). *The Cg Tutorial: The definitive guide to programmable real-time graphics*. Addison-Wesley Longman Publishing Co., Inc.

Goldreich, O. (1987). Towards a theory of software protection and simulation by oblivious RAMs. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pages 182–194. ACM.

Goldreich, O. and Ostrovsky, R. (1996). Software Protection and Simulation on Oblivious RAMs. *J. ACM*, 43(3):431–473.

Kirk, D. et al. (2007). NVIDIA CUDA software and GPU parallel computing architecture. In *ISMM*, volume 7, pages 103–104.

Krawiecki, A., Hołyst, J., and Helbing, D. (2002). Volatility clustering and scaling for financial time series due to attractor bubbling. *Physical review letters*, 89(15):158701.

Liu, C., Hicks, M., and Shi, E. (2013). Memory trace oblivious program execution. In *Computer Security Foundations Symposium (CSF), 2013 IEEE 26th*, pages 51–65. IEEE.

Mitchell, J. C. and Zimmerman, J. (2014). Data-Oblivious Data Structures. In *STACS*, pages 554–565.

Pinkas, B. and Reinman, T. (2010). Oblivious RAM Revisited. In *Proceedings of the 30th Annual Conference on Advances in Cryptology*, CRYPTO'10, pages 502–519, Berlin, Heidelberg. Springer-Verlag.

Preis, T., Virnau, P., Paul, W., and Schneider, J. J. (2009). Accelerated fluctuation analysis by graphic cards and complex pattern formation in financial markets. *New Journal of Physics*, 11(9):093024.

R.Ostrovsky (1990). Efficient computation on oblivious RAMs. *STOC*.

Stefanov, E., van Dijk, M., Shi, E., Fletcher, C., Ren, L., Yu, X., and Devadas, S. (2013). Path ORAM: An Extremely Simple Oblivious RAM Protocol. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer &#38; Communications Security*, CCS '13, pages 299–310, New York, NY, USA. ACM.

Wang, X., Chan, H., and Shi, E. (2015). Circuit oram: On tightness of the goldreich-ostrovsky lower bound. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pages 850–861. ACM.

Wang, X. S., Nayak, K., Liu, C., Chan, T., Shi, E., Stefanov, E., and Huang, Y. (2014). Oblivious data structures. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 215–226. ACM.

Implementation of a Pairing-Based Identity-Based Signature on iPhones

Wai-Choong Wong[1], **Tiong-Sik Ng**[1], and **Ji-Jian Chin** *[1]

[1]*Faculty of Engineering, Multimedia University*

E-mail: jjchin@mmu.edu.my
**Corresponding author*

ABSTRACT

Many cryptographic protocols that based on bilinear pairings were introduced over the years with various implementations such as identity-based identification (IBI) and identity-based signature (IBS). While Pairing Based Crypto (PBC) library is available as one of the most well-known open source pairing-based library based on elliptic curve cryptography (ECC), Tan et al. (2010) had developed a Java-based pairing library that functions in a similar manner as well as provides comparative performance based on pairing time. However, Tan et al. (2010)'s library does not support the development on Apple's iOS platform due to the distinct programming language it is written in. Recently Cheah et al. (2015) successfully simulated a pairing-based IBI scheme in iOS based on Tan et al. (2010)'s library by translating it but did not manage to extend his work as a mobile application due to linker error. In this work, we managed to overcome that problem and successfully implemented a pairing-based IBS scheme with Tan et al. (2010)'s library for mobile devices such as iPhones.

Keywords: Bilinear pairings, identity-based signature, iOS, Apple, iPhone

1 INTRODUCTION

According to the statistic shown by Pew Reserch Center, the ownership for mobile devices such as smartphones and tablets have been steadily growing over the recent years (Pew-Research-Center, 2015). People nowadays are relying on these smart devices to communicate with each other. Wireless connections between the mobile devices are convenient but have much to be desired in terms of security. Thus, identity authentication and verification processes between both parties are mandatory to ensure a secured environment for wireless communication. Cryptography plays a huge role in providing confidentiality and integrity in these situations.

Based on International Data Corporation (IDC)'s analysis in 2015, Googles Android OS is holding 82.8% of market share, followed by 13.9% for Apples iOS, 2.6% for Windows Phone, and 0.7% for others (International-Data-Corporation, 2015). Even though iOS is having a large user base, it appears that pairing-based cryptographic libraries are almost non-existent and has yet to receive widespread support.

1.1 Motivations and related work

Our primary motivation for this work is Cheah et al. (2015)'s successful work that showed an efficient identity-based identification scheme using ECC pairing-based libraries can be implemented in iOS by translating the Java-based library by Tan et al. (2010). Their work is significant as it shows a working pairing-based security scheme on iOS platform which could be further implemented for various security applications in the industry. Also, pairing-based implementations for the iOS platform are still rare at this time. Our goal for conducting this work is add to the research and application of pairing-based cryptography on iOS mobile platform.

1.2 Contributions

In this work, the IBS scheme implemented was proposed by Cha and Cheon (2003). We simulate the scheme by translating Tan et al. (2010)'s library and develop a mobile application for iOS to implement the scheme mentioned. Our achievement is an extension of Cheah et al. (2015)'s results, where the authors only managed to implement a simulator for the Mac. In contrast, our implementation successfully runs on iPhones. To our knowledge, this is the first successful implementation of a pairing-based IBS scheme on iOS platform for mobile devices using Tan et al. (2010)'s library. The advantage of running pairing-based cryptosystems is that the secret key length can be much shorter than pairing-free alternatives, such as RSA or DSA. We observe the running time for different stages of the scheme by using both simulator and mobile devices with efficient results as well. We describe the methods and procedures we used for conducting this work in detail.

The rest of the paper is organized as such. In section 2, we will first describe the Cha-Cheon IBS scheme and its stages in detail. In section 3, we will show how we developed the mobile application using available tools. In section 4, we will show our results on both simulator and iOS devices. Section 5 is the conclusion of our work.

2 THE CHA-CHEON IBS SCHEME

2.1 Preliminaries

Notation-wise let $\{0,1\}^*$ denote the set of all bit strings while $\{0,1\}^n$ the set of bit strings of length n. If a string $s \in \{0,1\}^*$ then $|s|$ denotes the length of s. If S is a set then $|S|$ denotes the size of S. Let $x \overset{\$}{\leftarrow} S$ denote a randomly and uniformly chosen element x from a finite set S. Lastly let Z_p denote the set of positive integers modulo a large prime p.

Let G be a group of prime order p and let q be a large prime where $q = p/2 - 1$. G is a group in which Computational Diffie-Hellman problem (CDHP) is considered intractable to be solved. Let P be a random generator in G. The DDHP problem is defined as: Given (P, aP, bP) as a Diffie-Hellman tuple, where $a, b \overset{\$}{\leftarrow} Z_q$, calculate abP. The Cha-Cheon IBS scheme is proven to be secure if the CDHP is intractable.

2.2 Bilinear Pairings

Secondly, Cha-Cheon IBS scheme runs using a bilinear pairing function mapping elements from group G to group G_T, i.e. $e : G \times G \to G_T$. The bilinear pairing function e requires the following properties:

1. Bilinearity: $e(aP, bP) = e(P, P)^{ab}$.
2. Non-degeneracy: $e(P, P) \neq 1$
3. e is efficiently computable.

2.3 The Cha-Cheon IBS Scheme

An IBS scheme consists of 4 stages: Setup, Extract, Sign and Verify. Following this model, the Cha-Cheon IBS scheme is defined in detail as follows:

1. **Setup**: Choose a prime generator $P \overset{\$}{\leftarrow} G$ and a master secret key $s \overset{\$}{\leftarrow} Z_q$, then set $P_{pub} = sP$. Select hash functions $H_1 : \{0,1\}^* \times G \to Z_q$ and $H_2 : \{0,1\}^* \to G$. Lastly establish the pairing function $e : G \times G \to G_T$. Publish the system parameters as $(G_1, G_T, q, e, P, P_{pub}, H_1, H_2)$ and keep s secret.
2. **Extract**: Given an identity ID (e.g. email address, domain), calculate $Q = H_2(ID)$ and $D_{ID} = sQ$. Q_{ID} plays the role of the associated public key whereas D_{ID} is the user secret key.
3. **Sign**: Given a message m and user secret key D_{ID} as input, choose a random number $r \overset{\$}{\leftarrow} Z_q$, compute $U = rQ_{ID}, h = H_1(m, U)$ and $V = (r + h)D_{ID}$. The signature is generated as $\sigma = (U, V)$.
4. **Verify**: To verify the signature $\sigma = (U, V)$ of a message m with identity ID, check the validity of the tuple $(P, P_{pub}, U + hQ_{ID}, V$ by resolving $e(P, V) = e(P_{pub}, U + hQ_{ID})$.

For correctness, the following check equation should hold:

$$\begin{aligned} e(P_{pub}, U + hQ_{ID}) &= e(sP, rQ_{ID} + hQ_{ID}) && (1)\\ &= e(P, s(r + h)Q_{ID}) && (2)\\ &= e(P, (r + h)D_{ID}) && (3)\\ &= e(P, V) && (4) \end{aligned}$$

3 METHODOLOGY

In this section we discuss about the methodology of applying Tan et al. (2010)'s library in iOS mobile application development. Objective-C has been the native programming language for developing an iOS application ever since Apple acquired NeXT in 1996 (Cox and Love, 2015). Long before Apple gained its ground of market share, it has always been providing support for Java, for example the ”Java Bridge” which is the binding between Java and iOS's native application programming interface (API) known as Cocoa.

However, the unpopularity of ”Java bridge” among the Cocoa developers and incompatibility of Cocoa's key features with Java forced Apple to officially deprecate their support for ”Java bridge” in 2005 (Apple, 2015). Thus, all the new features Cocoa introduced later than Mac OS X version 10.4 were not available for Cocoa-Java programming interface. In 2014, another programming language developed by Apple known as Swift was introduced to replace the Objective-C language.

To apply the libraries introduced by Tan et al. (2010), translation from Java to Objective-C is a must. Hence, a different approach other than ”Java Bridge” must be made to link the libraries for both Java and Objective-C such as using third party translations tools. Besides that, the mobile application development also requires the knowledge of using the development tools provided by Apple.

3.1 Translation of Java libraries using J2ObjC

Java source code can be translated using a tool called J2ObjC which is an open-source command-line tool developed by Google (Ball, 2015). This tool allows developers to implement their Java source to be part of iOS application's build by translating them to Objective-C. It supports most of the Java language and runtime features such as generic types, threads, reflection and exceptions. The goal of this tool is shared the application and data models written in Java to other platforms such as GWT web apps, Android, and iOS applications.

However, platform-independent user-interface (UI) toolkit is not provided, so developers are required to write their own iOS UI code using the related Software Development Kit (SDK) provided by the platform they working on. In this work, we designed our UI using Xcode, a native iOS Application Development Tool (ADT) with code written in either Objective-C or Swift 2.1. The requirements of J2ObjC were stated in (Ball, 2015), where users need a Mac workstation installed with Mac OS X version of 10.9 or higher. Besides that, users are also required to install Java Development Kit (JDK) with version 1.7 or higher on their Mac (Oracle, 2015). Lastly, users need to install the Xcode version 6 or higher on their Mac.

Before we get started, we need to download and unzip the distribution provided by Ball (2015), the version of J2ObjC we used for this work was 0.9.8.2.1. After we tested the translation with some example code we wrote ourselves following the guide located here, we started working on the translation of pairing libraries.

There are a total of four libraries introduced by Tan et al. (2010): *CpxBiginteger*, *Line*, *Point* and *Curve*. Our first attempt at direct translation did not succeed due to the library's dependencies of data types from each other. Hence, our following attempts were done by merging all four libraries into a single one and the translation was successful by following the J2ObjC guide. Two files were generated in *.m* and *.h* format respectively. *.h* file is the header file while *.m* file contains all the method declarations. Both are needed for our implementation of the IBS scheme using Xcode.

We used some exclusive Java objects (data types) in our original implementation in Java platform such as the *BigInteger* from *java.math* and *SecureRandom* from *java.security*. To reduce the complexities of our implementation, we decided to combine our IBS scheme written in Java with the merged library instead of finding third-party alternative Objective-C libraries that can provide similar functions. Now, we are only required to call the methods and pass in the desired input to test our IBS scheme.

The size of the translated library is much larger compared to the original Java libraries due to syntax differences between two programming languages. However, the beauty of J2ObjC is that no additional editing is needed to run our code.

3.2 Mobile application development using Xcode

Xcode is an integrated development environment (IDE) developed by Apple as a native software development tools for OS X and iOS. First of all, before we start to use our translated library, we need to link the J2ObjC to Xcode following the guide provided by Cheah et al. (2015), details as below:

1. Replace the J2ObjC's distribution directory to *$distribution-path*.
2. Select the project target in Xcode and click on the Build Rules tab.
3. Click the Add Build Rule button located at the bottom right of the panel.
4. Select "Java source files" in the new rule's Process option.
5. Add the following script in the Custom script text box:
   (a) *$distribution-path/j2objc -d ${DERIVED_FILES_DIR}*
   (b) *-sourcepath ${PROJECT_DIR}/$source-root*
   (c) *–no-package-directories ${INPUT_FILE_PATH};*
6. Click the + button in the Output Files panel, and add
   ${DERIVED_FILES_DIR}/${INPUT_FILE_BASE}.h
7. Click the + button again, and add
   $DERIVED_FILES_DIR/${INPUT_FILE_BASE}.m
8. Update the Build Settings:

   (a) In User Header Search Paths, add
      i. *$distribution-path/include and*
      ii. *${DERIVED_FILES_DIR}*
   (b) In Library Search Paths, add *$distribution-path/lib*
   (c) In Other Linker Flags, add *-ljre_emul*

9. Select Build Phases tab, open the Link phase and add:
   (a) The *libz.dylib* library
   (b) The Security Framework
   (c) The *libicucore.dylib* library [1]

10. Add the *.m* and *.h* files generated with J2ObjC to Xcode's project directories

The updated build settings for our project allows Xcode to apply the translated libraries by simply importing the *.h* header file. The JRE emulation library from J2ObjC emulates a subset of Java runtime libraries and is essential for the translated library to work properly.

Since our main goal is to test the simulation of IBS scheme by implementing it as an iOS mobile application, our UI design is relatively simple as shown in the left part of Figure 1. The right part of Figure 1 illustrates the completed UI. First of all, four methods were implemented to represent the four stages of the BLS-IBS scheme respectively. Execution of each method is represented by two buttons. One of the buttons executes the respective method a single time and shows the output in string format including the time taken for running a single stage. Another button will execute the respective method multiple times to calculate the average time taken. Figure 2 shows some blocks of code that represent the functions of the buttons of our application.

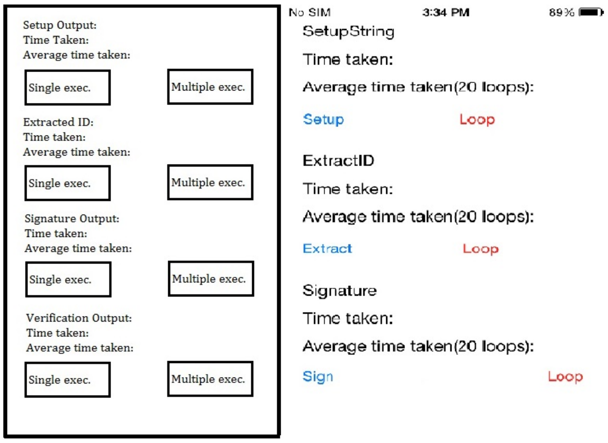
The flow of the program is following the IBS scheme discussed in Section 2. Thus, executing a specific method without following the scheme is not possible. For example, the execution of Extraction method is impossible without executing the Setup method beforehand because the parameters required for Extraction method are taken from the output of Setup method.

4 SIMULATION RESULTS

In this section, we will present the simulation results for running the BLS-IBS scheme using translated ECC pairing libraries as a mobile application.

*J2objcCurve_IBS_BLSSetupWithInt_withInt_* is the method called to run the Setup stage for IBS scheme. The passed arguments are 160 and 512 to represent the order and modulus bits of our elliptic curve respectively. This is equivalent to 1024-bit of RSA or discrete logarithms (DLOG) security.

