MS-Excel based application for security of English texts with 4×4 matrix Hill cipher

Md. Kamal Hossain[1], Md. Ferdous Wahid[2]

Abstract— This paper aims to protect the confidential English texts access by the parties that do not have permission. The process of encryption and decryption of English text using 4×4 matrix Hill cipher provides more security.

Keywords— Encryption, Decryption, Hill cipher, English texts, Matrix, Plaintext, MS EXCEL.

1 Introduction

In cryptography, the one of the most polyalphabetic substitution cipher is the Hill cipher, developed by the mathematician Lester Hill in 1929. In Hill cipher, the plaintext is divided into equal-size of blocks. The blocks are encrypted one at a time in such a way that each character in the block contributes to the encryption of other characters in the block. For this reason, Hill cipher is also called block cipher. [2, 3, 5]

The core of the Hill cipher is matrix manipulations. In Hill cipher, each character is assigned a numerical value like A=0, B=1, C=2……. Y=24, Z=25 and a number modulo 26 [8, 9]. For encryption, algorithm takes m successive plaintext letters and instead of that substitutes m cipher letters. The substitution of ciphertext letters in the place of plaintext letters leads to m linear equation.

In Hill cipher, the key is a square matrix of size m×m in which m is the size of the block. To encrypt a message, each block of m letters is multiplied by an invertible matrix of size m×m (modulo 26). In order to decrypt the message, each block is multiplied by the inverse matrix of the matrix used for encryption. The matrix should be chosen randomly from the set of invertible matrices of size m×m (modulo 26) [2, 3, 4, 5, 6, 7]. If P is the plaintext, C is the ciphertext, and K is the key. The encryption algorithm $E_k(x)$ creates the ciphertext from the plaintext. The decryption algorithm $D_k(x)$ creates the plaintext from the ciphertext. We assume that $E_k(x)$ and $D_k(x)$ are inverses of each other. They cancel the effect of each other.

Encryption algorithm for Hill cipher

$$C = E(K, P) = PK \bmod 26 \dots (1)$$

Decryption algorithm for Hill cipher

$$P = D(K, C) = Ck^{-1} \bmod 26 = PKK^{-1} = P \dots (2)$$

where, P is the plaintext, K is the key matrix and C is the ciphertext.

Example: Let the plaintext message is "*COMPUTER*" and the encrypted key is

$$\boldsymbol{K} = \begin{pmatrix} 1 & 1 & 0 & 7 \\ 2 & 2 & 1 & 0 \\ 25 & 0 & 25 & 22 \\ 24 & 2 & 5 & 13 \end{pmatrix}$$

Then the plaintext is divided into two blocks. One is "COMP" and the other is "UTER". Then the numeric position of the plaintext is represented by the following two quadruples:

$$COMPUTER \rightarrow (C\ O\ M\ P), (U\ T\ E\ R)$$
$$\rightarrow (2\ \ 14\ \ 12\ \ 15), (20\ \ 19\ \ 4\ \ 17)$$

The encryption is performed is as follows:

$$C = PK$$

$$(2\ \ 14\ \ 12\ \ 15).\begin{pmatrix} 1 & 1 & 0 & 7 \\ 2 & 2 & 1 & 0 \\ 25 & 0 & 25 & 22 \\ 24 & 2 & 5 & 13 \end{pmatrix} = (690\ \ 60\ \ 389\ \ 473)\ mod\ 26$$
$$\equiv (14\ 8\ 25\ 5)$$
$$(14\ 8\ 25\ 5) = (O\ I\ Z\ F) \dots (3)$$

$$(20\ \ 19\ \ 4\ \ 17).\begin{pmatrix} 1 & 1 & 0 & 7 \\ 2 & 2 & 1 & 0 \\ 25 & 0 & 25 & 22 \\ 24 & 2 & 5 & 13 \end{pmatrix} = (326\ \ 72\ \ 154\ \ 319)\ mod\ 26$$
$$\equiv (14\ \ 20\ \ 24\ \ 7)$$
$$(14\ \ 2\ \ 24\ \ 7) = (O\ U\ Y\ H) \dots (4)$$

$$(14\ 8\ 25\ 5), (14\ \ 20\ \ 24\ \ 7) \rightarrow (O\ I\ Z\ F), (O\ U\ Y\ H)$$
$$\rightarrow OIZFOUYH \dots (5)$$

So, the ciphertext is "$\boldsymbol{OIZFOUYH}$"

The key matrix K is invertible and the invertible key matrix is $\boldsymbol{K^{-1}}$ such that K. $\boldsymbol{K^{-1}}$= $\boldsymbol{K^{-1}}$.K=$\boldsymbol{I_4}$, where $\boldsymbol{I_4}$ is the identity matrix of size 4×4. The inverse matrix of the key matrix

