JOURNAL OF NATURAL SCIENCES AND MATHEMATICS RESEARCH

Available online at http://journal.walisongo.ac.id/index.php/jnsmr

Algorithm *Hill Cipher* for Securing Arabic Text Documents

Ulin Nihayah[1], Any Muanalifah[2], and Aini Fitriyah[2]

[1]Department of Mathematics Education , Universitas Islam Negeri Walisongo Semarang, Indonesia
[2]Department of Mathematics, Universitas Islam Negeri Walisongo Semarang, Indonesia

Corresponding author: ulinnihayah94@gmail.com
Received: 07 November 2017, Revised: 20 November 2017, Accepted: 01 Desember 2017.

Abstracts

This research is an application of Hill Cipher cryptographic techniques in the process of encoding messages in the form of Arabic text. Hill Cipher is a classic cryptographic algorithm that uses modulo matrix and arithmetic in its application. The Arabic text used in the Hill Cipher cryptographic technique includes hijaiyyah letters $= 0$ا, ...، $= 1$ب، $= 28$ي، $= 29$أ، $= 30$إ، $= 31$ؤ، $= 32$ئ، $= 33$ى، $= 35$ space، $= 34$ة، $= 36$،، $= 37$؛، $= 38$؟, umbers 0, 1, and Arabic n, $= 39$! 9, so the modulo used is $mod\ 50$. This study discusses the steps in encrypting messages with examples of their application in securing Arabic text documents.

Keywords: *Cryptographics: Hill Cipher, Encoding messages, algorithm.*

1. Introduction

The exchange of information is one of the triggers for technology that continues to develop. The rapid development of technology and communication makes it easy for humans to access information quickly. In the past before computers existed, the process of exchanging information could only be done face-to-face, so that the confidentiality of information content could be kept safe. Today, technology makes it easy for everyone to access any information freely and quickly without being bound by time and place. The safety and confidentiality of information is very important, but on the other hand it will also have a negative impact. One negative impact is the level of information security that is increasingly vulnerable to the existence of tapping of information or data theft by hackers or irresponsible parties. Therefore the coding process with certain codes is very important.

One of the methods used to maintain the confidentiality of a message to be safe to the recipient is cryptography. Cryptography is a technique to hide messages by encoding the contents of the message into codes that are incomprehensible, meaningless and illegible. Cryptography not only deals with hiding messages but rather leads to techniques that provide information security (Sadikin, 2012). Hill Cipher algorithm is one of the

cryptographic techniques. Hill Cipher was created by Lester S. Hill in 1929. The Hill Cipher algorithm is the application of modulo arithmetic and matrix in cryptographic activities. The Hill Cipher algorithm uses matrix multiplication and matrix inverse as key in its application. In the Hill Cipher cryptography technique, not all key matrices can be used for encryption and decryption processes. The key to the Hill Cipher is the square matrix and the invertible matrix. Square matrix is a matrix that has n rows and $n$columns, and is often written as a matrix with an order $n \times n$ (Anton & Rorres, 2014). Whereas an invertible matrix is an inverse matrix (Larson & Falvo, 2009).

Now encoding has been widely used and used in various languages. Almost all the languages of the countries of the world are stated in Latin alphabet letters. Therefore it is not surprising that encoding with various methods is often found by applying these letters. In this research an encoding attempt is made for documents that do not use Latin alphabet letters, namely Arabic letters. In addition to making the conversion more complete, researchers added Arabic numerals and some punctuation characters that are often used.

2. Methodology

This research is a library research by studying, analyzing, and investigating documents, writings or literature about Hill Cipher's cryptographic algorithm, modulo arithmetic, and basic theory about matrix. Furthermore, the Hill Cipher cryptographic algorithm is implemented on Arabic text documents.

3. Result and Discussion

The application of the Hill Cipher algorithm in the security of Arabic text documents is done by modifying the existing Hill Cipher algorithm in general. The Hill Cipher algorithm in general (Setyaningsih, 2015), namely:

Encryption process:

$$C = K \cdot P \bmod r$$

with:
C = *ciphertext*
K = key matrix
P =*plaintext*
r = many characters

Decryption process:

$$P = K^{-1} \cdot C \bmod r$$

with:
K^{-1} = key matrix

Hill Cipher cryptographic techniques in this study include hijaiyyah letters, some punctuation characters and Arabic numbers. The encryption and decryption process of Hill Cipher is represented using modular arithmetic by first changing the letters into numbers according to the conversion rules that have been used. Conversion of Arabic script characters used in this study is shown in Table 3.1.

