Journal of Computing and Applied Informatics (JoCAI) Vol. 4, No. 2, 2020 | 100 – 110

talenta PUBLISHER

DATA SCIENCE
Journal of Computing and Applied Informatics

Comparison Encryption of How to Work Caesar Cipher, Hill Cipher, Blowfish and Twofish

Wasis Haryono

Pamulang University, Information Technology, Indonesia

Abstract. Security is the level of confidentiality of data stored using cryptography. There are many ways you can do to improve security. In this study, the writer will use a method by encrypting the database with the Caesar Cipher Algorithm, Hill Cipher and Blowfish. Caesar Cipher, Hill Cipher and Blowfish are part of the symmetric algorithm, which means that the encryption and decryption process have the same key. The encryption and decryption process in Caesar Cipher, Hill Cipher and Blowfish Algorithms each has one key. algorithm encryption techniques using symmetric passwords have 2 types, namely block ciphers and stream ciphers. Caesar Cipher, Hill Cipher and Blowfish and Twofish Algorithms are the encryption of the block cipher that breaks or creates blocks to encrypt and obtain cipher text. Result in this paper In Caesar Cipher, it is carried out like 3 blocks of encryption. Whereas in Hill Cipher a word is divided into several blocks and each block is encrypted. In Blowfish, several iterations are performed to get the text cipher, the input is 64 bits of data that can be done as many as 16 iterations. In Twofish the input is 128 bits, in contrast to Blowfish which is only 64 bits, Twofish can also accept 256 bits long and do 16 iterations to get the cipher text. Twofish has stronger security than the 3 algorithms above, Twofish also takes up more memory and takes longer to encrypt.

Keyword: Caesar Cipher, Hill Cipher, Blowfish, Twofish

Received 3 June 2020 | Revised 26 June 2020 | Accepted 31 July 2020

1 Introduction

Along with the times that caused human needs to increase. Including the need for information. One example is the internet. In its development, the internet is no longer only monopolized by several industry elements, but is also used by most small and medium-sized industries to help their businesses. In the midst of increasingly sophisticated information technology development, the internet no longer guarantees the provision of private information. Various search engines and e-commerce services are also developing. Plus the virus and spam attacks are always growing and lurking.

*Corresponding author at: Pamulang University, Information Technology, Indonesia

E-mail address: @gmail.com

ISSN: 2580-6769 | e-ISSN: 2580-829X | DOI: https://doi.org/10.32734/jocai.v4.i2-4004
Journal Homepage: https://jocai.usu.ac.id

The more developed information technology, the more it causes data security to require quite good security. Now anyone can easily exchange information in any case, including sharing knowledge to illegally access data. Data security is a very important aspect of an information system[1]. A database is a place for storing data and information that must be kept safe and confidential. To maintain the security and confidentiality of data stored in the database using cryptography. The data warehouse in the database table has a login system installed with a password. But bad people find other ways to access the data by accessing directly to the database table without going through the application system. With the possibility of illegal data access that accesses directly to the database table, better security is needed for the database.

There are many ways you can do to improve security. In this study, the writer will use a method by encrypting the database with the Caesar Cipher Algorithm, Hill Cipher and Blowfish. Caesar Cipher, Hill Cipher and Blowfish are part of the symmetric algorithm, which means that the encryption and decryption process have the same key. The encryption and decryption process in Caesar Cipher, Hill Cipher and Blowfish Algorithms each has one key.Literature Review

2 Literature review

2.1 Cryptography

The word cryptography comes from Greek, krypto (hidden) and grafh (written) which means hidden writing. Cryptography is the study of mathematical techniques related to aspects of information security such as data confidentiality, data validity, data integrity, and data authentication. In general, cryptography consists of two main parts, namely the encryption section and the decryption section.