$\boldsymbol{K} = \begin{pmatrix} a_{11}a_{12}a_{13}a_{14} \\ a_{21}a_{22}a_{23}a_{24} \\ a_{31}a_{32}a_{33}a_{34} \\ a_{41}a_{42}a_{43}a_{44} \end{pmatrix}$ can be calculated by the following formula:

$$\boldsymbol{K^{-1}} = \frac{1}{\det K}\begin{pmatrix} A_{11}A_{12}A_{13}A_{14} \\ A_{21}A_{22}A_{23}A_{24} \\ A_{31}A_{32}A_{33}A_{34} \\ A_{41}A_{42}A_{43}A_{44} \end{pmatrix} \dots (6)$$

where det K is the determinant of a key matrix Kwith dimension 4×4 defined by equation (7) and $\boldsymbol{A_{ij}}$ is the cofactor of

the element $\boldsymbol{a_{ij}}$ in the i-th row and the j-column of the key matrix K:

$$\begin{aligned}\det \boldsymbol{K} = {} & a_{11}.a_{22}.a_{33}.a_{44} + a_{11}.a_{23}.a_{34}.a_{42} \\ & + a_{11}.a_{24}.a_{32}.a_{43} + a_{12}.a_{24}.a_{33}.a_{41} \\ & + a_{12}.a_{21}.a_{34}.a_{43} + a_{12}.a_{23}.a_{31}.a_{44} \\ & + a_{13}.a_{21}.a_{32}.a_{44} + a_{13}.a_{22}.a_{34}.a_{41} \\ & + a_{13}.a_{24}.a_{31}.a_{42} + a_{14}.a_{23}.a_{32}.a_{41} \\ & + a_{14}.a_{21}.a_{33}.a_{42} + a_{14}.a_{21}.a_{31}.a_{43} \\ & - a_{11}.a_{24}.a_{33}.a_{42} - a_{11}.a_{22}.a_{34}.a_{43} \\ & - a_{11}.a_{23}.a_{32}.a_{44} - a_{12}.a_{21}.a_{33}.a_{44} \\ & - a_{12}.a_{23}.a_{34}.a_{41} - a_{12}.a_{24}.a_{31}.a_{43} \\ & - a_{13}.a_{24}.a_{32}.a_{41} - a_{13}.a_{21}.a_{34}.a_{42} \\ & - a_{13}.a_{22}.a_{31}.a_{44} - a_{14}.a_{21}.a_{32}.a_{43} \\ & - a_{14}.a_{22}.a_{33}.a_{41} \\ & - a_{14}.a_{23}.a_{31}.a_{42} \quad \dots (7)\end{aligned}$$

The cofactor $\boldsymbol{A_{ij}}$ of the element $\boldsymbol{a_{ij}}$is calculated by equation (8) [1]:

$$\boldsymbol{A_{ij}} = (-1)^{i+j}\boldsymbol{D_{ij}} \dots (8)$$

where D_{ij} is the additional minor, In the case the additional minors D_{ij} are determinants of the third order and use the triangle's rule to calculate the determinant of matrix with size 3×3 that is given below (9):

$$\boldsymbol{D_{ij}} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a.e.i + b.f.g + c.d.h - c.e.g - a.f.h - b.d.i \dots (9)$$

After the modular reduction equation (6) is valid if the modular multiplicative inversion is used for calculating$(\det \boldsymbol{K})^{-1}$. Therefore, in the case, after a preliminary calculation of cofactor $\boldsymbol{A_{ij}}$ according to the equation (8) and (9) transformations led to the following equation for the inverse matrix:

$$\boldsymbol{K^{-1}} = \begin{pmatrix} 16 & 9 & 15 & 22 \\ 15 & 18 & 23 & 1 \\ 16 & 25 & 2 & 6 \\ 18 & 11 & 2 & 19 \end{pmatrix} mod\ 26 \dots (10)$$

To recover the plaintext "COMPUTER" the following operation can be performed:

$$OIZFOUYH \rightarrow (O\ I\ Z\ F), (O\ U\ Y\ H) \rightarrow (14\ 8\ 25\ 5), (14\ 20\ 24\ 7) \dots (11)$$

For the ciphertext OIZF:

$$(14\ 8\ 25\ 5).\begin{pmatrix} 16 & 9 & 15 & 22 \\ 15 & 18 & 23 & 1 \\ 16 & 25 & 2 & 6 \\ 18 & 11 & 2 & 19 \end{pmatrix} = (834\ 950\ 454\ 561) \equiv (2\ 14\ 12\ 15)\ mod 26 \dots (12)$$

and for OUYH

$$(14\ 20\ 24\ 7).\begin{pmatrix} 16 & 9 & 15 & 22 \\ 15 & 18 & 23 & 1 \\ 16 & 25 & 2 & 6 \\ 18 & 11 & 2 & 19 \end{pmatrix} = (1034\ 1163\ 732\ 605) \equiv (20\ 19\ 4\ 17) mod 26 \dots (13)$$

So the recovery plaintext is "COMPUTER"

$$(2\ 14\ 12\ 15), (20\ 19\ 4\ 17) \rightarrow (C\ O\ M\ P), (U\ T\ E\ R) \rightarrow COMPUTER$$

2 MS-Excel based application for security of English texts with 4×4 matrix Hill cipher

English texts are encrypted and decrypted by using the Hill cipher based on the matrix with dimension 4×4 [2, 4, 5, 6, 7, 10]. Based on the MS EXCEL application the process of encryption and decryption was developed.

