ALGORITHM ANIMATION FOR CRYPTANALYSIS OF CAESAR AND HILL CIPHERS

Sapiee Haji Jamel and Giuseppina Sherry Sayan
(Faculty of Information Technology and Multimedia,
Universiti Tun Hussein Onn Malaysia (UTHM) 86400 Parit Raja,
Batu Pahat JOHOR sapiee@uthm.edu.my, gusche85@hotmail.com)

ABSTRACT

Cryptographic algorithms are usually kept secret and the complexity of each algorithm is based on mathematical or statistical analysis. However, in modern cryptography, the strength of each cryptographic algorithm should only rely on the secret key(s) used to ensure that there is no trap door embedded in the algorithm. This new trend gives an advantage to cryptanalyst since types of algorithm(s) used are no longer a secret. Cryptanalysis steps can be easily explained using algorithm animation that can be easily integrated with any e-learning platform. This paper presents an algorithm animation for cryptanalysis of Caesar Cipher using ciphertext only attack and Hill Cipher using Brute force technique. Each of the steps involved in the cryptanalysis is shown using animation technique developed using Macromedia Flash. This application is suitable for e-learning activities for students interested in the cryptanalysis of block cipher.

1. INTRODUCTION

Cryptanalysis is a branch of Cryptology that deals with finding weaknesses in cryptographic algorithm. Cryptanalysts reveal secret data (plaintext or message) from scramble messages (ciphertext) without knowledge of keys or types of algorithms used. In the past, every elements of the cryptographic algorithms are kept secret and the complexities of each cryptographic algorithm are usually based on mathematical formulation to prove that it is secured from any simple mathematical manipulation or statistical analysis. An algorithm animation combines multimedia tools with other disciplines to simulate how specific algorithm or process work. It offers an alternative method for better understanding of complicated process with a help of computer software (Animation, 2007). Cryptology (study of cryptography and cryptanalysis) has evolved from an art in which it was mastered by several people into a science where it has significance importance especially in the digital communication (Henk et al., 2005). Previously, cryptographers and cryptanalysts (or hackers) worked in isolation in order to protect their own safety. Since the year 2000, both parties have openly developed and discussed any strengths or weaknesses of each algorithm as shown in the process of selecting the United States Advanced Encryption Standard (AES) new encryption standard for encrypting sensitive (unclassified) Federal information (NIST, 2001). Developing a new block cipher and tests it using several known cryptanalysis techniques has become a new trend (Junod, 2005; Piret, 2005; Dunkelman, 2006). Twofish developers (Schneier et al., 1999) cryptanalyzed their own algorithm more than one thousand man-hours with ten known cryptanalysis techniques similar to those described by Standaert (2003) as a proof of its strength. Even though several cryptanalysis of Caesar are available on the Internet, this project attempts to intergrate elements of multimedia (text, sound and animation) in the description of the cryptanalysis process for better understanding (Savarese et al., 2006).This paper presents cryptanalytic process of ciphertext from Caesar and Hill Cipher using algorithm animation to allow better understanding of how the actual process is performed. Software visualization approaches are used in order to understand a complex and time consuming cryptanalysis process (Price et al., 1992). An algorithm animation creates an abstraction of both the ciphertext and the operations of the cryptanalysis process. It also maps the data into an image, which then gets animated based on the operations of the cryptanalysis process. The rest of the paper is organized as follows. Section 2 describes the concept of cryptanalysis. Section 3 discusses the types of cryptanalysis techniques which can be used on any cryptographic algorithms. Section 4 outlines in detailed Caesar and Hill Cipher cryptanalysis animation algorithm. Section 5 concludes this paper with future work of this research.