Terms in cryptography

1. Message (Plaintext and Ciphertext): Message (message) is data or information that can be read and the meaning is understood. The original message is called plaintext or cleartext. While encoded message is called ciphertext (ciphertext)
2. Sender and Receiver: Data communication involves the exchange of messages between two entities. Sender (sender) is an entity that sends messages to other entities. The receiver is an entity who received a message.
3. Eavesdroppers are people who try to capture messages during transmission.
4. Cryptanalysis and Cryptology: Cryptanalysis (cryptanalysis) is the science and art of solving ciphertext becomes plaintext without knowing the key used. The culprit is called cryptanalyst. Cryptology (cryptology) is the study of cryptography and cryptanalysis.
5. Encryption and Decryption: The process of encoding plaintext into ciphertext is called encryption (encryption) or enciphering. While the process of returning the ciphertext into the original text called decryption (decryption) or deciphering.

6. Cipher and Key: Cryptographic algorithms are also called ciphers, which are rules for enchying and dechipering, or mathematical functions used for encryption and decryption. The key is the parameters used for enciphering and dechipering transformations. The key is usually in the form string or row of numbers. [2]

2.2 Encryption

Encryption is the process of transforming information or plaintext into other forms so that the actual message content cannot be understood or often called ciphertext, this is so that information remains protected from those who are not entitled to receive it. While decryption is the reverse process of encryption, i.e. transformation of data to the original form data.

A. Symmetric Key Encryption

Symmetric algorithm or often called the conventional cryptographic algorithm is an algorithm that uses the same key for the encryption process and the description process. Symmetric cryptographic algorithms are divided into two categories, namely Stream Algorithm (Stream Ciphers) and Block Algorithm (Block Ciphers). Where in the flow algorithm, the coding process will be oriented to one bit / byte of data. Whereas in the block algorithm, the coding process is oriented to a set of bits / bytes of data (per block). Examples of symmetrical key algorithms are DES (Data Encryption Standard), Blowfish, Twofish, MARS, IDEA, 3DES (DES applied 3 times), AES (Advanced Encryption Standard) whose real name is Rijndael.

B. Caesar Cipher Algorithm

In cryptography, Caesar's password, or sliding password, Caesar's code or Caesar Clash is one of the most well-known encryption techniques. This password includes a substitution password where every letter in the light text (plaintext) is replaced by another letter that has a certain position difference in the alphabet. This is the cryptographic algorithm which was first used by the Roman emperor, Julius Caesar (so called caesar cipher), to encode the message he sent to his governors. The trick is to replace (substitute or substitute) each character with other characters in alphabetical order. For example, each letter is substituted with the next third letter of the alphabetical order. In this case the key is the number of letter shifts (ie k = 3) [3]. By coding each letter of the alphabet with an integer as follows: A = 0, B = 1, ..., Z = 25, then mathematically caesarean cipher encodes plaintext pi into ci with the rules:

$$p_i = D(c_i) = (c_i - k) \bmod 26$$

and decrypt ciphertext ci into pi with the rules:

$$c_i = E(p_i) = (p_i + 3) \bmod 26$$

Since there are only 26 letters of the alphabet, the possible letter shift is from 0 to 25. In general, for letters shifting as far as k (in this case k is the encryption and decryption key), the encryption function is

$$c_i = E(p_i) = (p_i + k) \bmod 26$$

and the decryption function is

$$p_i = D(c_i) = (c_i - 3) \bmod 26$$

Each of the same letters is replaced by the same letters throughout the message, so the Caesar code is classified as monoalfabetic substitution.

C. Hill Cipher Algorithm

Hill Cipher is the application of modulo arithmetic in cryptography. This cryptographic technique uses a square matrix as a key that is used for encryption and decryption. The Hill Cipher was created by Lester S. Hill in 1929. This cryptographic technique was created with the intention of creating a cipher (code) that could not be solved using frequency analysis techniques. The Hill Cipher does not replace each of the same alphabet in the plaintext with the other alphabet in the same ciphertext because it uses matrix multiplication on the basis of encryption and decryption. Hill Cipher which is a polyalphabetic cipher can be categorized as a block cipher because the text to be processed will be divided into blocks of a certain size. Each character in one block will influence each other in the encryption and decryption process, so that the same characters are not mapped into the same characters. Hill Cipher belongs to the classic cryptographic algorithm which is very difficult for cryptanalysts to solve if it is done only by knowing the ciphertext file. However, this technique can be solved quite easily if cryptanalysts have ciphertext files and plaintext file fragments. This cryptanalysis technique is called known-plaintext attack.