Table 3.1. Convert characters (arabic) to numbers

ا	ب	ت	ث	ج	ح	خ	د	ذ	ر	ز	س	ش	ص	ض	ط	ظ
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
ع	غ	ف	ق	ك	ل	م	ن	و	ه	ء	ي	أ	إ	ؤ	ئ	ى
17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
ة	Spasi	.	،	؟	!	0	1	2	3	4	5	6	7	8	9	
34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	

Based on Table 3.1, the Arabic text object used in this study amounted to 50, so the modulo used in this study is $mod\ 50$. Thus, the Hill Cipher algorithm for Arabic text can be written as follows:

Encryption process:

$$C = K \cdot P \bmod 50$$

with:
C = *ciphertext*
K = key matrix
P =*plaintext*
r = many characters

Decryption process:

$$P = K^{-1} \cdot C \bmod 50$$

with:
K^{-1} = key matrix

Example (Encryption) Use the Hill Cipher matrix to find the ciphertext from the following plaintext:

رقم تليفوني 085747824973

If the key matrix is known:

$$K=\begin{bmatrix}1&2&3\\0&1&4\\5&6&0\end{bmatrix}$$

The following steps can be taken to convert plaintext to ciphertext in the Hill Cipher algorithm:

a. convert each character from plaintext to numbers according to Table 3.1. and dividing them into blocks according to the size of the key matrix. If the number of characters from the plaintext that has been divided into blocks does not match the size of the matrix or lacks the characters, it is necessary to add additional characters to the number of characters needed to match the size of the matrix.

ر	ق	م	Spasi	ث	ل	ي	ف	و	ن	ي	Spasi	0	8	5
9	20	23	35	2	22	28	19	25	24	28	35	40	48	45
B1			B2			B3			B4			B5		

7	4	7	8	2	4	9	7	3
47	44	47	48	42	44	49	47	43
B6			B7			B8		

b. multiplying the key matrix K with the numbers in each block B and the encrypted matrix modulated 50 so that the numbers of the matrix can be converted with the letters according to Table 3.1.

$$E1=\begin{bmatrix}1&2&3\\0&1&4\\5&6&0\end{bmatrix}\begin{bmatrix}9\\20\\23\end{bmatrix}=\begin{bmatrix}118\\112\\165\end{bmatrix}mod\,50=\begin{bmatrix}18\\12\\15\end{bmatrix}=\begin{bmatrix}غ\\س\\ط\end{bmatrix}$$

$$E2=\begin{bmatrix}1&2&3\\0&1&4\\5&6&0\end{bmatrix}\begin{bmatrix}35\\2\\22\end{bmatrix}=\begin{bmatrix}105\\90\\187\end{bmatrix}mod\,50=\begin{bmatrix}5\\40\\37\end{bmatrix}=\begin{bmatrix}ج\\0\\ء\end{bmatrix}$$

$$E3=\begin{bmatrix}1&2&3\\0&1&4\\5&6&0\end{bmatrix}\begin{bmatrix}28\\19\\25\end{bmatrix}=\begin{bmatrix}141\\119\\254\end{bmatrix}mod\,50=\begin{bmatrix}41\\19\\4\end{bmatrix}=\begin{bmatrix}1\\ف\\ج\end{bmatrix}$$

$$E4=\begin{bmatrix}1&2&3\\0&1&4\\5&6&0\end{bmatrix}\begin{bmatrix}24\\28\\35\end{bmatrix}=\begin{bmatrix}185\\168\\288\end{bmatrix}mod\,50=\begin{bmatrix}35\\18\\38\end{bmatrix}=\begin{bmatrix}Spasi\\غ\\؟\end{bmatrix}$$

$$E5=\begin{bmatrix}1&2&3\\0&1&4\\5&6&0\end{bmatrix}\begin{bmatrix}40\\48\\45\end{bmatrix}=\begin{bmatrix}271\\228\\488\end{bmatrix}mod\,50=\begin{bmatrix}21\\28\\38\end{bmatrix}=\begin{bmatrix}ك\\ي\\؟\end{bmatrix}$$

$$E6=\begin{bmatrix}1&2&3\\0&1&4\\5&6&0\end{bmatrix}\begin{bmatrix}47\\44\\47\end{bmatrix}=\begin{bmatrix}276\\232\\499\end{bmatrix}mod\,50=\begin{bmatrix}26\\32\\49\end{bmatrix}=\begin{bmatrix}ه\\ئ\\9\end{bmatrix}$$

