# THE HYBRID CUBES ENCRYPION ALGORITHM (HiSea)

SAPI'EE BIN JAMEL

A thesis submitted in
fulfillment of the requirement for the award of the
Doctor of Philosophy

Faculty of Computer Science and Information Technology
Universiti Tun Hussein Onn Malaysia

JUNE 2012

# CHAPTER 1

# INTRODUCTION

## 1.1 Background

The emergence of the Internet (network without physical boundary) in 1980s has increased the dependency of organizations on computer and network systems which carry significant data for their daily business transactions (Callie, 1994; Schneier, 2006). Any information which can be deducted from these data can lead to disaster if these messages reached unintended parties. Information such as financial transactions, top secret government information, medical records and personal information transmitted through the Internet can be made available by people with adequate knowledge of the Internet infrastructure such as Cloud Computing (Abawajy, 2011). Even physical security (for example, placing computers and network in a secure building) is no longer adequate to protect data which can travel outside the company parameters as the business transactions are moving from conventional approaches to online transactions where it relies heavily on the Internet as the main medium of transmission. Information security enforces the security of data through confidentiality, integrity and availability so that only intended party obtains the information (message) from the data.

Confidentiality mechanisms are useful to enforce the security of data. A variety of techniques have been developed in these three areas to protect data: access control mechanism, file security control and cryptography. Access control mechanism prevents unauthorized individual from accessing the computer and network system. File security control prevents authorized individual from accessing data beyond their authorization. Cryptography which is the last line of defense protects data using cryptographic algorithm in the situation where access control and file system control fails (Chapple, 2006).

Cryptography involves with the design and implements secure cryptographic algorithm or cipher while cryptanalysts tries to break any available cryptosystem. The idea of building a secure system for data protection is not new. It can be traced back to as early as the 16th and 17th Century (Leary, 1996). The importance of cryptography is apparent with the advent of complex network technologies (such as Internet) where data are no longer confined by physical boundaries (Schneier, 2006). Unfortunately, a totally secure cryptographic algorithms are difficult to build due to the existence of continuous challenges from cryptanalysts whose objectives are to find ways to defeat any known cryptographic algorithms. Furthermore, every cryptographic algorithm must pass the test of time and general acceptance by cryptographic communities as shown in the selection of Advanced Encryption Standard (AES) in 2000 (NIST, 2001).

Before the year 2000, most ciphers are kept secret by its designer to avoid attacks by cryptanalysts. However, after that, designing and publishing the ciphers to the public have become a recent trend in the cryptographic community. Publishing the algorithm reveals many advantages compared to the traditional approach of total secrecy of the whole algorithm design to the finished product. Callie (1994) outlines several advantages of this new trend. The advantages include overcoming the problem of maintaining the secrecy of an algorithm and the cost of developing the algorithm. Callie (1994) also pointed out that the cryptographic community usually tests the strength of the algorithm since the cost of developing a secure proprietary algorithm is very high. Trappe (2002) highlighted the important of openness in the algorithm where any suspicious codes (trapdoor) that is known only to the developer of the algorithm must be avoided.

With the wider use of ciphers to protect data, indirectly give continuous needs for new ciphers which meet openness criteria and support by mathematical formulation as shown in Rijndael (Daeman, 2000). Cryptographic algorithms developed based on strong mathematical formulation has an advantage where it can be evaluated systematically using mathematical modeling. Furthermore, proven result based on earlier finding can be useful to ensure that at least the model is correct. Continuous research in cryptographic communities to find suitable mathematical model for a new cryptographic model motivates this research.

Current block cipher development can be divided further into binary block ciphers and non binary block cipher. Binary block ciphers mostly dominate the development of the existing cryptographic standards such as Data Encryption Standard (DES) and Advanced Encryption Standard (AES). Research into non binary block cipher on the other hand is still new where the standard criteria is not available and only two ciphers (Granboulan, 2006; Baigneres, 2007) are available in the literature on this topic.

Adoption of elegant mathematical properties such as orthogonal Latin squares in Balanced Block Mixer (Ritter, 2004) and magic cube transformation as in Image Encryption Algorithm (Shen, 2005) open up opportunities for further development of new transformation based on permutation of integer numbers in development of non binary cipher.

## 1.2 Problem Statement

Permutation of finite set of numbers or symbols plays an important role in the development of ciphers as shown in simple transposition ciphers (Menenzes, 1996). In this cipher, permutation is used to mix up symbols or numbers to create ciphertext. This technique preserved the number of symbols or number of a given type within a block which makes it easy to be analyzed by cryptanalyst if the block size is small.

Development of efficient computer hardware and software make *permutation only algorithm* more prone to attack by cryptanalysts. Modern cryptographic algorithms such as Rijndael, Twofish, SAFER+ (AES, 2000) used combination of substitution and transposition to enhance the complexity of the ciphertext. Substitution creates confusion in the ciphertext to avoid attack just by applying simple permutation of symbols to get the encryption or decryption keys. Substitution Boxes or SBoxes are used as confusion element in Rijndael (Daeman, 2000) and Twofish (Schneir, 1999). Transposition added diffusion to the overall ciphertext block which disperse the bits in the message so that any redundancy in the message is spread out throughout the ciphertext. Maximum Distance Separable (MDS) matrix and Pseudo-Hadamard Transformation (PHT) are two diffusion primitives used in Rijndael and Twofish respectively.

Permutation of integer as appear in Orthogonal Latin square has been used as a fundamental concept in mixing data in Balanced Block Mixer (Ritter, 2004). The application of magic cubes of order *n* as part of the design of cryptographic algorithm appears in Image Encryption Algorithm (Shen, 2005; Li, 2005), it uses the rule of the magic cube game where each entries of the sub-cubes surfaces form six $\left(n \times n\right)$ matrices. The transformation is performed by changing the order of the sub-cubes by shifting and rotating rows and columns along its cube surface. However, permutation using simple combination of row-shift and column-shift was proven to be weak against Differential Chosen-Plaintext attack as highlighted by Li *et al*. (2008). This weakness provides an opportunity for further development of complex permutation based on a cube which can be used as part of cipher design. Granboulan et al. (2006) emphasis the need for pseudo-random permutation of an arbitrary set is needed. For example in the system that stores data in decimal values. Therefore, a new non binary block cipher is introduced based on permutation of integer numbers in a new hybrid cubes transformation.