[1] This is an extra step to link the JRE emulation library in step 8 above , thereby solving the linker error mentioned by Cheah et al. (2015)



Setup Output:
Time Taken:
Average time taken:

Extracted ID:
Time taken:
Average time taken:

Signature Output:
Time taken:
Average time taken:

Verification Output:
Time taken:
Average time taken:

SetupString

Time taken:

Average time taken(20 loops):

ExtractID

Time taken:

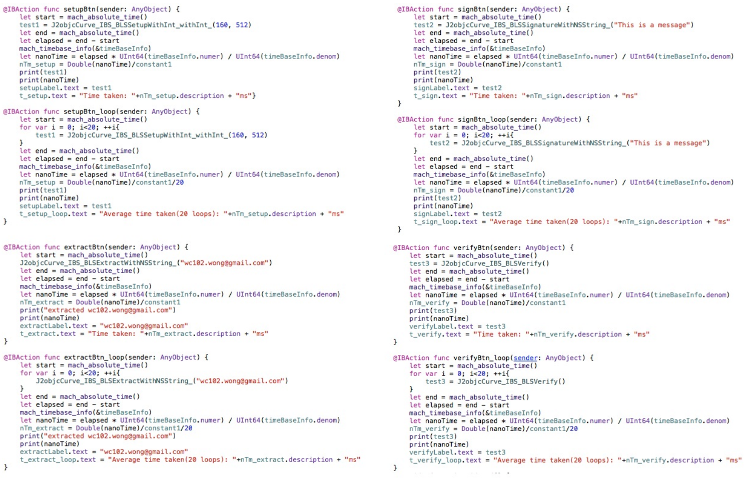
Average time taken(20 loops):

Signature

Time taken:

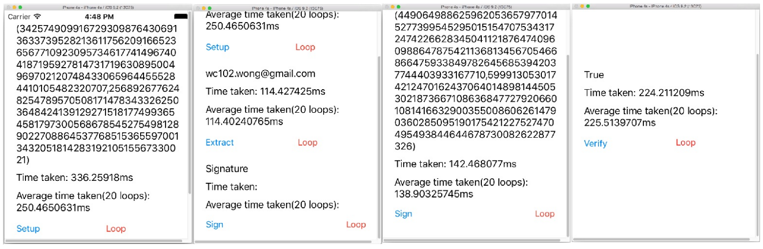
Average time taken(20 loops):

Figure 1: Mobile application UI design (left) and screenshot of iOS mobile application(right).



```
@IBAction func setupBtn(sender: AnyObject) {
    let start = mach_absolute_time()
    test1 = J2objcCurve_IBS_BLSSetupWithInt_withInt_(160, 512)
    let end = mach_absolute_time()
    let elapsed = end - start
    mach_timebase_info(&timeBaseInfo)
    let nanoTime = elapsed * UInt64(timeBaseInfo.numer) / UInt64(timeBaseInfo.denom)
    nTm_setup = Double(nanoTime)/constant1
    print(test1)
    print(nanoTime)
    setupLabel.text = test1
    t_setup.text = "Time taken: "+nTm_setup.description + "ms"}

@IBAction func setupBtn_loop(sender: AnyObject) {
    let start = mach_absolute_time()
    for var i = 0; i<20; ++i{
        test1 = J2objcCurve_IBS_BLSSetupWithInt_withInt_(160, 512)
    }
    let end = mach_absolute_time()
    let elapsed = end - start
    mach_timebase_info(&timeBaseInfo)
    let nanoTime = elapsed * UInt64(timeBaseInfo.numer) / UInt64(timeBaseInfo.denom)
    nTm_setup = Double(nanoTime)/constant1/20
    print(test1)
    print(nanoTime)
    setupLabel.text = test1
    t_setup_loop.text = "Average time taken(20 loops): "+nTm_setup.description + "ms"
}


@IBAction func extractBtn(sender: AnyObject) {
    let start = mach_absolute_time()
    J2objcCurve_IBS_BLSExtractWithNSString_("wc102.wong@gmail.com")
    let end = mach_absolute_time()
    let elapsed = end - start
    mach_timebase_info(&timeBaseInfo)
    let nanoTime = elapsed * UInt64(timeBaseInfo.numer) / UInt64(timeBaseInfo.denom)
    nTm_extract = Double(nanoTime)/constant1
    print("extracted wc102.wong@gmail.com")
    print(nanoTime)
    extractLabel.text = "wc102.wong@gmail.com"
    t_extract.text = "Time taken: "+nTm_extract.description + "ms"
}

@IBAction func extractBtn_loop(sender: AnyObject) {
    let start = mach_absolute_time()
    for var i = 0; i<20; ++i{
        J2objcCurve_IBS_BLSExtractWithNSString_("wc102.wong@gmail.com")
    }
    let end = mach_absolute_time()
    let elapsed = end - start
    mach_timebase_info(&timeBaseInfo)
    let nanoTime = elapsed * UInt64(timeBaseInfo.numer) / UInt64(timeBaseInfo.denom)
    nTm_extract = Double(nanoTime)/constant1/20
    print("extracted wc102.wong@gmail.com")
    print(nanoTime)
    extractLabel.text = "wc102.wong@gmail.com"
    t_extract_loop.text = "Average time taken(20 loops): "+nTm_extract.description + "ms"
}

@IBAction func signBtn(sender: AnyObject) {
    let start = mach_absolute_time()
    test2 = J2objcCurve_IBS_BLSSignatureWithNSString_("This is a message")
    let end = mach_absolute_time()
    let elapsed = end - start
    mach_timebase_info(&timeBaseInfo)
    let nanoTime = elapsed * UInt64(timeBaseInfo.numer) / UInt64(timeBaseInfo.denom)
    nTm_sign = Double(nanoTime)/constant1
    print(test2)
    print(nanoTime)
    signLabel.text = test2
    t_sign.text = "Time taken: "+nTm_sign.description + "ms"
}

@IBAction func signBtn_loop(sender: AnyObject) {
    let start = mach_absolute_time()
    for var i = 0; i<20; ++i{
        test2 = J2objcCurve_IBS_BLSSignatureWithNSString_("This is a message")
    }
    let end = mach_absolute_time()
    let elapsed = end - start
    mach_timebase_info(&timeBaseInfo)
    let nanoTime = elapsed * UInt64(timeBaseInfo.numer) / UInt64(timeBaseInfo.denom)
    nTm_sign = Double(nanoTime)/constant1/20
    print(test2)
    print(nanoTime)
    signLabel.text = test2
    t_sign_loop.text = "Average time taken(20 loops): "+nTm_sign.description + "ms"
}

@IBAction func verifyBtn(sender: AnyObject) {
    let start = mach_absolute_time()
    test3 = J2objcCurve_IBS_BLSVerify()
    let end = mach_absolute_time()
    let elapsed = end - start
    mach_timebase_info(&timeBaseInfo)
    let nanoTime = elapsed * UInt64(timeBaseInfo.numer) / UInt64(timeBaseInfo.denom)
    nTm_verify = Double(nanoTime)/constant1
    print(test3)
    print(nanoTime)
    verifyLabel.text = test3
    t_verify.text = "Time taken: "+nTm_verify.description + "ms"
}

@IBAction func verifyBtn_loop(sender: AnyObject) {
    let start = mach_absolute_time()
    for var i = 0; i<20; ++i{
        test3 = J2objcCurve_IBS_BLSVerify()
    }
    let end = mach_absolute_time()
    let elapsed = end - start
    mach_timebase_info(&timeBaseInfo)
    let nanoTime = elapsed * UInt64(timeBaseInfo.numer) / UInt64(timeBaseInfo.denom)
    nTm_verify = Double(nanoTime)/constant1/20
    print(test3)
    print(nanoTime)
    verifyLabel.text = test3
    t_verify_loop.text = "Average time taken(20 loops): "+nTm_verify.description + "ms"
}
```

Figure 2: Blocks of codes for button functions.



(342574909916729309876430691
3633739528213611756209166523
6567710923095734617741496740
4187195927814731719630895004
969702120748433065964455528
4410105482320707,256892677624
8254789570508171478343326250
3648424139129271518177499365
4581797300568678545275498128
9022708864537768515365597001
3432051814283192105155673300
21)

Time taken: 336.25918ms

Average time taken(20 loops):
250.4650631ms

Average time taken(20 loops):
250.4650631ms

wc102.wong@gmail.com

Time taken: 114.427425ms

Average time taken(20 loops):
114.40240765ms

Signature

Time taken:

Average time taken(20 loops):

(449064988625962053657977014
5277399545295015154707534317
2474226628345041121876474096
0988647875421136813456705466
866475933849782645685394203
7744403933167710,599913053017
4212470162437064014898144505
3021873667108636847727920660
1081416632900355008606261479
0360285095190175421227527470
495493844644678730082622877
326)

Time taken: 142.468077ms

Average time taken(20 loops):
138.90325745ms

True

Time taken: 224.211209ms

Average time taken(20 loops):
225.5139707ms

Figure 3: Result screenshots taken from iPhone 4s(iOS 8.4).

The method's running time is calculated by calling *mach_absolute_time()*. Results were taken in nanoseconds but converted to milliseconds afterwards by dividing the double type constant of 1,000,000.0. The loop button functions in a similar way except it is calling the methods 20 times to calculate the average time taken.

Devices	OS Version	Setup(ms)	Extract(ms)	Sign(ms)	Verify(ms)
Simulator(Mac)	OS X 10.11	250.4650631	114.4024077	138.9032575	225.5139707
iPhone 4s	iOS 8.4	2760.469677	1059.206292	1343.33111	2261.52975
iPhone 4s	iOS 9.2.1	2569.235804	1001.158285	1228.748246	2224.738048
iPhone 5s	iOS 9.2.1	998.4872354	301.9295063	313.1813958	588.9880083
iPhone 6	iOS 9.2.1	469.7018542	165.122175	186.0620521	353.6549333

Table 1: Runtime Results for Cha-Cheon IBS Scheme.

Other stages of the IBS scheme were implemented in a similar way as shown in Figure 2. For Extraction stage, we use an email as the ID input whereas some text that represent the message input were passed in for Signing stage. Inputs of both stages were passed in as strings. For Verification stage, the verify button will call the *J2objcCurve_IBS_BLSVerify()* method which returns a string of ”True” or ”False” depending on the validity of the tuple as explained in Section 2.

This simulation experiment was conducted on multiple iOS devices with different hardware specifications and also on the Xcode simulator from the Mac Mini we were using. The results of the experiment are listed in Table 1. Figure 3 consists of screenshots taken from various iPhone devices after running the experiment.

The translated library has been proven to have the same functionality as Tan et al. (2010)'s library. We also managed to obtain efficient results from the implementation as iOS mobile application. Our results also show that the simulator on Mac has the best results whilst the efficiency on mobile devices is determined by their hardware specifications. The upgrade on OS version only brings a slight improvement in performance.

5 CONCLUSION

In this work, we successfully developed a working mobile application simulator for Cha and Cheon's BLS-IBS scheme by using the translated Java based libraries from Tan et al. (2010). The translated library was proved functional by running the Setup, Extract, Sign and Verification algorithms. Reasonable running times were observed by running the simulation across multiple devices. Our future works include the extension of our implementation as a client-server signature verification mobile application.

ACKNOWLEDGMENTS

The authors are grateful to the Ministry of Education of the Government of Malaysia for partially funding this research under The Fundamental Research Project Scheme (No.: FRGS/2/2013/ ICT07/MMU/03/5) .

REFERENCES

Apple (2015). *Cocoa API - Implementations and Bindings*. https://en.wikipedia.org/wiki/Cocoa_(API).

Ball, T. (2015). *J2ObjC - Java to Objective-C Translator and Run-Time*. https://github.com/google/j2objc/releases.

Cha, J.-C. and Cheon, J.-H. (2003). An identity-based signature from gap diffie-hellman groups. In *ICCSA 2010*, volume 2567, pages 18–30.

Cheah, Z.-Y., Teh, T.-Y., Lee, Y.-S., and Chin, J.-J. (2015). Simulation of a pairing-based identity-based identification scheme in ios. In *ICSIPA 2015*, Pullman Hotel, Bangsar, Malaysia.

Cox, B. and Love, T. (2015). *Objective-C - Apple Development and Swift*. https://en.wikipedia.org/wiki/Objective-C.

International-Data-Corporation (2015). *International Data Corporation - Smartphone OS Market Share for Q2 2015*. http://www.idc.com/prodserv/smartphone-os-market-share.jsp.

Oracle (2015). *Oracle - Java SE Downloads*. http://www.oracle.com/technetwork/java/javase/downloads/index.html.

Pew-Research-Center (2015). *Technology Device Ownership Statistic*. http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015/.

Tan, S.-Y., Heng, S.-H., and Goi, B.-M. (2010). Java implementation for pairing-based cryptosystems. In *ICCSA 2010*, volume 6019, pages 188–198.

mPBC: An Efficient Pairing-Based Cryptography Library for Android

Syh-Yuan Tan[1], **Chee-Siang Wong**[2], and **Hoon-Herk Ng**[1]

[1]*Faculty of Information Science and Technology, Multimedia University, Malaysia*
[2]*Faculty of Information and Communication Technology, Universiti Tunku Abdul Rahman, Malaysia*

E-mail: sytan@mmu.edu.my, wongcs@utar.edu.my, malco_ng@live.com.my

ABSTRACT

There are numerous pairing-based cryptography (PBC) libraries available for desktop-based applications. However, these libraries are mostly not compatible or not optimized for mobile phone operating systems (OS) such as Android. In this paper, we show the inconsistency on benchmarking result for Java-based PBC libraries between Java Virtual Machine (JVM) and Dalvik Virtual Machine (DVM). Identifying the root cause, we present a new PBC library for Android, namely, mobile-PBC (mPBC) which combines the strengths of several Java PBC libraries and optimized for Android's DVM. The mPBC library outperforms the existing Java-based PBC libraries in DVM, yet as efficient as the fastest PBC library to date in the JVM. In particular, the pre-processed Tate pairing operation in mPBC uses 426.11ms in DVM (Samsung GT-N7000) and 4.50ms in JVM (Sager NP5160).

Keywords: Pairing-Based Cryptography, Elliptic Curve cryptography, Java, Android, Library

1 INTRODUCTION

There has been much interest in recent years in reducing the processing power of cryptographic schemes. Due to the rise of mobile platform and networking ability, a cryptography library which is suitable for mobile phone applications may help in increasing the efficiency of cryptographic computations. Even though there are a lot of cryptography libraries well developed for desktop-based applications, not much were done for mobile devices which have limited battery life and processing power.

Up to date, there are only a few elliptic curve cryptography libraries that support pairing operations written in Java (De Caro and Iovino, 2011, Dong, 2010, Tan et al., 2010) which work

Platform	JVM1	JVM2	DVM
Device	Sager NP5160	RaspBerry Pi II	Samsung GT-N7000
CPU	Quad-core 2 GHz Intel Core i7-2630M	Quad-core 900Mhz ARM Cortex-A7	Dual-core 1.4 GHz ARM Cortex-A9
RAM	4 GB	1GB	1 GB
OS	Windows 8.1	Raspbian Wheezy	Android 4.2.2
JDK/SDK Version	JDK 1.8.0.25	JDK 1.8.0.25	SDK 23.0.5

Table 1: Test Environment.

in both Dalvik Virtual Machine (DVM) and Java Virtual Machine (JVM). In 2014, Liu et al. (2014) ported the PBC library (Lynn, 2010) from C to Java by using Java Native Interface (JNI) and Android Native Development Kit (NDK). Compared to jPBC by De Caro and Iovino (2011), Liu et al. (2014) is a more complete Java version of PBC. However, Liu et al.'s library runs on Android 4.0.3 or higher and they did not manage to provide performance comparison with jPBC which can only be run on Android 2.2 or lower. This is the main drawback of using C libraries on Android in which they loss the platform independent feature. For instance, even though works by De Caro and Iovino (2011) and Liu et al. (2014) performed significantly faster than the other Java-based pairing libraries, their compilations are tied to the underlying hardware, e.g., the library compiled for ARM architecture cannot run on x86 architecture.

In short, one can view performance and platform independent as the trade-off to each other. On the other hand, we prefer to maintain the platform independent feature of the pairing libraries, so that pairing cryptosystems can run on any device that supports Java.

1.1 Motivation

Our motivation of this work came after the benchmarking on the Java pairing libraries in JVM and DVM. The benchmarks were done in a loop of 1100 times for JVM tests with the first 100 times neglected to avoid the caching of processor. The specification of our machines to carry out the target benchmarks are shown in Table 1.

Table 2 shows the timing performance (in nanoseconds) of the libraries under JVM1, JVM2 and DVM on the Type-A curve $y = x^3 + x$ mod p where p is a 512-bit prime with 160-bit Solinas prime order q (Lynn, 2010). If an operation is not available under a library, we mark it with the symbol '-'. For example, times of pre-processing multiplication and pre-processing pairing are recorded for jPBC which is the only library provides such features.

We discovered that existing Java PBC libraries perform differently in DVM (a register based virtual machine) as compared to JVM (a stack based virtual machine). For instance, Jpair (Dong, 2010) scores the highest among others (De Caro and Iovino, 2011, Tan et al., 2010) in point scalar multiplication under JVM1 and JVM2 but scores the lowest under DVM. Moreover, THG-

Table 2: Timing (ns) of Group Operations for Java PBC Libraries

Operations		JVM1			JVM2			DVM		
		THG-PBC	jPBC	Jpair	THG-PBC	jPBC	Jpair	THG-PBC	jPBC	Jpair
Negation		2,183	2,005	5,904	36,352	13,936	33,871	39,430	30,004	65,571
Addition		73,658	63,513	74,477	1,178,562	1,177,218	1,272,717	693,915	671,874	768,134
Doubling		74,737	63,094	75,467	1,250,657	1,255,628	1,368,083	1,107,691	1,081,931	1,200,287
D&A	Affine	15,394,156	-	-	282,096,335	-	-	269,322,363	-	-
	k-Affine	-	13,143,754	-	-	243,576,846	-	-	282,808,693	-
	Jacobian	-	-	5,853,007	-	-	177,507,191	-	-	523,765,047
P.P Mult.		-	1,902,335	-	-	34,925,697	-	-	34,246,110	-
Pairing		24,737,596	9,037,093	9,047,344	542,602,143	291,230,173	289,206,823	1,037,003,196	930,931,258	898,596,388
P.P. Pairing		-	4,426,967	-	-	143,932,848	-	-	466,375,265	-

PBC is slower than jPBC in JVM1 and JVM2 but faster in DVM despite the fact that jPBC uses k-bit Windows method which is theoretically faster. The results under JVM2 rule out the possibility of processor architecture as both RaspBerry Pi II and Samsung GT-N7000 use ARM-Cortex processors. The remaining causes of the inconsistencies may come from many aspects, ranging from the class structures to the virtual machine architecture.

1.2 Contribution

These inconsistencies indicate that the existing Java libraries (De Caro and Iovino, 2011, Dong, 2010, Tan et al., 2010) are not optimized for Android platform and inspired us to build an optimized PBC library primarily for Android's Dalvik Virtual Machine (DVM) with optimizations on Java Virtual Machine (JVM) come as a by-product. In this paper, we realize an optimized PBC library in Java, namely, mobile-PBC (mPBC), which outperforms the existing libraries (De Caro and Iovino, 2011, Dong, 2010, Tan et al., 2010) in Android's DVM.

1.3 Organization

The rest of the paper are organized as follows. In Section 2, we briefly discuss the algorithms involved in the point operations fo pairing based cryptography. In Section 3, we present the optimization techniques for mPBC in terms of programming approach and cryptography approach. Finally, we show and discuss the benchmark results of mPBC.

2 PRELIMINARIES

2.1 Point

A point P on elliptic curve $E(\mathbb{F}_p)$ under the finite field $\mathbb{F}_p$ with prime modulus p in the affine coordinate (x, y) can be represented in the format of Jacobian coordinate $(x/z^2, y/z^3, z)$ to get rid of the calculation on multiplicative inverse during point addition and point doubling.

2.2 Point Scalar Multiplication

THG-PBC uses Double and Add (Tan et al., 2010) while jPBC (De Caro and Iovino, 2011) and Jpair (Dong, 2010) use Double and Add with k-bit window as the point scalar multiplication algorithm.

2.2.1 Double and Add Algorithm

THG-PBC (Tan et al., 2010) and jPair (Dong, 2010) use the Double and Add algorithm as shown in Algorithm 1 as the point scalar multiplication. The difference between these two libraries is that the former use the algorithm in Affine coordinate while the latter is in Jacobian coordinate.

Algorithm 1 Double and Add

```
Require: $m, P$
Ensure: $Q = mP$
Q ← P
for all $i \leftarrow (\lg(m)) - 2$ to 0 do
    $Q \leftarrow 2Q$
    if $i$ is 1 then
        $Q \leftarrow Q + P$
    end if
end for
return $Q$
```

2.2.2 Double and Add Algorithm with *k*-bit Window

jPBC (De Caro and Iovino, 2011) uses an enhanced version of Algorithm 1, namely, Double and Add with k-bit Window. The parameter k defines the lookup table size which contains the precomputed points.

3 OPTIMIZATION TECHNIQUES

In this section, we examine the possible causes of the bizarre benchmarks and present the optimization techniques used in producing the mPBC library.

3.1 Global Methods

At the first glance, the performance of Jpair in JVM is benefited from its class structure which pushes all point operations to the `Curve` class and yields the global methods for `Point` objects. This may indicate that the execution of a global method is faster than a local method in JVM but the result is conversed in DVM. However, our quick experiments showed that this hypothesis is not true. The result showed that global methods work better than local methods in DVM by approximately 10% but no significant difference in JVM. This concludes that global variables help in reducing the execution time, but not as significant as shown in Table 2.

