iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** IV **Month of publication:** April 2022

DOI: https://doi.org/10.22214/ijraset.2022.41970

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue IV Apr 2022- Available at www.ijraset.com

Modified Hill Cipher Algorithm using Myszkowski Transposition to address Known-Plaintext attack

Raymond G. Barrieta[1], Aleksis S. Canlas[2], Dan Michael A. Cortez[3], Khatalyn E. Mata[4]
[1, 2, 3, 4] *College of Engineering and Technology, Pamantasan ng Lungsod ng Maynila*

Abstract: ***Hill Cipher algorithm is a Polygraphic cipher known to be vulnerable to the Known-Plaintext attack. It is a type of attack wherein the attacker has access to a pair of Plaintext and Ciphertext. Possible leaked information can be used to reveal critical data which can include the private key that is used in the encryption of the plaintext. In this study, the researchers propose an improved scheme by incorporating Myszkowski transposition in the algorithm to address the specified problem. Runtime Test and Statistical test such as the Avalanche effect were used to evaluate the performance of the algorithm. The test yielded a 25.70% faster runtime based on previous modifications and a 52.78% Avalanche Effect. Randomness test such as the Monobit test, Runs Test and Longest Run of Ones in a Block test were applied to evaluate the ciphertext produced. The criterion for the Randomness Test requires its p-value to be greater than 0.01, the evaluation showed a 0.249, 0.167 and 0.368 p-value, respectively. The proposed scheme proved that the algorithm is cryptographically secure and efficient as all tests yielded favorable results.***

Keywords: Cryptography, Hill Cipher, Encryption, Decryption, Key Generation

I. INTRODUCTION

The modern digital era ushered in the revolution of technology across the world. In line with this is the ever-increasing amount of information that is generated across private information channels. A primary consideration in the development of the channels is the need to secure the transportation and storage of private information to prevent the critical risk and threat of having a compromised system and having vital information ending up in the wrong hands. Cryptography is the field of study in science that emerged to ensure the protection of private information from unauthorized access, maintaining data integrity, authentication, and other tasks that concerns the protection and management of data in systems [1].

Various cryptographic algorithm was developed at a certain point of time, one of which is the Hill Cipher algorithm, developed by Lester S. Hill in 1929. The cipher relies on matrix and modulo operations to produce the ciphertext [2]. The formula,

$$C = KM \bmod R \quad (1)$$

where C is the matrix for ciphertext; K is the matrix key for encryption; M is a matrix composed of the plaintext and R is the range of possible value for key (in this research, R = 256). Hill cipher has a high performance, high speed, and simple structure, but it succumbs to known plaintext attack and the complexity of decrypting the ciphertext when the key used in encryption is not invertible [3]. The cipher is vulnerable to known-plaintext attack where, an adversary gains access to a pair of plaintexts and ciphertext and use it to derive the key in the algorithm and consequently compromise the communication between different parties. Another issue that occurs, is the non-invertible matrix that makes decryption difficult. The previously stated problems have become the major point of consideration for various studies that modified or enhanced the algorithm [4]. This study, with the same consideration in mind, aims to address the vulnerability of Hill Cipher algorithm to Known Plaintext attacks by utilizing the Myszkowski Transposition Method to improve its confusion and diffusion properties.

II. REVIEW OF LITERATURE

The Hill Cipher Algorithm is categorized as a Polygraphic Cipher. The technique works on a set of letters of equal parts partitioned accordingly based on the order of the key matrix. The letters are encrypted using mathematical operations, the concept of matrix multiplication. The mathematical nature of the cipher makes the encryption process exact and simple [5]. A drawback that occurs due to the linear nature of the cipher is its vulnerability to known plaintext attacks [6]. To resolve this problem, various methods and studies were conducted to strengthen the safety of ciphertexts and reduce the risk of an attack in the key matrix and encryption method of the plaintext [7]. The study of Jie Liew and Nguyen [6] modified the original algorithm by generating a skew-symmetric key matrix with random integer entries. After the key generation, algebraic operations were initiated from the matrix to form an orthogonal matrix. Additionally, they also embedded a random sequence to the product matrix and the plaintext using a fixed seed number.