$$E7=\begin{bmatrix}1&2&3\\0&1&4\\5&6&0\end{bmatrix}\begin{bmatrix}48\\42\\44\end{bmatrix}=\begin{bmatrix}264\\218\\492\end{bmatrix}mod\,50=\begin{bmatrix}14\\18\\42\end{bmatrix}=\begin{bmatrix}ض\\غ\\2\end{bmatrix}$$

$$E8=\begin{bmatrix}1&2&3\\0&1&4\\5&6&0\end{bmatrix}\begin{bmatrix}49\\47\\43\end{bmatrix}=\begin{bmatrix}272\\219\\527\end{bmatrix}mod\,50=\begin{bmatrix}22\\19\\27\end{bmatrix}=\begin{bmatrix}ل\\ف\\ء\end{bmatrix}$$

So, ciphertext for plaintext

رقم تليفوني

which has been encrypted using a key $K=\begin{bmatrix}1&2&3\\0&1&4\\5&6&0\end{bmatrix}$ is :

غشطح0،افج غ؟كي؟هئ9ضغ2لفء

Example (Decryption**)** Use the Hill Cipher matrix to find the plaintext of the following ciphertext:

غشطح0،افج غ؟كي؟هئ9ضغ2لفء

If known the key matrix:

$$K=\begin{bmatrix}1&2&3\\0&1&4\\5&6&0\end{bmatrix}$$

Steps to convert plaintext to ciphertext in the Hill Cipher algorithm:

a. convert each character from ciphertext to numbers according to Table 3.1. and dividing them into blocks according to the size of the key matrix.

غ	ض	ط	ح	0	،	ا	ف	ج	Spasi	غ	؟	ك	ي	؟	ه	ئ	9
18	12	15	5	40	37	41	19	4	35	18	38	21	28	38	26	32	49
B1			B2			B3			B4			B5			B6		

ض	غ	2	ل	ف	ء
14	18	42	22	19	27
B7			B8		

b. determine the inverse of the key matrix K. The inverse of a matrix can be obtained using the formula:

$$K^{-1} = \frac{1}{\det(K)} adj(K)$$(Anton & Rorres, 2005)

$$K^{-1} = \frac{1}{\det(K)} adj(K)(mod\ 50)$$

$$= \frac{1}{1}\begin{bmatrix} -24 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{bmatrix}(mod\ 50)$$

$$= \begin{bmatrix} -24 & 18 & 5 \\ 20 & -15 & -4 \\ -5 & 4 & 1 \end{bmatrix}(mod\ 50)$$

$$= \begin{bmatrix} 26 & 18 & 5 \\ 20 & 35 & 46 \\ 45 & 4 & 1 \end{bmatrix}$$

c. multiplying the inverse of the key matrix K with each block B number and the decryption matrix is modulated 50 so that the number of the matrix can be converted with the letters according to Table 3.1.

$$D1 = \begin{bmatrix} 26 & 18 & 5 \\ 20 & 35 & 46 \\ 45 & 4 & 1 \end{bmatrix}\begin{bmatrix} 18 \\ 12 \\ 15 \end{bmatrix} = \begin{bmatrix} 759 \\ 1470 \\ 873 \end{bmatrix} mod\ 50 = \begin{bmatrix} 9 \\ 20 \\ 23 \end{bmatrix} = \begin{bmatrix} ر \\ ق \\ م \end{bmatrix}$$

$$D2 = \begin{bmatrix} 26 & 18 & 5 \\ 20 & 35 & 46 \\ 45 & 4 & 1 \end{bmatrix}\begin{bmatrix} 5 \\ 40 \\ 37 \end{bmatrix} = \begin{bmatrix} 1035 \\ 3202 \\ 372 \end{bmatrix} mod\ 50 = \begin{bmatrix} 35 \\ 2 \\ 22 \end{bmatrix} = \begin{bmatrix} Spasi \\ ت \\ ل \end{bmatrix}$$

$$D3 = \begin{bmatrix} 26 & 18 & 5 \\ 20 & 35 & 46 \\ 45 & 4 & 1 \end{bmatrix}\begin{bmatrix} 41 \\ 19 \\ 4 \end{bmatrix} = \begin{bmatrix} 1428 \\ 1669 \\ 1925 \end{bmatrix} mod\ 50 = \begin{bmatrix} 28 \\ 19 \\ 25 \end{bmatrix} = \begin{bmatrix} ي \\ ف \\ و \end{bmatrix}$$

$$D4 = \begin{bmatrix} 26 & 18 & 5 \\ 20 & 35 & 46 \\ 45 & 4 & 1 \end{bmatrix}\begin{bmatrix} 35 \\ 18 \\ 38 \end{bmatrix} = \begin{bmatrix} 1424 \\ 3078 \\ 1685 \end{bmatrix} mod\ 50 = \begin{bmatrix} 24 \\ 28 \\ 35 \end{bmatrix} = \begin{bmatrix} ن \\ ي \\ Spasi \end{bmatrix}$$

$$D5 = \begin{bmatrix} 26 & 18 & 5 \\ 20 & 35 & 46 \\ 45 & 4 & 1 \end{bmatrix}\begin{bmatrix} 21 \\ 28 \\ 38 \end{bmatrix} = \begin{bmatrix} 1240 \\ 3148 \\ 1095 \end{bmatrix} mod\ 50 = \begin{bmatrix} 40 \\ 48 \\ 45 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 5 \end{bmatrix}$$

$$D6 = \begin{bmatrix} 26 & 18 & 5 \\ 20 & 35 & 46 \\ 45 & 4 & 1 \end{bmatrix}\begin{bmatrix} 26 \\ 32 \\ 49 \end{bmatrix} = \begin{bmatrix} 1497 \\ 3894 \\ 1347 \end{bmatrix} mod\ 50 = \begin{bmatrix} 47 \\ 44 \\ 47 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ 7 \end{bmatrix}$$

$$D7 = \begin{bmatrix} 26 & 18 & 5 \\ 20 & 35 & 46 \\ 45 & 4 & 1 \end{bmatrix}\begin{bmatrix} 14 \\ 18 \\ 42 \end{bmatrix} = \begin{bmatrix} 898 \\ 2842 \\ 744 \end{bmatrix} mod\ 50 = \begin{bmatrix} 48 \\ 42 \\ 44 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \\ 4 \end{bmatrix}$$

$$D8 = \begin{bmatrix} 26 & 18 & 5 \\ 20 & 35 & 46 \\ 45 & 4 & 1 \end{bmatrix}\begin{bmatrix} 22 \\ 19 \\ 27 \end{bmatrix} = \begin{bmatrix} 1049 \\ 2347 \\ 1093 \end{bmatrix} mod\ 50 = \begin{bmatrix} 49 \\ 47 \\ 43 \end{bmatrix} = \begin{bmatrix} 9 \\ 7 \\ 3 \end{bmatrix}$$

So, plaintext for ciphertext

غشطح0،افج غ؟كي؟هئ9ضغ2لفء

which has been encrypted using the key $K = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 5 & 6 & 0 \end{bmatrix}$ is: رقم تليفوني

Based on these examples, it shows that each letter in the same plaintext and ciphertext does not produce the same letter if the encryption or decryption process has been carried out. The encryption function on the Hill Cipher is obtained only by multiplying key matrices with plaintext that has been represented in the form of a matrix. While the decryption function is obtained only by multiplying the key matrix inverse with the ciphertext that has been represented in the form of a matrix.

4. Conclusion

The Hill Cipher algorithm is a cryptographic technique using square matrices as the key and modulo arithmetic to the matrix. The key matrix must be an invertible matrix. The implementation of the Hill Cipher algorithm in Arabic text documents involves both cryptographic processes, namely encryption and decryption.

Acknowledgment

The authors wish to thank Department of Mathematics, Universitas Islam Negeri Wallisongo Semarang for supporting this work.

References

[1] H. Anton, C. Rorres, *Elementary Linear Algebra: Application Version* (9^{th} ed). New York: John Wiley & Sons, Inc. (2005).
[2] _____ *Elementary Linear Algebra: Application Version* (11^{th} ed). New York: John Wiley & Sons, Inc. (2014).
[3] R. Larson, D. C. Falvo, *Elementary Linear Algebra* (6^{th} ed). USA: Houghton Mifflin Harcourt Publishing Company. (2009).
[4] R. Sadikin, *Kriptografi untuk Keamanan Jaringan dan Implementasinya dalam Bahasa Java*. Yogyakarta: Andi Offset (2012).
[5] Setyaningsih, Emy. *Kriptografi dan Implementasinya Menggunakan Matlab.* Yogyakarta: Andi Offset (2015).