3.2 Point Operations

The optimizations on point operations can be categorized into two main categories, namely, mathematical enhancements and coding optimizations. For the latter, we code the library in such a way that global variables will be utilized whenever it is possible to avoid the creation of local variables; for the former, we make use of bit operation functions supported in `java.math.BigInteger` to replace some simple yet repetitive mathematical operations;

3.2.1 Point Negation

Point negation is a simple point operation that negates the value y under $\mathbb{F}_p$. The implementations of most of the libraries are $-y \bmod p$, but this can be done more efficiently by calculating $p - y$. In the programming aspect, the point negation function `negate(BigInteger p)` can be optimized by removing the point validations (i.e. check $y < p$) because this function is always called from a valid `Point` object generated by the `Curve` object.

3.2.2 Point Addition and Doubling

In mPBC, we optimize the computation through the use of methods from `BigInteger` libraries to reduce the execution time in DVM. For example, we replace `equals(BigInteger.ZERO)` with `signum()` method which is able to check for zero values with reduced time.

For the algorithms which execute point addition and point doubling frequently, we set the temporary variables as `static` where these variables will be constructed only once. The reuse of temporary variables will save the constructions of new variables at each round of point addition and point doubling. Furthermore, the point addition and point doubling were placed under the `Curve` class as global methods. Additionally, the multiplication and exponentiation operation of `BigInteger` objects are also optimized by using bitwise operation such as `shiftLeft(int n)` whenever possible.

3.2.3 Point Scalar Multiplication

It is well known that point operation in Jacobian coordinate is faster than that of Affine coordinate but inconsistency of benchmark results were identified as well between Double and Add algorithm with k-bit Window and the original Double and Add algorithm, i.e., without k-bit Window in Table 1.1. Under JVM, Jpair's point multiplication is the fastest followed by jPBC's and lastly THG-PBC's; while in DVM, the sequence is totally reversed. It is obvious that the number of temporary variables is directly proportional to the speed of scalar multiplication algorithm in JVM, but inversely proportional to that in DVM.

In view of this, we modified point addition algorithm in Jacobian coordinate in such a way

Algorithm 2 Optimized Point Addition in Jacobian Coordinate

```
Require: $P, Q$
Ensure: $R = P + Q$
$t_1 \leftarrow x_1 z_2^2$
$t_2 \leftarrow x_2 z_1^2$
$t_3 \leftarrow t_1 - t_2$
$t_4 \leftarrow y_1 z_2^3$
$t_5 \leftarrow y_2 z_1^3$
$t_1 \leftarrow t_1 + t_2$
$t_2 \leftarrow t_4 - t_5$
$x_3 \leftarrow t_2^2 - t_1 t_3^2$
$y_3 \leftarrow (t_1 t_3^2 - 2x_3 t_2 - t_4 + t_5 t_3^3)/2$
$z_3 \leftarrow z_1 z_2 t_3$
return $R = (x_3, y_3, z_3)$
```

that the first three `BigInteger` variables are taken from the global variables which are shared by all the algorithms in the `Curve` class for calculation purposes. Moreover, by manipulating the algorithm steps, we can reduce four temporary variables from nine to five as shown in Algorithm 2. Therefore, in the implementation, we only need to create two temporary variables instead of nine as in the original algorithm.

3.2.4 Pre-Processing Point Scalar Multiplication

Pre-processing point multiplication is only found in jPBC (De Caro and Iovino, 2011) which uses more memory to construct a pre-processing table, on top of the lookup table of k-bit Window. In mPBC, the pre-processing algorithm used is similar to that of jPBC with optimization done in the programming aspect as discussed.

3.3 Bilinear Pairing

The bilinear pairing of mPBC relies on the Type-1 pairing algorithm used by Jpair (Dong, 2010) which is the fastest in DVM compared to the other two. Thus, we optimize the pairing operation by using the faster point addition and point doubling algorithm as discussed previously.

3.3.1 Pre-Processing Bilinear Pairing

Given that the point addition and point doubling were executed in the pre-processing phase, it does not help in reducing the actual computation time of the final pairing operation. In mPBC, the pre-processing algorithm used is similar to that of jPBC with optimization done in the programming aspect as discussed.

3.4 Class Structure

mPBC consists of seven classes only to keep the library simple, namely, `Curve`, `CpxBigInteger`, `Point`, `JcbPoint`, `PointMulPreProcessing`, `Pairing`, `PairingPreProcessing`. Among all, only `JcbPoint` inherits `Point` and there is no inheritance relationship among other classes.

We reassert that all optimizations are targeted on the performance in DVM, while the improvements in JVM come as byproduct.

4 BENCHMARKS

In this section, we show the benchmarks for mPBC in Table 3 with the same environment as in Section 1.1 where 'J' represents Jacobian and 'A' represents Affine.

Table 3: Timing (ns) of Group Operations for mPBC.

Operations		mPBC		
		JVM1	JVM2	DVM
Negation		899	13,870	17,811
J-Addition		12,475	479,858	603,285
J-Doubling		12,665	487,670	962,549
A-Addition		64,353	1,215,114	523,073
A-Doubling		63,444	1,248,487	592,711
D&A	Affine	15,591,693	283,056,439	310,566,204
	k-Affine	13,023,479	242,890,806	268,979,259
	Jacobian	5,199,897	169,289,050	443,377,476
	k-Jacobian	4,014,078	130,038,010	372,981,140
P.Proc. Mult.		1,937,780	34,740,619	26,003,816
Pairing		8,429,780	299,451,947	809,407,243
P.Proc. Pairing		4,501,117	162,697,997	426,115,736

4.1 Discussion

From Table 3, we can see that the more variables needed, the slower the algorithms execute in DVM. This explain why Jacobian coordinates system performs well in JVM but affine coordinates system performs well in DVM. Unlike JVM, Android's DVM uses garbage collector to dispose the unused variables (And) and the disposal cost is greater than speed-up gained from the gap of algorithm complexity between Jacobian coordinates and affine coordinates.

Secondly, simple class structure delays lesser than complex class structure in DVM. Although it is confirmed from Table 3 that Double and Add algorithm is faster with the presence of k-bit Window in DVM, some may noticed from Table 2 that THG-PBC uses Double and Add algorithm without k-bit Window but it is faster than that of jPBC. The cause of this situation is not due to the algorithm used, but the class structures of the two libraries. THG-PBC comprises of only four classes without any inheritance or interfaces but jPBC on the other hand, uses a relatively complex class structures.

Thirdly, the use of global methods can speed up the execution time in DVM. Every point operations in Jpair are the slowest among all but its pairing operation is still the fastest after mPBC. This is because the multiplication under $\mathbb{F}_{p^2}$ in Jpair is done by the `ComplexField` object, which provides all the field operations. So, instead of calling the multiplication method from a new $\mathbb{F}_{p^2}$ object, DVM calls the same `ComplexField` object in every round of the Miller algorithm.

ACKNOWLEDGEMENTS

This work is supported by the FRGS grant (FRGS/2/2014/ICT04/MMU/03/1).

REFERENCES

Android developers: Performance tips. `http://developer.android.com/training/articles/perf-tips.html`.

De Caro, A. and Iovino, V. (2011). jpbc: Java pairing based cryptography. In *Proceedings of the 16th IEEE Symposium on Computers and Communications, ISCC 2011*, pages 850–855, Kerkyra, Corfu, Greece, June 28 - July 1. IEEE.

Dong, C. (2010). Jpair: A quick introduction. `https://personal.cis.strath.ac.uk/changyu.dong/jpair/intro.html`.

Liu, W., Liu, J., Wu, Q., and Qin, B. (2014). Android pbc: A pairing based cryptography toolkit for android platform. In *Communications Security Conference (CSC 2014), 2014*, pages 1–6.

Lynn, B. (2010). The pairing-based cryptography library. `http://crypto.stanford.edu/pbc/`.

Tan, S.-Y., Heng, S.-H., and Goi, B.-M. (2010). Java implementation for pairing-based cryptosystems. In *Computational Science and Its Applications ICCSA 2010*, volume 6019 of *Lecture Notes in Computer Science*, pages 188–198. Springer Berlin Heidelberg.

P2P Botnets Detection Module through Hybrid Approach

Raihana Syahirah Abdullah^{*1}, **Faizal M.A.**^{1,2}, and **Zul Azri Muhamad Noh**^{2}

^{1}*Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka*

E-mail: raihana.syahirah@utem.edu.my, faizalabdollah@utem.edu.my,
zulazri@utem.edu.my
**Corresponding author*

ABSTRACT

Botnets or advance malware is a high-profile cyber-criminal. Arise of latest botnets or advance malware cause crisis and chaos to network security. The application of combining botnets and P2P technology is powerful but complicated. P2P botnets that accurately has abnormal traffic behaviours highly impact the networks operation, network security and cost financial losses. In order to detect these botnets, a complete flow analysis is necessary. In this paper, we proposed a new P2P botnets hybrid detection module which currently focuses on P2P based botnets. We consider both the coordination within a botnets and the malicious behaviour each bot exhibits at the host or network level and propose the hybrid detection module that combines host-based and network-based behaviour for making detection decisions. The hybrid approach have high detection accuracy and low false positive.

Keywords: Botnets, P2P Botnets, Hybrid, Host-based,Network-based

1 INTRODUCTION

Botnets is the latest scourge for the security crisis that criminally threat and attack the network to steal the data (Chandrashekar, 2009) (Tyagi and Aghila, 2011). At least 85% of overall traffic on the Internet sent through the spamming botnets have been reported from recent studies (Stringhini et al., 2011). The threat posed by botnets has become increasingly high profile criminal issue in the past several years. The issue of botnets has received high attention around the world as it becomes biggest threat to the internet stability and security (Diptamdutta, 2010). Meanwhile, P2P technology become more and more popular in the past decades and it has widely applied in file sharing and media streaming. Modern malwares also started shifting towards P2P architecture which provides a better resiliency against detection and take down. Due to its high efficiency and robustness, the P2P botnets is harder to locate, shutdown, monitor and hijack (Tyagi and Aghila, 2011).

Virtually, all organizations face increase threats to their networks and the services that they provide and this will lead to network security issues as mentioned by Li et al. (2010). This statement has been proven by the increasing number of computer security incidents related to malicious codes from 3,688 in 2009 to 11,427 in 2012 as reported by Computer Emergency Response Team, Cyber Security Malaysia (Malaysia, 2013) as illustrated in Table 1 (refer to Appendices), CSI (2011) also claimed that the botnets infections are the second major rank which is at 29% after the malware infections. Meanwhile, Cyber Security Malaysia (Malaysia, 2013) has also reported that the botnets drones has the third highest percentage of security incidents which is at 15.5%. The botnets drones has increase rapidly by more than one million case are reported. The vulnerability report by CERT/CC has visualized that botnets attack has generated significant worldwide epidemic to internet stability, network security and resulted in huge financial losses. The rest of paper is organized as follows. In Section II, we provides details background on the hybrid detection previous research. Section III will describes the methodology of overall process. Next, Section IV will provides details detection module of our proposed P2P botnets hybrid detection module and all its components. While, the results and discussion are discuses in Section V and at last our paper is concluded in Section VI.

2 BACKGROUND

Currently, the highlight problem in P2P botnets detection is to identify the continuous P2P traffic characterization for differentiate between the P2P normal and P2P botnets. The identification of malicious activity in P2P traffic is a tedious task due to network violations which make it more rigidity to recognize the P2P botnets characteristics and behaviours. Al-Hammadi and Aickelin (2010) claim botnets can be the latest challenge for IT personnel in analysing the P2P botnets network traffic on looking for signature which practically a tedious task because bots signatures can be either dynamic or encrypted. Hence, it is necessary to introduce a new technique to P2P botnets detection so that it can distinguish the normal and anomalous behaviours with low false alarm rates. In time, there are many of previous researches have concentrate on general P2P detection module and implement the P2P botnets detection. Through the critical review, the main weaknesses of previous modules are the failure to reveal bot server and classify the botnets (Zeidanloo, 2010),(Zeng et al., 2010), (Zhang et al., 2011), (Arshad et al., 2011), (Muthumanickam and Ilavarasan, 2012). Otherwise, the lack occurred when previous research is focused only on network traffic, ignoring the mined data process, ignoring the filtering agent, and does exclude the specific information for botnets detection. The entire weaknesses have been discussed cause to inefficiency of their module especially for powerful and robust bonnets.

Hence, this paper finds the previous modules are not comprehensively enough to recognize arise of latest botnets. Besides, there is no such module able to detect and classify the latest botnets. Yet, the previous module are still immature and have lots of drawbacks which require an improvisation. In order to bridge the gap, this paper proposes robust module using hybrid approach to detect and classify latest botnets. Differ with previous module, this propose module is not rely on a single detection analyser either in host analyser or network analyser. It provides the hybrid analyser (combination of host-based analyser and network-based analyser) and hybrid analysis (combination of static analysis and dynamic analysis). Moreover, lots of hybrid-based study has been done by previous researchers such as modules, approaches and techniques in order to enhance and empower the ability and effectiveness of their research. However, arise of botnets or advance malware make them harder to be detected and shut downed (Zeidanloo, 2010), (Broersma, 2007). In addition, botnets have been declared as quick evolving problems, the most emergent threat and high profile attack to the security (Diptamdutta, 2010). Zeidanloo et al. (2010) expects the crisis of continual botnets growth in future that requires update and powerful detection and protection techniques. Thus, this research aims to develop a new P2P botnets detection module using hybrid approach to detect the latest P2P botnets. Aware with limited resource where only one approach was practiced in previous research, this research will focuses on introducing the new technique with association of hybrid analyser, hybrid analysis and hybrid techniques. Another concern of this paper is to propose detection module for P2P botnets will help to solve the issue of limited detection technique in order to detect anomalous P2P botnets. So, the details process flow for P2P botnets detection module will be discovered in next section.

3 METHODOLOGY

This section discusses in detail the process involved for the P2P botnets detection as depicted in Figure 1 (refer to Appendices). The process started with capturing the raw of P2P traffic and read with TCP dump and Wireshark application. Both of the applications read the P2P traffic that has been captured with various variants used in the analysis and testing phase. As shown in Figure 1, the captured P2P traffic in monitoring module is passed thru filtering module for pre-processing process to select appropriate data to be used. In filtering module, the unnecessary data automatically will be removed and cleaned up. The output from filtering module then proceeded to analyser module to be analysed separately either in host or network analyser. Followed by that, the output from hybrid analyser is used to distinguish the behaviours on P2P normal and P2P botnets for each variant. Moreover, the attributes influence for P2P botnets is revealed in this module. After identification of the P2P botnets behaviours, it will be transferred into classification and detection module. Each of the variant has their own attributes to detect their malicious activities. If malicious activities are detected, then the activities will declared as containing the intrusion activities in report module. By referring to this module, the next section will extensively discusses in detail the implementation of improved P2P botnets module and technique.

4 PROPOSED P2P BOTNETS DETECTION MODULE AND COMPONENTS

This paper presents an approach to detect and classify P2P botnets activity through constructing a new hybrid module on combined host-based and network-based level. The implementation of the module is considered by preliminary analysis study that had been done by implementing the P2P botnets testbed environment. The study had analysing original codes of latest P2P botnets in the wild. The most important part in this preliminary analysis study is the dataset had been obtained from two different based data sources: first is host data (host-based analyser), second is network data (network-based analyser). In host-based, the data gathered from system command, system accounting, system log and security log. While the network-based get the data from full payload network packet that analyses focus on network segments and application of protocol activity.

This hybrid technique is used as the backbone in proposing the implementation of P2P botnets detection technique for modelling the intrusion report. As illustrates in Figure 2 (refer to Appendices), the process flow of P2P botnets detection began with mining the data through the pre-processing technique indicates by SVM classifying. Then, the host log and network traffic will be analysed through the signature-based technique by rule-based module. If a known pattern is matched, the attack alarm then will be generated. The statistical tests declare as anomaly-based revealed on anomalous volume that perform as second detection for the unknown intrusion events. The detection is combining the steps on Figure 2 that mainly indicate this procedure as hybrid detection technique. The hybrid technique combines the signature and anomaly based techniques that combines with data mining technique to detect the intrusions. The three layers of detection methods have offered their own advantages and disadvantages. The inspiration for using hybrid technique leads to prevail over the limitations of individual technique. Thus, the distinction has profound the positive implications for cyber security salvation. Additionally, optimization of accuracy and detection rate can be achieved by using the hybrid detection techniques (Bao et al., 2009).

The P2P botnets hybrid detection technique layer indicates the Signature Generation Module and Statistical Test Module. Theoretically, the hybrid technique is a mixture of techniques on improvised the weaknesses and limitation of previous theory which boosting up the ability to double the vulnerabilities towards evaluating the analyser and detection levels. In the other words, hybrid has been chose to multiply their abilities. The implementation of hybrid technique used to achieve maximum accuracy, higher effectiveness and greater efficiency in detection rate evaluation. The hybrid technique has using the combination on data mining-based, signature-based and anomaly-based as depicted in Figure 3 (refer to Appendices). This hybrid technique has encountered the attributes relate on behaviour features either in host-level or network-level. Then, the hybrid technique introduces significant improvement in P2P botnets detection technique. According to our opinion, the module is divides into three main stages: Mining data, Signature-based Detection and Anomaly-based Detection. The proposed hybrid module is introduced in Figure 3 (Appendices). This module incorporates with three main components as a hybrid. This combination successfully complement each other in the real world of P2P botnets traces. The components are:

4.1 Stage 1: Mining Data

A data pre-processing technique is applied to ensure the smoothness operation of the experiments (Chitrakar and Chuanhe, 2012). It is set as a data mining that exclusively denote by pre-processing stage. The selection has been done in selecting the useful attributes and consumed most of the time process. The dataset from both of combined host-based and network-based are mined through the pre-processing stage. The data reduction and data discretization that have been applied in pre-processing stage able to reduce the representation in analytical and numerical results. Overall, the major task implicates in mining data layer is SVM classifying concept that capable to classify every new unseen event by building the classification model for normal and abnormal events based on label training data. Thus, the mined data using SVM classification is composed by Algorithm 1 (refer to Appendices).

4.2 Stage 2: Signature-based Detection

Generally, signature-based is a supervised learning method in detecting the malicious behaviour on the basis of previously seen malicious events. The second stage of P2P botnets detection module as depicted in Algorithm 2 (refer to Appendices) is the signature-based detection that make detection on known attack. The detection signature has been composed through analysis part. The P2P botnets detection technique has been developed as signature module. The network traffic will be analysed through this stage by rule-based module, thus generating an attack alarm if a known pattern is matched. This technique has the capability to analyse the malicious activities described as variant behaviours, make classification on P2P variant types and sub-attack types and generate the conclusion either the submitted files is a P2P normal or P2P botnets indeed. This technique also able to produce the report on activities performed by P2P botnets.

