

Text Encryption: Hybrid cryptographic method using Vigenere and Hill Ciphers.

Hamza TOUIL

LISAC, Faculty of Sciences, Dhar-Mahraz (FSDM)

Sidi Mohamed Ben Abdellah University

Fez, Morocco

hamza.touil@usmba.ac.ma

Nabil EL AKKAD

LISAC, Faculty of Sciences, Dhar-Mahraz (FSDM).

Laboratory of Engineering, Systems and Applications (LISA), National School of Applied Sciences (ENSA)

Sidi Mohamed Ben Abdellah University

Fez, Morocco

nabil.elakkad@usmba.ac.ma

Khalid SATORI

LISAC, Faculty of Sciences, Dhar-Mahraz (FSDM)

Sidi Mohamed Ben Abdellah University

Fez, Morocco

khalidsatori@gmail.com

Abstract— *Cryptography is a method of controlling and protecting communications, which has been used exclusively in areas that require confidentiality. Today it is undergoing a considerable evolution, and computer networks require a phase of cryptography as a fundamental mechanism to ensure the privacy of digital information. It was beginning with the first Cesar encryption algorithm or those that appeared just afterwards, such as mono-alphabetic substitution encryption, which has a weakness against statistical attacks. In this paper, we will present a hybridization of Vigenere and Hill encryption that belongs to this family. By exploiting the enhancements, they have already implemented in Hill encryption, to hide the weak point of the Vigenere algorithm, represented in the ease to detect the size of the key to start a statistical attack, as well as the weakness of the algorithm to encrypt two same letters located in the same place in different blocks. This cohesion between these two methods will provide us with a reliable hybrid algorithm, resistant to varying attacks, including statistical attacks.*

Keywords— *Vigenere cipher, Hill Cipher, Text encryption, Statistical attack.*

I. INTRODUCTION

One way to hide information from those who are not authorized to read it is to encrypt it [1-2-3]. Not only on text formats but in different multimedia information sources such as images, videos, reconstruction and 3D recognition through cameras [4,14] To encrypt or decrypt a message, the encryption of information involves two elements, the algorithm, and the key. [15][16] The encryption algorithm consists of a series of mathematical processes that transform plaintext information into unintelligible information. Another round of treatments will, from the encrypted data, restore the plaintext information. The encryption key is information that will enable the encryption algorithm to encrypt the message in such a way that only the holder of the corresponding decryption key will be able to obtain the unencrypted message from the encrypted message. The algorithm is not a secret unless, of course, it is a proprietary algorithm. Most of the algorithms used in civil law are published standards such as DES, 3DES, AES, RSA. The secret element is the key. Thus, for a known algorithm, trying all possible keys to decipher a message, a so-called brute force attack, inevitably leads to the result, the decoded signal. Getting the result in clear text is just a matter of computing power and processing time. [17-18]

Taking, for example, the cryptography of images, which has experienced remarkable growth in recent years. Several researchers are focusing on image encryption as a solution to security problems and have developed several approaches, such as "watermarking" [19-20-21-22-23].

The transfer of confidential documents is no longer a simple task. There is an urgent need to improve the security and effects of the digital image encryption system. In 1949, C. E. Shannon gave in [24-25] the initial interpretation of the secret operation. He believed that in the right secret order, He believed that in a real secret system, encrypted information should be circulated in a public environment, and let the encryption and decryption algorithms do their work with confidence. He also pointed out the importance of protected keys in encryption schemes [26][27].

II. PROPOSED METHOD

This method combines two algorithms belonging to the same family (Substitution), but different categories. Hill belongs to the block cipher type, however vigenere stream cipher, with two different mechanisms. Beginning by introducing our text, the first phase consists of eliminating all spaces and tabulations existing between words or letters, and we will obtain a document in the form of a single homogeneous entity that is composed of several letters. In the second phase, we will apply the "hill" encryption [28-29-30-31]. First of all, each character is coded by a number between 0 and $n - 1$, so that the range $[0, n-1]$ represents the alphabetical numbering, which decreases by one unit after each increment on our case a 4×4 matrix was used. The characters are then grouped in blocks of 4 characters in size, forming a vector $H(X_1, X_2, X_3, X_4)$.