## 1.3 Motivation

Introduction of many new computer applications and increasing numbers of organization relaying on the Internet for their daily transactions, indirectly provide many opportunity for the development of new ciphers. Even though there are many ciphers available in the market for protecting secrecy of our message or data, cryptographic communities are actively creating an alternative cipher suitable for ever increasing new applications. There are three main reasons for this. Firstly, export restriction of ciphers by the United States of America indirectly restrict Internet community out sides the United States to use their ciphers for protecting message without any pre-arranged agreement.

Secondly, many ciphers developed by various governments and organizations around the world are kept secret since cost for developing them can be expensive if the intended security level is high. For this reason, not many ciphers are available for free to be used by the Internet users. Thirdly, commercial applications such as banking transactions still rely on binary block cipher to ensure the security of their transactions. The need for ciphers which use integer numbers were introduced by Granboulan et al. (2005) in their cipher called TOY100 and Baigneres et al. (2007) with a cipher called BEAN16. These type of ciphers do not require message to be converted to binary number in the encryption and decryption algorithm. Thus calculation based on integer number can be performed on message, key and the ciphertext.

Besides these three reasons, researchers are trying to introduce various techniques for increasing the complexity of their mathematical approach which lead to higher security in the overall implementation. Earlier cipher such as Caesar cipher (Leary, 1996) was developed based on permutation and combination of integer numbers. This technique is adequate during that time because the cipher was kept secret. Modern ciphers such as Rijndael, Twofish, SAFER (AES, 2000) use two dimensional (2D) matrices in their internal calculation to increase the complexity in the internal operation of the cipher. This technique can reduce the possibility of attack by cryptanalysts. Furthermore, matrices provide an elegant ways to represent numbers and various mathematical transformations can be performed easily using computer.

The application of a three-dimensional (3D) cube for solving problems increase the complexity further since it is a combination of 2D matrices. It has been used in several disciplines such as in data replication in Distributed Databases and Steganography. In Distributed Databases, faces of cube have been used to calculate nodes for solving shortest-path data replication problem (Mat Deris, 2008). Similar concept also appears in Steganography where faces of the cube is used to create confusion in the encryption algorithm (Shen, 2005; Li , 2005).

This thesis presents an extension to Magic Cube Transformation used in Image Encryption Algorithm (Shen, 2005; Li, 2005) which uses the face of magic cube as the method to create confusion in the ciphertext. A new hybrid cube transformation is introduced which uses entries of magic cube layers to develop new three-dimensional hybrid cubes of order 4. This new model used entries in the hybrid cube layers instead of surfaces of a cube as in Image Encryption Algorithm. Possibility of getting complex combination of invertible matrices based on the combination using layer of magic cubes has motivated us to investigate and develop a new non binary cipher which has displayed good diffusive characteristics. Our new non binary block cipher called The Hybrid Cubes Encryption Algorithm (HiSea) was developed based on concepts of orthogonal Latin square, magic squares and magic cubes and a newly introduced Hybrid Cube Transformation. Thus, HiSea can be used as an alternative non binary cipher for encryption and decryption of 64 character messages.

## 1.4 Objectives

The objectives of this research are as follows:

i) To propose a hybrid cube transformation based on the entries of magic cube layers of order 4 instead of its surfaces as proposed in Image Encryption Algorithm (Shen, 2005; Li, 2005).

ii) To propose a new key schedule algorithm based on combination of hybrid cubes of order 4 layers as the matrix generator.

iii) To propose a new non binary block cipher which consist of three components: key schedule algorithm, encryption algorithm and decryption algorithm.

iv) To implement a new non binary block cipher based on the proposed approach in (iii).

v) To analyze the strength of the cipher based on test tailored for non binary block cipher.

## 1.5 Contributions

This research is to explore on solving the following four issues. Firstly, the research needs to come out with all possible set of Latin squares of order 4 which is required for the generation of orthogonal Latin square of order 4. Secondly, all possible combination of orthogonal Latin squares of order 4 need to be generated using Latin squares of order 4. Thirdly, the research needs to propose a new transformation based on magic cube layers. Finally, a new non binary block cipher based on the newly introduced transformation for encrypting and decryption 64 character messages is formed. Thus the main contributions in this study are in the area of block cipher development in cryptography. The following are contributions of the research:

i) A complete set of Latin squares of order 4 which was generated using combination and permutation of row and column of set $A = \{1, 2, 3, 4\}$.

ii) A complete set of orthogonal Latin square of order 4 generated using combinations of Latin squares from (a) above.

iii) A new hybrid cubes of order 4 transformation based on combinations of magic cubes layers of order 4. These invertible hybrids layers were form using inner matrix multiplication from layers of magic cubes. These layers are used to generate encryption and decryption keys for the newly proposed block cipher.

iv) A new non binary block cipher which consists of encryption algorithm, decryption algorithm and key schedule algorithm. These encryption and decryption algorithms can be used to encrypt and decrypt 64 character messages.

## 1.6 Scope of Study

This research concentrates on the development of new transformation based on three dimension (3D) cube which can be used in the development of new non binary block cipher.

## 1.7 Thesis Organization

This thesis will discuss various aspects involved in the design of the new non binary block cipher. The following is the outlines of the thesis.

Chapter 2 describes related literature reviews which are relevant in the design, development and suitable security analysis for Hybrid Cubes Encryption Algorithm. Earlier part of this chapter will discuss general components of symmetric block cipher and two important principles involved in the design of new block cipher. Examples of binary and non binary block ciphers are briefly reviewed and explained to show their characteristics in meeting intended design principle. Various techniques for evaluating the security on non binary block cipher are also presented later in this chapter.

Chapter 3 presents theoretical concepts related to the design of the new non binary block cipher and the overall design framework of this research. This chapter is divided into two sections. The first section outlines concepts and design of hybrid cubes and the second section discusses the design of key schedule algorithm, encryption algorithm and decryption algorithm.