4.3 Stage 3: Anomaly-based Detection

Otherwise, if the pattern is not recognized in Stage 2, it will be processed in this stage through raise an "anomaly' alarm that allow as second detection for the unknown intrusion events. Even though the signature-based detection has been completely done but, several of undetectable P2P botnets are noticeably existed. This situation happens due to the capability of signature-based where it can only detecting the known attack instead of the unknown attack. Standing on the fact, the anomaly-based detection is alternately necessary to conquer this problem. The combination of signature-based and anomaly-based are technically complement of each other weaknesses.

As a result, this paper presents an anomaly detection technique in Stage 3 as portray in Algorithm 3 (refer to Appendices) based on the chi-square statistic. In this cases, if the pattern is not recognized in Stage 2 which is the signature-based module, then it will be processed through raise an "anomaly' alarm that allowed as second detection for the unknown intrusion events. This technique is tested for defining its performance in distinguishing normal events from intrusive events in each variant. The study also reveals that the multivariate statistical technique based on the chi-square test statistic indicates the intrusive events are detected as unknown attack. In the above context of detection, the developing techniques based on Hybrid-SAM, the combination of signature-based, anomaly-based detection technique with data mining-based proposed by Robiah et al. (2009) has been comprehensive approach to fight against botnets threat in the real world situation. It is because the combination of this two techniques have complementary each other in dealing both known and unknown botnets including detection on encrypted bot, reduce false positive and negative alert, real-world detection and reveal the bot C&C servers.

The principal step in data mining is the selection of appropriate features from the data. The selection has been done in selecting the useful attributes. This selection is done at the pre-processing stage of data mining process. Pre-processing stage consumes most of the time process. Signature-based has the ability to immediate detection and impossibility of false positives. But signature-based is only capable to be used for detection of well-known botnets. More important, very similar bots with slightly different signature may be missed-out to be detected. However, the anomaly-based technique faced with the problem of detecting unknown botnets through show existence of bots in the network. Anomaly-based technique also has the extra capabilities in terms of reducing false negative alert and detecting multistep attack (Robiah et al., 2009). Nevertheless, it cannot reduce the false positive alert which can only be reduced by using signature-based technique. Hence, this has given an implication that there are complement each other weaknesses. The fully results are briefly discusses in the next section.

5 RESULT AND DISCUSSION

In order to validate the propose module efficiency and effectively detects the P2P botnets, we respectively validate the results by three evaluation metrics; (i) Accuracy, (ii) Detection Rate and (iii) False Alarm Rate. The evaluation metrics that have been generated are recorded respectively. This is because the accuracy of detection, detection rate and false alarm rate are the performance criterion that need to be evaluated. Then, the results for each of evaluation metrics are calculated to find the most appropriate detection technique for P2P botnets activities. The percentage for each metrics is calculated to observe the performance of the systems. The lower and higher percentage in these evaluation metrics will directly reflect to the effectiveness and efficiency of the new P2P botnets detection technique.

Refer to Table 2 (refer to Appendices) showed that the overall detection rate are exceeded more than 90%. The higher value of percentage means the classification test has better prediction and capabilities to distinguish the difference between the attack and normal P2P traffic. The percentage result of overall detection rate indicated that the test is suitable, fit and good in predicting the outcome of attack. However, the percentage of false negative is very high indicated dangerously many attacks are still failed to detect. As a conclusion, a comprehensive improvement will be done with another detection method in the next stage to detect these anonymous attack accurately. The details of hybrid detection method will be discussed further in the next sub-section within the detection layer module. While, the result from Table 3 showed that the signature-based detection has the capabilities to predict 100% correctly for the overall detection rate with 0% of false alarm rate of the P2P network traffic. The improvement of overall detections in the signature-based module from classification table in data mining module are indicated that this signature-based system technically effective for outcome attack detection. Therefore, it can be summarized that this signature-based detection has better prediction and capabilities to distinguish between the normal and attack events reached for thousands of dataset for each variant.

Inclusively, this signature-based detection system in this paper promises the most sophisticated enhancement in P2P botnets detection technique. The entire six variants have fully detected as the P2P botnets based on the detection result. But, Table 3 (refer to Appendices) shows the detail of the result where the False Negative (FN) emphasize some of the undetectable attributes or undetectable P2P botnets values as the attack declares as normal. Alternately, this problem can be tackled by conducting the anomaly-based detection. In the next of detection stage, the chi-square statistical test with multivariate process has been perform. The tabulated of false negative that indicates undetectable P2P botnets has been proves can be successfully detected through the statistical approach. The statistical approach that applies in anomaly-based detection has proved that the undetectable P2P botnets in signature-based module can be detected through this approach. The result in Table 4 (refer to Appendices) shows the P2P botnets can be detected effectively in the anomaly-based rather than signature-based result. The false negative concerns as the undetectable numbers of attack that fail to be detected in signature-based module. This situation happened when the unknown attack has been detected normal. Significantly, the unknown attack is tackles by conducting the anomaly-based detection whereby the chi-square statistical test with multivariate process has been performed. The classification of attack that has been selected through the tremendous anomalous volume on the dataset which alerting of P2P botnets symptom. The classification of attack practically perform at the analysis part locates on host log and network packet.

Additionally, the result shows that detection not only capable to identify the undetectable value in signature-based but also the statistical test able to detect more than predictable value in anomaly-based. This outcome demonstrates that the others unknown attack also has been successful detected. The incremental of the effectiveness towards the combination of detection techniques known as hybrid technique with the hybrid approach help on boosting the detection values. At the same time, the hybrid technique with the hybrid approach will complementary the weakness and integrate the best result. Other than that, result proves that the correlation between anomaly-based and signature-based are essentially needed and relevantly to be used in detecting the P2P botnets.

6 CONCLUSION

Currently, the technique or approach that has been chosen by most of researchers are not comprehensive enough. This study presents a new module to detect P2P botnets. The proposed detection module is based on combination of data mining-based, signature-based and anomaly-based. The result shown the proposed detection module have high detection accuracy with ability to detect unknown P2P botnets and produce a high detection rate with low false alarm rate. Hence, the developing detection module based on hybrid-based technique with combined host and network level has been the most promising approach to fight against botnets threat in the real world situation.

ACKNOWLEDGMENTS

The authors would like to express the appreciation to Inforsnet Group Research of Universiti Teknikal Malaysia Melaka (UTeM) and MyBrain15 Programme by Ministry of Higher Education Malaysia (MoHE) in encouraging the authors to publish this paper. This work was supported by the MoE of Malaysia under Grants FRGS/1/2014/ICT04/FTMK/02/ F00212.

REFERENCES

Al-Hammadi, Y. and Aickelin, U. (2010). Behavioural correlation for detecting p2p bots. In *Future Networks, 2010. ICFN'10. Second International Conference on*, pages 323–327. IEEE.

Arshad, S., Abbaspour, M., Kharrazi, M., and Sanatkar, H. (2011). An anomaly-based botnet detection approach for identifying stealthy botnets. In *Computer Applications and Industrial Electronics (ICCAIE), 2011 IEEE International Conference on*, pages 564–569. IEEE.

Bao, X., Xu, T., and Hou, H. (2009). Network intrusion detection based on support vector machine. In *Management and Service Science, 2009. MASS'09. International Conference on*, pages 1–4. IEEE.

Broersma, M. (2007). Botnets getting harder to kill.

Chandrashekar, J. (2009). The dark cloud: Understanding and defending against botnets and stealthy malware. *Managing Editor*, 13(2):130.

Chitrakar, R. and Chuanhe, H. (2012). Anomaly detection using support vector machine classification with k-medoids clustering. In *Internet (AH-ICI), 2012 Third Asian Himalayas International Conference on*, pages 1–5. IEEE.

CSI (2011). 15th annual 2010/2011 csi computer crime and security survey. *Computer Security Institute*, pages 1–44.

Diptamdutta (2010). Botnets.

Li, X., Duan, H., Liu, W., and Wu, J. (2010). The growing model of botnets. In *Green Circuits and Systems (ICGCS), 2010 International Conference on*, pages 414–419. IEEE.

Malaysia, C. S. (2013). Mycert incident statistics.

Muthumanickam, K. and Ilavarasan, E. (2012). P2p botnet detection: combined host-and network-level analysis. In *Computing Communication & Networking Technologies (ICCCNT), 2012 Third International Conference on*, pages 1–5. IEEE.

Robiah, Y., Rahayu, S. S., Zaki, M. M., Shahrin, S., Faizal, M., and Marliza, R. (2009). A new generic taxonomy on hybrid malware detection technique. *International Journal of Computer Science and Information Security (IJCSIS)*, 5.

Stringhini, G., Holz, T., Stone-Gross, B., Kruegel, C., and Vigna, G. (2011). Botmagnifier: Locating spambots on the internet. In *USENIX Security Symposium*.

Tyagi, A. K. and Aghila, G. (2011). A wide scale survey on botnet. *International Journal of Computer Applications*, 34(9):10–23.

Zeidanloo, H. R. a. A., A. (2010). Botnet detection by monitoring similar communication patterns. *International Journal of Computer Science and Information Security*, 7.

Zeidanloo, H. R., Shooshtari, M. J. Z., Amoli, P. V., Safari, M., and Zamani, M. (2010). A taxonomy of botnet detection techniques. In *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on*, volume 2, pages 158–162. IEEE.

Zeng, Y., Hu, X., and Shin, K. G. (2010). Detection of botnets using combined host-and network-level information. In *Dependable Systems and Networks (DSN), 2010 IEEE/IFIP International Conference on*, pages 291–300. IEEE.

Zhang, J., Perdisci, R., Lee, W., Sarfraz, U., and Luo, X. (2011). Detecting stealthy p2p botnets using statistical traffic fingerprints. In *Dependable Systems & Networks (DSN), 2011 IEEE/IFIP 41st International Conference on*, pages 121–132. IEEE.

APPENDICES

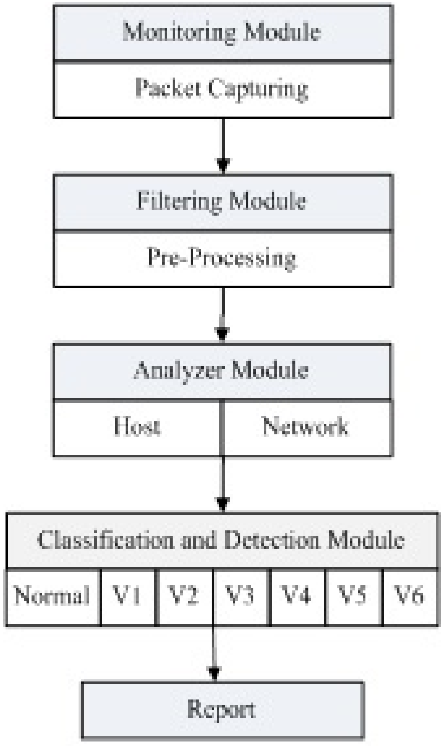


Figure 1: Process Flow of P2P Botnets Detection

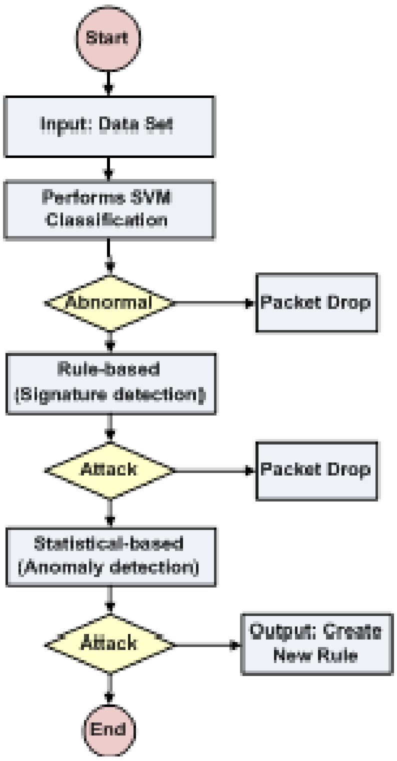

Start
Input: Data Set
Performs SVM Classification
Abnormal
Packet Drop
Rule-based (Signature detection)
Attack
Packet Drop
Statistical-based (Anomaly detection)
Attack
Output: Create New Rule
End

Figure 2: Process Flow of Overall P2P Botnets Detection

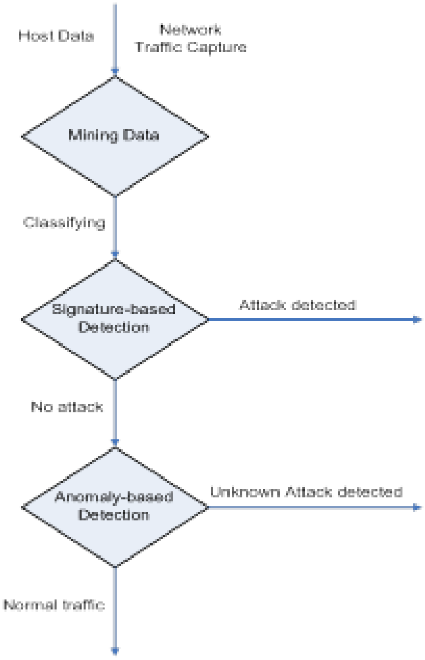
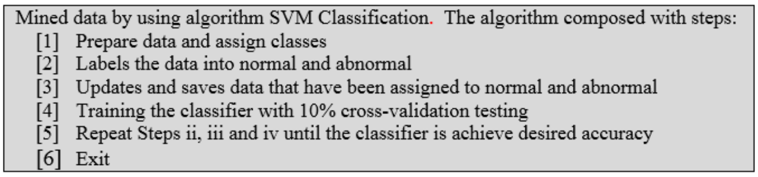
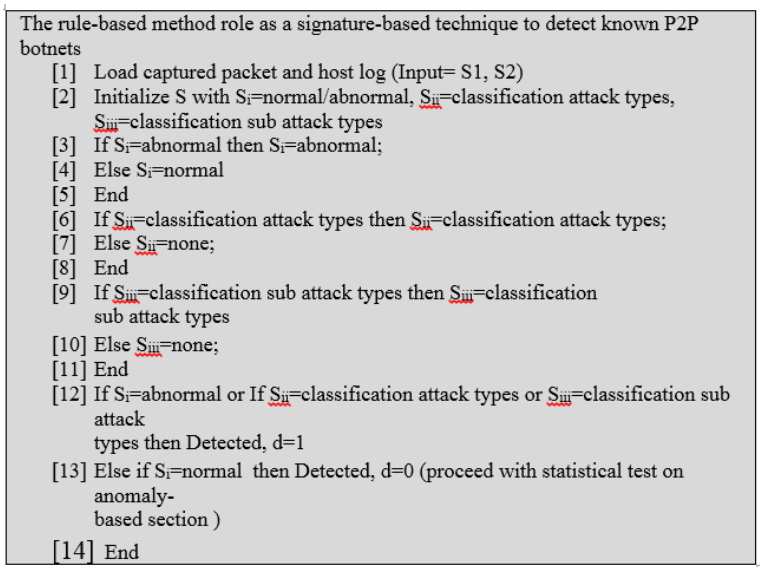


Figure 3: Proposed P2P Botnets Hybrid Detection Module Combined Host-Network Based



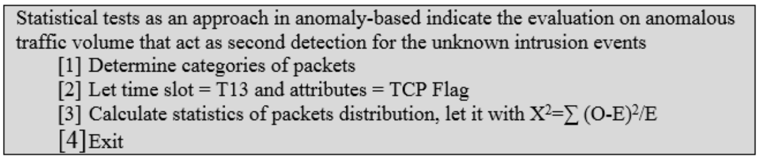
```
Mined data by using algorithm SVM Classification. The algorithm composed with steps:
    [1] Prepare data and assign classes
    [2] Labels the data into normal and abnormal
    [3] Updates and saves data that have been assigned to normal and abnormal
    [4] Training the classifier with 10% cross-validation testing
    [5] Repeat Steps ii, iii and iv until the classifier is achieve desired accuracy
    [6] Exit
```

Figure 4: Algorithm 1- SVM Classification



```
The rule-based method role as a signature-based technique to detect known P2P
botnets
    [1]  Load captured packet and host log (Input= S1, S2)
    [2]  Initialize S with $S_i$=normal/abnormal, $S_{ii}$=classification attack types,
         $S_{iii}$=classification sub attack types
    [3]  If $S_i$=abnormal then $S_i$=abnormal;
    [4]  Else $S_i$=normal
    [5]  End
    [6]  If $S_{ii}$=classification attack types then $S_{ii}$=classification attack types;
    [7]  Else $S_{ii}$=none;
    [8]  End
    [9]  If $S_{iii}$=classification sub attack types then $S_{iii}$=classification
         sub attack types
    [10] Else $S_{iii}$=none;
    [11] End
    [12] If $S_i$=abnormal or If $S_{ii}$=classification attack types or $S_{iii}$=classification sub
         attack
         types then Detected, d=1
    [13] Else if $S_i$=normal  then Detected, d=0 (proceed with statistical test on
         anomaly-
         based section )
    [14] End
```

Figure 5: Algorithm 2- Rule-based Detection



```
Statistical tests as an approach in anomaly-based indicate the evaluation on anomalous
traffic volume that act as second detection for the unknown intrusion events
    [1] Determine categories of packets
    [2] Let time slot = T13 and attributes = TCP Flag
    [3] Calculate statistics of packets distribution, let it with $X^2=\sum (O-E)^2/E$
    [4]Exit
```

Figure 6: Algorithm 3 - Statistical Test Detection

Year	Total of security Incidents
2012	11,427
2011	10,460
2010	6,554
2009	3,668

Table 1: Number of Security Incidents for 2009-2012 excluding spam

Variant	% Detection Rate	% False Positive	% False Negative	% Overall Rate
Invalid Hash	95.2	2.1	20	93.7
Allaple.L	93.9	6.1	5.4	94.1
RBot	95.8	4.2	19.6	93
Palevo	100	1.7	0	98.4
Srvcp	95	3.6	14.67	93.5
Tnnbtib	95.8	4.5	24.7	92.6

Table 2: Summary of Detection Rate for Data Mining

Variant	% False Negative	% Accuracy	% Detection Rate	% False Alarm Rate
Invalid Hash	0	100	100	0
Allaple.L	2	99.99	100	0
RBot	6	99.98	100	0
Palevo	3	100	100	0
Srvcp	2	99.99	100	0
Tnnbtib	0	100	100	0

Table 3: Signature-based Module Detection Result

Variant	Undetectable in Signature-based	Detectable in Anomaly-based
Invalid hash	0	0
Allaple.L	2	6
RBot	6	8
Palevo	3	7
Srvcp	2	6
Tnnbtib	0	0

Table 4: Combination of Data Mining-based, Signature-based and Anomaly-based Detection Result

Enhancement Taxonomy and Analysis on Android Malware Detection

Halizah Saad[*1], **Najiahtul Syafiqah Ismail**[1], **Faizal M.A**[1], **Robiah Yusof**[1], and **Raihana Syahirah Abdullah** [1]

[1]*Faculty of Information Communication and Technology, Universiti Teknikal Malaysia Melaka (UTeM)*

E-mail: azil.liza89@gmail.com
[*]*najiahtul.ismail@gmail.com, faizalabdollah@utem.edu.my, robiah@utem.edu.my, raihana.syahirah@utem.edu.my*
Corresponding author

ABSTRACT

The increasing of smartphone used have increase the popularity of Android operating system (OS) but it comes also with the growth of the Android malware threat. Therefore, an effective Android Malware Detection (AMD) is needed as a precautions from damages. To design an effective AMD firstly, we need to understand what is the criteria from existing AMD has offered to identify and prevent the Android malware. This is important to give a better understanding to provide a more effective AMD in future.