Paragas, Sison, and Medina [1] proposed a new simplified method by implementing a new key generation scheme that addresses the non-invertible matrix problem that limits the cipher and showed that the proposed method offers a significant improvement as compared to the original and modified algorithm. The authors also employed additional logical XOR and shift operations to the blocks of characters in the encryption process paired with Radix64 for the encoding and decoding of the data. In another study of Paragas, Sison, and Medina [7], a method using Ciphertext Block Chaining and Hexadecimal S-box was used to resolve the problems introduced in the previous studies. Qazi, Khan, and Agha [8], proposed a modification by making use of transposition, substitution, and left-right shift methods to the vectors before proceeding with the encryption and decryption process of the cipher. In another study, Munzhelele and Chibaya [9] argued that the security of the cipher can be increased by making use of high order matrix keys. Their study showed an improvement in the security of the conventional cipher and faster encryption speed due to the larger blocks of data being encrypted. Emad, Mohammed, and Norrozila [10] enhanced the cipher by incorporating a pseudo-random number generator with a seed number that is reset programmatically for the key generation and repeating the encryption process for N time. The study kept the simplicity of the method by changing the multiplication step into summation during the encryption process. The paper showed a technique to resolve the linear relationship between the plaintext and the ciphertext.

III. PROPOSED METHOD

To address the vulnerability of the Hill Cipher Algorithm to known-plaintext attack, a modified scheme is proposed. The proposed scheme is a block cipher made up of three phases namely, key generation, encryption, and decryption. The input is a plaintext of arbitrary length within the accepted values of the lookup table. A detailed discussion of each step is shown in the next section.

A. Key Generation

In the key generation process, the proposed method is to use a random number generator to generate a 3x3 matrix with values ranging from 0 to 255 and then, transform the matrix to produce the keys to be used in the Myszkowski Transposition part of the encryption process. The step-by-step process is as follows,

1) Generate a 3x3 matrix within values $0 - 255$:
2) if matrix is invertible:
3) for i = 0 to order of matrix:
4) myszkowski[i] ← random subset from matrix
5) myszkowski[i] ← partition into single digits
6) else:
7) repeat 1
8) end

Following the proposed algorithm, the process first generates a random 3x3 matrix to be used as a key both for the Myszkowski transposition and Hill Cipher process.

$$\text{3x3 Matrix} = \begin{vmatrix} 180 & 75 & 111 \\ 15 & 98 & 133 \\ 77 & 72 & 64 \end{vmatrix}$$

After the generation of an invertible matrix, a two-dimensional array of length 3 is created. Each partition in the array is transformed to contain a single number for each index. These 3 partitions are then used as keys for the 3 passes of Myszkowski transposition.

Partition 1 (K2) = [180, 75, 111] = [1, 8, 0, 7, 5, 1, 1, 1]
Partition 2 (K3) = [15, 98, 133] = [1, 5, 9, 8, 1, 3, 3]
Partition 3 (K4) = [77, 72, 64] = [7, 7, 7, 2, 6, 4]

The length of a partition is dependent on the number of digits in the key space of the cipher, in this case 256, since each partition has 3 random number from the hill cipher then it can be represented as $3 \leq P_i \leq 9$, where P_i is the length of a partition. The key partition in this scheme provides a layer of security that aims to resolve the vulnerability of cipher which occurred due to linear relationship of the plaintext and ciphertext. After generating the needed keys, the encryption process is performed. After generating the needed keys, the encryption process is then performed.

B. Encryption

The proposed encryption process for the modified Hill Cipher algorithm is a 4-step process. The proposed Myszkowski Transposition Method is applied by arranging the input into rows of same length with the key and combined by reading the input column-wise with respect to the key, if the final block is not filled completely, padding is applied. For key with repeating values, columns are read from left to right. The scheme will take a plaintext (P) as an input and would go through a round of the transposition method using key (K_2) to get the 1st Modified Plaintext (MP_1). A 2nd round of Myszkowski is employed using another unique key (K_3) to arrive at the final Modified Plaintext (MP_2).

