# Chaos Based Cryptography An Alternative to Algebraic Cryptography

Muhammad Rezal bin Kamel Ariffin

Laboratory of Theoretical Studies Institute for Mathematical Research Universiti Putra Malaysia

> Mathematics Department Faculty of Science Universiti Putra Malaysia

> rezal@putra.upm.edu.my

#### Abstract

In this paper we will first discuss cryptography from a historical point of view in order to understand the motivation behind this science. We will discuss the earliest form of cryptography before fast forwarding to the current modern forms of cryptography.

We will then mention issues surrounding current modern cryptographic methods before introducing chaos based cryptography. The relationship between chaos and cryptography makes it natural to employ chaotic systems to design new cryptosystems. It is based on the facts that chaotic signals are usually noise-like and chaotic systems are very sensitive to initial conditions. Their sensitivity to initial conditions and their spreading out of trajectories over the whole interval seems to be a model that satisfies the classic Shannon requirements of confusion and diffusion [1].

From 1989 onwards, many different chaotic encryption systems have been proposed. The most celebrated chaotic cryptosystem is based on the ergodicity property of chaotic maps [2] and has received more and more attentions in the past literature [3-17]. Introduced by Baptista in 1998, it is able to produce different ciphers for the same plaintext.

It was cracked by Alvarez in 2003 via the one-time pad attack. In 2008, M.R.K.Ariffin and M.S.M.Noorani [22] engaged the attack and devised a counter measure against it.

#### Introduction

Cryptography which is the science of scrambling messages has its roots since ancient times. Differing from the method of steganography (the science of hiding information—does not involve a key), cryptography does not hide messages. Cryptography changes the message to hide the

meaning and involves a key. In general, cryptography can be divided into two categories, namely the substitution cipher and the transposition cipher. The substitution cipher can be categorized into two. That is the cipher by coding (replacing words) and by ciphering (replacing letters). The earliest known transposition cipher can be traced to the device known as the scytale. The scytale is a device that transposes whole sentences. Figure 1 below is a pictorial representation of a scytale.

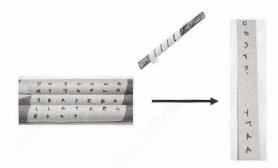


Figure 1: The Scytale

The above device was widely used during the military campaign by the Spartan against the Greeks. A piece of paper would be wrapped on a cylindrical object (eg: a stick) and a message would be written on it. When loosen from the cylinder, the message would appear to be scrambled.

In principle the security of a transposition cipher would depend on the length of the sentence. For example there are 3! possible ways to rearrange the letters BUS. And to rearrange the sentence "For example consider this short sentence" there are  $1.0333148 \times 10^{40}$  possible ways. At a glance this would provide very high security. We will see later that this is not the case.

The earliest known substitution cipher can be traced back to Julius Ceaser. This method which is now known as the Ceaser's cipher is an ingenious method which substitutes each letter in the alphabet

with another alphabet 3 place in front. Let us substitute the alphabets in the English language via the following numeric representation: A=1, B=2, C=3,..., Z=26. For the English alphabet which consists of 26 alphabets, this cipher method can be formalized via the following simple equation:

$$C = (P+3)mod26 \tag{1}$$

where C is the ciphertext and P is the plaintext. An example is as pictured below:



Figure 2: Ceaser's Cipher

In principle, if Ceaser's cipher were to be implemented on the English alphabet, for a cryptanalyst to actually derive the exact substitution method through brute force, it would be a huge effort. There are  $26! = 4.03291461 \times 10^{26}$  possibilities of shifting.

It is only natural for those on the other side of the divide to analyze and find weaknesses in ciphers of others. The first successful formal attack on ciphers was established by Al-Kindi (801-873). He cryptanalyzed monoalphabetic substitution ciphers using frequency analysis.



**Figure 3:** The first page of al-Kindi's manuscript *On Deciphering Cryptographic Messages* 

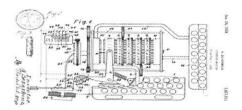
Al-Kindi's method exposed weaknesses on the first two ciphers that we discussed earlier.

This was the start of a long "love hate" relationship between cryptographers and cryptanalysts. Ciphers evolved and new techniques were derived to overcome the ciphers. For example the The Vigenère cipher developed in 1553 remained secure until 1854.



Figure 4: The Vigenère square or Vigenère table

World War 1 and 2 saw rapid development in cryptography in order to secure messages from one's enemy. From encrypted telegraph messages in World War 1, the world saw the birth of the Enigma machine in World War 2.