Keywords: Mobile Operating System (OS), Android Malware, Detection Technique, Detection Analysis, Android Malware Detection (AMD)

1 INTRODUCTION

Nowadays, there is an amazing growth in smartphone used where as it is more convenient for today lifestyle. There are many types of smartphone operating system (OS) available in market like Apples iOS, Nokias Symbian, Blackberry OS and etc. According to Statista (2016) statistic in their website, it clearly shows that Android is leading with huge differences compared to other OS from 2011 until 2015 and this prove that Android have become top choice by Smartphone user among other platforms. Unfortunately, it also attract a dangerous threat to Android platform. Plus, its features as an open source OS and the openness of Android apps are some of the reason that attract both the developer and attacker to choose Android platform. Figure 1 shows the increasing of Android malware threat.

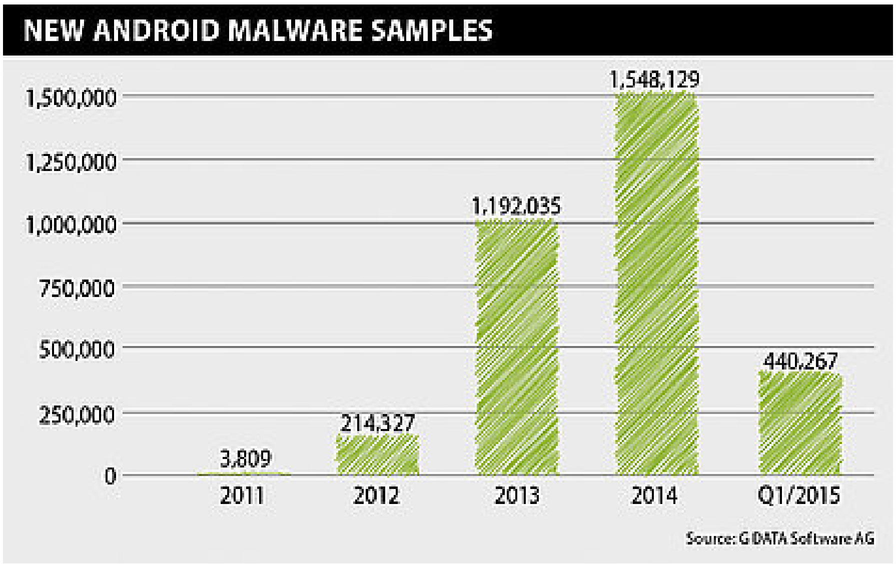


Figure 1: New Android Malware threat from 2011 to 2015 (Q1)

Based on Figure 1, it shows that the Android malware threat continues to increase every year. Furthermore, AG (2015) reported within the first quarter of 2015 there are 440,267 new Android malware threats addition. To address the problem, a better Android Malware Detection (AMD) is needed in order to identify and prevent the Android malware from harm the user.

Therefore, we will discuss about the existing AMD to provide a glimpse idea about what have been done by earlier researcher before proceed to design an improvement for the AMD in future work. The remainder of this paper is organized as follows: Section 2 describes related work. In Section 3 we discuss and explain about the analysis we have been done, and finally the paper is concluded in Section 4.

2 RELATED WORK

2.1 Android Architecture

Android is basically a stack of software which can be divided into four major layers. Figure 2 shows the Android architecture.

Based on Figure 2,the top layer of the stack consists of pre-installed application provided by original equipment manufacturer (OEM) and the third party application installed by the user.Second layer contain a set of services that collectively construct the environment in which Android applications operate and are managed.Next layer consist of two parts which are Libraries and Android Runtime. The libraries are guide the device in handling different types of data. Meanwhile, Android runtime is made of two components that provides the core API for the Java language, and the Dalvik Virtual Machine.Lastly, the bottom layer provides a level

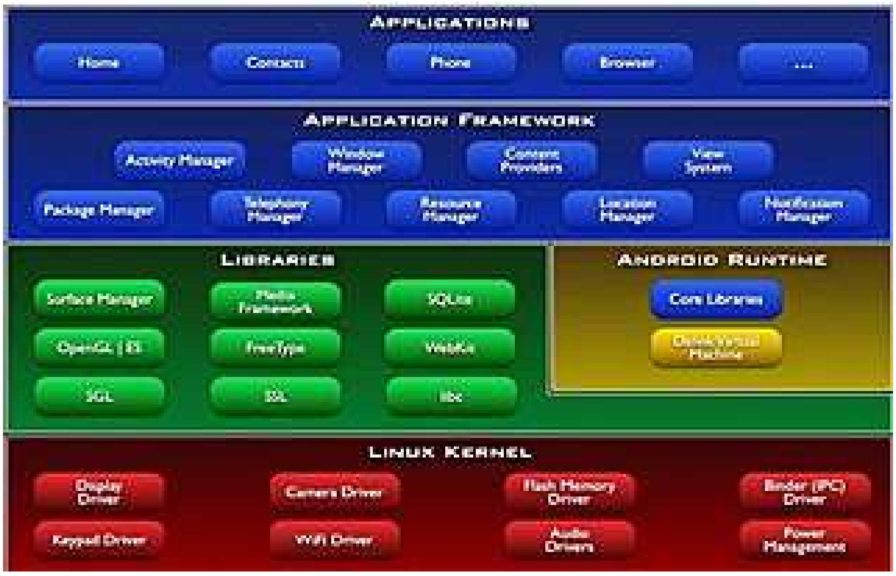


Figure 2: Architecture of Android

abstraction between device hardware and software.

2.2 Android Malware

In Introduction section, we have discussed Android is the most popular smartphone OS. Due to its usage in today lifestyle unfortunately, it also motivate the attacker which harm the user. Malware is referred to several names such as malicious software or malevolent software, and malicious code. According to Robiah et al. (2009), malware is a program that been written by inventor with intention to disrupt or damage a computer system. Therefore, android malware can be described as a program that been embedded in Android platform application with intent to disrupt, breach user privacy and confidentiality by impose as legitimate application. Besides that, android malware continue emerging every year thus, there is a requirement for android malware detection (AMD) to identify and prevent from andorid malware attack.

2.3 Android Malware Detection (AMD)

Android malware detection can be define as a framework that has been created to detect android malware. As stated by Mas'ud et al. (2014), the android malware detection system can be categorized into four categories which are detection techniques, detection analysis, detection platform and detection audit data source.

2.3.1 Detection technique

Based on Baraiya and Hiteishi (2015), usually malware detection technique can be divided into two major categories which are signature-based (SB) and anomaly or behavior-based (AB). Besides, there is another technique that also can be used which known as specification-based (SPB). Signature-based (SB) use a matching pattern signature with a sets of policies or rules to detect malware. MeanWhile, anomaly or behavior-based (AB) technique use its knowledge to learn what is the normal malware behavior to decide whether the inspection application is malicious. Plus, it also usually applies machine learning algorithm to learn about the malware behavior and predicting the unknown malware. Specification-based used certain rule set that considered as normal to decide whether the program under inspection is violating the rule set. Hence, the program that violates the predefined rule set is considered as malicious program.

2.3.2 Detection analysis

According to Alazab et al. (2012), AMD is similar to desktop malware where it consists of two common approaches which consists of static and dynamic analysis. Static analysis is the process of dissamble or dissecting code to analyze the programs code without truly executing it (Moser et al., 2007). Plus, it parses instruction of binary image to understand and detect the malicious function or shell code that has been embedded . On the other hand, dynamic analysis involves running the sample in controlled and isolated environment for analyzing the malware behavior based on their execution traces. Besides that, there are several researches have decide to use hybrid analysis which a combination of both static and dynamic approach.

2.3.3 Detection platform

The android malware detection can be employed on two type of platform which are on the host itself or on the remote server. Usually the android malware detection will be deployed on host to get quicker result because the monitoring and analysis process will be executed on the device itself. There are also another platform which is to deploy it on the remote server or cloud. As for the remote server or cloud deployment, the monitoring process will be executed on host but the analysis process will be done on the server or cloud for example, Crowdroid.

2.3.4 Detection audit data source or detection input data

The input data collected for analysis process is different for each layer and all the information needed can be gathered from all the related layer in the android stack. Plus, it can be collected using suitable approach such as static or dynamic. The type of data can be classified into an application package data, network traffic, system call, hardware performance such as CPU and memory usage, and etc.

Based on the discussion earlier, Figure 3 shows the types of malware detection techniques.

According to Baraiya and Hiteishi (2015), malware detection techniques can be classified into two types: signature-based and anomaly-based. The authors also classified specification based as a special type of anomaly-detection which clarify the two main types of detection techniques.

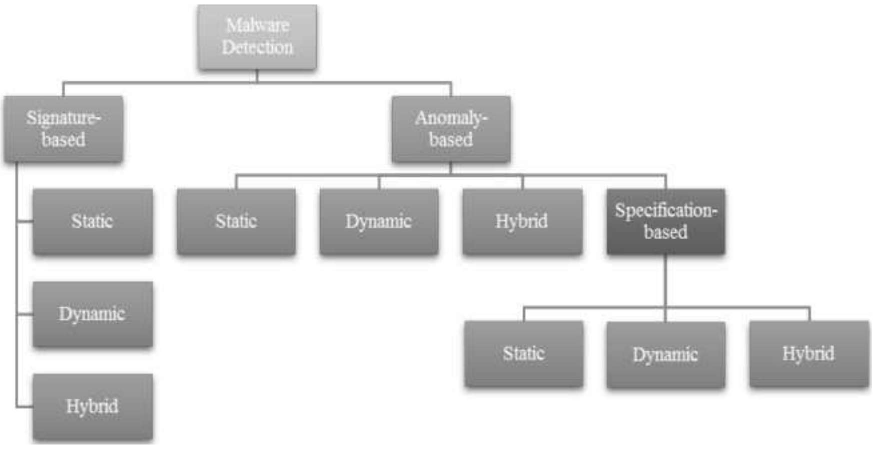


Figure 3: Types of Malware Detection Techniques (Baraiya and Hiteishi, 2015)

3 DISCUSSION AND ANALYSIS

There are 23 Android malware detection researches have been discussed. From B, we can summarize that 15 researchers: (Eder et al., 2013), (Chan and Song, 2014), (Dini et al., 2012),(Wu et al., 2012), (Aafer et al., 2013), (Arp et al., 2014), (Enck et al., 2010), (Burguera et al., 2011), (Shabtai et al., 2011), (Rastogi et al., 2013), (Sanz et al., 2014), (Almin and Chatterjee, 2015), (Sanz et al., 2013), (Peiravian and Zhu, 2013), (Yu et al., 2013) and Li et al. (2014) have choose anomaly-based detection technique while the others 5 researcher: (Yang and Yang, 2012), (Faruki et al., 2013), (Feng et al., 2014), (Qadir et al., 2011), (Chan et al., 2012) and Sheen et al. (2015) choose to use signature-based technique because it can detect known and unknown malware effectively while the signature-based is insufficient in detecting unknown malware where the malware use more sophisticated obfuscated technique to distinguish them self. Normally, anomaly-based will using dynamic analysis while signature-based will be use a static analysis approach. However, there is an exception when choosing to combine static analysis with machine learning method therefore it is known as anomaly-based technique.

There are 8 researcher: (Feng et al., 2014), (Dini et al., 2012), (Enck et al., 2010), (Burguera et al., 2011), (Rastogi et al., 2013), (Yu et al., 2013), Dini et al. (2013) and Li et al. (2014) have choose to use dynamic approach whereas the rest use a static analysis because the dynamic analysis consume more resources to process the analysis and can cause power draining to smartphone even though it provide more efficient result . In contrast, static analysis give a quicker analysis and not very resource consuming.Therefore, (Eder et al., 2013) decided to applied both approaches in their framework to overcome both static and dynamic analysis to produce better result and more code coverage.

Majority researchers decided to deploy in host platform except (Feng et al., 2014),(Burguera et al., 2011) and Li et al. (2014) which decide to deploy their framework on remote server or cloud. Remote server or cloud platform require an internet connection for analyzing process on server as the host only manage the monitoring process. However, researcher need to consider the confidentiality of information while transferred the data into the server although it helps to reduce storage space and complex processing.

Next, we will discuss about the selection of audit source data that been used as an input for analysis process. 12 researchers: (Faruki et al., 2013), (Chan and Song, 2014), (Feng et al., 2014), (Qadir et al., 2011), (Wu et al., 2012), (Aafer et al., 2013),(Arp et al., 2014), (Chan et al., 2012), (Sanz et al., 2014), (Almin and Chatterjee, 2015), (Sanz et al., 2013) , (Peiravian and Zhu, 2013) and Sheen et al. (2015) choose Application package as their source input data. There are several different features can be used like permissions, java code and intent filters. The permissions is the most popular feature because it is the first barrier to attacker. On the contrary, (Enck et al., 2010) and (Burguera et al., 2011) choose to use kernel data which is system call as the selected feature while Li et al. (2014) decided to use network traffic data. The system call and network traffic are dynamic features so it is collect using dynamic analysis which different than permission: static feature. Other researches selected several sources as their input data: (Dini et al., 2012) and (Shabtai et al., 2011) choose combination of all the dynamic features such as system calls, network traffic, system components and user features while (Eder et al., 2013) decided to use combination of static and dynamic data like network traffic, kernel and application package data as their source data. Both researchers: (Yang and Yang, 2012) and (Enck et al., 2010), choose to look for sensitive data as their input for detecting the leak of privacy information. Choosing appropriate source data or features is an important step in conducting experiment where it will determine the effectiveness results for future research.

Lastly, we will discuss the methods that has been used to conduct the experiment. There are various methods that has been used which suites the selected approach, either static, dynamic or hybrid. Several researcher (Chan and Song, 2014), (Wu et al., 2012), (Aafer et al., 2013), (Almin and Chatterjee, 2015), (Peiravian and Zhu, 2013) and (Yu et al., 2013) choose different methods in their experiment such as different machine learning algorithm. On the other hand, only few researcher using single methods in their research. The methods in Appendices A have been classified according to the approach that has been analyzed and summarized as in Appendices B. Based on Appendices A, we have distinguished the common methods that has been used in existing Android Malware Detection (AMD).

4 CONCLUSION

The continually growth of android malware have motivate authors to do research in android malware detection in order to prevent the android malware attack. This research will give a brief enlightenment on existing studies for android malware detection system. Based on the generated taxonomy of existing Android Malware Detection which using appropriate features and methods will be guiding to produce more accurate result and increase the effectiveness of the AMD.

ACKNOWLEDGMENTS

The authors would like to express the appreciation to Inforsnet Group of University Teknikal Malaysia Melaka (UTeM) and MyBrain15 Programme by Ministry of Higher Education Malaysia (MoHE) in encouraging the authors to publish this paper. This work was supported by the MoE of Malaysia under Grants FRGS/1/2014/ICT04/FTMK/02/F00212.

REFERENCES

Aafer, Y., Du, W., and Yin, H. (2013). Droidapiminer: Mining api-level features for robust malware detection in android. *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering Security and Privacy in Communication Networks*, page 86103.

AG, G. D. S. (2015). *G DATA MOBILE MALWARE REPORT*. THREAT REPORT: Q1/2015. G DATA Software AG Germany.

Alazab, M., Monsamy, V., Batten, L., Lantz, P., and Tian, R. (2012). Analysis of malicious and benign android applications. *2012 32nd International Conference on Distributed Computing Systems Workshops*.

Almin, S. B. and Chatterjee, M. (2015). A novel approach to detect android malware. *Procedia Computer Science*, 45:407417.

Arp, D., Spreitzenbarth, M., Hbner, M., Gascon, H., and Rieck, K. (2014). Drebin: Effective and explainable detection of android malware in your pocket. *Proceedings 2014 Network and Distributed System Security Symposium*.

Baraiya, D. and Hiteishi, D. (2015). A Survey on Android Malware and Malware Detection Techniques. *International Journal for Scientific Research & Development*, 3(01):143–147.

Burguera, I., Zurutuza, U., and Nadjm-Tehrani, S. (2011). Crowdroid: Behavior-Based Malware Detection System for Android. *Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices - SPSM '11*, page 15.

Chan, P. P., Hui, L. C., and Yiu, S. M. (2012). Droidchecker. *Proceedings of the fifth ACM conference on Security and Privacy in Wireless and Mobile Networks - WISEC '12*.

Chan, P. P. K. and Song, W.-K. (2014). Static detection of android malware by using permissions and api calls. *2014 International Conference on Machine Learning and Cybernetics*.

Dini, G., Martinelli, F., Saracino, A., and Sgandurra, D. (2012). Madam: A multi-level anomaly detector for android malware. *Lecture Notes in Computer Science Computer Network Security*, page 240253.

Dini, G., Martinelli, F., Saracino, A., and Sgandurra, D. (2013). Probabilistic contract compliance for mobile applications. *2013 International Conference on Availability, Reliability and Security*.

Eder, T., Rodler, M., Vymazal, D., and Zeilinger, M. (2013). Ananas - a framework for analyzing android applications. *2013 International Conference on Availability, Reliability and Security*.

Enck, W., Gilbert, P., Chun, B.-G., Cox, L. P., Jung, J., McDaniel, P., and Sheth, A. N. (2010). TaintDroid: An Information-Flow Tracking System for Realtime Privacy Monitoring on Smartphones. *9th USENIX Symposium on Operating Systems Design and Implementation (OSDI10*, 49:1–6.

Faruki, P., Ganmoor, V., Laxmi, V., Gaur, M. S., and Bharmal, A. (2013). Androsimilar. *Proceedings of the 6th International Conference on Security of Information and Networks - SIN '13*.

Feng, Y., Anand, S., Dillig, I., and Aiken, A. (2014). Apposcopy: semantics-based detection of android malware through static analysis. *Proceedings of the 22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering - FSE 2014*.

Li, J., Zhai, L., Zhang, X., and Quan, D. (2014). Research of android malware detection based on network traffic monitoring. *2014 9th IEEE Conference on Industrial Electronics and Applications*.

Mas'ud, M. Z., Sahib, S., ., ., Abdollah, M. F., Selamat, S. R., and Yusof, R. (2014). Android malware detection system classification. *Research J. of Information Technology Research Journal of Information Technology*, 6(4):325341.