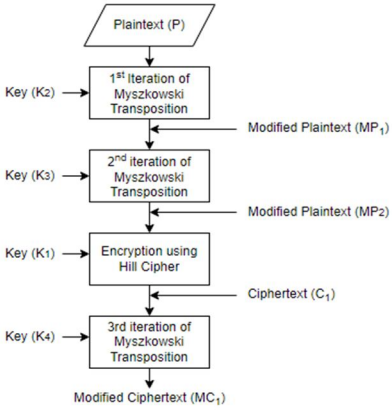


Fig. 1. Proposed Modified Encryption Process

The encryption process of Hill Cipher is then applied using an invertible key (K_1). MP_2 is divided into blocks of letters of equal length and is operated on using a modified formula,

$$C = K1MP_2 mod 256 \quad (2)$$

Where C is the ciphertext, K is the invertible key matrix and MP_2 is the modified plaintext. A final round of Myszkowski Transposition is finally applied to further increase the security of the scheme and to resolve the linear relationship between the plaintext and ciphertext.

A simulation of the encryption process based on the proposed scheme is performed using the previously generated keys,

Plaintext = CRYPTOCURRENCY

1) Myszkowski transposition first iteration:

K_2 = [1, 8, 0, 7, 5, 1, 1, 1]
MP_1 = YECOCURYXXTCPNRR

2) Myszkowski transposition second iteration:

K_3 = [1, 5, 9, 8, 1, 3, 3]
MP_2 = YCYCRXURPNXXEXROTXCXX

3) Hill Cipher Encryption:

K*1* = Generated 3x3 Matrix
C1 = ¦Áf¢¥'b+µØê*-÷ãp€EüQ¥

4) Myszkowski transposition third iteration:

K4 = [7, 7, 7, 2, 6, 4]
MC*1* = ¢ØpX'*EX¥ê€X¦Áfb+µ-÷ãüQ¥

Final Encryption Result: ¢ØpX'*EX¥ê€X¦Áfb+µ-÷ãüQ¥

C. Decryption

The decryption process for the proposed scheme is performed by reversing the order of the encryption step. As shown in Figure 2, the decryption is accomplished by taking MC_1 and K_4 and feeding it to the decryption procedure of Myszkowski Method. The same procedure is followed in the 2nd and 1st iteration. In the decryption for Hill Cipher, the inverse of K_1 is used for the modulo matrix multiplication with the Ciphertext using the modified formula,

$$P = CK_1^{-1} mod 256 \quad (3)$$

Where P is the plaintext, C is the Ciphertext, and K^{-1} is the inverse of the matrix.

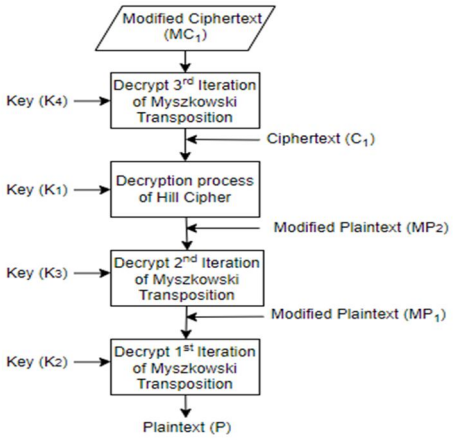


Fig. 2.Proposed Modified Decryption Process

Fig. 3.

A simulation of the Decryption process is performed as shown, getting the inverse of the generated 3x3 matrix is necessary as it will be used as key for the Decryption process of Hill Cipher. But to get the matrix inverse, we must first find for the Determinant.

Determinant value = 9

3x3 Matrix inverse = $\begin{vmatrix} 146 & 54 & 16 \\ 127 & 167 & 105 \\ 9 & 5 & 117 \end{vmatrix}$

The same keys used for the Myszkowski transposition during the encryption process would still be used for its decryption process.
Modified Ciphertext = *¢ØpX'*EX¥ê€X¦Áfb+µ-÷ãüQ¥*

1) Decryption of third iteration of Myszkowski transposition:

K4 = [7, 7, 7, 2, 6, 4]
C1 = ¦Áf¢¥'b+µØê*-÷ãp€EüQ¥

2) Hill Cipher Decryption:

K*1*$^{-1}$ = 3x3 Matrix inverse
MP*2* = YCYCRXURPNXXEXROTXCXX

3) Decryption of second iteration of Myszkowski transposition:

K*3* = [1, 5, 9, 8, 1, 3, 3]
MP*1* = YECOCURYXXTCPNRR

4) Decryption of first iteration of Myszkowski transposition:

K*2* = [1, 8, 0, 7, 5, 1, 1, 1]
P = CRYPTOCURRENCY

Final Decryption Result: CRYPTOCURRENCY

IV. PERFORMANCE AND SECURITY ANALYSIS

A. Known-Plaintext Attack

In this study, the cipher is modified by implementing a Transposition technique to mitigate the chances of having the scheme subjected to the cryptographic attack. Modified partitions of the hill cipher key are utilized in the transposition technique and applied to the encryption and decryption process of the classic cipher to improve its confusion properties.