Chapter 4 provides the implementation of the new non binary block cipher. This chapter is also divided into two sections. The first section describes the implementation of Hybrid Cubes based on Latin squares, orthogonal Latin squares and Magic cubes. The second section discusses the implementation of Key Schedule algorithm, Encryption and Decryption algorithm.

Chapter 5 deals with security analysis of the new cipher based on selected techniques suitable for non binary block cipher. Several binary block cipher evaluation techniques will be modified to suit the non binary implementation style. The results from the analysis are presented and discussed to verify the strength of the proposed algorithm.

Chapter 6 provides conclusion and direction for further research in this research topic.

# CHAPTER 2

# LITERATURE REVIEW

## 2.1 Introduction

This chapter describes related mathematical terms, concepts and principles, existing ciphers and its evaluation criteria which are relevant in the development of the new non binary cipher.

## 2.2 Overview of Mathematical Notations

This sub-section describes some of relevant mathematical terms used in the thesis.

### 2.2.1 Modulo Arithmetic

Modulo arithmetic plays an important role in the development of block cipher which use positive integer number $(Z)$ as it basic element. It is used to ensure that the integer numbers are within a finite set as defined in Definition 2.1 (Gilbert, 2005).

***Definition 2.1***

Let m be a fixed positive integer. If $a, b \in Z$, we say that "a is **congruent** to b **modulo** m" and write

$$a \equiv b \pmod{m} \quad \text{whenever} \quad m \mid (a-b) \qquad (2.1)$$

The condition for $a$ to be congruent to $b$ modulo $m$ is equivalent to the condition that

$$a = b + km \quad \text{for some } k \in \mathrm{Z}. \qquad (2.2)$$

Congruences occur in everyday life. The short hand of a clock indicates the hour modulo 12, while the long hand indicated the minute modulo 60. For example, 20 hours after midnight, the clock indicates 8 o'clock because $20 \equiv 8 \pmod{12}$. The congruence class modulo $m$ of the integer is defined in Definition 2.2

***Definition 2.2***

The congruence class modulo $m$ of the integer $a$ is the set of integers

$$[a] = \{x \in \mathrm{Z} \mid x \equiv a \bmod m\}. \qquad (2.3)$$

The set of congruence class of integers, under the congruence relation modulo $m$, is called the set of integer modulo $m$ and is denoted by $\mathrm{Z}_m$. The set $\mathrm{Z}_m$ is the quotient set of $\mathrm{Z}$ defined by the congruence relation modulo $m$ and

$$\mathrm{Z}_m = \{[0], [1], \cdots, [m-1]\}. \qquad (2.4)$$

This congruent relation is explained in Example 2.1.

**Example 2.1**

$\mathrm{Z}_4 = \{[0], [1], [2], [3]\}$, where the four congruence classes are

$$[0] = \{\ldots, -8, -4, 0, 4, 8, 12, \ldots\} = \{4k \mid k \in \mathrm{Z}\}$$

$$[1] = \{\ldots, -7, -3, 1, 5, 9, 13, \ldots\} = \{4k+1 \mid k \in \mathrm{Z}\}$$

$$[2] = \{\ldots, -6, -2, 2, 6, 10, 14, \ldots\} = \{4k+2 \mid k \in \mathrm{Z}\}$$

$$[3] = \{\ldots, -5, -1, 3, 7, 11, 15, \ldots\} = \{4k+3 \mid k \in \mathrm{Z}\}.$$

The suitability of modulo arithmetic as the basis for cipher development is because addition and multiplication are well defined in $\mathrm{Z}_m$.

### 2.2.2 Combination

A binomial is a sum of two quantities, such as $a+b$. The Binomial Theorem yields the expansion of $(a+b)^n$ for each positive integer exponent. The coefficients that occur in the binomial expansions are called binomial coefficients and can be conveniently written in terms of factorials as defined in Definition 2.3 (Gilbert, 2005).

***Definition 2.3***

If $0 \le r \le n$ then the binomial coefficient $\binom{n}{r}$ or n choose r is defined by

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!} \tag{2.5}$$

where 0! is defined to be 1, so that $\binom{n}{n} = 1$.

This binomial coefficient is to determine the number of combination ($C$) on $n$ objects chosen $r$ at a time in the construction of Latin squares of order 4.

### 2.2.3 Permutations

A permutation ($P$) of a set $S$ with $n$ elements is a listing of the elements of $S$ in some orders where each element in the set can occur only once. A permutation of a set can be defined as a bijective function as in Definition 2.4 (Gilbert, 2005).

***Definition 2.4***

Let $S$ be a finite nonempty set. A permutation $\sigma$ on the set $S$ is a bijection

$$\sigma : S \to S.$$

If $S = \{a, b, c, d\}$, then one permutation of $S$ is defined by $\sigma(a) = b,\ \sigma(b) = d,$

$\sigma(c) = a,$ and $\sigma(d) = c.$ We can think of the permutation $\sigma$ as a rearrangement of the elements $a, b, c, d$ to form $b, d, a, c.$ A convenient way of writing this permutation is

$$\sigma = \begin{pmatrix} a & b & c & d \\ b & d & a & c \end{pmatrix} \tag{2.6}$$

where the elements of $S$ are written in the top row, and their corresponding images under $\sigma$ are written below. The total number of permutation can be calculated using the following equation:

$$nPr = \frac{n!}{(n-r)!} \tag{2.7}$$

where 0! is defined to be 1.

If $S = \{1, 2, 3, 4\}$, the set $S_4 = \{1, 2, 3, 4\}$ has twenty four different permutations. The total number of permutation can be calculated using Equation 2.1. where $4P4$ is 4!

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}, \sigma_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$$

$$\sigma_5 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}, \sigma_6 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}, \sigma_7 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}, \sigma_8 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}$$

$$\sigma_9 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}, \sigma_{10} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}, \sigma_{11} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}, \sigma_{12} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$$

$$\sigma_{13} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}, \sigma_{14} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \sigma_{15} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}, \sigma_{16} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

$$\sigma_{17} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}, \sigma_{18} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}, \sigma_{19} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}, \sigma_{20} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}$$

$$\sigma_{21} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}, \sigma_{22} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}, \sigma_{23} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix} \sigma_{24} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}.$$

This permutation is applied in the construction of Latin squares of order 4.