Moser, A., Kruegel, C., and Kirda, E. (2007). Exploring multiple execution paths for malware analysis. *2007 IEEE Symposium on Security and Privacy (SP '07)*.

Peiravian, N. and Zhu, X. (2013). Machine learning for android malware detection using permission and api calls. *2013 IEEE 25th International Conference on Tools with Artificial Intelligence*.

Qadir, M., Jilani, A., and Sheikh, H. (2011). Automatic Feature Extraction, Categorization and Detection of Malicious Code in Android Applications. *International Journal of Information & Network Security (IJINS)*, 3(1):12–17.

Rastogi, V., Chen, Y., and Enck, W. (2013). Appsplayground. *Proceedings of the third ACM conference on Data and application security and privacy - CODASPY '13*.

Robiah, Y., Rahayu, S. S., Zaki, M. M., Shahrin, S., Faizal, M. A., and Marliza, R. (2009). A New Generic Taxonomy on Hybrid Malware Detection Technique. *arXiv preprint arXiv:0909.4860*, 5(1):6.

Sanz, B., Santos, I., Laorden, C., Ugarte-Pedrero, X., Nieves, J., Bringas, P. G., and Maran, G. l. (2013). Mama: Manifest analysis for malware detection in android. *Cybernetics and Systems*, 44(6-7):469488.

Sanz, B., Santos, I., Ugarte-Pedrero, X., Laorden, C., Nieves, J., and Bringas, P. G. (2014). Anomaly detection using string analysis for android malware detection. *Advances in Intelligent Systems and Computing International Joint Conference SOCO13-CISIS13-ICEUTE13*, page 469478.

Shabtai, A., Kanonov, U., Elovici, Y., Glezer, C., and Weiss, Y. (2011). Andromaly: a behavioral malware detection framework for android devices. *Journal of Intelligent Information Systems*, 38(1):161–190.

Sheen, S., Anitha, R., and Natarajan, V. (2015). Android based malware detection using a multifeature collaborative decision fusion approach. *Neurocomputing*, 151:905912.

Statista (2016). Smartphone operating systems: Global market share 2011-2015 — forecast.

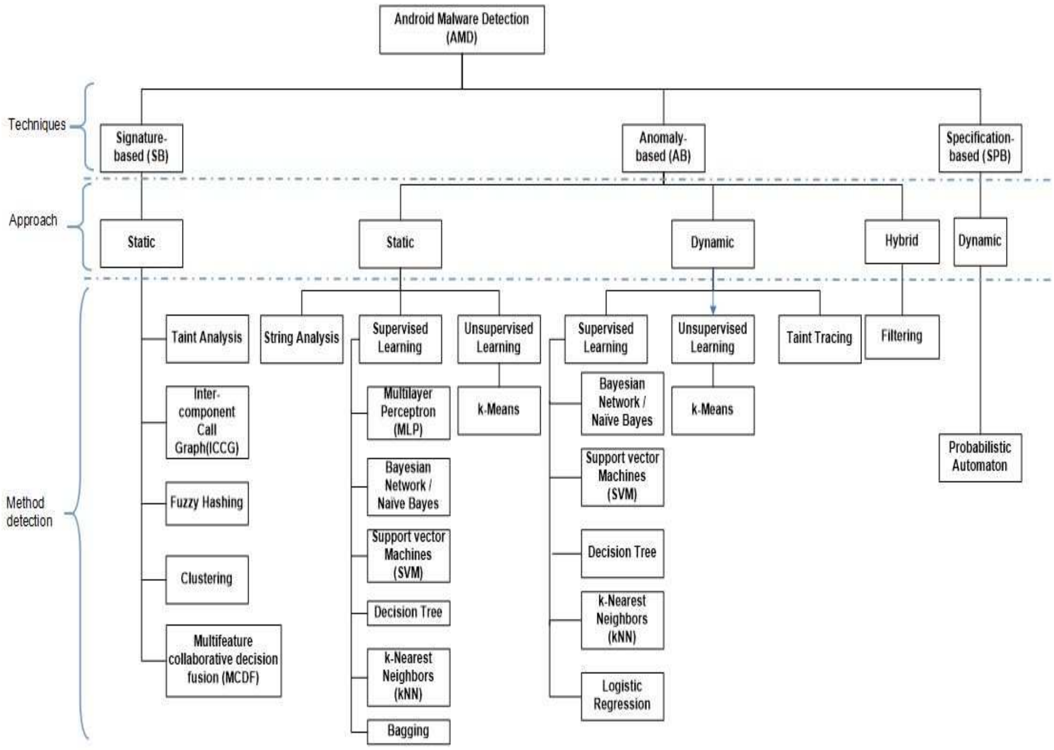
Wu, D.-J., Mao, C.-H., Wei, T.-E., Lee, H.-M., and Wu, K.-P. (2012). Droidmat: Android malware detection through manifest and api calls tracing. *2012 Seventh Asia Joint Conference on Information Security*.

Yang, Z. and Yang, M. (2012). Leakminer: Detect information leakage on android with static taint analysis. *2012 Third World Congress on Software Engineering*.

Yu, W., Zhang, H., Ge, L., and Hardy, R. (2013). On behavior-based detection of malware on Android platform. *GLOBECOM - IEEE Global Telecommunications Conference*, pages 814–819.

A APPENDICES

Taxonomy of Existing Android Malware Detection (AMD)



B APPENDICES

Summary of Android Malware Detection (AMD)

References	Technique			Approach		Platform		Source Input Data	Methods
	SB	AB	SPB	Static	Dynamic	H	S		
Yang and Yang (2012)	√			√		√		Sensitive Data	Taint Analysis
Eder et al. (2013)		√		√	√	√		Network traffic, Kernel & Application Package	Filtering
Faruki et al. (2013)	√			√		√		Application Package	Fuzzy Hashing
Chan and Song (2014)		√		√		√		Application Package	Naive Bayes, SVM with SMO, Liblinear, J48 decision tree, RBF network, MLP & Random Forest
Dini et al. (2012)		√			√	√		All except Application Package	KNN
Feng et al. (2014)	√			√			√	Application Package	Taint Analysis & Inter-Component Call Graph (ICCG)
Qadir et al. (2011)	√			√		√		Application Package	Clustering using C++
Wu et al. (2012)		√		√		√		Application Package	k-Means, KNN & Naive Bayes
Aafer et al. (2013)		√		√		√		Application Package	ID3 DT, C4.5 DT, KNN & linearSVM
Arp et al. (2014)		√		√		√		Application Package	Linear Support Vector Machines
Enck et al. (2010)		√			√	√		Sensitive Data	Taint tracking
Burguera et al. (2011)		√			√	√	√	Kernel	k-Means
Shabtai et al. (2011)		√			√	√		All except Application Package	k-Means, Logistic Regression, Histograms, Decision Tree, Bayesian Networks & Naive Bayes
Rastogi et al. (2013)		√			√	√		Kernel, network traffic	Taint tracing
Chan et al. (2012)	√			√		√		Application Package	Taint Analysis & Inter-Component Call Graph (ICCG)
Sanz et al. (2014)		√		√		√		Application Package	String Analysis
Almin and Chatterjee (2015)		√		√		√		Application Package	k-Means & Naive Bayes
Sanz et al. (2013)		√		√		√		Application Package	KNN, Random Forest, SVM & Bayesian Network
Peiravian and Zhu (2013)		√		√		√		Application Package	SVM, Decision Tree (J48) & Bagging
Yu et al. (2013)		√			√	√		Kernel	Support Vector Machine (SVM) and Naive Bayes
Dini et al. (2013)			√		√	√		Kernel	Probabilistic Automaton
Li et al. (2014)		√			√		√	Network Traffic	SVM
Sheen et al. (2015)	√			√		√		Application Package	MCDF

*Legend = Hybrid(Static+Dynamic), Signature-based(SB), Anomaly-based(AB), Specification-based(SPB), Host(H), Server/Cloud(S)

An Analysis Technique for Cost Estimation in Information Security

Ahmed Yaser Mohd Zabawi^{*1}, **Rabiah Ahmad**^{1,2}, and **Shekh Faisal Abdul-Latip**^{1,2}

^{1}*INFORSNET Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka 76100 Durian Tunggal, Melaka*

E-mail: yaserzabawi@gmail.com
**Corresponding author*

ABSTRACT

In this paper we propose a technique for cost estimation in information security risk analysis based on known threats. The motivation of this study emerges based on the fact that, to the best of our knowledge, there is no existing analysis tools that can provide cost estimation for each risk that had been identified. Most of the current risk analysis tools only use conventional techniques which can be divided into two categories, namely qualitative and quantitative techniques. Therefore, in our work a risk analysis tool comprises of cost estimation and risk assessment known as CARA (cost and risk analysis) was developed for leader. The development of this tool is a part of the preliminary work in data collection for threat analysis and cost estimation in which right now there is no particular tool that can furnish with those features. However the result from our preliminary work shows that this tool can be used as one of threat analysis tool. In the future we aim at analysing the data and improving the tools that may lack of perfection in any aspects.

Keywords: Qualitative, Quantitative, Cost Estimation, CARA

1 INTRODUCTION

Presently, threat analysis has become one of the remarkable topics to be discussed by many people especially in the field of computer network and information security. In Margaret (2014) defined Information security as a set of business process that protects information assets regardless of how the data is formatted or whether it is being handled, is in transit or is being stored. It is not a technology; rather consider it as a technique involved the process, tools and strategies required to prevent, identify, report and counter threats to digital and non-digital information (Margaret, 2014). Consequently, the rapid growth of technology may lead to the emergence of possible threats to information security.

Threats could be anything that can cause interference, intrusion and destruction to any valuable items or services that can be recognized as the company assets. Whether due to human or nonhuman origin, the analysis must examine every component that may cause possible security risk (Dimitar, 2014, Institute, 2002). In addition threats must also be looked in relation to the business environment and what influence they have to the organization (Institute, 2002). Management is better ready to comprehend the ramifications of the threat and vulnerabilities when they are quantifiable and measurable.

Addition to that, threat assessment is the initial step in risk management process, where it considers the full range of threats for a given facility or location. The assessment ought to analyse for supporting information to assess the relative probability of circumstance for each threat. There are four example assessment fall under threat assessment which are (Nancy A. Renfroe and Joseph L. Smith, 2015);

- Defined (Man-made) - There are aggressors who utilize this strategy who ought to know be focusing this facility or the association. Specific threats have been perceived by law enforcement agencies. Natural: Events of this nature happen in the quick region on a frequent basis.

- Credible (Man-made) - There are aggressors who apply this strategy who are known to target this type of facility. No specific threat has been known by law enforcement agencies. Natural: Events of this nature happen in the immediate vicinity intermittently (i.e. once every 10 years).

- Potential (Man-made) - There are aggressors who work with this strategy, but they are not known to target this type of facility. Natural: Events of this nature happen in the region on a periodically basis.

- Minimal (Man-made) - No aggressors who practise this strategy are recognized for this facility. Natural: There is no history of this type of event in the area.

The main focus of this paper is to propose an enhanced approach in risk analysis tool. Most of current tools only use conventional techniques, namely, qualitative or quantitative techniques. It is important to have an enhanced tool that can perform risk assessment and cost estimation for each possible risk which may help to improve existing tools. For instance, most of existing tools do not provide cost estimation. Moreover most of current tools are also unable to perform assessment based on qualitative and quantitative data in a single tool. This paper is organized as follow; Section 2 provides an explanation about steps of cost estimation in details. Next, in Section 3 we describe methods used to analyse quantitative and qualitative data. The description covers the advantages and disadvantages of both methods. The multi-factors of threats with description and explanation using table will further discussed in Section 4. In section 5 we introduce enhanced model and soft computing inclusive with technical description that will be developed to improve the current risk analysis tools. The final section provides a conclusion of the overall study.

2 RISK ANALYSIS TOOLS

This leaves the intangible assets involved, such as client confidence and experience. These things, while important, are not so easily priced and will not be included in the cost estimation but it must be remembered that they are present and will be included in deciding what risks have been reduced to acceptable level. Risk analysis can be defined as the base of information security, risk management, and risk during process of information protection. Risk analysis is a part of risk management, where the assessments are conducted. Risk analysis comprises process such as identification of activity, threat analysis, vulnerability analysis and guarantees (Betina A. and A., 2012, Lee, 2014). Betina A. and A. (2012), Palisade (2014) had state the purpose of having risk analysis which is to provide decision maker with best likely information about the chance of loss and offer a better thoughtful of the possible consequences that could happen.

In a wide sense, the approaches of risk analysis can be categorised into two major group which are qualitative and quantitative technique. Both techniques are important for evaluating the influences of risk on decisions. These two categories of risk analysis can be conducted simultaneously or in an order, and even within a distinct period gap (D., 2011). Quantitative technique is a method uses two basic fundamentals: the probability of an event occurring and the losses that may be incurred, while qualitative method rates the magnitude of the potential impact of a threat as high, medium, or low. Qualitative methods are the most common to measure of the impact of risks (Betina A. and A., 2012).

The advantages and disadvantages of risk analysis assessment for both quantitative and qualitative methods are shown in Table 1 and Table 2 respectively.

Quantitative methods use a mathematical approach and statistical tools to represent risk in risk analysis (Wawrzyniak, 2006). As an example, ”there are a lot of viruses attacking managers computer”. In this situation, we can see a lot of virus attacking computer, so the quantitative methodology will give results; it is risky to be a threat if many viruses attack computers. However, Wawrzyniak (2006) has identified risk analysis tool that uses quantitative methods are not efficient for the intensive use of information security management. Therefore, this method is rarely used in the field of business. Risk analysis tool that uses quantitative methods are IS-RAM, CORA, IS, RISKWATCH and etc.

Qualitative methods risk assessed with the help of adjectives instead of mathematic (Wawrzyniak, 2006). Currently, most of developer use qualitative approach as their methodology to develop new analysis tools (M., 2010). It is because qualitative method is more flexible and more suitable then quantitative method (C. and S, 2003). However, qualitative method does not provide complete output information to be used in the risk management process. Risk analysis tool that uses quantitative methods are OCTAVE, OCTAVE-S, CORAS, CRAMM, FRAP and etc.

Most of the risk analysis tools regardless qualitative or quantitative implementing information security attributes which are confidentiality, availability and integrity. However, some studies should be conducted to prove the risk analysis tools have characteristics such as confidentiality, availability, integrity, reliability and others. This is necessary as to help users to

Risk Analysis	**Quantitative Methods**
Advantages	- It gives more accurate image of risk. - It allows for determination of consequences of incidents occurrence in quantitative way, what facilitates realization of costs and benefits analysis during selection of protections. - It applies mathematical and statistical tools to represent risk.
Disadvantages	- Not suitable for intensive analysis nowadays. - In complicated environment it is more difficult to use mathematical models. - Quantitative measures depend on the scope and accuracy of defines measure scales. - Results of analysis may be precise and even confusing. - Analysis conducted with the applications of this method is generally more expensive, demanding greater experience and advanced tools.
Example Of Tools	- ISRAM, CORA, IS, RISKWATCH and etc.

Table 1: Quantitative Methods

Risk Analysis	**Qualitative Methods**
Advantages	- Analysis is relatively easy and cheap. - It allows to prioritize the risks. - It allows for determination of areas of greater risk in a short time without bigger expenditures. - Perform risk analysis with the help of adjectives, not mathematical models. - It is more suitable for complicated risk analysis nowadays.
Disadvantages	- Unstable results - It depends on the ideas of those who undertake risk analysis. - It does not allow to determine the probability and results using numerical measures. - Cost-benefits analysis is more difficult during the selection of protections.
Example Of Tools	- OCTAVE, OCTAVE-S, CORAS, CRAMM, FRAP and etc.

Table 2: Qualitative Methods

select the best risk analysis tools to be used to solve different problems faced by individuals and organizations.

3 METHODOLOGY

Risk assessment is a structured and systematic procedure (N. and L., 2009), which is depending upon correct identification of threats and proper measures to evaluate possible risks arising from the threat. Nowadays, there are many types of computer crimes; money theft 44%, damage of software 16%, theft of information 16%, alteration of data 12%, theft of services 10%, trespass 2% (S., 2003).

3.1 Scope Statement

In CARA, the scope statement is a primary key to carry out cost estimation. This scope will be analyzed by the 1st level of information security committee in the organization. The scope statement should have these criteria;

∗ Specify exactly what is to be evaluated.

∗ State what kind of risk analysis will be performed.

∗ Provide the expected results.

As example, a quantitative risk assessment will be performed on the organization XY to reduce the risks to their system to an acceptable level using benefit-cost analysis methodologies for determining applicable controls (James W. Meritt, 1999).

3.2 Assets

The information that was described in the scope statement will be broken down into its components which will then set its own value. While it is possible to break down the systems into functional units, (James W. Meritt, 1999) has proposed a solution to be carried out which is much easier to disassemble the overall system into its tangible components which may be more easily priced;

This leaves the intangible assets involved, such as client confidence and experience. These things, while important, are not so easily priced and will not be included in the cost estimation but it must be remembered that they are present and will be included in deciding what risks have been reduced to acceptable level (James W. Meritt, 1999).

All assets are not equally vulnerable to every risk. Therefore, James W. Meritt (1999) has made research and produce vulnerability values for every assets according their impact

Assets	Descriptions
Equipment : PC	This category includes all information processing equipment maintained by the organization. It contains, but is not limited to, PCs, front-end processors, fileservers, mainframe computers and workstations.
Equipment : Printers	This category contains items of information technology used to impress information upon paper. It includes things such as a variety of printers (varying from dot matrix through laser printers) and plotters.
Equipment : Other	This category contains items of equipment not covered by other designated categories. It contains, but obviously is not limited to, such things as external Hard disk, Smartphone and power supplies.
Data Information : Other	This category includes all information sources not readily identifiable as belonging in one of the other two.
Facilities	This may be the entire building itself and its supplied services or simply the table the system is on. It depends, of course, on the system being analyzed.

Table 3: Major Assets

on threats. These values were derived using the combined experience and skills of a number of experts in the arena of information systems security. They are suggested values and do not take the local threat environment or existing countermeasure effectiveness into account (James W. Meritt, 1999). Table 4 shows some of the data obtained by James W. Meritt (1999).

Expected losses are based on the expected impact of the threat on the asset. The amount of loss depends on the vulnerability of the asset. The vulnerability factor (0.0 to 1.0) of an asset with respect to a threat is the ratio of:

- The expected loss from a single impact of the threat on the asset
- The loss potential of the asset.

3.3 Risks

Risk is something that can cause harm, or reduce the value of an asset in information technology. As highlighted by J. (2009) in his article Digital Threat.

Assets and List of Threats	**Values**
Software: Operating System	
Power Loss	0.20
Communication Loss	0.10
Data Integrity Loss	0.00
Accidental Errors	0.10
Computer Virus	0.80
Abuse of Access Privileges by Employees	0.20
Natural Disasters	0.50
Attempted Unauthorized System Access by Outsider	1.00
Theft or Destruction of Computing Resource	1.00
Destruction of Data	0.00
Abuse of Access Privileges by Other Authorized User	0.00
Successful Unauthorized System Access by Outsider	0.60

Table 4: Impact/Exposure Coefficient

Risk is the potential that a given threat will exploit vulnerabilities of an asset or group of assets and thereby cause harm to the organization.