2. CONCEPT OF CRYPTANALYSIS

The main objective of cryptanalyst is to find the cryptographic key used to encrypt and decrypt message. An indication of weaknesses can be in the form of the recovery of cryptographic keys used, recovery of plaintext from ciphertext or even the key generation procedure for an undisclosed cipher. Before cryptanalyst can start performing his or her work, the type of data available must be identified whether it is a ciphertext or keys to be used for encryption and decryption. Even though several cryptographic algorithms are published and evaluated openly in journals and conference papers (NIST, 2001), they represent only the tip of the ice berg since most algorithms used by the Governments or military institutions classified properties world wide. One reason for the possibilities of cryptanalysis to be performed on any block cipher is that plaintext (messages), ciphertext and key combination must be finited to ensure unique encryption and decryption of messages. Modern symmetric (similar key for encryption and decryption) cryptographic algorithms use key schedule algorithm using master key to generate different sub-keys for each iterated rounds.The following section discusses basic elements of cryptographic algorithm required before any cryptanalysis can be performed.

2.1 Message

Messages are written using 27 standard alphabets (a..z, A..Z) and then translated into suitable 128 printable ASCII characters. Later, they get converted into 8- bit binary numbers before entering encryption process. Messages are usually grouped into a block of either 128-bit, 192-bit or 256-bit depending on the key length used in the algorithm.

2.2 Encryption and decryption elements

Software implementation of modern cryptographic algorithm utilizes basic Boolean algebra operators (OR, Exclusive-OR (XOR), AND and Modulo arithmetic) for performing their basic routines. Confusion process (DES, Rijndael, Twofish) which translates plaintext into intermediate values (hexadecimal) is implemented using non-linear substitution tables (S-Boxes). Diffusion is implemented using Maximum Distance Separable (MDS) matrices or Pseudo-Hadamard Transform (PHT) as in Twofish.

2.3 Ciphertext

Ciphertext usually consists of a string of binary sequence, which is a mixture of scrambled messages, key sequences and extra information (whitening) added by the encryption process. A common ciphertext block length for modern cryptographic algorithms are 128- bit, 192-bit and 256-bit.

2.4 Key Schedule algorithm

Modern cryptographic algorithms integrate complex procedure to avoid related-key cryptanalysis. This type of cryptanalysis is designed specifically to find any weaknesses in key generation method. Regardless of any improvements implemented by cryptographers, cryptanalyst still can perform on going analysis on cryptographic algorithm offline using all of the above information.

3. CRYPTANALYSIS TECHNIQUES ADOPTED IN THIS PROJECT

The method for performing cryptanalysis can be classified into several areas based on the model use:

3.1 Brute Force Method

This method is still the most efficient attack on any cryptographic algorithms and researchers are actively trying to find the best solutions to this problem (Schneier, 1999; Clayton, 2001). In this cryptanalysis technique, cryptanalyst will try to recover cryptographic key used in an algorithm to encipher or decipher message.

Any successful trial less than the maximum number of possibilities are considered a successful attack on any cryptographic algorithm. The strength of modern cryptographic algorithms is dependent on the property that the resources for performing analysis on the 128-bit key are still limited. Rijndael, Twofish, RC6, MARS and Serpent (AES, 2000) can accept multiple key sizes such as 128, 192 and 256-bit key which might take more than 10x1012 years to get the appropriate key. Maximum number of possible attempts for key recovery allowed for these algorithms is 2128 times. Machine design specifically to perform Brute Force on key space of DES also available in the literature (Wiener, 1993). With Rijndael, this Brute Force technique is time consuming and required vast computing power to get the appropriate result. Most cryptographic algorithms are designed to ensure it is impossible for cryptanalyst to do reverse engineering in order to predict the correct key generated by the algorithm. Keys for encrypting messages are generated from master key. The key generation procedure is usually embedded with confusion and diffusion technique to ensure keys are unique and randomly generated.

3.2 Ciphertext only attack

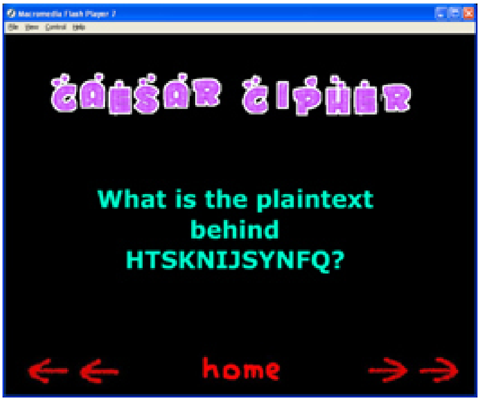
In this technique, cryptanalyst have the ciphertext only for finding the appropriate plaintext or the cryptographic key used. Statistical analysis of the ciphertext can be helpful.

4. CRYPTANALYSIS OF CAESAR AND HILL CIPHER

4.1 Cryptanalysis of Caesar Cipher using ciphertext only attack

For Caesar cipher, cipher text is generated by shifting the alphabet according to an agreed value (key) and Mathematic modulo operator is used to ensure the total number of possible ciphertext is 26. For example, if the key is equal to 3, every alphabet in the plaintext will be shifted 3 positions to generate the appropriate ciphertext. Snap shots of cryptanalysis process for finding plaintext from Caesar ciphertext are shown from figure 1 until figure 4.

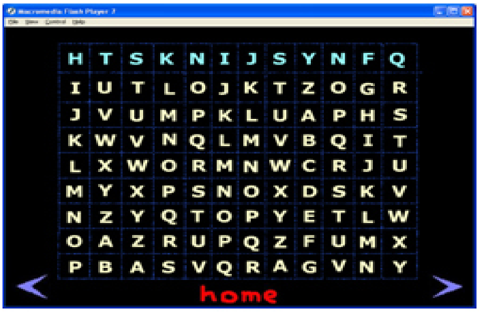
Figure 1 shows ciphertext “HTSKNIJSYNFQ” that is used in the cryptanalysis of Caesar Cipher using ciphertext only attack. Each possible letters are then rearranged according to the next alphabet as shown in Figure 2. “H” is followed by “I”, “J”, “K” and other consecutive alphabets. This will represent all the possible combinations of plaintext for this ciphertext using Brute Force method. Toward the end of this process, a familiar English word will appear as shown in Figure 3 where the word CONFIDENTIAL appear at line 5 of the screen. Result from the cryptanalysis process is shown in Figure 4. Once the meaning of the ciphertext is known to the cryptanalyst, he or she can find the key used for decrypting the message. As for this example, the key used is 3.



CAESAR CIPHER

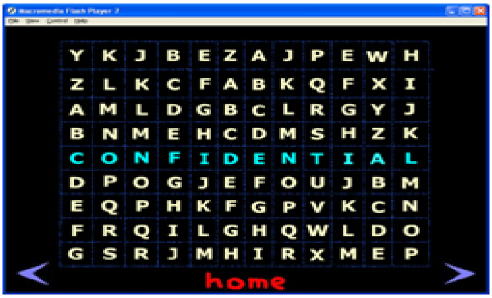
What is the plaintext behind HTSKNIJSYNFQ?

Figure 1: Finding plaintext from ciphertext



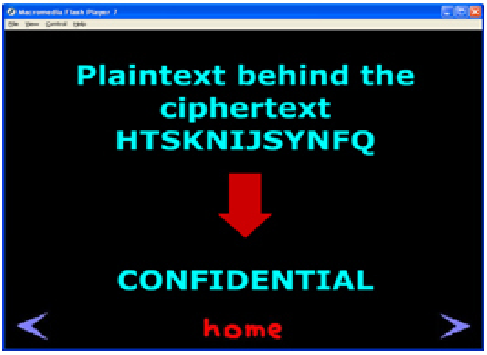
H	T	S	K	N	I	J	S	Y	N	F	Q
I	U	T	L	O	J	K	T	Z	O	G	R
J	V	U	M	P	K	L	U	A	P	H	S
K	W	V	N	Q	L	M	V	B	Q	I	T
L	X	W	O	R	M	N	W	C	R	J	U
M	Y	X	P	S	N	O	X	D	S	K	V
N	Z	Y	Q	T	O	P	Y	E	T	L	W
O	A	Z	R	U	P	Q	Z	F	U	M	X
P	B	A	S	V	Q	R	A	G	V	N	Y

Figure 2 :All possible combination of plaintext



Y	K	J	B	E	Z	A	J	P	E	W	H
Z	L	K	C	F	A	B	K	Q	F	X	I
A	M	L	D	G	B	C	L	R	G	Y	J
B	N	M	E	H	C	D	M	S	H	Z	K
C	O	N	F	I	D	E	N	T	I	A	L
D	P	O	G	J	E	F	O	U	J	B	M
E	Q	P	H	K	F	G	P	V	K	C	N
F	R	Q	I	L	G	H	Q	W	L	D	O
G	S	R	J	M	H	I	R	X	M	E	P

Figure 3: Meaningful word from the ciphertext



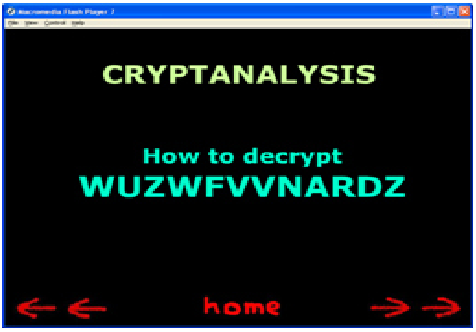
Plaintext behind the ciphertext HTSKNIJSYNFQ

CONFIDENTIAL

Figure 4: Result from cryptanalysis process

4.2 Cryptanalysis of Hill Cipher using Brute Force attack

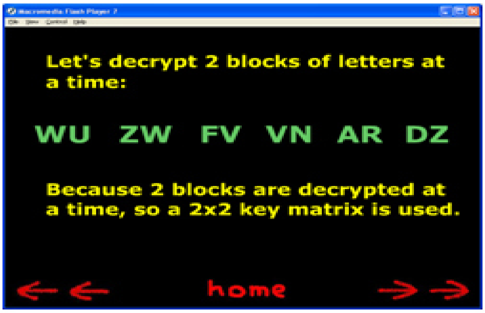
Hill Cipher use matrix as the foundation for converting message and the key to create the ciphertext. The process of cryptanalysis for Hill Cipher ciphertext is shown in Figure 5. The ciphertext WUZWFVVNARDZ is then separated into a block of size 2 as shown in Figure 6. With the assumption that the block size is two for this ciphertext, the key must be of two by two matrix. Suitable key for encryption and decryption process must meet two conditions where it must be invertible modulo 26 and a square matrix. Invertible matrix can be easily tested by finding the Greatest Common Divisor (GCD) between the value of matrix determinant and 26. If the GCD value is 1, the matrix is a possible candidate for the encryption key as shown in Figure 8. The next process is to try this matrix with the ciphertext until a meaningful English word appears from the ciphertext. Figure 8 shows a successful outcome from the cryptanalysis process which revealing the message COFIDENTIAL from the ciphertext WUZWFVVNRDZ.



CRYPTANALYSIS

How to decrypt WUZWFVVNARDZ

Figure 5: Cryptanalysis of Hill Cipher

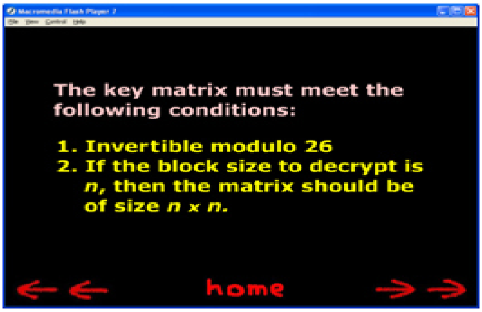


Let's decrypt 2 blocks of letters at a time:

WU ZW FV VN AR DZ

Because 2 blocks are decrypted at a time, so a 2x2 key matrix is used.