**Figure 5:** Arthur Scherbius's Enigma Patent —U.S. Patent 1,657,411, granted in 1928





Figure 6: The Enigma Machine (L) being used during WW2 (R)

The Enigma machine was fully utilized by the Germans and was invincible (or they thought so). Enigma was designed to defeat basic cryptanalytic techniques by continually changing the substitution alphabet. Like other rotor machines, it implemented a polyalphabetic substitution cipher with a long period. With single-notched rotors, the period of the machine was  $16,900 \ (26 \times 25 \times 26)$ . This long period helped protect against overlapping alphabets.

Cryptanalysis started before World War 2 in December 1932, by a 27-year-old Polish mathematician, Marian Rejewski, who had joined the Polish Cipher Bureau in September that year.

The effort evolved and it took a large number of mathematicians based in Bletchley Park, in Buckinghamshire England to finally cryptanalyze the machine.

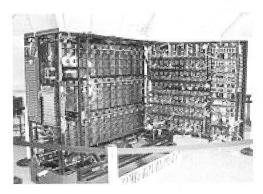


**Figure 7:** During World War 2, codebreakers at Bletchley Park decrypted and interpreted messages from a large number of Axis code and cipher systems, including the German Enigma machine



**Figure 8:** Alan Turing (23 June 1912 – 7 June 1954) was an English mathematician, logician and cryptographer. He is also attributed with the title "father of modern Computers"

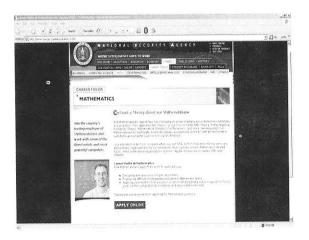
During the Second World War, Alan Turing was a main participant in the efforts at Bletchley Park to break German ciphers. Building on cryptanalysis work carried out in Poland by Marian Rejewski, Jerzy Różycki and Henryk Zygalski from Cipher Bureau before the war, he contributed several insights into breaking both the Enigma machine and the Lorenz SZ 40/42 (a Teletype cipher attachment codenamed "Tunny" by the British).



**Figure 9:** Within weeks of arriving at Bletchley Park, Turing had designed an electromechanical machine which could help break Enigma faster than bomba from 1932, the bombe, named after and building upon the original Polish-designed bomba

Today, historians believe that the work of the code breakers at Bletchley Park shortened the war by two years.

The British Government still operates a code breaking department, at "Government Communication Headquarters", (GCHQ) in Cheltenham. And to this day they rely on mathematicians for their problem solving abilities and logical thinking: GCHQ boasts the highest concentration of pure mathematicians in the country. The National Security Agency of USA also employs a large number of mathematicians for the above purpose.



**Figure 10:** The NSA website recruiting mathematicians

# Modern cryptography

With the invention of the computer, bit and bytes came into play. Encryption now worked on 0's and 1's.

A=01000001, B=01000010,..., Z=01011010,....

Above is an example of binary representation

There are 256 elements represented by 8 bits (1 byte), beginning from the byte 00000000 to 111111111.

In modern cryptography, the length of a key must be of reasonable length and hard to be generated again. In principal, if a key is only 1-bit long there are only 2 possible choices. If it is k-bit long there are  $2^k$  choices.

Assuming if one has a supercomputer that can try a million keys per second, a 56-bit key can be determined in 2285 years. If it is 64-bit long it would take 585,000 years. If it is 128-bit long it would take  $10^{25}$  years, and the universe is only  $10^{10}$  years old.

But before one tries to attempt to create a cryptosystem with a 8-kilobyte key other factors must be taken into consideration especially the practicality and time consumption of the system.

Shannon in 1949 through his celebrated paper stressed out the importance of diffusion and confusion for a cryptosystem to be accepted. Diffusion means spreading out of the influence of a single plaintext digit over many ciphertext digits so as to hide the statistical structure of the plaintext. While confusion is a means to complicate dependence of ciphertext statistics compared to plaintext statistics. Modern cryptography could be divided into two categories. First, the symmetric cryptosystem. This cryptosystem utilizes the same key. Second, the asymmetric cryptosystem which uses different key to encrypt and decrypt with the encryption key made known to public while the private key for decryption is kept private. The relation between the public and private key is some hard mathematical problem which will make it impossible to extract the private key from the mathematical formula even if one knows the public key.

# Symmetric cryptosystems

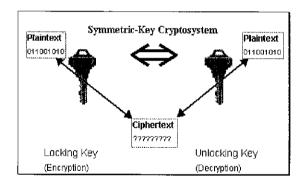
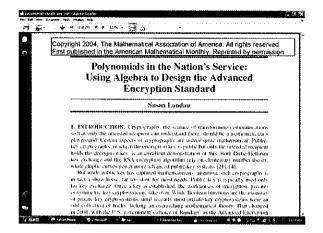


Figure 11: The Symmetric Cryptosystem

An example of symmetric cryptosystems is the Data Encryption Standard, DES (with minimum mathematics involved). Commissioned in 1975 by the U.S Government, it initially utilized 56-bit keys. In the 1990's DES was being questioned, and was replaced by the highly mathematical cryptosystem known as the Advanced Encryption Standard (AES) in 2001. The AES invented by two Belgian inventors Joan Daemen and Vincent Rijmen, is a Block cipher, which means that it works on fixed-length group of bits, which are called blocks. Unlike DES, which is based on a Feistel network, AES is a substitution-permutation network, which is a series of mathematical operations that use substitutions (also called S-Box) and permutations (P-Boxes) and their careful definition implies that each output bit depends on every input bit.



**Figure 12:** Front page "Polynomial's in the Nation's Service"

# Birth of public key infrastructure (asymmetric cryptosystem)

Initially, keys for symmetric cryptosystems (in 1970's) were distributed by trusted employees manually all over the world. They would race across the world with padlock briefcases and distributing keys to everyone who would receive messages from the bank. As business grew this was a nightmare. Then there was an idea, to encrypt using a key that is different from the key that is going to be used to decrypt.

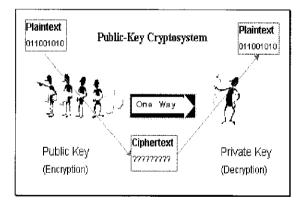


Figure 13: The Asymmetric Cryptosystem

The idea was mooted by Whitfield Diffie (mathematician) and Martin Hellman (electrical engineer) at Stanford, California, USA in 1976 via their paper New Directions in Cryptography [18].

It was until 1977, that a real application was found by Ron Rivest (computer scientist), Adi Shamir (computer scientist) and Leonard Adleman (mathematician) at MIT, USA. The algorithm which is now popularized by the jargon RSA is the core engine in many encryption applications (eg: SSL). Based upon the hard factorization problem, a system implementing this algorithm will have to employ a 1024-bit prime number as a key. However, in May 2007, it was published that researchers from the University of Lausanne, the University of Bonn, and NTT DoCoMo have factorized a 307 digit composite Mersenne number. The 307-digit number itself was not an RSA key the number was 2<sup>(1039)</sup>-1, a special-form number called a Mersenne number which permits an efficient variant of the factoring algorithm in question, the so called Special Number Field Sieve (SNFS) to be used. RSA keys are typically generated by multiplying together two very large prime numbers, each at around 150 digits apiece, and require more labor-intensive General Number Field Sieve (GNFS) to factor. But the project shows that given enough time and computer power, the 1024-bit encryption kevs used on many ecommerce sites could also be cracked in the not-sodistant future. A 2048-bit number is not practical in application since the time needed to encrypt would rise exponentially.

And with patent issues surrounding Elliptic Curve Cryptography, ECC (another PKI which was suggested independently by Neal Koblitz and Victor S. Miller in 1985) the race for another PKI algorithm is imminent. ECC which is now more or less patented by almost every aspect by Certicom Corporation, Canada is an encryption method utilizing the elliptic curve discrete log problem. ECC uses smaller key size which is equivalent to RSA.

RSA and DH	ECC
(key size -bits)	(key size -bits)
1024	160
2048	224
3072	256
7680	384
15360	521

Figure 14: Equivalence table

#### Chaos based cryptography

With issues as mentioned above surrounding algebraic cryptography, new methods must be explored to ensure sustainable methods in securing information.

The relationship between chaotic dynamical systems and cryptography makes it natural to employ chaotic systems to design new cryptosystems. It is based on the facts that chaotic signals are usually noise-like and chaotic systems are very sensitive to initial conditions. Their sensitivity to initial conditions and their spreading out of trajectories over the whole interval seems to be a model that satisfies the classic Shannon requirements of confusion and diffusion [1].

A non-linear system is chaotic if it has a positive Lyapunov exponent. That is,

$$\lambda = \lim_{\substack{t \to \infty \\ |\Delta x_0 \to 0|}} \frac{1}{t} \ln \frac{|\Delta x(x_0, t)|}{|\Delta x_0|} > 0$$
 (2)

Graphically,

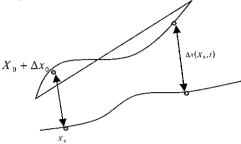


Figure 15

Let us consider a one dimensional dynamical system  $f: I \to I$ . When  $\lambda > 0$  we have

$$\forall \varepsilon > 0, \exists n_1, n_2, \exists x \in U_{n_1}, v_2, \forall n \in [n_1, n_2], \forall z_1, z_2 \in U_{n_1}, v_2$$

$$\exp\{(\lambda - \varepsilon)n\}|z_1 - z_2| < |f''(z_1) - f''(z_2)| < \exp\{(\lambda + \varepsilon)n\}|z_1 - z_2|$$

This means that the initial distance  $|z_1 - z_2|$ between 2 arbitrary points  $|z_1, z_2|$  (which are elements of the neighborhood  $U_{n_1,n_2}$  of point x) after n iterations will increase at least  $exp\{(\lambda-\varepsilon)n\}$  times.

A chaotic dynamical system which is to be considered as a candidate for a cryptosystem must also posses the ergodicity property. Ergodicity implies that the phase space X cannot be nontrivially divided into several parts. That is, if some trajectory starts from any point  $x_0 \in X$ , it never localizes in a smaller region, and knowing the final state of the system we can never restore the region (smaller than X) where the trajectory started.

The chaotic dynamical system must also posses the mixing property. Basically, this means if we start our trajectory at some point  $x_0 \in X$  then after sufficiently many iterations we reach any region in the space X with the same probability.

The idea of using Chaos for data encryption can be traced to the paper by Shannon. Shannon suggested of using measure preserving transformations which depend on their arguments in a 'sensitive' way. A dynamical chaotic system which is to be built upon it a cryptosystem must be guaranteed (either with or without restrictions) that all the periodic points of chaotic f are repelling. That is, if a trajectory happens to come close to a periodic cycle for some k, it will separate from it for indices greater than k. Although chaotic dynamical systems are linearly unstable and unpredictable, they can synchronize which makes them promising candidates for constructing cryptographic systems. It is widely accepted that chaotic dynamical systems have good diffusion and confusion properties.

The following is a simple text book example.

### **Example**

Let  $\gamma$  be a one-dimensional chaotic map with a positive Lyapunov exponent  $\lambda$ ,  $\gamma:[0,1] \to [0,1]$  and  $p \in (0,1)$  be a message to be encrypted. Fix a natural number n (number of iterations (determine the best first through research)) and choose a secret key  $k \in (0,1)$ . Let  $\overline{C}$  be some selected pre-image of P under the map  $\gamma^n$ :

$$\begin{split} \overline{C} &= \gamma^{-n} \left( P \right) \\ \gamma^{n} \left( \overline{C} \right) &= \gamma^{n} \left( \gamma^{-n} \left( P \right) \right) = P \end{split}$$

Then we calculate the cipher text P as

$$C = \overline{C} + k \pmod{l}$$

Decryption is the inverse operation

$$P = \gamma^n (C - k)$$

Let's say an eavesdropper tries to approximate the secret key k by assuming the value as  $k_1$  such that  $|k-k_1|<10^{-20}$ . Then he calculates the value of the plain text  $P_I=\gamma^n\left(C-k_I\right)$ . For  $n=30,\ |\lambda-\varepsilon|\approx 1.558$  (which is reasonable for many dynamical systems). Due to chaos we have:

$$\left|P-P_{1}\right|=\left|\gamma^{n}\left(C-k\right)-\gamma^{n}\left(C-k_{1}\right)\right|{}^{3}\left|e^{n\left(\lambda-\epsilon\right)}\left|k-k_{1}\right|\gg0.5$$

The above statements demonstrates how the chaos property protects the system against a brute force attack (i.e. for n=30, accuracy of guessing equal to  $10^{-20}$  and the Lyapunov exponent  $\lambda \approx 1.6$ .

The following are some techniques used to design cryptosystems from chaotic dynamical systems.

# 1. The Masuda & Aihara Technique [19]

This technique re-defines (discretize) the chaotic map (form a one-to-two mapping to a one-to-one mapping). Discretization is the process of redefining f on the

continuum I to a finite number of intervals. This is due to the fact that the computer is a finite machine.

Let  $f: I \to I$  be a mapping on a continuous interval. During the process of redefining  $f^*: I' \to I'$  where I' is the discrete interval, it must be observed that the condition  $|f - f^*| \le M^{-1}$  remains (M is the number of discrete intervals).

Security analysis includes:

i. Sensitive dependence on plain texts

ii. Sensitive dependence on keys

iii. Relative prime conditions (i.e. possibility to decrease number of iterations n)

iv. Exponential decay of information

v. Bitwise independence

vi. Statistical security analysis

### 2. Yi, Tan & Siew Technique [20]

This technique redefines the chaotic map by adding an extra condition for cases when the mapping falls outside the intended interval. For example, the authors redefined the tent map:

$$F_{\sigma}(x_{i}) = \begin{cases} \frac{x_{i-1}}{a} & 0 \le x_{i-1} \le a \\ \frac{1 - x_{i-1}}{1 - a} & 0 < x_{i-1} \le 1 \end{cases}$$

as follows:

$$G_{a,b}(x_i) = \begin{cases} F_a(x_{i-1}) & 0 < x_{i-1} < 1 \\ b & \text{otherwise} \end{cases}$$

where the secret keys are (a,b,c,K),  $0 < a,b,c < 1,a^{T}b$ ,  $a \ne b$ ,  $F_a(b) \ne a$ , and K is 4n-bit long.

Security analysis includes:

i. Choosing suitable a values

ii. Statistical tests

iii. Other standard cryptographic tests

#### 3. Baptista technique

This cryptosystem is based on the property of ergodicity of chaotic systems (i.e. the eventual visit of the trajectory to all partitions in the phase space as the number of iteration grows). In this work, Baptitsta utilizes the logistic equation, X(n+1) = rX(n)(1-X(n)) and  $X(n) \in [0,1]$ 

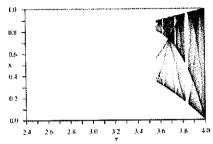


Figure 16: The bifurcation diagram of the logistic map

This technique has the ability to produce different ciphers even for the same plaintext using the same key. This fact will actually leave the attacker helpless to crack the encrypted message. For example let us assume that the secret keys are  $X_0 = 0.232323$  and r = 3.78 and we will use the interval [0.2,0.8]. For simplicity we will use a 4-symbol source  $S_4 = \{s_1, s_2, s_3, s_4\}$ . The plaintext  $s_1$  can have the following ciphers representing it  $E_1 = (5,9,5,3,4,5,...)$ 

## The attack on Baptista's technique

However in 2003, Alvarez [6] designed an ingenious method, now known as the one-time pad attack to crack this cryptosystem. The method could actually decrypt all ciphers encrypted with this method without actually having to know the secret keys. He proved that the ergodic cipher put forward by Baptista behaves as a one-time pad which reuses its key, and as a result, is easy to break. The method of attack is based on the symbolic dynamics of one dimensional quadratic map.

# Counter measures against the one-time pad attack on Baptista's technique

In 2008 M.R.K.Ariffin and M.S.M.Noorani [21] and [22] successfully developed a countermeasure to overcome the one-time pad attack.

A formal treatment for the one-time pad attack was formulated. From the observation of the mathematical behavior of this cryptosystem when under attack via the one-time pad attack, definitions were derived and mathematical explanations for this phenomenon were given.

A theorem was formulated, and if satisfied by a "counter measure" method, would result in the Baptista type chaotic cryptosystem being invulnerable against the one-time pad attack. Finally, an example was provided to simulate an example of this countermeasure method.

#### Conclusion

The science of cryptography has gone a long way since man decided to keep certain information secret from others. This vibrant area of research which combines the sciences of mathematics, computer, engineering, management etc is ever becoming more relevant when we enter the era of information technology.

Thomas L. Friedman in his famous book "The World is Flat" categorized the globalization into 3 vital eras. The first era shrank the world from a large size to medium size while the second era shrank it further from medium size to smaller size. It was during this era that we really saw the birth and maturization of a global economy through the diffusion of the telegraph, telephones, the PC, satellites, fiber optic cable and the early version of the World Wide Web. While the third era which is the current phase, which provides even more platform for individuals to collaborate and compete globally in various kinds of field and industry.

When the world becomes flatter and the wall thinner, we need to be more aware of the security aspect which will be the only barrier that can prevent the globalized world from being out of control. We can't stop globalization but we can still control it. In this era, cryptology will become a vital tool and discipline for the world society to establish a stable and equally prosperous life.

#### References

- [1] Shannon, C.E. 1949. Communication Theory of Secrecy Systems. *Bell Systems Tech. J.* **28**: 656-715.
- [2] Baptista, M.S. 1998. Cryptography with Chaos. *Phys. Lett. A*, **240**: 50-54.
- [3] Alvarez, E., Fernandez, A., Garcia, P., Jimenez, J., and Marcano, A. 1999. New Approach to Chaotic Encryption. *Phys. Lett. A*, **263**: 373-375.
- [4] Alvarez, G., Montoya, F., Romera, M., and Pastor, G. 2000. Cryptanalysis of a Chaotic Encryption System. *Phys. Lett. A*, **276**: 191-196.

- [5] Jakimoski, G., and Kocarev, L. 2001. Analysis of Some Recently Proposed Chaos-Based Encryption Algorithms. Phys. Lett. A, 291: 381-384.
- [6] Alvarez, G., Montoya, F., Romera, M. and Pastor, G. 2003(b). Cryptanalysis of an Ergotic Chaotic Cipher. *Phys. Lett. A*, 311: 172-179.
- [7] Alvarez, G., Montoya, F., Romera, M., and Pastor, G. 2004. Cryptanalysis of Dynamic Look-Up Table Based Chaotic Cryptosystems. *Phys. Lett. A*, **326**: 211-218.
- [8] Garcia, P., Jimenez, J. 2002. Communication through Chaotic Map Systems. *Phys. Lett. A*, **298**: 35-40.
- [9] Pareek, V., Patidar, N.K. and Sud, K.K. 2003. Discrete Chaotic Cryptography Using External Key. *Phys. Lett. A*, **309**: 75-82.
- [10] Alvarez, G., Montoya, F., Romera, M., and Pastor, G. 2003(a). Cryptanalysis of a Chaotic Secure Communication System. *Phys. Lett. A*, **306**: 200-205.
- [11] Alvarez, G., Montoya, F., Romera, M. and Pastor, G. 2003(b). Cryptanalysis of an Ergotic Chaotic Cipher. *Phys. Lett. A.* 311: 172-179.
- [12] Alvarez, G., Montoya, F., Romera, M. and Pastor, G. 2003(c). Cryptanalysis of a Discrete Chaotic Cryptosystem Using External Key. Phys. Lett. A, 319: 334-339.
- [13] Li, S. and Mou, X., Cai, Y. 2001. Improving Security of a Chaotic Encryption Approach. *Phys. Lett. A*, **290**: 127-133.
- [14] Wong, W.K., Lee, L.P. and Wong, K.W. 2001. A Modified Chaotic Cryptographic Method. *Comput. Phys. Commun*, **138**: 234-236.
- [15] Wong, K.W. 2002. A Fast Chaotic Cryptographic Scheme with Dynamic Look-Up Table. *Phys. Lett. A*, **298**: 238-242.
- [16] Palacios, A., Juarez, H. 2002. Crytography with Cycling Chaos. *Phys. Lett. A* (303): 345-351.

- [17] Wong, K.W. 2003. A Combined Chaotic Cryptographic and Hashing Scheme. *Phys. Lett. A*, **307**: 292-298.
- [18] Diffie, W. and Hellman, M.E. 1976. New Directions in Cryptography. *IEEE Transactions on Information Theory*, (IT-22): 644-654.
- [19] Naoki Masuda and Kazuyuki Aihara. 2002. Cryptosystems with Discretized Chaotic Maps. *IEEE Trans. On Circuits and Systems*, **49**(1): 28-40.
- [20] Yi, T. and Siew. A New Block Cipher Based on Chaotic Tent Maps. 2002. *IEEE Trans. On Circuits and Systems*, **49**(12): 1826-1829.
- [21] Ariffin, M.R.K., Noorani, M.S.M. 2008. Modified Baptista Type Chaotic Cryptosystem Via Matrix Secret Key. Phys. Lett. A, 372: 5427-5430.
- [22] Ariffin, M.R.K., Noorani, M.S.M. 2008. Conditions for Counter Measures against One Time Pad Attack on Baptista type chaotic cryptosystem. Special Edition of the Malaysian *Journal of Mathematical Sciences*—to appear.