While threat is an event that cannot be expected that can damage the organisation assets and prejudice personal data protection (S., 2015). Threats to information and information systems can include purposeful attacks, environmental disruptions, and human/machine errors and result in great harm to organisation (G. and Gallagher, 2011). Then, to produce a tool that able to evaluate cost estimation, we used the collection of the threat data generated by Ahmad R. (2009). Below is part of the threats data we used in CARA;

- Unsecured wireless network
- Network software failure
- Network congestion
- Switch port problems
- Network Connection failure
- Communications infiltration - Hackers due to unsecured network
- Spoofing/impersonation due to unsecured network
- Out-dated application software (Office applications and Antivirus software)
- Malware attacks due to an introduction of damaging or disruptive software

After that, the dataset of threats will be categorized as below according to major assets contained in table 5.

Assets	List Of Threats
Network: Modems	Unsecured wireless network Network software failure
Network: Other	Server down due to power failure Spoofing/impersonation due to unsecured network
Software : Application	Out-dated application software Out-dated system software Application software failure
Equipment : PC	Out-dated hardware Various kinds of malware attacks Hardware maintenance error

Table 5: Qualitative Methods

Based on table 5, the threats is categorized by assets to determine estimation cost faced by users based on a predetermined value of the assets.

3.4 Calculation

Evaluation and determination of intangible assets value is an open problem in determining the size of security risk. The fact that some information is more important or interesting does not explain much to the manager who needs to invest in security. Because information value needs to be determined more precisely it is necessary to understand its appearance, manifestation, activity methods and structure of its value. In CARA, user must select potential threats on their assets. After that, they need to provide price for every asset they want to analyse.

Cost Estimation Formula :

X= Value of threats (Table 4)
Y= Assets price
N= Estimation Cost

$$N=Y(X)$$

For example, price of printer own by company XY is RM1000, and the value of threat for that printer is 0.4, estimation costs of loss that may be experienced by company XY is RM400.

3.5 Controls

Controls is a counter measures to avoid, counteract or minimize loss or unavailability due to threats acting on their matching vulnerability such as security risk (S., 2009). Examples of possible controls are James W. Meritt (1999);

- Develop, document, and test backup procedures
- Develop, document, and test continuity of operations procedures
- Implement and access control mechanism
- Implement user authentication mechanism
- Implement encryption mechanism
- Implement a configuration management process for software
- Implement a version control process for documentation
- Procedures concerning the security of the system operation
- Develop user documentation on proper use of the system

When implementing controls to reduce the risk it is usually the case that the cost of the treatment should be lower than the cost of the impact. This is one factor which will determine how appropriate the control is along with the overall security requirements (members, 2013).

3.6 Results

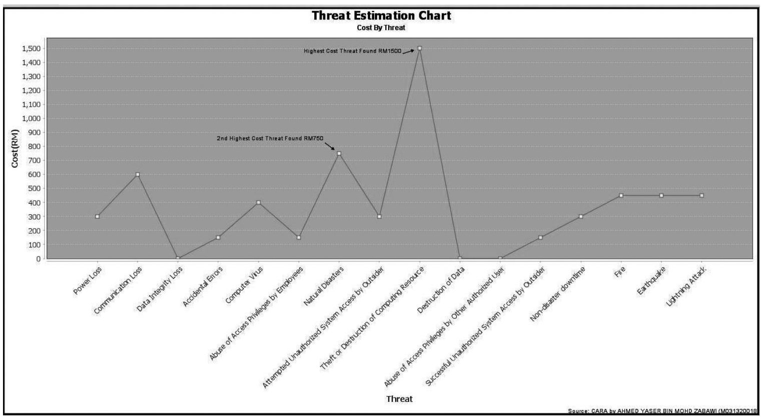
To carry out cost estimation, user need to have information or data like assets want to be evaluate and value of each assets need to be in rinngit malaysia (RM).

The assessment will be done like what have been explained in calculation part at calculation section. Each assets has respective value that have been collected and compiled by (Ahmad R., 2009, James W. Meritt, 1999).Below are the screenshot of the tool.

Figure 1 above shows results of cost estimation and user can view the most risky threats that might happened to assets with different type of threats. based on a given result, the user can decide how to control or mitigate the effects of the threat. Steps to control threats have been explained in the control part.

4 CONCLUSIONS

Threats become an enemy that exist in the project that aware people to find solutions to overwhelm them. They are increased by percentage in line with the rapid growth of technology. With the implementation of risk analysis tool (loss cost estimation), it is hoped that threat can be detected faster and give some information to people about loss cost estimation for each risk. A enhanced analysis tool is developed by combining both techniques, qualitative and quantitative



Source: CARA by AHMED YASER BIN MOHD ZABAWI (M031320018

Figure 1: Cost Estimation

techniques where it can offers people with many specialities that can assist people in handling risk analysis.

The attempt through this study has indicated that a enhanced analysis tool is practicable to be used as risk analysis tool as it has the loss cost estimation feature where very crucial in risk assessment. Hence, it is paramount to have a enhanced analysis tool that combining both methods, and expected can counter risk in early phase.

5 ACKNOWLEDGEMENT

The authors would like to thank Universiti Teknikal Malaysia Melaka (UTeM) and Ministry of Higher Education for FRGS fund, FRGS/2/2013/ICT07/FTMK /02/3/F00186 (FRGS - FTMK, Malaysia for supporting this research. Not to forget the two most important people in my life, Father and Mother who have given me support, patience. I owe you both gratitude. Again thank you for being there when I needed you most.

REFERENCES

Ahmad R., Samy G. N., I. N. K. B. P. A. I. Z. (2009). *Threats Identification in Healthcare Information Systems Using Genetic Algorithm and Cox Regression.*

Betina A., A. R. R. and A., C. J. (2012). A survey of information security risk analysis methods.

C., F. D. and S, E. M. D. (2003). *Cyberwar-Netwar Security in the information age*.

D., A. (2011). Qualitative and quantitative risk analysis.

Dimitar, K. (2014). Cyber threat analysis.

G., L. and Gallagher, P. D. (2011). Managing information security risk. http://csrc.nist.gov/publications/nistpubs/800-39/SP800-39-final.pdf..

Institute, S. (2002). An overview of threat and risk assessment.

J., M. (2009). Threat vs vulnerability vs risk, digital threat.

James W. Meritt, C. (1999). A method for quantitative risk analysis.

Lee, M.-C. (2014). Information security risk analysis methods and research trends: Ahp and fuzzy comprehensive method.

M., R. (2010). Risk analysis definition.

Margaret, R. (2014). Information security (infosec) definition.

members, U. (2013). Risk assessment. https://www.it.ox.ac.uk/policies-and-guidelines/is-toolkit/risk-assessment..

N., B. and L., R.-D. (2009). *Risk Assessment of information technology systems*. IISIT.

Nancy A. Renfroe, P. and Joseph L. Smith, P. (2015). Threat/ vulnerability assessment and risk analysis.

Palisade (2014). *The Worlds Most Popular Risk Analysis Tool*.

S., B. (2003). *IT Security cookbook. Boran Consulting*. Boran Consulting.

S., B. (2015). Threat (information security). http://www.privacycommission.be/en/glossary/threat..

S., N. (2009). Security control. http://www.sans.edu/research/security-laboratory/article/security-controls.

Wawrzyniak, D. (2006). *Information Security Risk Assessment Model For Risk Management*.

Determinants of Information Privacy Concern and Privacy Protection Behaviour Strategies in Social Networking Sites among Undergraduates in Malaysia

Nur Fadzilah Othman[*1], **Rabiah Ahmad**[1], and **Muliati Sedek**[2]

[1]*Information Security and Networking Research Group (InForSnet), Center for Advanced Computing Technology,Universiti Teknikal Malaysia Melaka*
[2]*Center for Teaching and Learning, Universiti Teknikal Malaysia Melaka*

E-mail: nurfadzilah132@gmail.com
[*]*Corresponding author*

ABSTRACT

This paper attempts to investigate the determinants of privacy protection behaviour strategies and information privacy concerns in using social networking sites. The factors, as gathered is based on the protection motivation theory. Hence, this paper highlights the roles of information privacy concerns in social networking sites by identifying the determinants as well as the behavioural strategies that individuals employ in ensuring that their privacy is protected. An empirical analysis that involved 488 undergraduates from a public Malaysian university was interpreted. Structural Equation Modelling (SEM) technique was used in analysing the data and the results were based on SEM outputs, which demonstrate the acceptance and confirmation of all factors. From the results, it demonstrates that information privacy concerns among users contribute to privacy protection behaviour strategies. Perceived severity, perceived vulnerability, self-efficacy and response efficacy are found to be the determinants for privacy protection behaviour strategies.

Keywords: social networking sites, privacy protection behaviour strategies, information privacy concern, protection motivation theory

1 INTRODUCTION

Social Networking Sites (SNSs) have become a phenomenon amongst Malaysians. Statistics from the Malaysian Communication and Multimedia Commission (MCMC) in their Pocket Book of Statistics Q1 2014 has demonstrated that 45.5 percent of the population or 13.3 million users are registered Facebook users (MCMC, 2014). The increase of users can be contributed to the variety of tools that are offered by SNSs that enable and facilitate communication and information sharing. SNSs allow individuals to stay in touch with their friends, reconnect with old friends and create new relationships with other people through the plethora of activities provided, such as sharing photos, videos, archiving events, updating others on activities, sending messages privately and posting public testimonials (Boyd, 2008, Vithessonthi, 2010). Therefore, the nature of SNSs that offer an attractive way of online interaction and communications encourage users to use it to its zenith. Unfortunately, information sharing and activities performed while accessing SNSs in an uncontrolled manner can lead to a privacy breach on the users behalf

Recently, privacy issues related to personal information has been widely discussed and deliberated by various academic researchers (Nemec Zlatolas et al., 2015). Due to the rapid development of technology that facilitates communication such as increased Internet access and smartphones, privacy has become a serious concern. Users willingly share their private information in SNSs subconsciously without a clear idea of who is allowed access to their personal information and what portion of it is really accessed. Even though SNSs themselves have been equipped with systematic safety features, it still cannot guarantee that one's privacy is fully protected (Salleh et al., 2012). Hence, there is an urgent need for an assessment mechanism that can detect threats from engaging in risky situations that users have access to so that they can determine how much and what type of personal information should be shared and disclosed.

This study aims to investigate the determinants of the privacy protection behaviour strategies that users employ in SNSs. Understanding the determinants of privacy protection behaviour in SNSs has the ability of generating awareness that can protect users and allow them to confidently impose their self-control through the execution of privacy protection behaviour strategies.

2 PRIVACY PROTECTION BEHAVIOUR STRATEGIES

Literature has given various definitions of privacy. The concept of privacy ranges from a ”right to be alone” as from the perspective of law (Warren and Brandeis, 1890) , ”state of limited access” in the aspect of philosophy (Schoeman, 1984), to the ”control over information about one's self” as given from the views of social science (Westin, 1970).

Privacy protection behaviour can be defined as specific computer-based actions that individuals take to keep their information safe. Rogers (1975), has discussed that individuals are motivated to rely on protection behaviour in order to cope and adopt behaviour to control risk, threat and danger. Coping strategies can be divided into two dimensions, which are approach and avoidance (Amirkhan, 1990, Endler and Parker, 1990, Piko, 2001). Approach strategies include fabricating personal information and seeking social support whereas avoidance strategies include withholding personal information. In the context of SNS, fabricating information as based on approach strategies refers to a user covering up or disguising their identity by using fake or false information in SNSs. Seeking social support means that users ask for advice and read privacy statements from SNS providers in order to gain knowledge that would help them adopt privacy protection behaviour strategies alongside increasing their privacy. As for avoidance strategies, refraining information means that users will refuse to provide their personal information to SNSs and instead even begin to patronize them over it.

Information privacy concern is the ”extent to which an individual is concerned about organizational practices related to the collection and use of his or her personal information (Smith et al., 1996). Previous research has shown that information privacy concern has an impact on privacy protection behaviour strategies (Feng and Xie, 2014, Jiang et al., 2013, Mohamed and Ahmad, 2012). Within the protection motivation theory, information privacy concern is considered to be a mediating variable that explains the relationship between the factors involved and privacy protecting behaviour strategies (Lwin et al., 2007).

3 THEORETICAL FRAMEWORK AND HYPOTHESES

3.1 *Protection Motivation Theory*

Protection Motivation Theory (PMT) as introduced by Rogers (1975) postulates that an individual’s motivation to protect from risk and threat comes from: (1) perceived severity, (2) perceived vulnerability and (3) response efficacy. The model was modified to explain failures concerned in protection behaviour by including (4) self-efficacy, (5) response cost and (6) rewards associated with risky behaviour (Rogers, 1975, 1983). PMT has been principally used in the health industry (Floyd et al., 2000, Fruin et al., 1992) and according to Grindley et al. (2008) , this theory has also been used in more than 20 different health-related areas in order to study about intentions and behaviours. In the field of Information System (IS), PMT has also been widely used to examine protection behaviour in online transactions (Lee et al., 2008, Mohamed and Ahmad, 2012) awareness of employees in organizational information security policies Vance et al. (2012) and individual use of security software (Johnston and Warkentin, 2010)

3.1.1 *Perceived severity*

Perceived severity refers to an individual’s belief that the judgement of severity significance results from a threatening event (LaRose and Rifon, 2007). Perceived severity evaluated how severe an individual believes that that threat will interrupt their life. Individuals will adopt recommended action when they seriously perceive the negative consequence (Zhang and McDowell, 2009). Moreover, Youn (2009) found that an individual’s motivation for engaging in risk-reducing behaviour is increased by perceived severity. For the purpose of this study, users will develop a perceived severity after losing information privacy to SNSs. They will significantly associate this loss with information privacy concern and in this form, indirectly motivate them to adopt privacy protection strategies in SNSs.

3.1.2 *Perceived vulnerability*

Perceived vulnerability explains an individuals perception in experiencing possible negative effects that stem from performing risky behaviour (Lee et al., 2008). Based on the findings of Fuller et al. (2014), it can be argued that perceived severity is found to have increased students intention to perform malware avoidance behaviour. Subsequently, Mohamed and Ahmad (2012) agreed that one of the factors that contribute to users increasing information privacy concerns in SNSs is perceived vulnerability. Conversely, perceived vulnerability had an insignificant impact on employees intention to comply with IS security policies (Vance et al., 2012). Thus, for this study, it is suggested that individuals who perceive the risk and threats of losing information privacy through SNSs will increase their information privacy concerns, which will motivate them to use privacy protection behaviour strategies.

3.1.3 *Self-efficacy*

Self-efficacy can be defined as an individual's belief that they have the capability to implement protective behaviour (Compeau et al., 1999). Several studies provide evidence that self-efficacy plays an important role in a user's choice to perform risky online behaviour.Vance et al. (2012) identifies that employees believe that they can successfully comply with security policies and enhance compliance with policies and procedures while Lee et al. (2008), proves that self-efficacy should be influential factors of stimulation in order to perform protection behaviour. Hence, this study suggests that individuals who are self-efficacious in using SNSs are more likely concerned with their information privacy and as such are motivated to use privacy protection behaviour strategies.

3.1.4 *Response efficacy*

Response efficacy is the belief that a recommended coping response is effective in avoiding a threat (Woon and Tan, 2005). Research identifies that response efficacy is a significant predictor behaviour that determines the decision of home wireless network users in implementing security features on their networks (Woon and Tan, 2005), increase intentions to use anti-spyware software as protective technology (Chenoweth et al., 2009), predicts backing up data on personal computers (Crossler, 2010) and influences behaviour intentions in performing malware avoidance behaviours when using personal mobile devices (Dang-pham and Pittayachawan, 2014). Therefore, the study posits that recommended protective action in SNSs could help them avoid losing information privacy and could also motivate them to use privacy protection behaviour strategies.

3.1.5 *Rewards*

Reward refers to an individual's expectation in getting benefits when keeping with selective behaviour (Lee et al., 2008). Previous study has given significant negative influence and suggested that individuals who find great enjoyment and satisfaction from sharing personal information are less inclined to make adaptive change for protection (Marett et al., 2011). Additionally, individuals that are willing to disclose their information may experience a sense of being close to their friends and family (Baren et al., 2003) and as such get satisfaction from the feeling of togetherness (IJsselsteijn et al., 2009). So, for the purpose of this study, rewards refer to the extent of benefits that are received from using SNSs and as such are significantly associated with information privacy concern.

3.2 *Information Privacy Concern*

Information privacy concern is the ”extent to which an individual is concerned about organizational practices related to the collection and use of his or her personal information (Smith et al., 1996). Previous research has shown that information privacy concern had an impact on privacy protection behaviour strategies (Feng and Xie, 2014, Jiang et al., 2013, Mohamed and Ahmad, 2012) . Within the protection motivation theory, information privacy concern is considered to be a mediating variable that explains the relationship between the factors involved and privacy protecting behaviour strategies (Lwin, Wirtz, and Williams, 2007). In this study, we hypothesize five aspects of PMT and information privacy concern. We also propose investigating the effects of information privacy concern towards privacy protection behaviour strategies. The following hypotheses are as follows:
H1: Information privacy concern is significantly associated with privacy protection behaviour strategies.
H2: Perceived severity is significantly associated with information privacy concern.
H3: Perceived vulnerability is significantly associated with information privacy concern.
H4: Self-efficacy is significantly associated with information privacy concern.
H5: Response efficacy is significantly associated with information privacy concern.
H6: Reward is significantly associated with information privacy concern.

Because existing theories and empirical evidence do not hint at a clear causal relationship between response cost towards information privacy concern and privacy protection behaviour strategies, we do not hypothesize on them. The proposed research model is presented in Figure 1.