The MS EXCEL based application scheme the following modules:

1) Module "Matrix"
2) Module "Inverse Matrix"
3) Module "Encryption"
4) Module "Decryption"

2.1 English Alphabet Representation

The numeric position of each English alphabet is represented by a number modulo 26 using the simple scheme A=0, B=1, C=2…, Z=25 (Table 1). Whereas some references used the following scheme: A=25, B=24, C=23…, Z=0 but this does not affect the encryption and decryption processes. Using the function =CODE(address)-66 the transformation is realized. The built-in function CODE determines the ASCII code of the symbol written in the cell, whose address is specified as an argument to the function (address). The ASCII code of the symbol B is 42 (hexadecimal), i.e. 66 (decimal). Therefore, to obtain the numeric code 1 for the letter B it is required to remove the number 66 by its ASCII code (for the uppercase letters of the English alphabet).

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

Table 1: Numerical Position of the English Alphabet.

2.2 Modulo Matrix

The key matrix K is selected. Then the key matrix K must be satisfied the following two conditions. Firstly, the matrix is invertible, i.e. its determinant is different from zero. Secondly, its determinant has no common divisors with number 26. i.e. is not divisible by 2 and 13.

In MS EXCEL the function =MOD(address; 26), is used to determine the determinant of the key matrix K, where address is the address of the cell containing the value in the key matrix K. After determining the value of the determinant of the key matrix K in modulo 26, two conclusions are displayed.

Firstly, If the calculated value is different from zero, then display "The determinant of the matrix is different from 0 - correct choice of the matrix!". In MS EXCEL function: =IF(address<>0; "The determinant of the matrix is different from 0 – correct choice of the matrix!"; "The determinant of the matrix is equal to 0 – wrong choice of the matrix!") is used to choice the correct matrix, where address is the address of the cell containing the value of the determinant of the matrix K modulo 26.

Secondly, In MS EX CEL function: =IF(GCD(address;26)<>1; "The greatest common divisor of the determinant of the matrix and the number 26 is different from 1 – wrong choice of the matrix!"; "The greatest common divisor of the determinant of the matrix and the number 26 is equal to 1 – correct choice of the matrix!") is also used to choice the correct matrix, where address behaves the address of the cell containing the value of the determinant of the matrix K modulo 26 [10].