D. Blowfish Algorithm

The symmetrical blowfish cryptographic algorithm is a modern symmetric key algorithm

block cipher [4].Blowfish was created by a Cryptanalyst named Bruce Schneier, President of Counterpane Internet Security, Inc. (a consulting firm on cryptography and computer security) and was published in 1994. Created for use on computers that have large microposesors (32-bit and above with large data caches) . Blowfish is a non-patented and licensefree algorithm, and is available free of charge for a variety of uses. When Blowfish is designed, it is expected to have the following design criteria:

1. Fast, Blowfish encrypts data on 32-bit microprocessors with a rate of 26 clock cycles per byte.
2. Compact (lightweight), Blowfish can run on memory less than 5K.
3. Simple, Blowfish only uses simple operations: additions, XORs, and lookup tables on 32-bit operands.
4. Having varying levels of security, the key length used by Blowfish can vary and can be as long as 448 bits. In its application, this algorithm often becomes not optimal. Because the

implementation strategy is not right. Blowfish algorithm will be more optimal if used for applications that do not frequently change keys, such as communication networks or automatic file encryption. In addition, because this algorithm requires a large amount of memory, this algorithm cannot be applied to applications that have small memory such as smart cards. The key length used also affects the security of the application of this algorithm.

E. Twofish Algorithm

Twofish is a cryptographic algorithm that operates in block cipher mode. Twofish was one of the finalists in the Advanced Encryption Standards (AES) competition held by the National Institute of Standards and Technology (NIST). Twofish is a 128-bit block cipher that can accept keys up to 256 bits long [5].

1. Feistel Network is a general method for transforming a function into a form of permutation. The most fundamental part of the Feistel Network is 16 F functions, namely a keydependent mapping from an input string to an output string. In Twofish, Feistel Network was conducted 16 times. In twofish, the feistel network consists of Input Whitening, S-boxes, Pseudo Hadamard Transformation, and Output Whitening.
2. S-Boxes are matrices that contain simple substitutions that map one or more bits with one or more other bits. In most Chiper block algorithms, S-boxes map m input bits and n output bits (m x n). Twofish uses four bijective, key-dependent and 8-by-8-bit S-boxes. This Sboxes is made using two 8-by-8-bit permutations and key material.
3. Code Maximum Distance Separable (MDS) The four results of the s-boxes are interpreted as a vector whose length is 4 and multiplied by the 4x4 MDS matrix. To produce a vector interpreted as 32-bit words as a result of g:

$$x_i = \lfloor X / 2^{8i} \rfloor \bmod 2^8 \qquad i = 0,\ldots,3$$
$$y_i = s_i[x_i] \qquad i = 0,\ldots,3$$

$$\begin{pmatrix} z_0 \\ z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} . \\ .MDS \\ . \end{pmatrix} . \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$Z = \sum_{i=0}^{3} z_i . 2^8$$

4. Pseudo-Hadamard (IPM) transformation, which is by dividing blocks into two equal blocks and applying simple arithmetic operations.

   Pseudo-Hadamard encryption: a ’= a + b (mod 2 ^ n) b’ = a + 2b (mod 2 ^ n)

   Decryption of pseudo-Hadamard: b = b ’- a’ (mod 2 ^ n) a = 2a ’- b’ (mod 2 ^ n)

5. Whitening is a technique to XOR key material before the first round and after the last round. In the attack on Twofish, it was proven that whitening increased the difficulty of attacking the Chipper, by hiding specific input for the beginning and end of the round from twofish.

3 Methodology

The method in this study consisted of

3.1 Literature review

Data collection using or collecting written sources, by reading, studying and recording important matters relating to the problem being discussed in order to obtain a theoretical picture.

3.2 Field study

Field study is a method of collecting data by direct observation of research objects to obtain data by observation, namely collecting data by making observations directly to the object of research online, by noting important things related to research, in order to obtain complete and accurate data.

4 Result

4.1 Comparison of Caesar Cipher, Hill Cipher and Blowfish Algorithms

Of the three Algorithms above, the Caesar Cipher Hill Cipher Algorithm, and Blowfish use the Block Cipher, which encrypts the plaintext blocks and produces text cipher blocks. Here are the encryption and decryption comparisons of the three Algorithms:

A. Caesar Cipher

1. Determine the magnitude of the character shift used to form the ciphertext to the plaintext.
2. Exchange characters in plaintext into ciphertext based on predetermined shifts.

Following is an example of using Caesar Cipher with a large shift of 3 characters. With the value of the shift, the Caesar Cipher value shift table is obtained as follows:

Substitution Table:

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

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

Examples of the process of using Caesar Cipher:

Message: INI ADALAH KATA SANDI YANG BENAR

Encryption result: LQL DGDODK NDWD VDQGL BDQJ EHQDU

B. At Hill Cipher

The encryption process in the Hill Cipher is done per block plaintext. The size of the block is the same as the size of the key matrix. Before dividing the text into rows of blocks, the plaintext is first converted to a number, each so that A = 1, B = 2, up to Y = 25. Z is given a value of 0.

A	B	C	D	E	F	G	H	I	J	K	L	M
1	2	3	4	5	6	7	8	9	10	11	12	13
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
14	15	16	17	18	19	20	21	22	23	24	25	0

Figure 1 Hill Cipher

Mathematically, the encryption process on Hill Cipher is:

C = K. P.

C = Ciphertext

K = Key

P = Plaintext

If there is a plaintext P:

P = STRIKE NOW

Then the plaintext is converted to:

P = 19 20 18 9 11 5 14 15 23

The plaintext will be encrypted with the Hill technique

Cipher, with key K which is a 2 × 2 matrix

$$K = \begin{bmatrix} 5 & 6 \\ 2 & 3 \end{bmatrix}$$

Because the key matrix K is 2, the plaintext is divided into blocks, each block having 2 characters. Because the last character does not have a partner, then given the same character pair, W. P becomes STRIKENOWW. The first block of plaintext P is:

$$P_{1,2} = \begin{bmatrix} 19 \\ 20 \end{bmatrix}$$

The plaintext block is then encrypted with the K key through the equation.

$$C_{1,2} = \begin{bmatrix} 5 & 6 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 19 \\ 20 \end{bmatrix} = \begin{bmatrix} 215 \\ 98 \end{bmatrix}$$

The results of the calculations produce numbers that do not correspond to letters,

$$C_{1,2} = \begin{bmatrix} 215 \\ 98 \end{bmatrix} = \begin{bmatrix} 7 \\ 20 \end{bmatrix} (\text{mod } 26)$$

then do modulo 26 on that result. So, C1.2 becomes: The characters corresponding to 7 and 20 are G and T. then S becomes G and T becomes T. After encrypting all the blocks in the P plaintext, the ciphertext C is generated as follows:

P = STRIKENOW

C = 7 20 14 11 7 11 4 21 19 11

C = GTNKGKDUSK

C. On blowfish

Consists of simple function iterations (Feistel Network) of 16 rotations (iteration), the input is 64-bit X data elements. Each cycle consists of key-dependent permutations and key- and datadependent substitution. All operations are addition and XOR to 32-bit variables. Other additional operations are just four indexed array table searches for each round. The steps are as follows.

1. Divide X into two parts, each consisting of 32 bits: XL, XR.
2. Perform the following steps For i = 1 to 16: XL = XL ⊕ Pi XR = F (XL) ⊕ XR Exchange XL and XR
3. After the 16th iteration, exchange XL and XR again to cancel the last exchange.
4. Then do XR = XR ⊕ P17 XL = XL ⊕ P18 5. Finally, recombine XL and XR to get the ciphertext.

D. On Twofish

Plaintext is divided into four parts of words with a size of 32 bits. Plaintext with the size of 16 bytes namely p0, ..., p15 is divided into four parts into P0, ..., P3 with a size of 32 bits each by using a little endian conversion.

$$P_i = \sum_{j=0}^{3} p_{(4i+j)} \cdot 2^{8j} \qquad i = 0, \ldots, 3$$

In the input whitening step, all four parts of the plaintext are performed XOR operations with four keywords for the expanded key.

$$R_{0,i} = P_i \oplus K_i \qquad i = 0, \ldots, 3$$

This step is then followed by 16 rounds. At each second round the first word is used as input for function F, which also accepts round number input. The third word is performed XOR operation with the first output function F and then rotated to the right by 1 bit. The fourth word is rotated to the right by 1 bit and then an XOR operation is performed with the second output of the function F. Then swap the two parts.

$$\begin{aligned}
(F_{r,0}, F_{r,1}) &= F(R_{r,0}, R_{r,1}, r) \\
R_{r+1,0} &= \mathrm{ROR}(R_{r,2} \oplus F_{r,0}, 1) \\
R_{r+1,1} &= \mathrm{ROL}(R_{r,3}, 1) \oplus F_{r,1} \\
R_{r+1,2} &= R_{r,0} \\
R_{r+1,3} &= R_{r,1}
\end{aligned}$$

Note: r = 0, ..., 15

ROR and ROL = functions that rotate the first argument to the right or left with the number of bits in accordance with both arguments

Duplicate whitening of the output is done by canceling the exchange process in the last round and performing an XOR operation of data words with four words of the expanded key.

$$C_i = R_{16,(i+2) \bmod 4} \oplus K_{i+4} \qquad i = 0, \ldots, 3$$

These four ciphertext words are then written as 16 bytes c0, ..., c15 using the little endian conversion, the same as those used in plaintext.

$$c_i = \left\lfloor \frac{C_{\lfloor i/4 \rfloor}}{2^{8(i \bmod 4)}} \right\rfloor \bmod 2^8 \qquad i = 0, \ldots, 15$$

The function F is a permutation that depends on keys and operates on 64-bit values. The F function asks for the input of three arguments, the two words input R0 and R1, and the round number r used to select the corresponding up-key. R0 is then passed to the function g which

produces T0. R1 is then rotated to the left 8 bits and then passed to the g function to produce T1. T0 and T1 results are then combined in the Pseudo Hadamard Transform (IPM) and two keys of the expanded key are added.

$$
\begin{aligned}
T_0 &= g(R_0) \\
T_1 &= g(\mathrm{ROL}(R_1, 8)) \\
F_0 &= (T_0 + T_1 + K_{2r+8}) \bmod 2^{32} \\
F_1 &= (T_0 + 2T_1 + K_{2r+9}) \bmod 2^{32}
\end{aligned}
$$

Keterangan:
F_0, F_1 = hasil keluaran dari fungi F

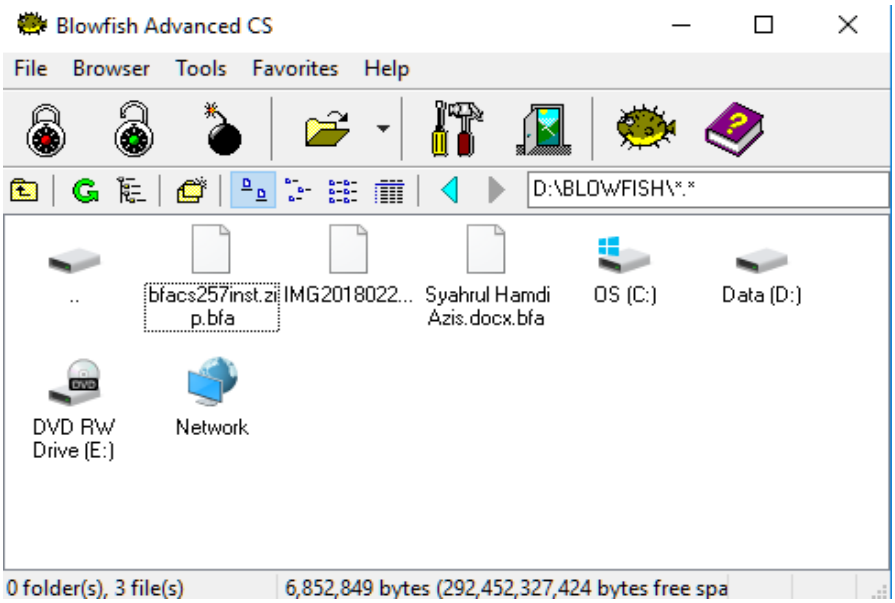

Figure 2 Encription Blowfish using Advanced CS

Figure 2 show The encrypted file will change to .bfa format, where the file cannot be opened or in other words is locked.

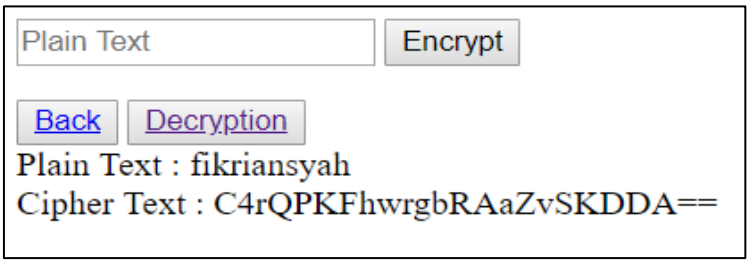

Figure 3 Encription Blowfish using PHP

Figure 3 is blowfish encryption using PHP by showing Plain Text Encryption from the name "fikriansyah"

5 Conclusions

algorithm encryption techniques using symmetric passwords have 2 types, namely block ciphers and stream ciphers. Caesar Cipher, Hill Cipher and Blowfish and Twofish Algorithms are the encryption of the block cipher that breaks or creates blocks to encrypt and obtain cipher text.

In Caesar Cipher, it is carried out like 3 blocks of encryption. Whereas in Hill Cipher a word is divided into several blocks and each block is encrypted. In Blowfish, several iterations are performed to get the text cipher, the input is 64 bits of data that can be done as many as 16 iterations. In Twofish the input is 128 bits, in contrast to Blowfish which is only 64 bits, Twofish can also accept 256 bits long and do 16 iterations to get the cipher text. Twofish has stronger security than the 3 algorithms above, Twofish also takes up more memory and takes longer to encrypt.

REFERENCES

[1] J. Jumrin, S. Sutardi, and S. Subardin, "Aplikasi Sistem Keamanan Basis Data dengan Teknik Kriptografi RC4 Stream Cipher," *semanTIK*, vol. 2, no. 1.

[2] S. Sitinjak, Y. Fauziah, and J. Juwairiah, "Aplikasi Kriptografi File Menggunakan Algoritma Blowfish," in *Seminar Nasional Informatika (SEMNASIF)*, 2015, vol. 1, no. 3.

[3] S. Althaf, "Implementasi Keamanan Data Menggunakan Algoritma Blowfish Pada Sistem Informasi Koperasi Rias," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 8, no. 1, pp. 251–264, 2017.

[4] E. D. Santosa, "Implementasi Algoritma Caesar Cipher dan Hill Cipher Pada Database Sistem Inventori TB Mita Jepara," *Dok. Karya Ilm. Progr. Stud. Tek. Inform. Univ. Dian Nuswantoro. Semarang*, 2015.

[5] A. Randy, "Studi dan Perbandingan Algoritma Blowfish dan Twofish." 2008.