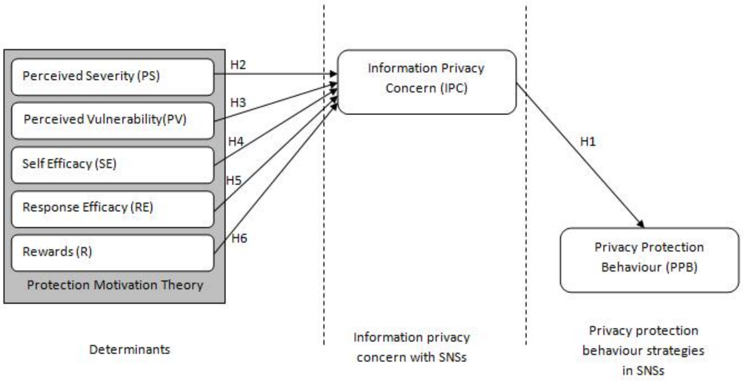

Perceived Severity (PS)
Perceived Vulnerability(PV)
Self Efficacy (SE)
Response Efficacy (RE)
Rewards (R)
Protection Motivation Theory
H2
H3
H4
H5
H6
Information Privacy Concern (IPC)
H1
Privacy Protection Behaviour (PPB)
Determinants
Information privacy concern with SNSs
Privacy protection behaviour strategies in SNSs

Figure 1: Research Model

4 RESEARCH METHODOLOGY

For the purpose of this study, a total of eleven hypotheses were tested and a quantitative approach was employed to them. Quantitative approach is the best method to use in order to test any existing theory as it involves a collection and statistical analysis of numerical data (Ary et al., 2010)

The instrument used in this study was a questionnaire that consisted of 44 items in total. Five items for perceived severity adapted from Crossler (2010), LaRose and Rifon (2007) and Woon and Tan (2005). Six items for perceived vulnerability adapted from Woon and Tan (2005) and Dinev and Hart (2004). Five items for self-efficacy adapted from Crossler (2010) ,LaRose and Rifon (2007) and Woon and Tan (2005). Five items for response efficacy adapted from Crossler (2010) , Zhang and McDowell (2009) and Lee et al. (2008) while six items for rewards adapted from Youn (2005). 10 items for information privacy concern adapted from Dinev and Hart (2004). Finally 7 items for privacy protection behaviour strategies adapted from Feng and Xie (2014). All the items used a five point Likert scale, where 5 represented strongly agree and 1 represented strongly disagree.

4.1 *Sample selection and Data collection*

For the sampling process, stratified random sampling was used in this study. From the data given by the universities' administration on the number of active undergraduates as of February 26, 2015, there were approximately 9,205 undergraduates in total.Sedek et al. (2012) recommended that the ideal number for sample size suitable for analysis using SEM should be approximately between 300 to 800 samples. Of the 550 distributed, 499 were returned. 485 were usable for the purpose of this study with a response rate of 88 percent. Table 1 shows the profile of the respondents.

Variable	Type	Frequency	% Percent
Gender	Male	254	52
	Female	231	48
Age	15-20	-	-
	21-25	449	93
	26-30	36	7
	31-32	-	-

Table 1: Profile of respondent.

5 RESULTS

The first step conducted in SEM analysis was Confirmatory Factor Analysis (CFA)(Hair et al., 2010). CFA was meant to identify the individual construct and was employed for three major purposes, which are (i) model fit, (ii) convergent validity and (iii) construct validity. Maximum likelihood estimate (MLE) was used to estimate the structural model. Table 2 presents the test of overall model fit. All the fit indices were above recommended values.

Construct	Convergent Validity	
	Composite Reliability (CR)	Average variance extracted (AVE)
Privacy protection behaviour (PPB)	0.874	0.578
Information privacy concern (IPC)	0.917	0.650
Perceived severity (PS)	0.833	0.555
Perceived vulnerability (PV)	0.845	0.579
Self-efficacy (SE)	0.882	0.882
Response efficacy (RE)	0.867	0.623
Rewards (R)	0.903	0.702

Table 2: Result of CFA model.

Table 3 shows the criteria for fit indices and recommended values.

Name of category	Name of Index	Level of acceptance	Source
Absolute fit	RMSEA	≤ 0.08	Zainudin (2012)
Incremental fit	GFI	≥ 0.8	Baumgartner and Homburg (1996) Doll et al. (1994)
	CFI	≥ 0.8	Baumgartner and Homburg (1996) (1996); Doll et al. (1994)
Parsimonious fit	Chisq/df	≤ 3.0	Zainudin (2012))

Table 3: Categories of model fit and their level of acceptance.

Table 4 is the result of fitness indexes for the research model. All required levels were achieved.

The root mean square error of approximation (RMSEA), which measures the discrepancy per degree of freedom, was 0.065. The goodness-of-fit index (GFI) was 0.851, comparative fit index (CFI) was 0.882 and the discrepancy Chi Square (Chisq/df) was 2.964.

Figure 2 below presents the detailed result of the structural model. The R2 values for information privacy concern and privacy protection behaviour is 0.55 and 0.08 respectively. As shown in Figure 1, the paths from rewards to information privacy concern were insignificant whereas all the other paths were significant. Thus, H6 is not supported while H1, H2, H3, H4 and H5 are supported. The summary of regression path coefficients, significance values and hypothesis statement for every path and its conclusion is as shown in Table 5.

Name of category	Name of index	Index value	Comments
Absolute fit	RMSEA	0.065	The required level is achieved
Incremental fit	GFI	0.851	The required level is achieved
	CFI	0.882	The required level is achieved
Parsimonious fit	Chisq/df	2.964	The required level is achieved

Table 4: The fitness Indexes for research model.

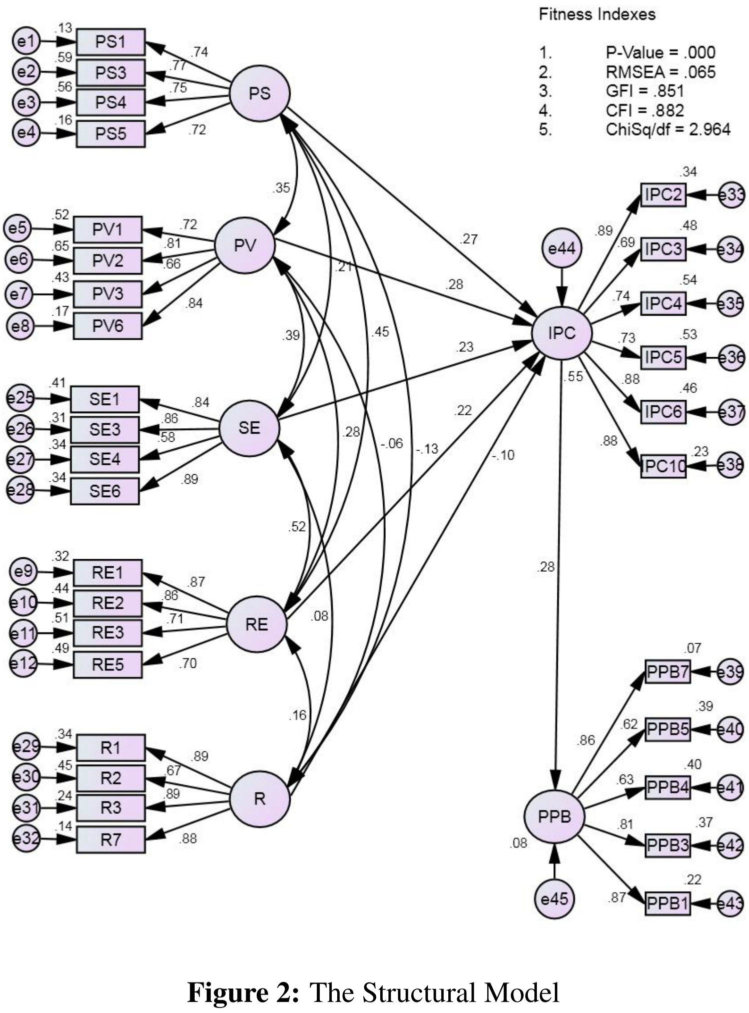


Figure 2: The Structural Model

Source	Destination	Hypothesis statement of path analysis	Estimates	P-value	Results on hypothesis
IPC	PPB	H1.Higher information privacy concern will increase privacy protection behaviour.	0.28	0.003	Supported
PS	IPC	H2.Higher perceived severity will increase information privacy concern	0.27	0.043	Supported
PV	IPC	H3.Higher perceived vulnerability will increase information privacy concern	0.28	0.033	Supported
SE	IPC	H4.Higher self-efficacy will increase information privacy concern	0.23	0.028	Supported
RE	IPC	H5.Higher response efficacy will increase information privacy concern	0.22	0.001	Supported
R	IPC	H3.Higher rewards will reduce information privacy concern	-0.10	0.605	Supported

Table 5: The regression path coefficients, significance values and hypothesis statement for every path and its conclusion.

6 CONCLUSION

The data analysis as executed from SEM reveals that there was a significant relationship of perceived severity, perceived vulnerability, response efficacy and self-efficacy towards information privacy concern as well as a significant relationship of information privacy concern and privacy protection behaviour strategies. Findings from this study are expected to enhance the understanding towards information privacy concern, its antecedents and privacy protection behaviour strategies in SNSs.

Similar with the results of prior research, individuals who are concerned with their information privacy in SNSs were found to use and adopt privacy protection behaviour strategies (Jiang et al., 2013, Mohamed and Ahmad, 2012). Therefore, to ensure that individuals use privacy protection strategies, awareness and concern toward information privacy must be provided. In order to materialise such concern and awareness, several determinants have been found. Perceived severity was found to be one such determinant. Users who feel that they will be seriously affected by the loss of information privacy will be more concerned with their information privacy, whereas those who think otherwise will not be as concerned. This is consistent with previous research (Chenoweth et al., 2009; Lee et al., 2008; Mohamed and Ahmad, 2012). The next determinant that is significantly associated with information privacy concern is perceived vulnerability. The finding suggests that individuals who have been or are exposed to the loss of information privacy are more concerned with information privacy, whereas those who have not been or are not exposed to such a loss are less concerned. This finding supports Mohamed and Ahmad (2012), Crossler (2010) and Lee et al. (2008). The next determinant proven in this study that contributes to information privacy concern is self-efficacy. Individuals who believe that they have the ability to use the protective strategies in SNSs will be more concerned with their information privacy as supported by prior research from Mohamed and Ahmad (2012), Milne, Labrecque, and Cromer (2009) and Lee et al. (2008). Finally, the last determinant that contributes to information privacy concern is response efficacy. As supported by previous research by Crossler (2010) and Chenoweth et al. (2009), individuals that use recommended protective action in SNSs can avoid losing information privacy. One determinant however, was found to not support the hypotheses as stated in this study. It was found that reward is not significantly associated with information privacy concern. Youn, (2009) and Salleh et al. (2012) favour this finding by saying that great rewards and benefits gained from online activities make them less concerned with their information privacy.

This research is crucial as it serves as a guide that provides instructions and guidelines that can help users of SNSs to keep their privacy intact. Besides that, this finding may prove useful to SNS providers via the accumulation and analysis of user data with regards to privacy concerns. It may also act as a trigger for a redesign of privacy protection strategies. For educators, this research may be used to encourage users to exercise more caution when utilizing SNSs and for institutions, the creation of appropriate awareness programmes in order to increase concerns towards information privacy in SNSs.

ACKNOWLEDGMENTS

The authors are grateful to Universiti Teknikal Malaysia Melaka (UTeM) for the financial support through grant scheme (FRGS/2/2013/ICT07/UTEM/02/1) and Ministry of Higher Education of Malaysia for the financial support through the Fellowship Scheme.

REFERENCES

Amirkhan, J. H. (1990). A factor analytically derived measure of coping: The Coping Strategy Indicator. *Journal of Personality and Social Psychology*, 59(5):1066–1074.

Ary, D., Jacobs, L. C., and Sorensen, C. (2010). *Introduction to Research in Education*. Wadsworth Publishing.

Baren, J., IJsselsteijn, W., Romero, N., Markopoulos, P., and Ruyter, B. (2003). Affective benefits in communication: The development and field-testing of a new questionnaire measure. In *PRESENCE 2003, 6th annual international workshop on Presence, Aalborg, Denmark*, page 48.

Baumgartner, H. and Homburg, C. (1996). Application of Structural Equation Modeling in Marketing and Consumer Research: a review. *International Journal of Research in Marketing*, 13(2):139–161.

Boyd, D. (2008). Facebook's Privacy Trainwreck: Exposure, Invasion, and Social Convergence. *Convergence: The International Journal of Research into New Media Technologies*, 14(1):13–20.

Chenoweth, T., Minch, R., and Gattiker, T. (2009). Application of protection motivation theory to adoption of protective technologies. In *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on*, pages 1–10. IEEE.

Compeau, D., Higgins, C. A., and Huff, S. (1999). Social cognitive theory and individual reactions to computing technology: A longitudinal study. *MIS Quarterly*, 23(2):145–159.

Crossler, R. E. (2010). Protection Motivation Theory: Understanding Determinants to Backing Up Personal Data. *2010 43rd Hawaii International Conference on System Sciences*, pages 1–10.

Dang-pham, D. and Pittayachawan, S. (2014). Comparing intention to avoid malware across contexts in a BYOD-enabled Australian university : A Protection Motivation Theory approach. *Computers & Security*, 48.

Dinev, T. and Hart, P. (2004). Internet privacy concerns and their antecedents - measurement validity and a regression model. *Behaviour & Information Technology*, 23:413–422.

Doll, W., Xia, W., and Torkzadeh, G. (1994). A confirmatory factor analysis of the end-user computing satisfaction instrument. *MIS Quarterly*, 18(4):357–369.

Endler, N. S. and Parker, J. D. (1990). Multidimensional assessment of coping: a critical evaluation. *Journal of personality and social psychology*, 58(5):844–854.

Feng, Y. and Xie, W. (2014). Teens' concern for privacy when using social networking sites: An analysis of socialization agents and relationships with privacy-protecting behaviors. *Computers in Human Behavior*, 33:153–162.

Floyd, D. L., Prentice-Dunn, S., and Rogers, R. W. (2000). A Meta-Analysis of Research on Protection Motivation Theory.

Fruin, D. J., Pratt, C., and Owen, N. (1992). Protection Motivation Theory and Adolescents ' Perceptions of Exercise '. *Journal of Applied Psychology*, 22:55–69.

Fuller, B. T., Fahrni, S. M., Harris, J. M., Farrell, A. B., Joan, B., Gerhart, L. M., Ward, J. K., Taylor, R. E., and Southon, J. R. (2014). Comparing intention to avoid malware across contexts in a BYOD-enabled Australian university: A Protection Motivation Theory approach. *Quaternary Geochronology*.

Grindley, E. J., Zizzi, S. J., and Nasypany, A. M. (2008). Use of protection motivation theory, affect, and barriers to understand and predict adherence to outpatient rehabilitation. Physical Therapy. *Jounal of American Physical Theraphy Association*, 88(12):1529–1540.

Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2010). *Multivariate Data Analysis (7th Ed)*. Pearson Prentice Hall, New Jersey.

IJsselsteijn, W., van Baren, J., Markopoulos, P., Romero, N., and Boris de Ruyter (2009). Measuring Affective Benefits and Costs of Mediated Awareness: Development and Validation of the ABC-Questionnaire Wijnand. *Human-Computer Interaction*, pages 187–206.

Jiang, Z. J., Heng, C. S., and Choi, B. C. F. (2013). Privacy Concerns and Privacy-Protective Behavior in Synchronous Online Social Interactions. *Information Systems Research*, 24(3):579–595.

Johnston, B. A. C. and Warkentin, M. (2010). Fear Appeals and Information security Behaviors: An Empirical Study. *Ministry of Education Official Website*, 34(3):549–566.

LaRose, R. and Rifon, N. J. (2007). Promoting i-safety: Effects of privacy warnings and privacy seals on risk assessment and online privacy behavior. *Journal of Consumer Affairs*, 41(1):127–149.

Lee, D., Larose, R., and Rifon, N. (2008). Keeping our network safe: a model of online protection behaviour. *Behaviour & Information Technology*, 27(5):445–454.

Lwin, M., Wirtz, J., and Williams, J. D. (2007). Consumer online privacy concerns and responses: a powerresponsibility equilibrium perspective. *Journal of the Academy of Marketing Science*, 35(4):572–585.

Marett, K., McNab, a. L., and Harris, R. B. (2011). Social Networking Websites and Posting Personal Information: An Evaluation of Protection Motivation Theory. *AIS Transactions on Human-Computer Interaction*, 3:170–188.

MCMC (2014). Communications & Multimedia Pocket Book of Statistic. Technical report, Malaysin Communications and Multimedia Commission.

Mohamed, N. and Ahmad, I. H. (2012). Information privacy concerns, antecedents and privacy measure use in social networking sites: Evidence from Malaysia. *Computers in Human Behavior*, 28(6):2366–2375.

Nemec Zlatolas, L., Welzer, T., Heričko, M., and Hölbl, M. (2015). Privacy antecedents for SNS self-disclosure: The case of Facebook. *Computers in Human Behavior*, 45:158–167.

Piko, B. (2001). Gender Differences and Similarities in Adolescents' Ways of Coping. *The Psychological Record*, 51(2):223–236.

Rogers, R. W. (1975). A Protection Motivation Theory of Fear Appeals and Attitude Change1. *The Journal of Psychology*, 91(November 2014):93–114.

Rogers, R. W. (1983). Cognitive and Physiological Process in fear Appeals and Attitude Change: A Revised Theory of Protection Motivation. In *Basic Social Psychophysiological Research*. Guildford Press.

Salleh, N., Hussein, R., Mohamed, N., Abdul, N. S., Ahlan, A. R., and Aditiawarman, U. (2012). Examining Information Disclosure Behavior on Social Network Sites Using Protection Motivation Theory , Trust and Risk. *Journal of Internet Social Networking & Virtual Communities*, 2012.

Schoeman, F. D. (1984). *Philosophical dimensions of privacy: An anthology*. Cambridge University Press.

Sedek, M., Mahmud, R., Jalil, H. A., and Daud, S. M. (2012). Types and Levels of Ubiquitous Technology use among ICT Undergraduates. *Procedia - Social and Behavioral Sciences*, 64:255–264.

Smith, H. J., Milberg, S. J., and Burke, S. J. (1996). Information privacy: measuring individuals' concerns about organizational practices. *MIS quarterly*, pages 167–196.

Vance, A., Siponen, M., and Pahnila, S. (2012). Motivating IS security compliance: Insights from Habit and Protection Motivation Theory. *Information & Management*, 49(3-4):190–198.

Vithessonthi, C. (2010). Knowledge sharing, social networks and organizational transformation. *The Business Review, Cambridge*, 15(2):99–109.

Warren, S. D. and Brandeis, L. D. (1890). The Right to Privacy. *Harvard Law Review*, 4(5):193–220.

Westin, A. F. (1970). *Privacy and Freedom*. The Bodley Head Ltd.

Woon, I. and Tan, G.-w. (2005). A Protection Motivation Theory Approach to Home Wireless Security. In *International Conference on Information Systems (ICIS)*.

Youn, S. (2005). Teenagers ' Perceptions of Online Privacy and Coping Behaviors : A Risk Benefit Appraisal Approach. *Journal of Broadcasting & Electronic Media*, 49(1):86–110.

Youn, S. (2009). Determinants of Online Privacy Concern and Its Influence on Privacy Protection Behaviors Among Young Adolescents. *Journal of Consumer Affairs*, 43(3):389–418.

Zainudin, A. (2012). *A handbook on SEM: Structural equation modelling using amos graphics*. Kelantan: University Technology MARA Press.

Zhang, L. and McDowell, W. C. (2009). Am I Really at Risk? Determinants of Online Users' Intentions to Use Strong Passwords. *Journal of Internet Commerce*, 8(3-4):180–197.