After performing the operation $H * A \text{ Modulo } (26)$ such that A is the key used to encrypt the text, we obtain an encrypted Text C . Continuing to group the letters of our text C in blocks of length 4 in the same way, we choose a key consisting of 4 numbers from 0 to 25: (n_1, n_2, n_3, n_4) . The ciphering consists of performing a Cesar ciphering, whose shift depends on the rank of the letter in the block. Towards the end of this operation, we obtain a new C cipher text, as shown in fig.1.

A. The encryption process

Below is an explanation of the different steps of Hybrid encryption.

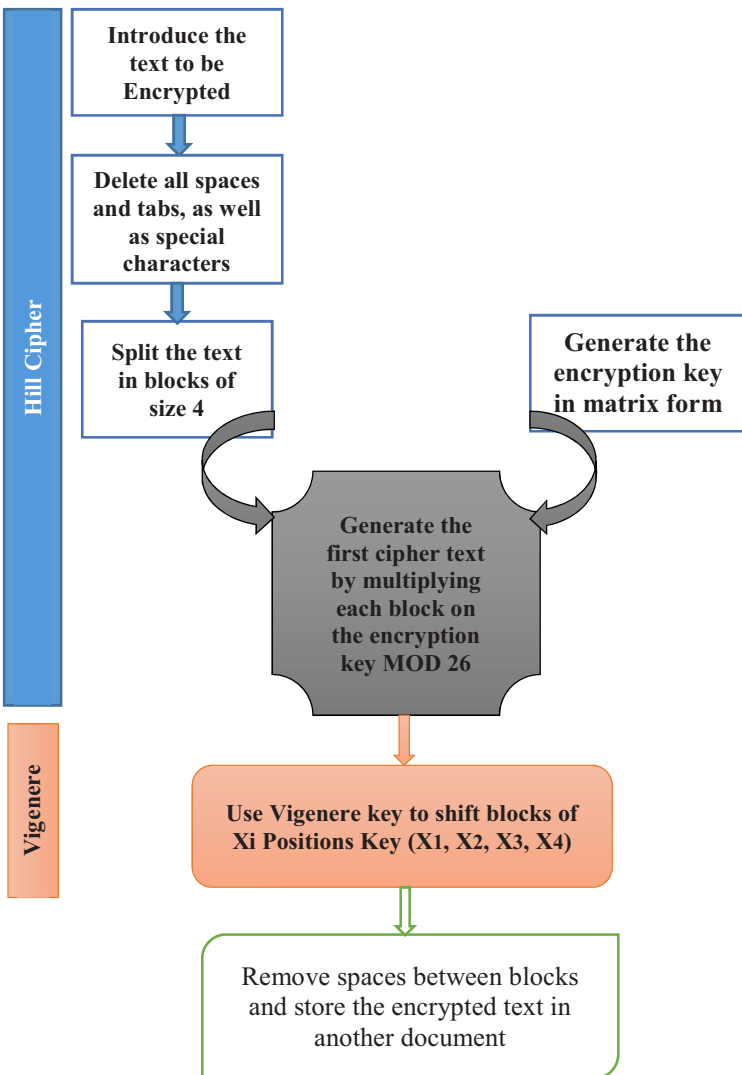


Fig. 1: Hybrid encryption steps.

B. Algorithm Explanation

In hill encryption, the source text is divided into blocks of equal size. The blocks are encrypted one at a time in such a way that each character in the block contributes to the encryption of the other characters in the block. For this reason, Hill belongs to the category called block cipher. The key is a square matrix of size $m \times m$, where m is the size of the block. Its components must, first of all, be positive integers. It must also have an inverse form in Z_{26} . This inverse matrix exists if $(ad-bc) \cdot -1 \pmod{26}$ exists, which is the case when $(ad-bc)$ and 26 are prime between them. We must, therefore, check that $(ad-bc)$ is odd and is not a multiple of 13 [32-33-34-35].

cryptanalysis only for Hill ciphertext is difficult. First, the brute force attack is extremely complex because the matrix key is $m \times m$. Each entry can have one of 26 values. First of

all, that means the size of the key $26^{m \times m}$. However, not all matrices have a multiplicative inversion. Therefore, the range of the keys is still not so huge.