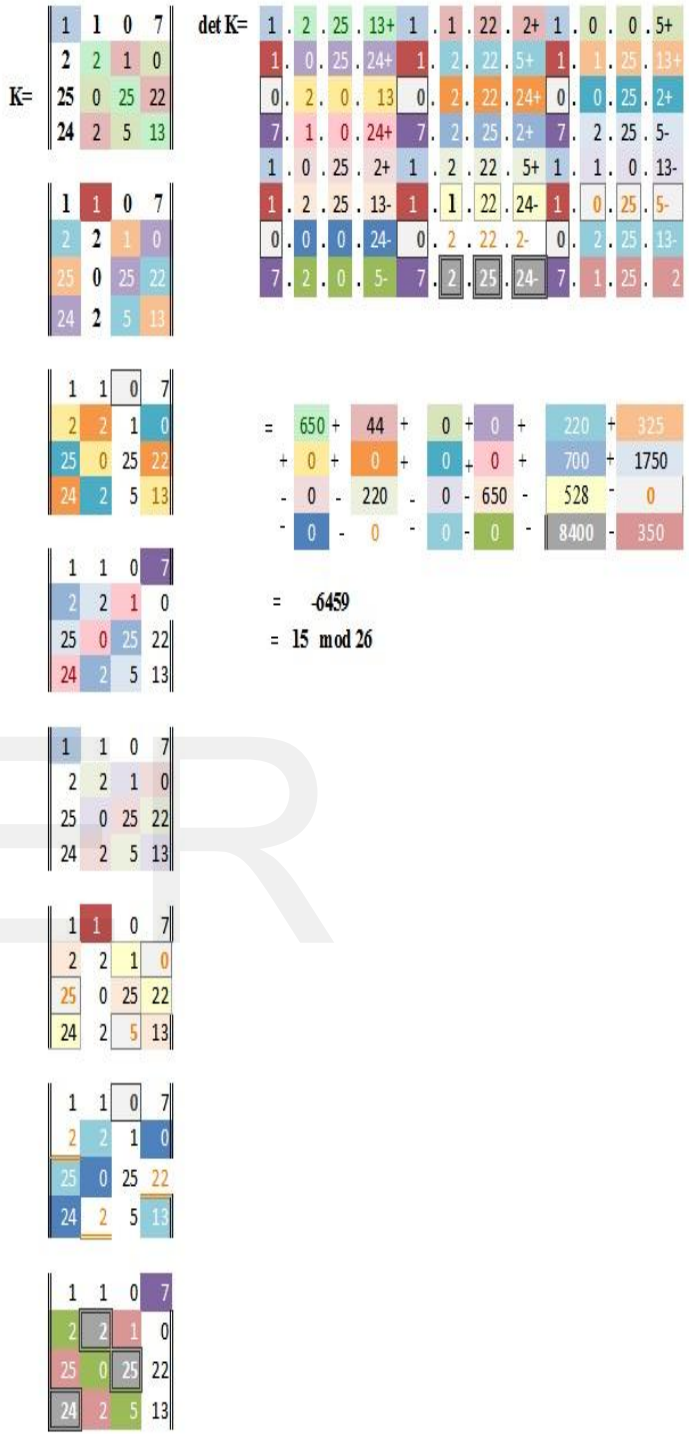
2.3 Inverse Matrix

The determinant of the matrix K is first calculated using the relationships (6), (7), (8) and (9) in it is presented with color codes. In determining the elements of the inverse matrix the modular reduction is used on the base of the modular multiplicative inversion to calculate $(\det K)^{-1}$. Table 2 shows the det K and $(\det K)^{-1}$ based on the the product of det K and $(\det K)^{-1}$ divided by 26 is 1.

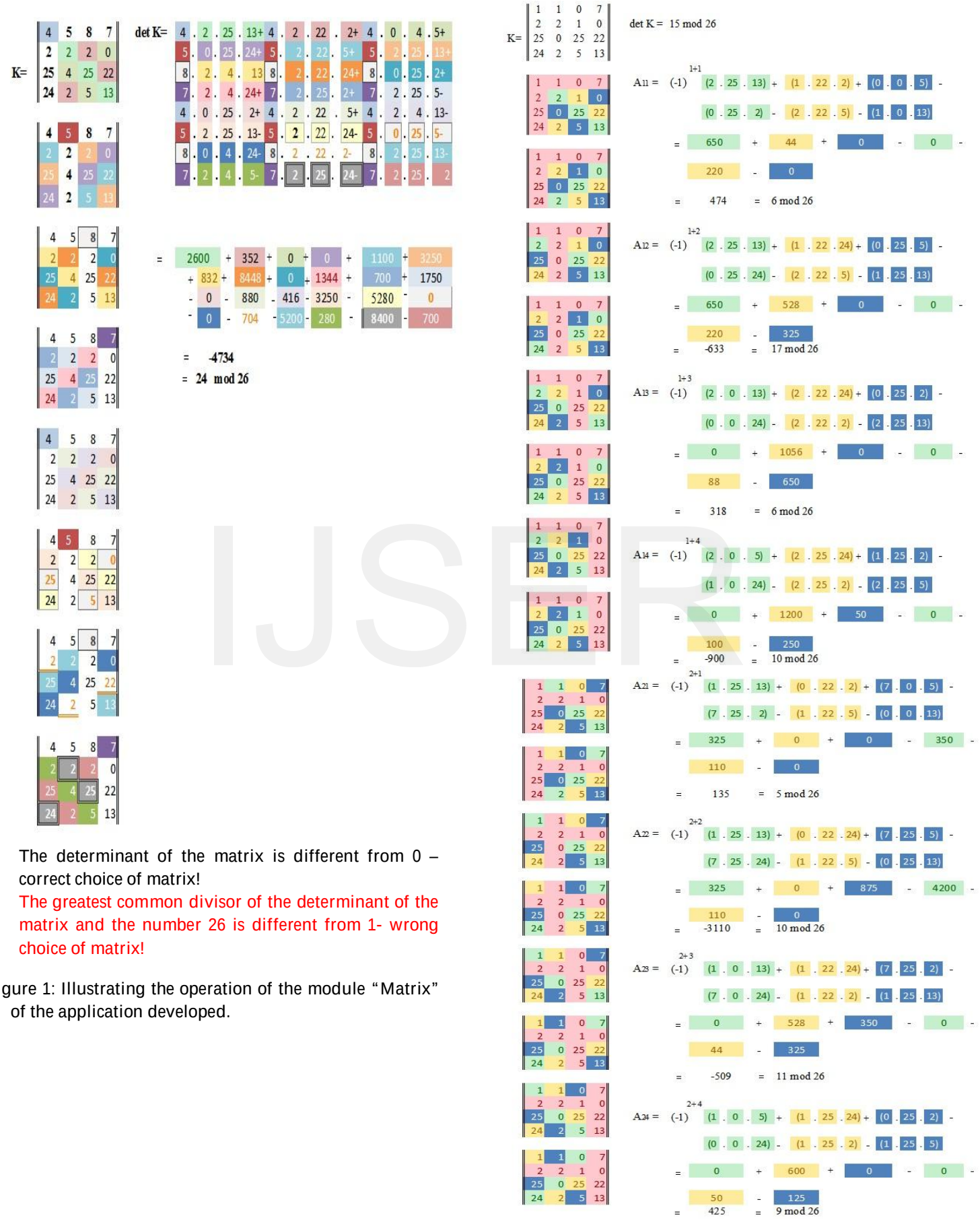
det K	1	3	5	7	9	11	15	17	19	21	23	25
(det K)-1	1	9	21	15	3	19	7	23	11	5	17	25

Table 2: Correspondence between det K and $(\det K)^{-1}$

In MS EXCEL function: =MDETERM(address 1:address 2) is used determined the determinant of the matrix, where address 1:address 2 indicates the "array". The modular multiplicative inverse of the det K in the application developed is determined using the pre-established correlation table in the case of English (Table 2) and several input calls to the built-in function IF:=IF(address=1;1;IF(address=3;9;IF(address=5;21;IF(address=7;15;IF(address=17;23;IF(addres=11;19;IF(address=25;25)))))))), where address is the address of the cell containing the determinant of the matrix K according to the equation (6) and address1 is the address of a remote [10].

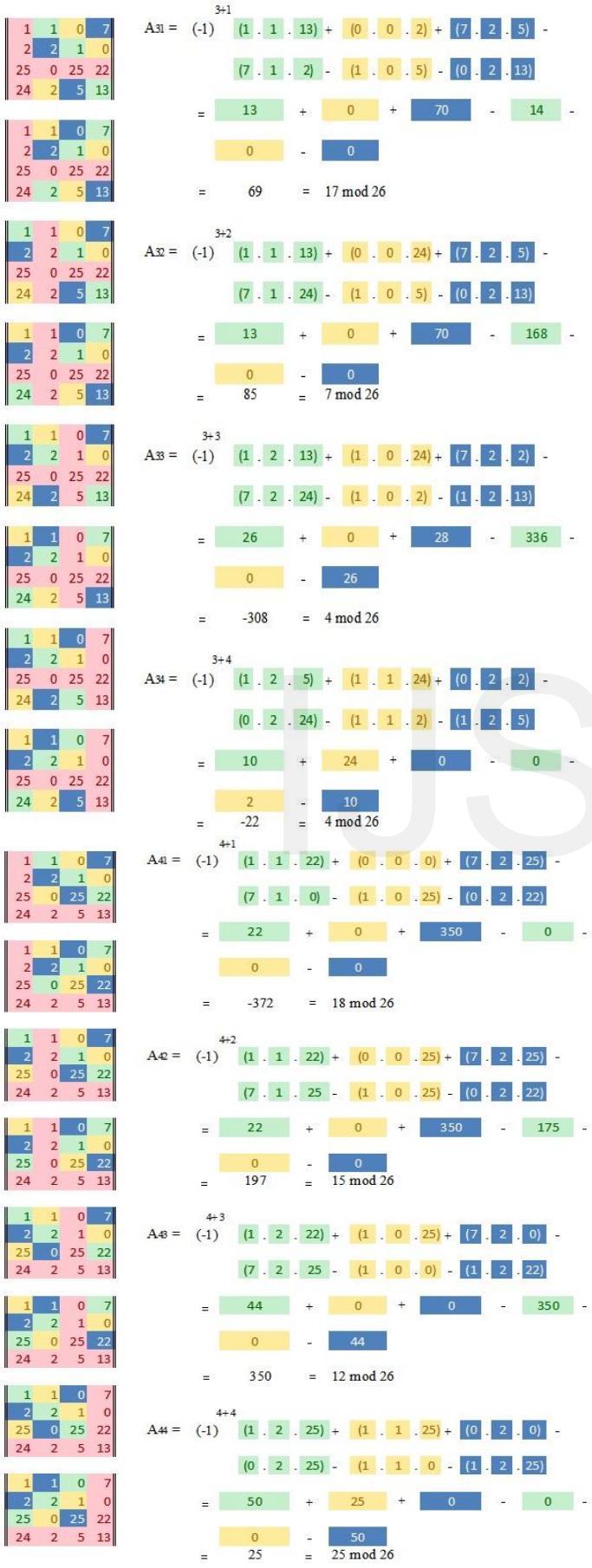


1. The determinant of the matrix is different from 0 – correct choice of matrix!
2. The greatest common divisor of the determinant of the matrix and the number 26 is equal to 1- correct choice of matrix!



3. The determinant of the matrix is different from 0 – correct choice of matrix!
4. The greatest common divisor of the determinant of the matrix and the number 26 is different from 1- wrong choice of matrix!

Figure 1: Illustrating the operation of the module "Matrix" of the application developed.



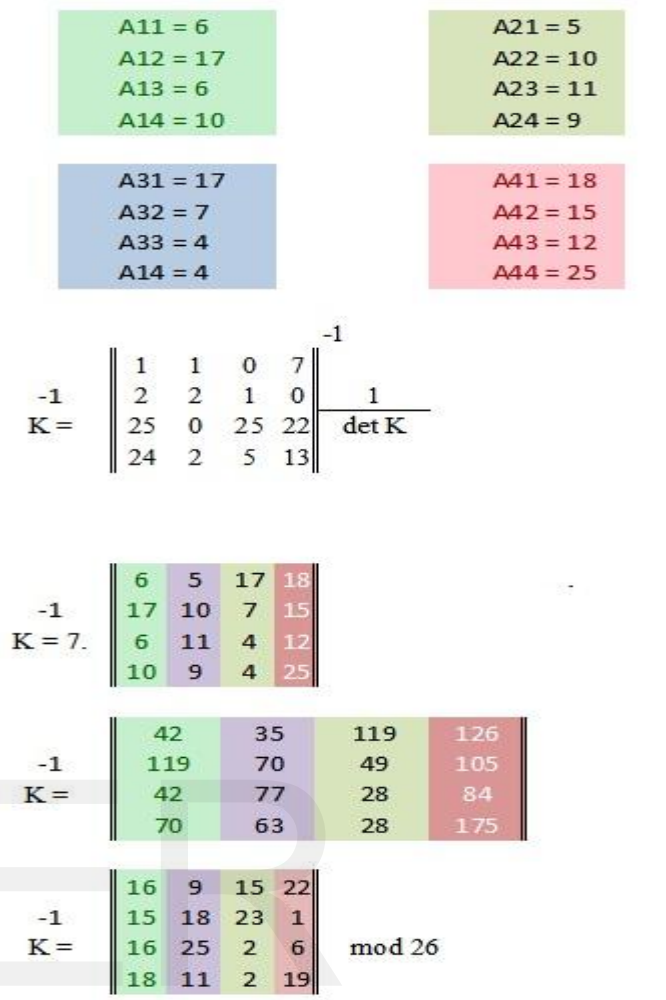


Figure 2: Illustrating the operation of the module "Inverse Matrix"

2.4 Encryption Process

The plaintext message breaks into blocks of 4 letters for encryption. Then each block of 4 letters is multiplied by the invertible matrix of with dimension 4×4 again modulo 26. A general equation is used to calculate the encryption that is previously discussed.

$$C = P.K = (p1\ p2\ p3\ p4).\begin{pmatrix} a11\ a12\ a13\ a14 \\ a21\ a22\ a23\ a24 \\ a31\ a32\ a33\ a34 \\ a41\ a42\ a43\ a44 \end{pmatrix}$$

$$= (p1.a11 + p2.a21 + p3.a31 + p4.a41 \quad p1.a12 + p2.a22 + p3.a32 + p4.a42 \quad p1.a13 + p2.a23 + p3.a33 + p4.a43 \quad p1.a14 + p2.a24 + p3.a34 + p4.a44) \ldots (13)$$

After the multiplication of the matrices using the relationship (13) the C matrix is obtained with dimension 1×4 which is used for building the ciphertext. The function =CHAR(address+65) is used in MS EXCEL to realized

thetransformation of the number from 0 to 25 in symbols based on table 1 to obtained the ciphertext.

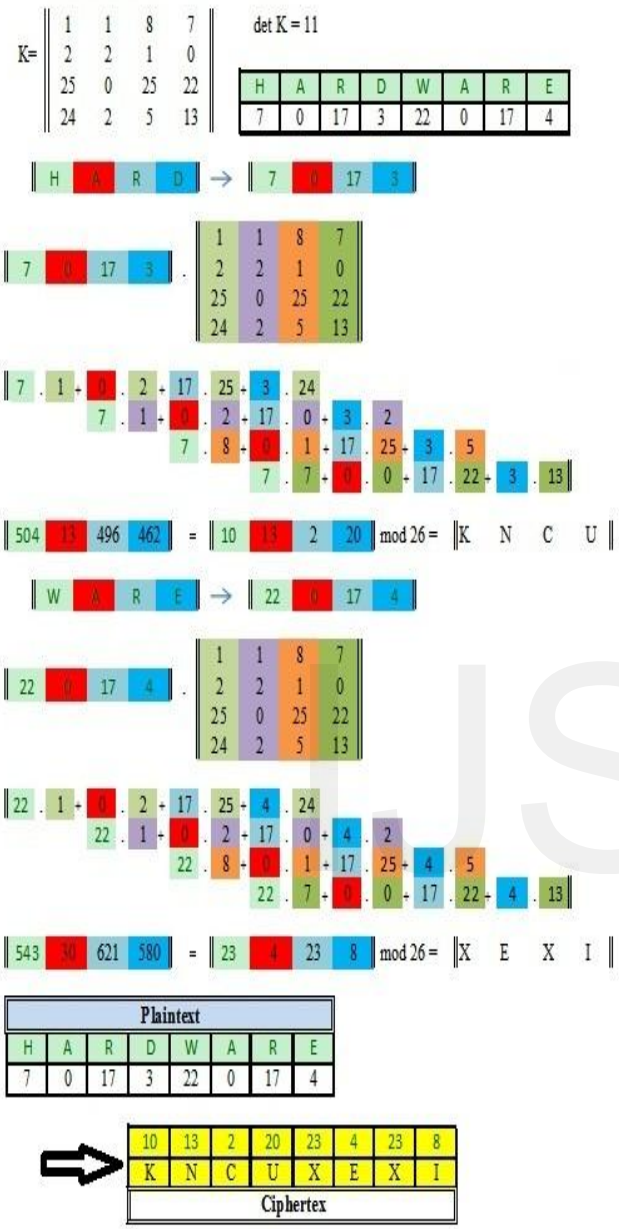


Figure 3: Illustrating the operation "Encryption" when selected matrix (Option 1) and text HARDWARE

2.5 Decryption Process

To recover the original message the ciphertext message is broken into 4 letters blocks for decryption. Then each block of 4 letters is multiplied by the inverse matrix of the encryption matrix again modulo 26. The decryption process is inverses to the encryption process and they cancel the effect of each other.The general equation of the decryption process is given below:

$$P = C.K^{-1} = (c1\ c2\ c3\ c4).\begin{pmatrix} k11 & k12 & k13 & k14 \\ k21 & k22 & k23 & k24 \\ k31 & k32 & k33 & k34 \\ k41 & k42 & k43 & k44 \end{pmatrix}$$

$$= (c1.k11 + c2.k21 + c3.k31 + c4.k41 \quad c1.k12 + c2.k22 + c3.k32 + c4.k42 \quad c1.k13 + c2.k23 + c3.k33 + c4.k43 \quad c1.k14 + c2.k24 + c3.k34 + c4.k44) \ldots (14)$$

where, k_{ij} are the elements of the inverse matrix K^{-1}.

After the multiplication of the two matrices by equation (14) the P matrix is obtained with dimension 1×4 which built the original plaintext. The plaintext is obtained by transforming the numbers from 0 to 25 in letters based on table 1.

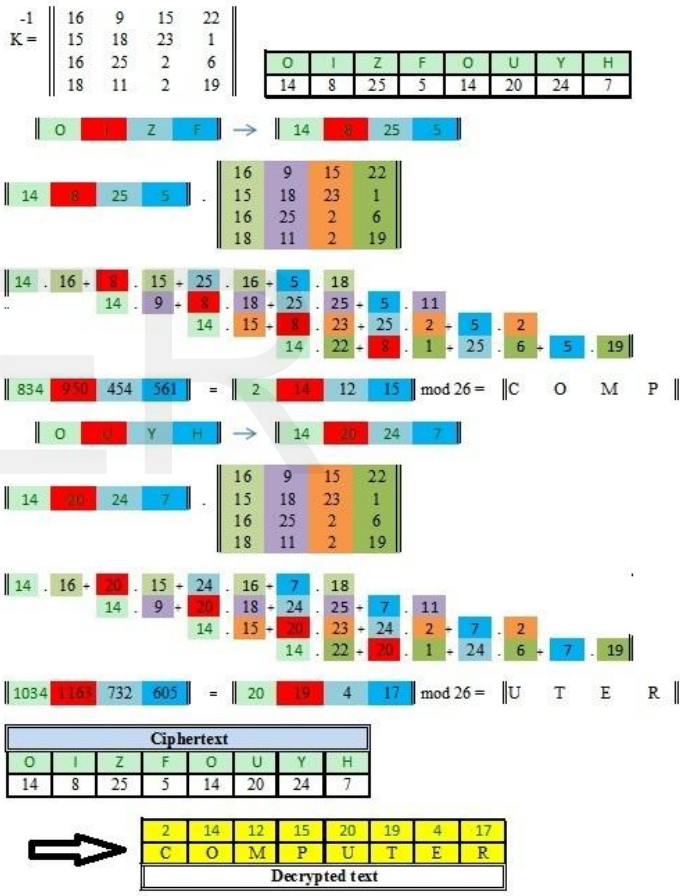


Figure 4: Illustrating the operation of "Decryption" when selected matrix (Option 1) and ciphertext OIZFOUYH

Five valid options are given for choosing of the key matrix K with dimension 4×4. In order to illustrate the reaction for both cases (valid and an invalid matrix)that shown in figure 1. The five selected matrices is calculated and all matrices are valid because determinant satisfied the both conditions (i) all determinant are different from zero (ii) all determinant has no common divisors with number 26. i.e. is not divisible by 2 and 13. In the module "Encryption" the 8 – letter English words are encrypted using the example

matrix and each of the five selected matrices in the module "Matrix" is implemented. The selected 8-letter words in the field of cryptography are COMPUTER, DOCUMENT, HARDWARE, SOFTWARE, SECURITY, ANALYSIS, DISCRETE, APPROACH (Figure 3). In the modulo "Decryption" the 8-letter encrypted English words are decrypted using the example matrix and each of the five selected valid matrices in the module "Matrix" is realized (Figure 4). The results of the operation of both modules are summarized in table 4.

Option	K	K^{-1}
1	$\begin{pmatrix} 1 & 1 & 8 & 7 \\ 2 & 2 & 1 & 0 \\ 25 & 0 & 25 & 22 \\ 24 & 2 & 5 & 13 \end{pmatrix}$ det K=11	$\begin{pmatrix} 10 & 11 & 11 & 10 \\ 11 & 2 & 3 & 19 \\ 10 & 1 & 24 & 20 \\ 8 & 23 & 4 & 25 \end{pmatrix}$
2	$\begin{pmatrix} 4 & 1 & 8 & 7 \\ 2 & 2 & 1 & 0 \\ 25 & 0 & 25 & 22 \\ 24 & 2 & 5 & 13 \end{pmatrix}$ det K=3	$\begin{pmatrix} 2 & 23 & 23 & 2 \\ 23 & 23 & 24 & 5 \\ 2 & 13 & 10 & 12 \\ 12 & 17 & 24 & 3 \end{pmatrix}$
3	$\begin{pmatrix} 4 & 5 & 8 & 7 \\ 2 & 2 & 1 & 0 \\ 25 & 0 & 25 & 22 \\ 24 & 2 & 5 & 13 \end{pmatrix}$ det K=19	$\begin{pmatrix} 14 & 9 & 5 & 8 \\ 5 & 5 & 12 & 9 \\ 14 & 25 & 18 & 18 \\ 6 & 11 & 20 & 13 \end{pmatrix}$
4	$\begin{pmatrix} 4 & 5 & 8 & 7 \\ 2 & 2 & 1 & 0 \\ 25 & 4 & 25 & 22 \\ 24 & 2 & 5 & 13 \end{pmatrix}$ det K=21	$\begin{pmatrix} 4 & 25 & 7 & 16 \\ 7 & 7 & 22 & 23 \\ 4 & 15 & 20 & 0 \\ 18 & 23 & 2 & 7 \end{pmatrix}$
5	$\begin{pmatrix} 1 & 1 & 0 & 7 \\ 2 & 2 & 1 & 0 \\ 25 & 0 & 25 & 22 \\ 24 & 2 & 5 & 13 \end{pmatrix}$ det K=15	$\begin{pmatrix} 16 & 9 & 815 & 22 \\ 15 & 18 & 23 & 1 \\ 16 & 25 & 2 & 6 \\ 18 & 11 & 2 & 19 \end{pmatrix}$

Table 3: Selection of Matrix K and its inverse matrix K^{-1}.

No	Plaintext (Decrypted text)	Ciphertext (Option 1)	Ciphertext (Option 2)	Ciphertext (Option 3)	Ciphertext (Option 4)	Ciphertext (Option 5)
1	COMPUTER	OIPEUOAH	UIPEWOAH	UQPFCQAH	UMPFCGAH	OIZFOUYG
2	DOCUMENT	PTGNVGAT	KTGNFGAT	YFGNFCAT	YNGNFCAT	PTINVGIT
3	HARDWARE	KNCUXEXI	FNCULEXI	FPCULOXI	FFCULEXI	KNYUXEDI
4	SOFTWARE	DGOPXEXI	FGOPLEXI	FAOPLOXI	FUOPLEXI	DGAPXEDI
5	SECURITY	KOMOSDLR	MOMORDLR	MIMORTLR	MQMORRLR	KOYOSDFR
6	ANALYSIS	EWQNQSGG	EWQNKSGG	EWQNKKGG	EWQNKQGG	EWQNQSWG
7	DISCRETE	JXYBYHLR	SXYBXHLR	SJYBXXLR	SDYBXVLR	JXABYHFR
8	APPROACH	HMHFYCPZ	HMHFOCPZ	HMHFOGPZ	HUHFOOPZ	HMHFYCHZ

Table 4: Results when encryption and Decryption of English texts using the application developed.

3Conclusion

This paper describes the classical Hill cipher to encrypt the English texts using 4×4 matrix. The application is developed based on MS EXCEL and the implementing processes of encryption and decryption of English texts using the Hill cipher is presented in details.

REFERENCES

[1]V. Gellert, H. Kestner, Z. Noyber. Mathematical Encyclopedic Dictionary. Translation from German: N. Bozhinov and D. Minev, Sofia, Publishing House "Science and Art", 1983.
[2]https://en.wikipedia.org/wiki/Hill_cipher
[3]http://www.practicalcryptography.com/ciphers/hill-cipher
[4]http://crypto.interactive-maths.com/hill-cipher.html
[5]http://writing.ufl.edu/3254/ExamplesandReadings/InstructionManualExample.doc
[6]http://www.nku.edu/~christensen/092mat483%20hill%20cipher.pdf
[7]M. Eisenberg. Hill Ciphers and Modular Linear Algebra, http://apprendre-en-ligne.net/crypto/hill/Hillciph.pdf
[8]Menezes, A.J., P.C. Van Oorschot, S.A. Van Stone. 1996. Handbook of Applied Cryptography. CRC press
[9]Stalling, W. Cryptography and Network Security. 2005. 4th edition, Prentice Hall.
[10]Adriana Naydenova Borodzhieva, MS Excel-Based Application for Encryption and Decryption of English Texts with the Hill Cipher on the Basis of 3×3-Matrix.

Author Biography

Md. Kamal Hossain Completed B.Sc in Telecommunication and Electronic Engineering from Hajee Mohammad Danesh Science and Technology University, Bangladesh. He is serving as a lecturer in the department of Electronics and Communication Engineering (Former Telecommunication and Electronic Engineering) at Hajee Mohammad Danesh Science and Technology University, Bangladesh. His research interest is Wireless Communication, Network Security.

Md. Ferdous Wahid Completed B.Sc in Electronics and Communication Engineering from Khulna University of Engineering &Technology.He is serving as a lecturer in the department of Electrical and Electronic Engineering at Hajee Mohammad Danesh Science and Technology University, Bangladesh.His research interest isOptical Fiber communication, Artificial Neural Network.