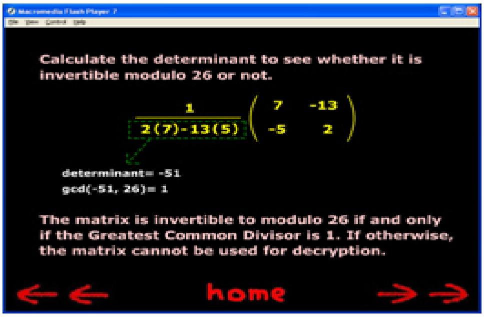
Figure 6: Ciphertext divided into block size of 2



The key matrix must meet the following conditions:

1. Invertible modulo 26
2. If the block size to decrypt is *n*, then the matrix should be of size *n x n*.

Figure 7: Finding Suitable Key



Calculate the determinant to see whether it is invertible modulo 26 or not.

$$\frac{1}{\boxed{2(7)-13(5)}}\begin{pmatrix} 7 & -13 \\ -5 & 2 \end{pmatrix}$$

determinant= -51

gcd(-51, 26)= 1

The matrix is invertible to modulo 26 if and only if the Greatest Common Divisor is 1. If otherwise, the matrix cannot be used for decryption.

Figure 8: Finding Suitable key for Hill Cipher

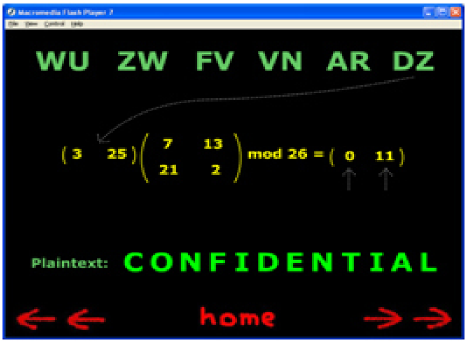

WU ZW FV VN AR DZ
$(3\ \ 25)\begin{pmatrix}7 & 13\\ 21 & 2\end{pmatrix} \bmod 26 = (0\ \ 11)$
Plaintext: CONFIDENTIAL
home

Figure 9: Meaningful message from the ciphertext

5. CONCLUSION

Algorithm animation has been accepted in research communities as a key element to understand complicated algorithms better. It has been widely used to explain the steps involved in algorithm via multimedia application. Even though this project only looks at the cryptanalysis of two simple block cipher, the process involved in designing the algorithm animation can be extended to modern cryptographic algorithms such as Rijndael and Twofish. Cryptanalysis animation developed using Macromedia Flash can be easily integrated to any e-learning platform for making learning basic cryptanalysis process of cryptographic algorithm fun.

REFERENCES

Animation 2007. Introduction to Algorithm Animation,URL:http://www.ickn.org/elements/hyper/cyb105.htm(Accessedon2February2007)

Biham E. and Shamir A., 1991. Differential Cryptanalysis of DES-like Cryptosystems. Journal of Cryptography, 4(1):3-72.
Clayton R., 2001. Brute Force Attacks on Cryptographic Keys, URL :http://www.cl.cam.ac.uk/~rnc1/brute.html Accessed on 4 June 2007.

Feistel H., 1973. Cryptography and Computer Privacy, Scientific American. 228(5): pp 15 -23.

Henk C. A. and Tilborg V., 2005. Encyclopedia of Cryptography and Security, Springer, United States of America.
Junod P., 2005. Statistical Cryptanalysis of Block Cipher, Ph. D. Thesis, Ecole Polytechnique Federale De Lausanne.

Matsui M., 1993. Linear Cryptanalysis Method For DES Cipher, Advances in Cryptology, EUROCRYPT 93: Workshop on the Theory and Application of Cryptographic Techniques, Loftthus, Norway, May 1993. Proceeding, volume 765 of LNCS, pages 386-397. Springer-Verlag.

National Institute of Standards (NIST) FIPS Pub 197: Advanced Encryption Standard (AES), 2001. http://csrc.nist.gov/publications/fips/fips197/fips-197.pdfAccessedon22March,2005

Price B. A., Small I. S. and Baecker M., 1992. A Taxonomy of Software Visualization, Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences.

Savarese C. and Hart B., 1999. The Caesar Cipher, URL : http://starbase.trincoll.edu/~crypto/historical/caesar.html , Accessed on 4 June 2007.

Standaert F.-X, Piret G. and Quisquater J.-J., 2003 Cryptanalysis of Block Ciphers: A Survey. UCL Crypto Group Technical Report Series. Technical Report CG-2003/2 http://www.dice.ucl.be/crypto/tech_reports/CG_2003_2.pdf(Accessedon17February2006).