Secondly, Hill ciphers do not store statistics in plain text. It is challenging to analyze the frequency of individual letters of two or three letters. An analysis of the frequency of words of size m could work, but very rarely does the source text have many identical lines of size m . However, the encryption can be attacked using the source text knowledge method if it knows the value of m and knows the "source text/encrypted text" pairs of at least m blocks. The blocks can belong to the same message or different messages but must be different. Two $m \times m$ matrices can be created, P (plain text) and C (ciphertext), in which the corresponding lines represent known single text/ciphertext pairs. Since $C = PK$, the relations $K = CP^{-1}$ can be used to find the key if P is reversible. If P is not reversible, then different sets of m single pairs/ciphertext must be used.

If we do not know the value of m , we can try different values, provided that m is not very large.

We want to encrypt the text with the hybrid algorithm, taking for example the message: HAMZA TOUILOU.

- The first step is to split the message into blocks of 4.

HAMZ ATOU ILOU

- Then we have to index the letters:
H(7)A(0)M(12)Z(25) **A(0)T(19)O(14)U(20)**
I(8)L(11) O(14)U(20).

- Choose a key in the form of a 4×4 matrix that meets the mentioned requirements.

$$\begin{pmatrix} 8 & 6 & 9 & 5 \\ 6 & 9 & 5 & 10 \\ 5 & 8 & 4 & 9 \\ 10 & 6 & 11 & 4 \end{pmatrix}$$

- Applying the operation: Text * Key Modulo (26), we obtain the following ciphered text.

C = DOWQCZYKSBAQ

If we use Vigenere key (3, 1, 3, 2), then for the first block "DOWQ":

- an offset of 3 for D gives G,
- a 1-shift for O Gives P,
- an offset of 3 for W gives Z,
- a 2-shift for Q offers S.

Continuing the operation until the end gives us the final text in figures:

C'=GPZSFABMVCD

The novelty of this hybrid algorithm is that it regroups Streaming and block ciphers. In practice, source blocks are encrypted individually, but they use stream keys to encrypt the entire message block by block. In other words, the encryption is a block when applied to individual blocks, but it also flows encryption when applied to the entire message, considering each block as a unit. Each block uses a different key that has been generated in advance or during the encryption process.

C. Key space, and possible attack.

The primary element is no longer a letter but a block, i.e., a grouping of letters. The encryption function associates to a block of length k, another block of length k, which gives, when mathematicising things [36-37-38] :

$$C_{N1,N2,...,Nk} : \begin{cases} Z/26Z \times Z/26Z \dots \times Z/26Z \longrightarrow Z/26Z \times Z/26Z \dots \times Z/26Z \\ (X_1, X_2, \dots, X_k) \longrightarrow (X_{1+n1}, X_{2+n2}, \dots, X_{k+nk}) \end{cases}$$

Each of the components of this function is Cesar encryption. The decryption function is $C^{-n1}, -n2, \dots, -nk$.

There are 26k possible choices of keys when the blocks are of length k. For blocks of length k = 4, this already gives 456,976, and even if a computer tests all possible combinations without any problem, it is not easy to browse this list to find the plaintext message [39-40-41], that is to say, the understandable one! There is still a weakness of the same order as that encountered in mono-alphabetic encryption: the letter A is not always encrypted by the same letter, but if two letters are located at the same position in two different blocks (such as "ATOU ILOU") then they will be encrypted by the same letter. A possible attack is, therefore, the following: we split our message into several lists, the first letters of each block, the second letters of each block... and we make a statistical attack on each of these groupings. This type of attack is only possible if the size of the blocks is small in front of the length of the text.

Taking example, our initial message: HAMZ ATOU ILOU

If the Vigenere encryption is applied directly with the key (3, 1, 3, 2), the encrypted message comes.

C = KBPBDURWLMRW.

We make a classification of the letters positioned in the same place, as shown in the table below.

Position 1	Position 2	Position 3	Position 4
K	B	P	B
D	U	R	W
L	M	R	W

R and W are necessarily encrypted in the same way, so a statistical attack can break the ciphertext and make it clear.