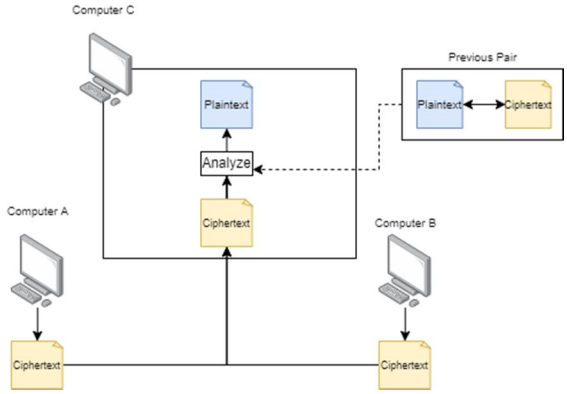


Fig. 4.Known-Plaintext Attack

The figure shows the flow of the cryptographic attack. In the proposed cipher, known-plaintext attack using previous pairs would be difficult, as compared to the classic cipher, because the plaintext and ciphertext are put through multiple rounds of the incorporated transposition technique, thus making the analysis of previous pairs more complex. The discussed vulnerability, which initially occurred due to the linear relationship of the plaintext input and the ciphertext output and its dependence to matrix multiplication is resolved using the presented modified encryption and decryption process in this paper.

B. Randomness Test

Randomness of a cipher is an important aspect that helps an algorithm become more random, secure from attacks and be unpredictable. It is a trait that must be considered when developing a cryptographically secure algorithm. National Institute of Standards and Technology Statistical Test Suite or NIST-STS is a set of tests used to verify such randomness, as well as evaluate the security provided by Cryptographic algorithms [14]. To be considered as truly random, and therefore a successful test, the computed P-value of a specific byte sequence must be greater than or equal to 0.01[16]. The first test, the Monobit Test which checks the proportion of “0” and “1” is used. This test must be carried out in priority as other test cannot be carried out if it does not pass the set P-value [15]. The Runs Test was then implemented to check the oscillation of values between sub blocks are either too fast or too slow and finally the Longest Run of Ones in a Block was applied to verify the longest consecutive sub sequence in the input data. The table below shows the result of the randomness test.

TABLE I. RANDOMNESS TEST

	N Total	P-Value	Conclusion
Monobit Test	288	0.249	PASSED
Runs Test	288	0.167	PASSED
Longest Run of Ones in a Block	288	0.368	PASSED
P-value is not less than 0.01. the sequence is random.			

The results shows that the computed P-value for the 288-byte sequence of the given Ciphertext is approximately 0.249 which is greater than 0.01. Similarly, the result for Runs Test yielded a 0.167 P-value which is also greater. Finally, the Longest Run of Ones in a Block also passed as it yielded 0.368 thus, we can say that the produced ciphers are truly random.

C. Runtime Test

A runtime testing in milliseconds is performed to measure how faster is a program, encryption in this case, compared to its other alternatives. A lower runtime speed means a faster execution for the said program, thus yielding to a much efficient usage [18,21].

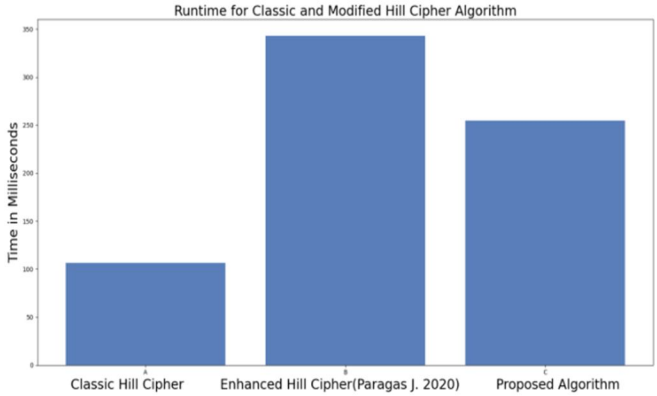


Fig. 5.Encryption time Test (Column based)

Multiple blocks were tested using the encryption process of three Hill Cipher methods. Table II and Figure I shows that the Classic Hill Cipher yielded the fastest result with an average of 106.27 millisecond runtime as it does not have any modifications to its algorithm. A previous study that tackled the same problem with known-plaintext attack showed a runtime of 343.09 millisecond runtime while the proposed scheme showed 254.90 millisecond runtime. The proposed algorithm provides a 25.70% faster execution as compared to a previously modified Hill Cipher algorithm.