### 2.2.4 Matrices

Application of Linear Algebra (i.e matrices) exist in many cipher development because it offers compact and elegent way of representing many numbers and equations.

#### 2.2.4.1 Square Matrices

A matrix that is $m \times m$ for some value of $m$ is said to be square. Square matrices such as $(4 \times 4)$ has been adopted in many cipher such as Rijndael (Daeman, 2000), Twofish (Schneier, 1999) and SAFER+ (Massey, 1999).

#### 2.2.4.2 Invertible Matrix

Square matrix which has an inverse is called invertible matrix. Only invertible matrices are suitable for encryption and decryption in a symmetric block cipher. One of the common methods to determine if a matrix has an inverse is by using the determinant.

For $(4 \times 4)$ matrix A, the determinant is calculated as follow:

$$A = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix},$$

$$\det |A| = (afkp + bglm + chin + dejo) - (cfip + belo + ahkn + dgjm).$$

Matrix A has an inverse if and only if the value of its determinant is not equal to zero or $\det |A| \neq 0$.

## 2.3 Overview of A Cipher System

A cipher consists of three algorithms: Encryption algorithm, decryption algorithm and key schedule algorithm. Cryptosystem can be further classified into deterministic encryption scheme and probabilistic encryption scheme (Goldwasser, 2008). In deterministic encryption scheme, encryption of messages with keys will always produce similar ciphertext regardless of how many execution the encryption processes are repeated. There is a strict rule that messages should have one-to-one relation with the ciphertext and the keys. Otherwise the decryption process will create more than one answer for a particular message. Probabilistic encryption scheme allows messages to have many different ciphertext using different keys. This scheme is semantically secure where it is difficult for cryptanalyst to derive useful information about the message given only ciphertext and corresponding encryption key or public key. Even though deterministic encryption scheme is less secure than probabilistic encryption scheme, one-to-one mapping of message with encryption key enhanced the search of data, for example, within the encrypted database files.

The development of encryption and decryption algorithms can be further divided into two general categories based on the characteristic of its keys: symmetric and asymmetric key or also known as public key (Schneier, 1996, Trappe, 2002). In symmetric algorithm, both of the encryption and decryption keys are kept secret and known to both parties who want to communicate. The encryption keys can be calculated from the decryption keys and vice versa. Thus the strength of symmetric algorithm is strictly based on the secret of encryption and decryption keys. Symmetric algorithm is simple to use. The sender and receiver only need to specify and share the secret key and then begin to encrypt and decrypt messages. Different keys can be allocated to different party which increases the overall message security.

Symmetric algorithm can be classified into two categories based on the grouping of message bits: block and stream cipher. The implementation of encryption and decryption algorithms can be divided into binary and non binary block cipher based on the final output of the message, key and the ciphertext. For binary block ciphers, message bits size are 64, 128, 192 and 256. A typical block size is 64 bits. Standard for non binary block cipher has not been set and varies according to the cipher implementation. In stream cipher, the block size is one character and is no longer suitable for software processing because the key must be as long as the message.

In asymmetric or public key cryptography, the key used for the encryption is totally different from the decryption key. Furthermore, the encryption key is made known to the public and can be used to encrypt message intended for the owner of the key. Only the owner of the encryption key can decrypt the message with the right pair of decryption keys. The strength of asymmetric cryptography is that it can provide the users with message authentication detection when used with digital signature. The user can verify that the message is from a genuine user.

In this research, symmetric block cipher is selected over asymmetric block cipher because it is simple to use for encrypting and decrypting messages. Further more, symmetric block cipher is the existing Advanced Encryption Standard (AES) for encrypting and decrypting messages adopted by The United States of America (MS ISO/IEC 18033-3, 2005).

The next sub-section will discuss briefly on the basic component of symmetric block cipher which is similar for binary and non binary cipher.

## 2.4 Components of Symmetric Block Cipher

Symmetric block cipher consists of five basic components: messages, encryption algorithm, scrambled message, decryption process and key schedule algorithm (Stalling, 2003). The original message (plaintext) will go to encryption process to be encrypted into scrambled messages (ciphertext) which is usually transmitted using secured or unsecured channel. The encryption process will convert the plaintext into ciphertext using secret key(s) generated from key schedule algorithm for encrypting the message. Then, the scrambled message will be transferred to the appropriate recipients which use the decryption process to get back the original message from the ciphertext. The whole process is shown in Figure 2.1.

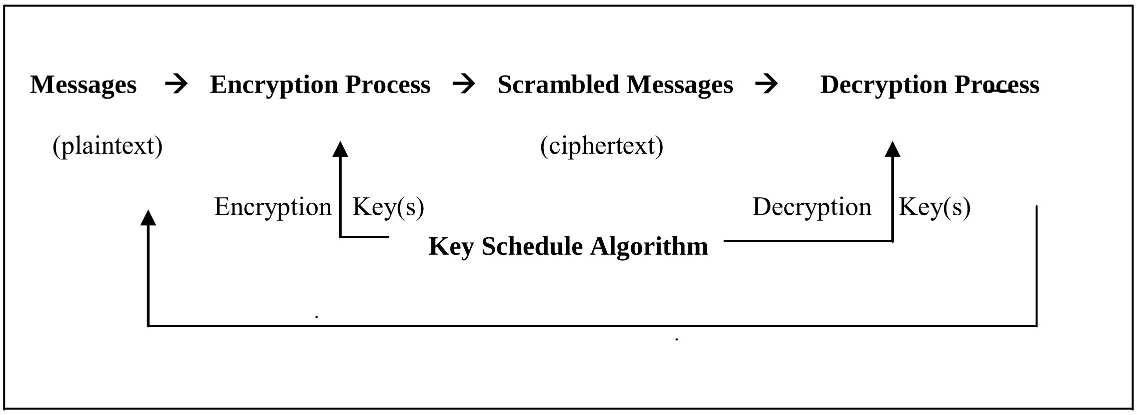

Figure 2.1. Basic Components of Symmetric Block Cipher

The detailed descriptions of each component in Figure 2.1 are discussed in the following sub-sections.

### 2.4.1 Messages or plaintext

Message is the original message (readable to human) or data that needs to be protected during transmission or storage and acts as the input to the encryption algorithm. The current standard for binary block cipher's message block length as required by Advanced Encryption Standard (AES, 2000) is 128-bit, 192-bit and 256-bit. Most algorithms split the data into word size of 32-bit (four bytes). Rijndael, Twofish and SAFER+ uses American Standard Characters II (ASCII) characters for the message and ciphertext. In contrast, A Large Block Cipher (Sastry, 2010) uses ASCII decimal for messages, keys and the ciphertext. For a non binary block cipher, decimal representation is commonly used as the message as in TOY100 (Granboulan, 2006) and DEAN18 (Baigneres, 2007).

### 2.4.2 Encryption Process

This is the process of converting the message into ciphertext using unique encryption key(s) generated from the key schedule algorithm. Moreover, most modern ciphers utilized several input mixing technique to enhance the overall security. Some common techniques include: Whitening (message are XORed with four key words) as in MARS (MARS, 1999), RC6 (Rivest, 1999) and Twofish (Schneir, 1999), Feistel Structure, Substitution and Permutation Network (SPN) in Data Encryption Standard (DES), Maximum Distance Separable Matrices (MDS) and Pseudo-Hadamard Transform (PHT) in Twofish (Schneir, 1998), Substitution Boxes (S-Boxes) and MDS in Rijndael (Daeman, 1999).

### 2.4.3 Scrambled Message (ciphertext)

Scrambled message or ciphertext is the output from the encryption algorithm. Two common format for ciphertext are hexadecimal and decimal ASCII characters. Hexadecimal format (0..F) which can be easily processed using binary boolean operators such as XOR, OR, AND and NOT. Decimal ASCII characters on the hand ease the basic matrices operation such as addition and multiplication as implemented in A Large Block Cipher (Sastry, 2010), TOY100 (Granboulan, 2006) and DEAN18 (Baigneres, 2007).

### 2.4.4 Key Schedule Algorithm

Key schedule algorithm generates secret keys and plays a very important role in the design of the overall encryption and decryption algorithm. The encryption keys also become an input to the encryption algorithm as shown in Figure 2.1. Poor key generation algorithm will generate weak keys for the encryption process which can be easily attack using Brute Force technique. In this attack, the attacker will try all possible keys to get the appropriate plaintext from the ciphertext.

It is considered a successful attack if the encryption key is proven to be correct with less than the total number of possible $2^{128}$ trials. For 128-bit key, the attacker will try all possible key combination. Any successful trial less than $2^{128}$ means that the encryption algorithm has been compromised. This is the primary reason why AES competition enforced that all cryptographic algorithms must support three different key lengths: 128-bit, 192-bit and 256-bit. The current accepted standard for encryption and decryption key size is 128-bit.

### 2.4.5 Decryption Process

Decryption process for symmetric block cipher is almost similar to the encryption process with the exception of the decryption keys and any operation on the intermediate ciphertext such as confusion and diffusion process are applied in reverse as in the encryption process.

## 2.5 Related Principles for Symmetric Block Cipher Design

In this subsection, several important principles and concepts which influence the development of block ciphers are discussed in brief.

### 2.5.1 Kerchoffs's Principle

Kerchoffs's Principle introduced by Auguste Kerckhoffs in 1883 is one of the most important assumptions which changed the way cryptographer build new cryptographic algorithms (Trappe, 2002, Menezes, 1996). According to this principle, the strength of the overall block cipher strength should rely only on the encryption and decryption keys and cryptanalysts can access to the encryption and decryption method. This principle ensure that they were no secret codes (trapdoors) embedded by the original designer of the cipher. Furthermore, cryptographic communities can openly evaluate the correctness of its fundamental mathematical formulation or the implementation aspect of it (test of time). Besides this principle, Shannon's Principles also plays an important role to be considered when developing a new block cipher.

### 2.5.2 Shannon's Principles

Two important principles introduced by Shannon's (1949) are principle of confusion and diffusion (Mel, 2001). These two principles, touches more on the functionality of modules inside encryption and decryption algorithm. The relationship between these two principles with the component of block cipher is described in Figure 2.2.

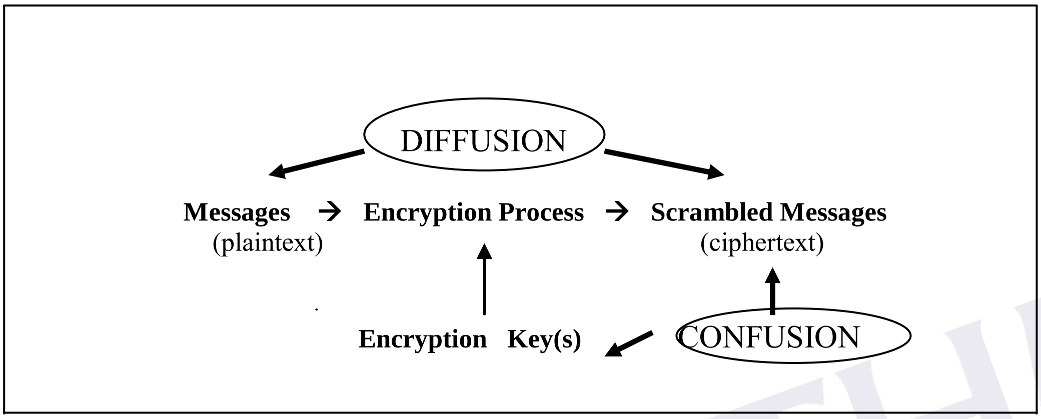

Figure 2.2 Confusion and diffusion in relation with cipher components

Based on Figure 2.2, the effect of diffusion is between messages and the ciphertext. Diffusion function will ensure that small changes made to the message will create an avalanched effect to the ciphertext. Confusion on the other hand, introduces a non-linear relationship between encryption keys and the ciphertext. Both of these principles are explained in the next sub-section.

### 2.5.2.1 Principle of confusion

Confusion hides the relationship between ciphertext and encryption or decryption keys. This is to prevent cryptanalyst from using ciphertext to guess the secret encryption keys. Gordon et al. (1982) introduce a technique to replace one byte to another using a fixed table called Substitution Boxes or S-Boxes. Several works for enhancement of confusion techniques have been proposed in the literature such as avalanche (Feistel, 1973), completeness of Substitution-Permutation (SP) cipher (Kam,1979), Strict Avalanche Criterion (SAC) (Webster, 1985), construction of nonlinear S-Boxes ( Pieprzyk, 1989), perfect nonlinearity of S-Boxes, new design criteria of S-Boxes based on information theory (Dawson, 1991), propagation criterion and Global Avalanche Criterion (Zhang, 1995). Confusion component for non binary cipher appear in TOY100 (Granboulan, 2006) and DEAN18 (Baigner, 2007). The main objective of S-Boxes is to remove any linear relationship between the message and the ciphertext.

### 2.5.2.2 Principle of diffusion

Diffusion is a technique or method for creating unpredictable cascading changes (random) made to the ciphertext when small modifications are made to the inputs (plaintext or cryptographic key). This principle hides the relationship between ciphertext and plaintext. Developing an efficient diffusion element for symmetric block (similar key for encryption and decryption) cipher started way back in 1949 as stated by Shannon (Shannon, 1949). Finding an effective technique to achieve perfect code diffusion is also an ongoing research. Schnorr et al. (1993), enhanced an earlier method of diffusion (using programming *nested loop* e.g Data Encryption Standard (DES)) using multi-permutation. Vaudenay (1995) later defines a linear multi-permutation is equivalent to Maximum Distance Separable (MDS) code which comply the Singleton Bound Criterion. MDS guarantees a certain degree of diffusion which can be used to evaluate the strength of any cipher.
The next section will discuss several issues that are relevant in the process of designing of next non binary block cipher.

## 2.6 Design Component of Block Cipher

The selection of the following components is based on the suitability and some designs are slightly modified for our proposed non binary ciphers.

### 2.6.1 Whitening

Whitening is the process of mixing the message with key material before mixed with session keys designed to increase the difficulty of key search attacks against the cipher. For binary cipher such as Twofish (Schneier, 1999), exclusive OR (XOR) is used for mixing messages. This process can be applied in non binary block cipher by mixing messages with key material using Addition or Multiplication operators.

### 2.6.2 Confusion Function

Digital Encryption Algorithm for Numbers or DEAN18 (Baigners, 2007) is a non binary cipher that encrypts blocks of 18 decimal digits. This cipher could be used to encrypt a credit card number and the internal structure is similar to Rijndael (Deaman, 2002). DEAN18 uses SubBytes function which applies a fixed bijective substitution box (S-Box) to the intermediate message. The application of S-Box is based on the mapping of an element $(a,b) \in Z_{10}^{2}$ which is represented as integer $10 * a + b \in [0,99]$ on each 2-digit element of the ciphertext.

Table 2.1. A fixed substitution box on $Z_{10}^{2}$ (Baigneres, 2007)

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
| **27** | **48** | **46** | **31** | **63** | **30** | **91** | **56** | **47** | **26** | **10** | **34** | **8** | **23** | **78** | **77** | **80** | **65** | **71** | **43** |
| 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 |
| **36** | **72** | **29** | **79** | **83** | **7** | **58** | **95** | **69** | **74** | **67** | **35** | **32** | **59** | **82** | **14** | **75** | **99** | **24** | **87** |
| 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 |
| **16** | **90** | **76** | **51** | **28** | **93** | **50** | **38** | **25** | **3** | **13** | **97** | **55** | **60** | **49** | **86** | **57** | **89** | **62** | **45** |
| 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 |
| **18** | **37** | **1** | **6** | **98** | **68** | **39** | **17** | **19** | **20** | **64** | **44** | **33** | **40** | **96** | **2** | **12** | **41** | **52** | **85** |
| 80 | 81 | 82 | 83 | 84 | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 |
| **81** | **5** | **0** | **15** | **54** | **88** | **92** | **21** | **84** | **22** | **53** | **11** | **4** | **94** | **42** | **66** | **70** | **9** | **61** | **73** |

The designer of DEAN18 claims that it will require $3.8\times10^{16} \approx 2^{61}$ samples to attack four rounds of DEAN18.

### 2.6.3 Diffusive Techniques for Rijndael, Twofish, SAFER+ and TOY100

In this sub-section, review of three diffusive techniques for binary ciphers (Rijndael, Twofish and SAFER+) and non-binary cipher TOY100 are made.

#### 2.6.3.1 Rijndael

Based on the Advanced Encryption Standard (AES) candidates in 1997, there are three common methods used as the building block to design the diffusive component of each algorithm: Maximum Distance Separable (MDS), Pseudo-Hadamard Transform (PHT) and Invertible Linear Transformation Matrix (M).

MDS matrix ia a linear transformation matrix ($\mathbf{M}_{nxn}$) which is used as one of the methods to mix the input or message. It is used as a perfect diffusion primitive in Rijndael (Daeman, 2002) and Twofish encryption algorithm (Schneier, 1999). Pseudo-Hadamard Transform (PHT) is also one of the technique to mix the intermediate output from two MDS matrix used in Twofish encryption algorithm to further enhanced the security of this algorithm. Invertible Linear Transformation Matrix (M) is square matrix used in SAFER+ encryption algorithm (Massey, 1999). This algorithm also uses PHT and Armenian Shuffle to provide the overall diffusion function.

Rijndael diffusion is performed using the following equation:

$$\mathrm{C}(x) = \mathrm{A}(x) \,.\, \mathrm{B}(x) \bmod \mathrm{M}(x). \qquad (2.8)$$

where

- A(x) is 128-bit message or plaintext represented in Hexadecimal.
- B(x) is the MDS and it is built with combination of values: 01 implies no processing; 02 represent multiplication with *x* which can be implemented with

dedicated routines; and multiplication with 03 is a combination of 02 and Exclusive-*OR (XOR)* with the operand or ($x$ +1).

- M(x) is the irreducible polynomial divisor ($x^8 + x^4 + x^3 + x + 1$).
- C(x) is the intermediate ciphertext values

The value of Rijndael's 4-by-4 MDS matrix $B(x)$ is described below:

$$B(x) = \begin{pmatrix} 02 & 01 & 01 & 03 \\ 03 & 02 & 01 & 01 \\ 01 & 03 & 02 & 01 \\ 01 & 01 & 02 & 03 \end{pmatrix}$$

The effect of this MDS matrix on the ciphertext is shown in example 2.2 below.

**Example 2.2**

For the Rijndael algorithm, the first input used to see the effect of MDS function is [15 15 15 15] which is similar to [F F F F] in Hexadecimal. The reason for using decimal values is to simplify the calculation in MATLAB. In the second row, the value of the input is reduced by 1 and the result should be different from the first row. For diffusive matrix to be optimum, changing a single bit to the input vector will produce different output values as shown in Table 2.2. In row 6 to 9, one byte is set to zero and the output generated is different from the first row which indicates that the MDS matrix is efficient in providing the diffusion function.

Table 2.2 Result from Rijndael MDS Matrix

| **Row** | **Input (Decimal)** | **Intermediate values(Decimal)** |
|---|---|---|
| 1 | [15 15 15 15] | 105 105 90 75 |
| 2 | [15 15 15 14] | 104 104 88 117 |
| 3 | [15 15 15 13] | 103 103 86 114 |
| 4 | [15 15 15 12] | 102 102 84 108 |
| 5 | [15 15 15 11] | 101 101 82 105 |
| 6 | [15 15 15 0] | 90 90 60 75 |
| 7 | [15 15 0 15] | 90 60 60 105 |
| 8 | [15 0 15 15] | 60 75 75 105 |
| 9 | [ 0 15 15 15] | 75 90 75 75 |

### 2.6.3.2 Twofish

Twofish encryption algorithm (Schneier, 1999) also adopts similar approach as in Rijndael except with different coefficients and irreducible polynomial value. The following 4-by-4 MDS matrix $V(x)$ is used with irreducible polynomial divisor of ($x^8 + x^6 + x^5 + x^3 + 1$).

$$V(x) = \begin{pmatrix} 01 & EF & 5B & 5B \\ 5B & EF & EF & 01 \\ EF & 5B & 01 & EF \\ EF & 01 & EF & 5B \end{pmatrix}$$

Reason for selection of 01, 5B and EF as for the MDS matrix is to ensure multiplication with these coefficients represents two Linear Feedback Shift Register (LFSR) right shift modulo irreducible polynomial and addition of a few bytes with XOR. The effect of this diffusion function is shown Table 2.3.

**Example 2.3**

Similar technique as in Example 2.2 is also was performed on Twofish's MDS matrix as shown in Table 2.3 The output from row 2 to 9 is different from row 1. This shows that the MDS matrix is optimum.

Table 2.3 Result form Twofish MDS Matrix

| Row | Input (Decimal) | Intermediate values(Decimal) |
|---|---|---|
| 1 | [15 15 15 15] | 247 247 247 193 |
| 2 | [15 15 15 14] | 8 246 8 102 |
| 3 | [15 15 15 13] | 130 245 130 11 |
| 4 | [15 15 15 12] | 252 244 252 281 |
| 5 | [15 15 15 11] | 13 243 13 190 |
| 6 | [15 15 15 0] | 272 232 272 272 |
| 7 | [15 15 0 15] | 272 326 232 218 |
| 8 | [15 0 15 15] | 326 272 272 178 |
| 9 | [ 0 15 15 15] | 232 272 326 272 |

# REFERENCES

AlHassan, H. A., Saeb, M. & Damed, H. D. (2005) The Pyramids Block Cipher, International Journal of Network Security (IJNS), Vol.1, No. 1, pp 52 – 60.

Abawajy, J.H. (2011) Computing Technical Trends, Seminar Series, Universiti Teknologi MARA (UiTM), 7 July 2011, Kuala Lumpur.

Baigneres, T. Stern J. & Vaudenay S. (2007) Linear Cryptanalysis of Non Binary Ciphers with an Application to SAFER, LNCS Volume 4876/2007, pp 184-211.

Block Cipher (2010) Retrieved August 4, 2010 from http://www.answers.com/topic/block-cipher

Biham, E. & Shamir, A. (1990). Differential Cryptanalysis of DES-like Cryptosystem. Advances in Cryptology – CRYPTO'90. Springer Verlag. 2-21.

Caelli, W. Longley, D. & Shain, M. (1994). Information Security Handbook, Macmillan Press Ltd. pp 2-4.

Chapple, M. (2010) The GSEC Prep Guide Mastering SANS GIAC Security Essentials Retrieved September 9, 2010 from http://www.coursehero.com/textbooks/47275-The-GSEC-Prep-Guide-Mastering-SANS-GIAC-Security-Essentials/

Daemen, J. & Rijmen, V. (2002). The Design of Rijndael *: AES* – The Advanced Encryption Standard, Springer Verlag.

Gilbert W.J. and Vanstone S.A.(2005) An Introduction to Mathematical Thinking, Algebra and Number Systems, Prentice Hall pp 64-66.

Granboulan, L. Levieil E. & Piret G. (2006) Pseudorandom Permutation Families over Abelian Groups. In Fast Software Encryption 2006, volume 4047 of LNCS, pp 57-77. Springer-Verlag.

Goldwasser S. and Bellare M. (2008) Lecture Notes on Cryptography, Cambridge, Massachusetts, Retrieved September 9, 2011 from
http://cseweb.ucsd.edu/~mihir/papers/gb.pdf

Jamel S., Tan S. L, Md Nasir N. F., Mat Deris M. (2007), Multimedia Application for cryptographic substitution boxes (s-boxes) and diffusion design Principles, Third International Conference on Research and Education in Mathematics (ICREM3), April 10 – 12, 2007 The Legend Hotel, Kuala Lumpur.

Jamel S., Mohamad K. M., Mat Deris M. (2007), Application of Linear Algebra and Modular Arithmetic in the Design of Cryptographic Algorithms, The 3$^{rd}$ IMT-GT Regional Conference on Mathematics, Statistics and Applications (IMT-GT RCMSA 2007) 5 -6/12/2007 The Gurney Hotel, Pulau Pinang.

Jamel S. and Mat Deris M. (2008), Diffusive Primitives in the Design of Modern Cryptographic Algorithms, International Conference on Computer & Communication Engineering (ICCCE'08), International Islamic University Malaysia on 13- 15 May 2008, Hotel Istana, Kuala Lumpur.

Jamel S., Herawan T. and Mat Deris M. (2009), Steps For Constructing Magic Cube Using Two Orthogonal Latin Squares and A Magic Square, Malaysian Technical Conference and Engineering and Technology (MUCEET 2009) on 20 ~ 22 June 2009, MS Garden, Kuantan. Pahang.

Jamel S., Herawan T. and Mat Deris M. (2010), A Cryptographic Algorithm using Hybrid Cubes, International Conference on Computational Science and Its Applications – ICCSA 2010, Sangyo University, Fukuoka, Japan, 23 ~ 26th March 2010 (Published by Springer Verlag).

Jamel S., Mat Deris M., Yanto I. T. R. and Herawan T. (2011a), The Hybrid Cubes Encryption Algorithm (HiSea), International Conference on Computational Science, Engineering and Applications – ICCSEA 2011, Dubai, 25 ~ 27th May 2011, Springer Verlag.

Jamel S., Mat Deris M., Yanto I. T. R. and Herawan T. (2011b), HiSea: A Non Binary Toy Cipher, Journal of Computing, Volume 3, Issue 6, June 2011.

Junod, P. (2004), Statistical Cryptanalysis of Block Ciphers, Ph.D. Thesis No. 3179, Ecole Polytechnique Federale De Lausanne (EPFL).

Leary, T. (1996). Cryptology in the 16th and 17th Centuries, Retrieved January, 17, 2006 from http://home.att.net/~tleary/cryptolo.htm

Lee, A. (1999) Guideline for implementing Cryptography in the Federal Government, NIST Special Publication 800-21, U.S. Department of Commerce, November 1999. Retrieved April 8,2005 from http://csrc.nist.gov/publications/mistpubs/8000-21/800-21.pdf

Li, C. Li, S. Chen, G. & Halang, W. A. (2008) Cryptanalysis of an Image Encryption Scheme Based on a Compound Chaotic Sequence, Image and Vision Computing.

Massey, J.L. (1999) On the Optimality of SAFER+ Diffusion, Cylink Corporation, Sunnyvale, CA, USA,

Mel, H.X. & Baker, D. (2001). *Cryptography Decrypted*. Addison-Wesley.

MS ISO/IEC 18033-3 : 2005 - Information technology - Security techniques - Encryption algorithms - Part 3 : Block ciphers, Sirim Berhad.

Mat Deris M., Abawajy J.and Mamat A. (2008) An efficient replicated data access approach for large-scale distributed systems, Future Generation Computer Systems 24 (2008)-19. Elsevier.

Newton, P. K. & Desalvo, S. A. (2010). *The Shannon Entropy of Sudoku matrices*, Proceedings of the Royal Society A, First Cite e-publishing .

National Institute of Standards (NIST). *FIPS Pub 197: Advanced Encryption Standard (AES*), (2001). http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf (Accessed on 22 March, 2005).

Oppliger, R. (2005). *Cotemporary Cryptography*, Artech House Inc.

Pattern Storm Web site, Cipher mechanisms with fencing and balanced block mixing Retrieved July 28, 2010 from http://www.patentstorm.us/patents/5623549.html

Rivest, R. L. (2001) The RC6 Block Cipher: A simple fast secure. AES proposal.

Sastry V.U.K., Mrrthy D.S.R. and Bhavani S.D. (2010). A Large Block Cipher Involving a Key Applied on Both the Sides fo the Plain Text, International Journal of Computer and Network Security (IJCNS), Vol. 2, No. 2, 2010.

Schneier, B. Kelsey, J. Whiting, D. Wagner, D. Hall, C. & Ferguson, N. (1999). *The Twofish Encryption Algorithm*, John Wiley and Sons, New York.

Schneier B. (2006). Keynote Speaker, Hack In The Box Security Conference (HITBC), Kuala Lumpur, Malaysia.

Shannon C. E (1949). *Communication Theory of Secrecy Systems* Retrieved August 22, 2006 from www.cs.ucla.edu/~jkong/research/security/shannon1949.pdf

Shen, J. Jin, X. & Zhou, C. (2005).: A Color Image Encryption Algorithm Based on Magic Cube Transformation and Modular Arithmetic Operation, PCM 2005, Part II, LNCS 3768 pp. 270-280.

Swenson, C. (2008). *Modern Cryptanalysis, Techniques for Advanced Code Breaking*, Wiley Publication Inc.

Stalling, W. (2003). *Cryptography and Network Security: Principles and Practices*, 3rd Edition, Prentice Hall, New Jersey.

Suzuki, M. (2001), Magic Squares, Retrieved September 9, 2010 from http://mathforum.org/te/exchange/hosted/suzuki/MagicSquare.html

The MARS cipher (2000). IBM submission to AES. Retrieved May 19, 2000 from http://domino.research.ibm.com/comm/research_projects.nsf/pages/security.mars.html

Trappe, W. & Washington L. C. (2002), *Introduction to Cryptography with Coding Theory*, Prentice Hall, New Jersey.

Trenkler, M. (1998). *Magic Cubes*. The Mathematical Gazeete-82, 56-61.

Trenkler, M. (2000). *A Construction of Magic Cubes.* The Mathematical Gazeete, 36-41.

Trenkler, M. (2005). An Algorithm for making Magic Cubes, The $\prod ME$ Journal, Vol. 12, No. 2, pp. 105-106.

Tuan Sabri Tuan Mat (2000), Design of New Block and Stream Cipher Encryption Algorithms for Data Security., Ph.D. Dissertation., Universiti Teknologi Malaysia (UTM).

Welsh, D. (1998). *Codes and Cryptography*, Oxford University Press,

Wu, S., Zhang, Y. & Jing, X. (2005). A Novel Encryption Algorithm based on Shifting and Exchanging Rule of Bi-Column Bi-row Circular Queue, International Conference on Computer Science and Software Engineering, IEEE 2005.

Zhang, L. Shiming, J. Xie, Y. Yuan, Q., Wan, Y. & Bao, G. (2005). Principle of Image Encrypting Algorithm Based on Magic Cube Transformation, LNAI 3802.