But if we keep the same text with the same parameters, we add the "Hill" encryption.

Position 1	Position 2	Position 3	Position 4
G	P	Z	S
F	A	B	M
V	C	D	S

Even though both letters indexed in the same position, the encryption is not identical. The thing that makes cracking the text almost impossible. We can deduce that our hybrid method gives more flexibility and can combine the simplicity of the Vigenere algorithm and the strengths of the Hill algorithm.

To see the effectiveness of this method, we will test it on a text with three letters in the same positions.

For example, the initial text: COUCCLOUD

If the Vigenere encryption is applied directly with the key (2, 4, 5, 3), the encrypted message becomes.

C = ESZFESZG

Arrange the letters positioned in the same location, as shown in the table below.

Position 1	Position 2	Position 3	Position 4
E	S	Z	F
E	S	Z	G

Through a statistical attack, a hacker can easily find the original text.

This time applying Hill encryption, using our usual key:

$$\begin{pmatrix} 8 & 6 & 9 & 5 \\ 6 & 9 & 5 & 10 \\ 5 & 8 & 4 & 9 \\ 10 & 6 & 11 & 4 \end{pmatrix}$$

C = EYMUJIVY

If the Vigenere encryption applied with the same key (2, 4, 5, 3), the encrypted message becomes:

C' = GCRXLMAB

Sorting letters positioned on the same location

Position 1	Position 2	Position 3	Position 4
G	C	R	X
L	M	A	B

Even though both letters indexed in the same position, the encryption is not identical. The thing that makes cracking the text almost impossible.

Our method is still useful even if we place identical letters in the same positions because they will be encrypted in the same way. Again, the Hill encryption will create an abstract layer between the plain and encrypted text.

D. The decryption process

Below (Fig.2) is an explanation of the different steps involved in deciphering the text [42][43]. What starts with a recursive shift using the Vigenere vector key.

Let n element of \mathbb{N}^* .

Let A element of $M_n(\mathbb{N})$ such as $\text{PGCD}(\det(A), 26) = 1$.

Let i be the multiplicative inverse of $\det(A)$ modulo 26, and let $B = i \times \text{com}(A)$ [26].

A block of n letters $Y = (Y_1 \dots Y_n)$ of the encrypted message will correspond to the block $X = (X_1 \dots X_n)$ of the original message verifying.

$$X \equiv A^{-1}Y \pmod{26}$$

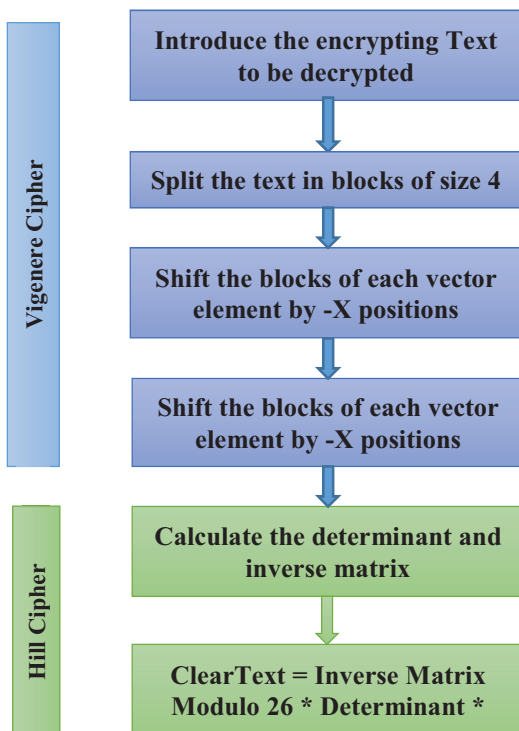


Fig. 2: Hybrid decryption steps.

E. Experimentation

we take a paragraph as entered without special characters

the estimated number of letters he has available in 66 letters; two Hill and vigenere encryption keys will be used, one from Hill and the other from vigenere the table below illustrates the operation

Text Encryption

PLAINTEXT
Text Encryption Hybrid cryptographic method using Vigenere and Hill Ciphers

- Divide the text in blocks of size 4 then complete the last one with blocks by XX

PLAINTEXT
Text Encr ypti onHy brid cryp togr aphi cmet hodu sing Vige nere andH illC iphe rsXX

- Indexed each letter starting with 0 for letter A

PLAINTEXT
T(19)e(4)x(23)t(19) E(4)n(13)c(2)r(17) y(24)p(15)t(19)i(8).....

- For each block does the operation Blocks * Matrix MOD 26.

PLAINTEXT	HILL KEY
T(19)e(4)x(23)t(19) E(4)n(13)c(2)r(17) y(24)p(15)t(19)i(8)	$\begin{pmatrix} 9 & 4 & 3 & 3 \\ 2 & 23 & 5 & 11 \\ 6 & 3 & 32 & 43 \\ 43 & 43 & 12 & 23 \end{pmatrix}$ MOD 26

- When it's finished, we'll get the first encrypted text.

PLAINTEXT	CIPHERED TEXT
Text Encryption Hybrid cryptographic method using Vigenere and Hill Ciphers	BMPMPKACTEXJLCB DGYADVEUUKFNKB APDFPFVGZAVRNAI ROUHONANEBCZIZ HJYVVLQDO

- Afterwards a strong encryption is applied with the key (3 20 22 25), shifting each letter by the number indicated on the key (3 20 22 25), and we have our encrypted text.

CIPHERED TEXT WITH HILL	VIGENERE KEY	CIPHERTEXT
BMPM TEXJ GYAD KFNK FPFV RNAI ONAN ZIZH LQDO	PKAC LCBD VEUU BAPD GZAV ROUH EBUC JYVV	3 20 22 25 EGLLSEWBW YTIOWXCJSW CYYQTNZJJEU LCIJBUIJTUU HWHUIQGRH WMHVQBCCV GMSRUOKZN

Decryption of ciphertext

- Apply a back off shifting by the vigenere key

CIPHERTEXT	VIGENERE KEY	FIRST PLAINTEXT
EGLLSEWBW YTIOWXCJSW CYYQTNZJJE ULCIJBUIJT UHHWHUIQG RHHMHVQB CCVGMSRUO KZN	{-3 -20 -22 -25}	BMPMPKAC TEXJLCBD GYADVEUU KFNKBAPD FPFVGZAV RNAIROUH ONANEUC ZIZHJYVV LQDO

- For each block does the operation Inverse Matrix MOD 26 * Determinant of matrix

CIPHERTEXT	HILL KEY	CIPHERTEXT
BMPMPKAC TEXJLCBD GYADVEUU KFNKBAPD FPFVGZAV RNAIROUH ONANEUC ZIZHJYVV LQDO	Determinant * Inverse Key MOD 26	Text Encrypti on Hybrid cryptogra phic method using Vigenere and Hill Ciphers

III. CONCLUSION

We have implemented a hybrid text encryption approach based on the simplicity of Vigenere encryption and Hill's

encryption complexity. We have seen that even if the number of combinations to tested with Vigenere encryption is quite large (estimated in 456,976), there is still a weakness that can be exploited on statistical attacks. Still, with Hill encryption, this weakness is no longer present. Therefore, our approach is strong enough to withstand any bridge attack, including statistical attacks. due to the abstraction mode used in the method

REFERENCES

- [1] Hacigümüş, H., Iyer, B., Li, C., & Mehrotra, S. (2002, June). Executing SQL over encrypted data in the database-service-provider model. In Proceedings of the 2002 ACM SIGMOD international conference on Management of data (pp. 216-227). ACM.
- [2] Wang, Zheng-Fei, et al. "Fast query over encrypted character data in database." International Conference on Computational and Information Science. Springer, Berlin, Heidelberg, 2004.
- [3] Overbey, Jeffrey, William Traves, and Jerzy Wojdylo. "On the keyspace of the Hill cipher." Cryptologia 29.1 (2005): 59-72.
- [4] El akkad N, El Hazzat S, Saaidi A and Satori K (2016). Reconstruction of 3D Scenes by Camera Self-Calibration and Using Genetic Algorithms. 3D Research, 6 (7): 1-17.
- [5] El Hazzat, S., Merras, M., El Akkad, N., Saaidi, A., Satori, K. (2018). 3D reconstruction system based on incremental structure from motion using a camera with varying parameters. Visual Computer. 34(10), pp. 1443-1460.
- [6] M. Merras, A. Saaidi, N. El akkad and K. Satori. Multi-view 3D reconstruction and modeling of the unknown 3D scenes using genetic algorithms. Soft computing (Springer). 22(19), pp. 6271-6289, 2017.
- [7] S. El hazzat, M. Merras, N. El akkad, A. Saaidi and K. Satori. Enhancement of sparse 3D reconstruction using a modified match propagation based on particle swarm optimization. Multimedia Tools and Applications (Springer), 78, pages14251–14276, 2019.
- [8] B. Boudine, S. Kramm, N. EL Akkad, A. Saaidi and K. Satori. A flexible technique based on fundamental matrix for camera self-calibration with variable intrinsic parameters from two view. Journal of Visual Communication and Image Representation (Elsevier). Vol 39, pp. 40–50, 2016.
- [9] N. El akkad, M. Merras, A. Saaidi and K. Satori. Camera Self-Calibration with Varying Intrinsic Parameters by an Unknown Three-Dimensional Scene. The Visual Computer (Springer). Vol. 30, No. 5, pp. 519-530, 2014.
- [10] M. Merras, N. El akkad, A. Saaidi A. G. Nazih and K. Satori. Camera Self-calibration with varying parameters by an unknown tree dimensional scene using the improved genetic algorithm. 3D research (Springer). Vol. 6, No. 1, pp. 1-14, 2015.
- [11] El Akkad N, Saaidi A, Satori K (2012). Self-calibration based on a circle of the cameras having the varying intrinsic parameters. In: Proceedings of IEEE International Conference on Multimedia Computing and Systems, pp 161–166
- [12] El akkad N, Merras M, Saaidi A and Satori K. Camera self-Calibration with Varying Parameters from Two views. Wseas Transactions on Information Science and Application, Vol. 10, No. 11, pp. 356-367, 2013.
- [13] El akkad N, Merras M, Saaidi A and Satori K (2013). Robust Method For Self-Calibration Of Cameras Having The Varying Intrinsic Parameters. Journal Of Theoretical And Applied Information Technology 50 (1): 57-67
- [14] El akkad N, Merras M, Saaidi A and Satori K Camera self-calibration having the varying parameters and based on homography of the plane at infinity. Multimedia Tools and Applications. DOI: 10.1007/s11042-017-5012-3, 2017
- [15] Bouganin, Luc, and Yanli Guo. "Database encryption." Encyclopedia of Cryptography and Security. Springer US, 2011. 307-312.

- [16] Garfunkel, S.L; "Public Key Cryptography", Computer, IEEE, Volume: 29, Issue: 6, June 1996.
- [17] W.Diffie; M.E.Hell man, "New Directions in Cryptography" IEEE Transactions Information Theory, Nov 2000, pp 644-654
- [18] Sion, Radu, Mikhail Atallah, and Sunil Prabhakar. "Watermarking relational databases." (2002).
- [19] Acharya, Bibhudendra, et al. "Image encryption using advanced hill cipher algorithm." International Journal of Recent Trends in Engineering 1.1 (2009): 663-667.
- [20] Ali Mansouri¹ · Xingyuan Wang^{1,2} ; Image encryption using shuffled Arnold map and multiple values manipulations; Springer-Verlag GmbH Germany, part of Springer Nature 2020
- [21] Hofmann, G.R.: The modelling of images for communication in multimedia environments and the evolution from the image signal to the image document. Vis. Comput. 9(6), 303–317 (1993).
- [22] Lin, C.-H., Chao, M.-W., Liang, C.-Y., Lee, T.-Y.: A novel semi-blind-and-semi-reversible robust watermarking scheme for 3D polygonal models. Vis. Comput. 26(6), 1101–1111 (2010).
- [23] Tu, S.-C., Tai, W.-K., Isenburg, M., Chang, C.-C.: An improved data hiding approach for polygon meshes. Vis. Comput. 26(9), 1177–1181 (2010).
- [24] Li, G., Wang, L.: Double chaotic image encryption algorithm based on optimal sequence solution and fractional transform. Vis. Comput. 35(9), 1267–1277 (2019).
- [25] HONGFENG GUO^{1,2}, XIN ZHANG¹, XINYAO ZHAO¹, HANG YU¹, AND LI ZHANG^{2,3}; Quadratic Function Chaotic System and Its Application on Digital Image Encryption; 2020
- [26] C. E. Shannon, "Communication theory of secrecy Systems," Bell Syst. Tech. J., vol. 28, no. 4, pp. 656–715, Oct. 1949.
- [27] R. Matthews, "On the derivation of a 'Chaotic' encryption algorithm," Cryptologia, vol. 13, no. 1, pp. 29–42, 1989.
- [28] J. Fridrich, "Symmetric ciphers based on two-dimensional chaotic maps," Int. J. Bifurcation Chaos, vol. 8, no. 6, pp. 1259–1284, Jun. 1998.
- [29] Es-Sabry, M., ElAkkad, N., Merras, M., Saaïdi, A., Satori, K. : A new image encryption algorithm using random numbers generation of two matrices and bit-shift operators. (2019).
- [30] Es-sabry, M., El Akkad, N., Merras, M., Saaïdi, A., Satori, K.: Grayscale image encryption using shift bits operations. In: International Conference on Intelligent Systems and Computer Vision, ISCV (2018)
- [31] M. Es-sabry, N. El akkad , M. Merras, A.Saaïdi and K.Satori. A Novel Text Encryption Algorithm Based on the Two-square Cipher and Caesar Cipher. The 3rd International Conference on Big Data, Cloud and Applications, BDCA (2018).
- [32] F. Elazaby, N. El Akkad and S. Kabbaj. A new encryption approach based on four squares and Zigzag. The 1st international conference on Embedded Systems and Artificial Intelligence, ESAI (2019).
- [33] M. Es-sabry, N. El akkad , M. Merras, A.Saaïdi and K.Satori. A New Color Image Encryption Using Random Numbers Generation And Linear Functions. The 1st international conference on Embedded Systems and Artificial Intelligence, ESAI (2019).
- [34] Eastaway R. et Wyndham J., Why do buses always come in threes? Flammarion, 2001, pp. 95-107
- [35] Hill Lester S., « Cryptography in an Algebraic Alphabet », American Mathematical Monthly, 36, 1929, pp. 306-312
- [36] Lewand Robert Edward, Cryptological Mathematics, published by The Mathematical Association of America, 2000.
- [37] Wilson, P., Garcia, M.A Modified Version of the Vigenère Algorithm (2006) Int. J. Comput. Sci. Netw. Secur, 6 (3), pp. 140-143.
- [38] Ravindra, P., Kallam, B., Kumar, S.U., Vinaya, A., Shravan, V. A Contemporary Polyalphabetic Cipher using Comprehensive Vigenere Table (2011) World Comput. Sci. Inf. Technol. J, 1 (4), pp. 167-171..
- [39] Kartha, R.S., Paul, V.A New Cryptosystem Based On Polyalphabetic Substitution Scheme with Multiple Number Of Cipher (2014) 6Th IRF International Conference, pp. 40-44.
- [40] Soofi, A.A., Riaz, I., Rasheed, U. An Enhanced Vigenere Cipher For Data Security (2016) Int. J. Sci. Technol. Res, 5 (3), pp. 141-145
- [41] Kester, Q.-A. A Hybrid Cryptosystem Based on Vigenère Cipher and Columnar Transposition Cipher (2013) Int. J. Adv. Technol. Eng. Res, 3 (1), pp. 141-147
- [42] Kester, Q.-A. A cryptosystem based on Vigenère cipher with varying key (2012) Int. J. Adv. Res. Comput. Eng. Technol, 1 (10), pp. 108-113
- [43] Md. Saiful Islam Chowdhury, Shoyeb Al Mamun Shohag and Md. Hasan Sahid, "A Secured Message Transaction Approach by Dynamic Hill Cipher Generation and Digest Concatenation", International Journal of Computer Applications, Vol. 23, June 2011.