TABLE II. AVERAGE RUNTIME ENCRYPTION TEST

	Classic Hill Cipher	Paragas J., 2020 [20]	Proposed Algorithm
Average Runtime in Milliseconds	106.27 ms	343.09 ms	254.90 ms

D. Avalanche Effect

Performing a Cryptanalysis is critically necessary to test the encryption strength of a Cryptographic algorithm. A desirable property of any Cryptographic algorithm is that a small change in either the plaintext or the key should produce a significant change in the ciphertext [11, 12]. This property is known as the Avalanche effect (AE), and this can be calculated using the equation below:

$$AE = (No.of\ flipped\ bits/\ Total\ bits)x100 \quad (4)$$

Original Plaintext: cryptocurrency is the future
Original Ciphertext: nÇÌi\<±YËÚU¬Xëau~±ÙÛI+P*w2´9úÓØþ`¸)ð
Original Ciphertext in Binary form:
01101110 11000111 11001100 01101001 01011100 00111100 10110001 01011001 11001011 11011010 01010101 10101100
01011000 11101011 01100001 01110101 01111110 10110001 11011001 11011011 01001001 00101011 01010000 00101010
01110111 00110010 10110100 00111001 11111010 11010011 11011000 11111110 01100000 10111000 00101001 11110000

Altered Plaintext: kryptocurrency is the future
Altered Ciphertext: [tÓ%üð¹÷Wè ×½tãI!;Öª EA:}£=m")§&C¯À0d
Altered Ciphertext in Binary form:
01011011 01110100 11010011 00100101 11111100 11110000 10111001 11110111 01010111 11101000 10100000 11010111
10111101 01110100 11100011 01001001 00100001 00111011 11010110 10101010 01000101 01000001 00111010 01111101
10100011 00111101 01101101 00100010 00101001 10100111 00100110 01000011 10101111 11000000 00110000 01100100

Avalanche effect = *(152/288)100 = 52.78%*

TABLE III. AVALANCHE EFFECT SCORE COMPARISON

Classic Hill Cipher	Mani, K. et. al, 2017 [13]	Paragas J., 2020 [20]	Proposed Algorithm
18.72%	32.41%	52.45%	52.78%

The proposed algorithm showed a higher percentage on Avalanche effect compared to the classical and previously modified Hill Cipher algorithms.

V. CONCLUSION AND RECOMMENDATION

The efficiency and performance of the proposed algorithmic method is evaluated using the Randomness Test which yielded a p-value greater 0.01 for the multiple tests conducted, thus displaying an improvement to the security it offers. Runtime Test in Milliseconds demonstrates a 25.70% faster encryption time based on previous studies. Finally, the cipher provides a 52.78% Avalanche Effect proving that a small change in the input results in a significant change in the output. The results show a significant improvement in the different tests conducted proving that the new scheme is cryptographically secure. Future works could enhance the security of the cipher by using high order matrix keys, introduction of different transposition or substitution techniques and a larger key space to accommodate a wider array of characters. Cryptanalysis of the proposed scheme can also be done by carrying out different tests to analyze the performance of the cipher against other cryptographic attacks.

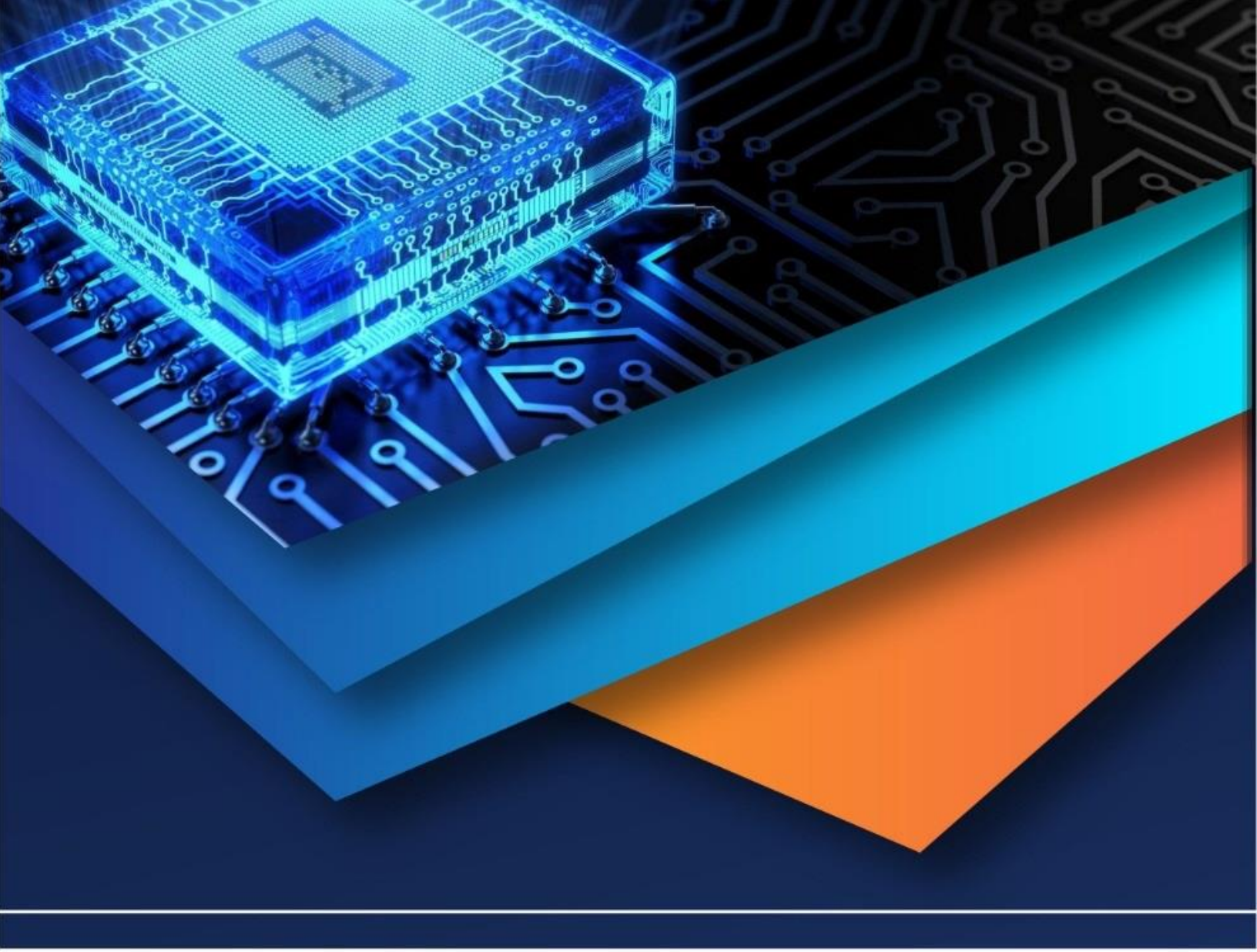
VI. ACKNOWLEDGMENTS

The researchers would like to express their deepest appreciation for the support and guidance of our Advisor, Dr. Dan Michael Cortez. To our professors, for encouraging us in this journey. Finally, the researchers are eternally grateful to God for bestowing us with the wisdom and strength to accomplish this endeavor.

REFERENCES

[1] Paragas, J. R., Sison, A. M., & Medina, R. P. (2019). Hill Cipher Modification: A Simplified Approach. 2019 IEEE 11th International Conference on Communication Software and Networks (ICCSN). doi:10.1109/iccsn.2019.8905360
[2] Tuti Alawiyah et al (2020). Journal of Physics: Conference Series Generation of Rectangular Matrix Key for Hill Cipher Algorithm Using Playfair Cipher: Conf. Ser. 1641 012094. doi.org/10.1088/1742-6596/1641/1/012094
[3] Stallings, W. (2017). Cryptography and Network Security Principles and Practice (7th Edition). Pearson.
[4] Munzhelele, R., & Chibaya, C. (2020). Generation of Invertible High Order Matrix Keys for the Hill Cipher. (2020) 2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC). doi:10.1109/imitec50163.2020.9334
[5] Mohan, M., Kavithadevi, M. K., & Jeevan Prakash, V. (2016). Improved Classical Cipher for Healthcare Applications. Procedia Computer Science, 93, 742–750. doi: 10.1016/j.procs.2016.07.285
[6] Jie Liew K. & Nguyen V. T. (2020). Hill Cipher Key Generation Using Skew-symmetric the 7th International Cryptology and Information Security Conference 2020.
[7] Paragas, J. R., Sison, A. M., & Medina, R. P. (2019). An Improved Hill Cipher Algorithm using CBC and Hexadecimal S-Box. 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE). doi:10.1109/ecice47484.2019.89427
[8] Qazi, Farheen & Khan, Fozia & Agha, dur-e-shawar & Rehman, Saqib. (2019). Modification in Hill Cipher for Cryptographic Application. 240-257. 10.17993/3ctecno.2019.specialissue2.240-257.
[9] Munzhelele, Rotondwa & Chibaya, Colin. (2020). Generation of Invertible High Order Matrix Keys for the Hill Cipher. 1-4. 10.1109/IMITEC50163.2020.9334140.
[10] Emad Taha Khalaf, Muamer N. Mohammed, and Norrozila Sulaiman. 2016. Iris Template Protection based on Enhanced Hill Cipher. In Proceedings of the 2016 International Conference on Communication and Information Systems (ICCIS '16). Association for Computing Machinery, New York, NY, USA, 53–57. doi: https://doi.org/10.1145/3023924.3023938
[11] Mandal, A., & Tiwari, M. (2012). Analysis of Avalanche Effect in Plaintext of DES using Binary Codes. International Journal of Emerging Trends and Technology in Computer Science (IJETTCS), 1(3), 166–177. Retrieved from http://www.ijettcs.org/Volume1Issue3/IJETTCS-2012-10-25-097.pdf.
[12] Dewangan, C., & Agrawal, S. (2012). A Novel Approach to Improve Avalanche Effect of AES Algorithm. Journal of Advanced Research in Computer, 1(8). Retrieved from http://ijarcet.org/wp-content/uploads/IJARCET-VOL-1-ISSUE-8-248-252.pdf
[13] M. Viswambari, & Mani, K. (2017). Generation of Key Matrix for Hill Cipher using Magic Rectangle. Proceedings - 2nd World Congress on Computing and Communication Technologies, WCCCT 2017, 10(5), 51–54. https://doi.org/10.1109/WCCCT.2016.22
[14] Rukhin, A., Soto, J., & Nechvatal, J. (2010). A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications Special Publication800-22R1a. (April)

[15] Mengdi, Z., Xiaojuan, Z., Yayun, Z., & Siwei, M. (2021). Overview of Randomness Test on Cryptographic Algorithms. Journal of Physics: Conference Series, 1861(1), 012009. doi:10.1088/1742-6596/1861/1/012009
[16] Cortez, D.M., Sison, A., Medina, R. (2020). Cryptographic Randomness Test of the Modified Hashing Function of SHA256 to Address Length Extension Attack. In Proceedings of the 2020 8th International Conference on Communications and Broadband Networking (ICCBN '20). Association for Computing Machinery, New York, NY, USA, 24–28. doi.org/10.1145/3390525.3390540
[17] Putrie, V. M., Sari, C. A., & Rachmawanto, E. H. (2019, October 14). Super encryption using transposition-hill cipher for Digital Color Image. IEEE Xplore. Retrieved from https://ieeexplore.ieee.org/document/8864361
[18] Serdano, A., Zarlis, M., & Nababan, E. B. (2021, April 28). Performance of Combining Hill Cipher Algorithm and Caesar Cipher Algorithm in Text Security. IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/document/9466039
[19] Khalaf, A. A. M., Abd El-karim, M. S., & Hamed, H. F. A. (2016). A triple hill cipher algorithm proposed to increase the security of encrypted binary data and its implementation using FPGA. 2016 18th International Conference on Advanced Communication Technology (ICACT), 752–759.
[20] Paragas, Jessie. (2020). An Enhanced Cryptographic Algorithm in Securing Healthcare Medical Records. 1-6. 10.1109/ICVEE50212.2020.9243228.
[21] Hossain, Md. Alam & Hossain, Md & Uddin, Md & Imtiaz, Shariar Md. (2016). Performance Analysis of Different Cryptography Algorithms. International Journal of Advanced Research in Computer Science and Software Engineering. 6.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY