United International Journal for Research & Technology

Volume 04, Issue 08, 2023 | Open Access | ISSN: 2582-6832

Enhancement of the 95x95 Vigenère Cipher Using a 3D Tabula Recta and a New Key Generation Technique in Application to Database Encryption

Clarck L. Tumazar[1], Benjamin D. Elevazo[2], Vivien A. Agustin[3], Jonathan C. Morano[4], and Mark Christopher R. Blanco[5]

[1,2]Student, College of Engineering - Pamantasan ng Lungsod ng Maynila

[3,4,5]Professor, College of Engineering - Pamantasan ng Lungsod ng Maynila

***Abstract*—** This study enhances the 95x95 Vigenère cipher by addressing its limitations. The current cipher struggles with encrypting non-English alphabets and is vulnerable to kasiski, brute force, and frequency analysis attacks. The researchers proposed an enhanced version by employing two keys generated from a cryptographically secure random number generator (CSPRNG), expanding the character set, and using a 3D tabula recta. The enhanced cipher successfully encrypts and decrypts various alphabets like Hangul, Arabic, Japanese, and Chinese, using a subset of printable characters from the Unicode Basic Multilingual Plane (BMP). It exhibited improved resistance to kasiski attacks, as the produced ciphertext lacked patterns such as bigram, trigram, up to n numbers of n-gram. The expanded character set and 3D tabula recta significantly increased the key permutations, making brute force attacks less likely to succeed. The algorithm also provided better defense against frequency analysis due to the utilization of two secure keys.

***Keywords*—** Brute Force Attack, Cryptographically Secure Random Number Generator (CSPRNG), Frequency Analysis, Kasiski Attack, N-grams, Permutation, Unicode, Vigenère Cipher

I. INTRODUCTION

One of the most common ways of protecting data in the digital world is through encryption. Encryption is a method of scrambling data to ensure that only authorized parties can access and understand the information. There are two main types of encryptions: asymmetric and symmetric encryption. Asymmetric encryption employs two different keys for encrypting and decrypting data, while symmetric encryption utilizes a single key for both encryption and decryption purposes [1].

An example of a symmetric encryption algorithm is the Vigenère cipher, which is one of the most popular symmetric encryption algorithms. The Vigenère cipher is a substitution cipher that conceals the plaintext structure by employing multiple monoalphabetic substitution ciphers [2]. Each letter of the plaintext is encrypted using a corresponding character from the key. Consequently, the key length must match the length of the plaintext. If the key length is shorter than the plaintext, the key is repeated until it matches the plaintext. Despite its popularity, Vigenère cipher has been limited in its usage due to inherent weaknesses, such as key repetition [3]. This leads to the following statement of the problem:

1. Due to the repeating nature of the key, its length can be derived from the ciphertext using attacks such as kasiski examination and frequency analysis.
2. The cipher cannot encrypt other alphabets besides the English alphabet.
3. When the character set is low, short ciphertexts are vulnerable to brute force attacks.

This study seeks to address the stated problem above by accomplishing the following objective:

1. Make the 95x95 Vigenère cipher less vulnerable to Kasiski Examination by implementing a new Vigenère key generation technique.
2. Increase the number of characters available for data encryption thus allowing multilingual encryption and making the cipher less vulnerable to frequency analysis.
3. Enhance the cipher’s complexity and mitigate its vulnerability against brute force attacks by implementing a 3D tabula recta.

II. RELATED WORKS

There are two main types of encryptions: symmetric-key encryption, which uses a single secret key, and asymmetric encryption or public-key encryption [4]. Symmetric-key encryption is known for its faster execution time compared to asymmetric-key encryption. 2048-bit asymmetric key and a 128-bit symmetric key provide a similar level of security [5].

One study focused on the flaws of the original Vigenère cipher. It lacked mathematical symbols, punctuation, and digits, which rendered it unable to encrypt certain characters. Additionally, the limited character set of the cipher's tabula recta made it vulnerable to frequency analysis attacks. To address these weaknesses, the researchers proposed an algorithm that increased the original 26x26 tabula recta to 95x95 and increased resistance to frequency attacks [6].

Another study highlighted the vulnerability of the Vigenère cipher's key to various cryptanalysis techniques, such as frequency analysis, brute force attacks, and Kasiski attacks. To mitigate these weaknesses, researchers proposed a modified key generation scheme that combined a pseudo-random number generator based on XOR shift with the Fisher-Yates Algorithm [7].

In a study concerning the Playfair cipher, researchers enhanced the encryption process by adding an extra dimension to the encryption table and increasing the character set. By reducing the number of rows and columns and introducing the third dimension, the proposed algorithm became more resistant to brute force attacks and frequency analysis [8].

Although seemingly advantageous for security, the repeating nature of the Vigenère cipher's key also provides an avenue for cryptanalysts to exploit. By employing the Kasiski and Friedman tests, cryptanalysts can derive the length of the key. A new key generation technique was proposed to address this vulnerability, resulting in higher entropy and a lower index of coincidence for the Vigenère cipher [9].

One study proposed an alternative arrangement for the Vigenère cipher's tabula recta. The researchers used a 512-bit Electronic Code Book (ECB) along with an arbitrary key to improve the arrangement of the cipher. This modification aimed to enhance the security of the Vigenère cipher and make it less susceptible to known cryptanalysis techniques [10].

III. METHODOLOGY

This chapter presents the design and methodology employed in the research, focusing on the processes and requirements needed to enhance the algorithm.

A. 95x95 Vigenère Cipher

The 95x95 Vigenère is an improved version of the original Vigenère cipher. It uses 95x95 tabula recta consisting of all possible characters, mathematical symbols, digits, and punctuations that are available on an ordinary QWERTY keyboard layout. Since the character set is increased from 26 to 95, the cipher's encryption equation is changed to Ci = (Pi + Ki) mod 95, and the decryption equation is changed to Pi = (Ci - Ki) mod 95. Here, C represents ciphertext, P represents plaintext, and K represents the key [6].

Pseudocode for 95x95 Vigenère Cipher:

```
1. START
2. Assign a numeric equivalent to every character in
   the character set.
3. Read the plaintext and convert each letter to its
   numeric equivalent.
4. Read the key. If the key length does not match the
   length of the plaintext, repeat the key until it
   matches the plaintext length.
5. Convert every character in the key to its numerical
   equivalent.
6. Apply the encryption equation using the numerical
   equivalent of the plaintext and key.
7. END
```

A. Unicode and the Basic Multilingual Plane (BMP)

Unicode is a universal character encoding standard that aims to represent all characters from all writing systems used in the world. It assigns a unique numerical value, known as the “code point” to every single character [11]. The Basic Multilingual Plane (BMP, or Plane 0) contains common-use characters for all the modern scripts of the world as well as many historical and rare characters. By far most of the Unicode characters for almost all textual data can be found in the BMP [12].

B. Cryptographically Secure Pseudo-Random Number Generator (CSPRNG)

A Cryptographically Secure Pseudo-Random Number Generator (CSPRNG) is a type of number generator that is applicable in security applications. When generating random numbers for security, the number should be random enough so that unauthorized parties would not be able to recreate/reproduce the number generated [13].

C. 3D Vigenère Tabula Recta

The 3D Vigenère Tabula Recta introduces an additional dimension, enabling the encryption of an already encrypted ciphertext to add another layer of security. To encrypt the ciphertext produced by the 2D tabula recta, a separate key must be employed. The utilization of an additional key significantly increases the number of attempts required for a successful brute-force attack.

D. Proposed Enhancement

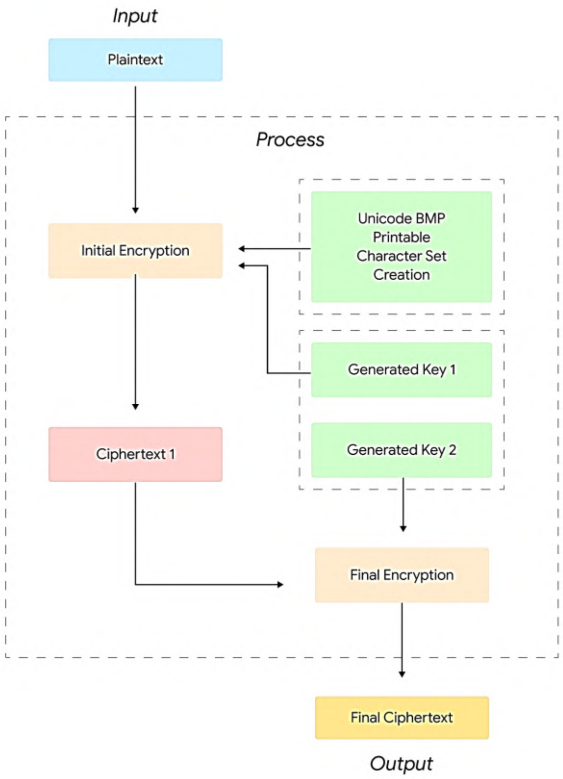


Figure I. *Conceptual Framework of the Enhanced Algorithm*

Fig. I show the conceptual framework of the enhanced algorithm. The plaintext is fed into the enhanced algorithm then the character expansion of the Vigenère tabula recta is executed. Next, key generation using a cryptographically secure pseudo-random number generator for key 1 and key 2 is executed. Key 1 will be used to encrypt the plaintext using the encryption formula of the enhanced algorithm. Key 2 will be used to encrypt Ciphertext 1 which is the result of the first encryption. The resulting ciphertext from the 2nd layer of encryption will be the final output of the enhanced algorithm.

Outlined below are the details of the proposed enhancements:

1. Expanding the Characters in the Tabula Recta

The researchers have extended the character set of the 95x95 Vigenère Cipher to encompass 56,981 characters under the Unicode Basic Multilingual Plane (BMP). This expansion enhances the cipher's capabilities for data encryption by increasing its versatility and adaptability. It also ensures compatibility with diverse languages, symbols, and special characters, making it suitable for database encryption in different countries.

2. Auto - generating the keys using a Cryptographically Secure Pseudo-Random Number Generator (CSPRNG)

The researchers proposed a modified key generation technique that auto-generates the keys and removes the ability of the user to provide an initial key. This eliminates the possibility of providing a weak key and ensures that every key is robust and strong. The researchers will employ a Cryptographically Secure Pseudo-Random Number Generator (CSPRNG) which is specifically designed to provide a higher level of cryptographic security by generating a sequence that closely approximates true randomness in contrast to a normal Pseudo-Random Number Generator.

3. Adding a 3rd Dimension to the Tabula Recta and introducing a 2nd key

In the problem statement, it was mentioned that short ciphertexts are vulnerable to brute-force attacks, particularly when the character set is limited. To address this vulnerability, the researchers introduced a second key. The second key will also be generated using a CSPRNG algorithm. By incorporating a second key, we effectively increase the overall key length, which in turn expands the number of possible key permutations to be guessed by a brute-force attack, where modern computers can guess passwords at a rate of 10,000 to 1 billion times per second [14]. This additional key introduces an additional layer of security, making it more challenging for an attacker to guess the correct permutation of the keys.

A. Evaluation Test

This section outlines the tests used to evaluate the proposed enhancement:

1. Kasiski Examination

The Kasiski examination, developed by Friedrich Kasiski, extracts the key length of the Vigenère cipher by identifying repeating patterns in the ciphertext, such as n-grams. The recorded distance between the occurrences of these patterns determines a factor of the key length [15].

2. Index of Coincidence

The index of coincidence (IoC) is a statistical analysis tool mostly used in ciphers that are based on substitution techniques. It measures the probability of two randomly selected letters in a text being the same [16].

3. Frequency Analysis

Frequency analysis is observance of patterns in a set of data. In the context of cryptography, it is most often used to check how frequently every character appears in the ciphertext.

4. Brute-force attack

A brute force attack is an attack method that systematically attempts to guess a key or password by trying every possible combination. The number of possible attempts depends on the length of the key and the size of the character set available for the password.

5. Monobit Test

Monobit test is a test that determines whether the number of ones and zeros in a sequence are approximately the same as would be expected for a truly random sequence. When the resulting P-value of the test is greater than 0.01, the sequence is said to be random [17].

V. RESULT AND DISCUSSION

This section will present the results and discussion of the evaluation tests conducted on the proposed enhancement of the Vigenère Cipher in comparison to the 95x95 Vigenère Cipher.

Table I. *Encryption Result of both baseline and enhanced algorithms*

	95x95 Vigenère Cipher	**Enhanced Vigenère Cipher**
Plaintext	THIS IS THE PLAINTEXT WE WOULD LIKE TO ENCRYPT PLEASE HACK ME NOW	
Initial key (Key inputted by the user)	KEY	(The key is auto generated)
Final Key	KEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKEYKE	**1st key:** 洋3ᵖघ□ 퓻ᅜ欧퓚쒻 ᆜ骴셀슭薳灶ᄀ 쉳闌ᄆ扢淽 ꣲK텡ᄇᄒ ᄀᄉ怋ঙ□ 둥糫跧趚ᄡ 辡wN̠芫萝ꝏ⑸侻蒓ৼ穅 멸멳抾立□蹟戌대ϛ緅闍熬樊ꠣ뙳□業□ **2nd key:** 맥燍錯絅쌔랪□鄄禋鵐儒ৠ珊絵–ỉC詾쏩騫 燏ৢ줘阑襪퇽ᮡ桁ⴳↄ籔ǫӜ□□叻蒲ꡦ畓껑趑橒쓻 ⁄ঌ 餺숖뙅됫鐖鑊椹첪□ɱ쑢联□ᆜĊ ⁄ũ閙バ
Ciphertext	dLgcDgcDrRIXZPYSRrObrJacJamePbJPgUIXdSXORabcndDnVIYcIXREaUDkODlYa	ᵛ爈楔骪졷곷덛철潦茸瑺啫嘿릍鰢蓱⊥ꝋ竟ᅷ□穰署骔 篖ƻ岈漾両탈뛹蓎ʒ鮔瞶횛玥ᇂ瘴à㒸僾뭉琞蘶辰🪙鸛摦 偬쯤颎ᑐ促奞ᢒ：孯譫팽□□乽敨疆

Table I shows the result of the encryption of the plaintext “THIS IS THE PLAINTEXT WE WOULD LIKE TO ENCRYPT PLEASE HACK ME NOW” of both the 95x95 Vigenère Cipher and the Proposed Enhanced Vigenère Cipher. The baseline algorithm inputs an initial key of “KEY” while the enhanced algorithm requires no input by the user. The final key of the baseline algorithm is shown to have a repeated key while the enhanced algorithm produced two keys consisting of different language characters and symbols.

Lastly, the ciphertext of the baseline algorithm shows a mixture of upper and lowercase letters of the alphabet while the enhanced algorithm again shows a wide range of characters and symbols from different languages.

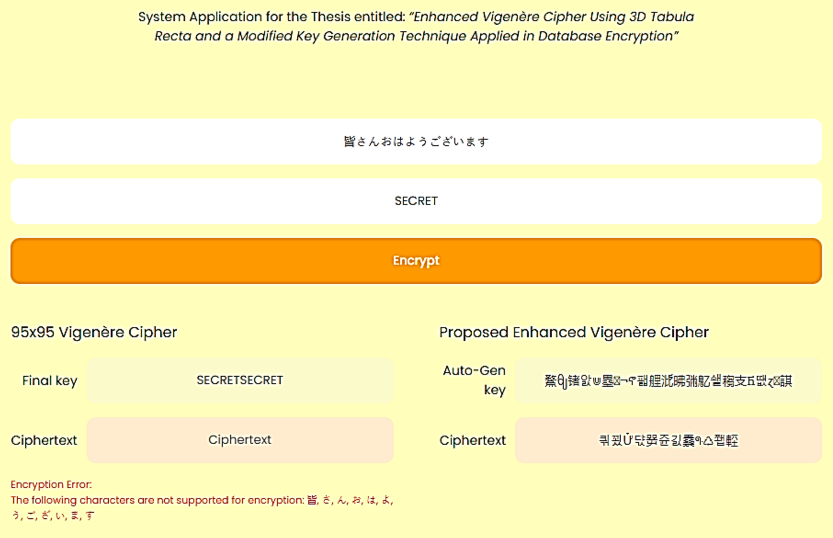
Table II. *Evaluation Result of both the Baseline and Enhanced Algorithms*

	Evaluation Results				
	Kasiski Examination	Index of Coincidence	Frequency Analysis	Monobit Test	Brute-Force Test
95x95 Vigenère Cipher	Bigram, Trigram	A probable key of 3 was derived	Shows uneven distribution of characters in the	0.00286	**Approx.** 1.4289583 minutes (Guessing every possible key permutation)

		from the ciphertext.	ciphertext. Multiple characters were repeated more than once.		
Enhanced Vigenère Cipher	No Pattern	No probable key was derived.	Shows even character distribution and no character was repeated.	0.02092	**Approx.** 1.534407080313600 06913246252225e29 centuries per key

Table II discusses the evaluation results of the different tests done to both the baseline and enhanced algorithm. Kasiski examination was conducted on both algorithms and the baseline ciphertext showed bigram and trigram patterns, while the enhanced algorithm showed no pattern. In using the Index of Coincidence to identify the probable key length used, a probable key length of 3 was derived from the baseline's ciphertext while the enhanced algorithm showed no probable key length. Frequency analysis was also done on both the algorithms and the baseline shows a result of uneven distribution of characters in the ciphertext where multiple characters were repeated more than once. Monobit test was conducted on both algorithms' ciphertext. The baseline algorithm showed a monobit p-value of 0.00286 which is lower than the threshold of 0.01 to pass the randomness test. The enhanced algorithm on the other hand, showed a p-value of 0.02092 which is higher than the 0.01 threshold thus passing the randomness test.

Finally, brute-force attack was done on both algorithms to test its resistance. The baseline algorithm key resulted in approximately 1.4289583 minutes to guess every possible key permutation while the enhance algorithm resulted to approximately 1.53440708031360006913246252225e29 centuries per key.



System Application for the Thesis entitled: *"Enhanced Vigenère Cipher Using 3D Tabula Recta and a Modified Key Generation Technique Applied in Database Encryption"*

皆さんおはようございます

SECRET

95x95 Vigenère Cipher

Final key: SECRETSECRET

Ciphertext: Ciphertext

Encryption Error:
The following characters are not supported for encryption: 皆, さ, ん, お, は, よ, う, ご, ざ, い, ま, す

Proposed Enhanced Vigenère Cipher

Auto-Gen key: [illegible]

Ciphertext: [illegible]

***Figure. II** GUI of the Simulation Application*

Fig. II shows the GUI of the simulation application created by the researchers. As shown in the figure, the researchers inputted a Japanese plaintext with a key of "SECRET". The 95x95 Vigenère Cipher is shown to have an encryption error and states that the following characters are not supported for encryption while the proposed Enhanced Vigenère Cipher have successfully encrypted the Japanese plaintext.

VII. CONCLUSION AND RECOMMENDATION

In this study, the researchers enhanced the 95x95 Vigenère cipher to improve its security and functionality. Firstly, they implemented a new Vigenère key generation technique using an auto-generated key generated by a CSPRNG algorithm. This eliminated the repetitive nature of the key and rendered the cipher less vulnerable to Kasiski examination attacks as the ciphertext produced no N-gram patterns. Furthermore, the produced ciphertext of the enhanced algorithm got a Monobit P-value of 0.02092 which is higher than the 0.01 threshold thus passing the randomness test. Additionally, the increased character set of 56,981 characters from the Basic Multilingual Plane (BMP) enabled multilingual encryption and strengthened the cipher's defense against frequency analysis. Lastly, the incorporation of a 3D tabula recta and a second key significantly increased the complexity of the cipher and mitigated its vulnerability to brute force attacks. The new algorithm's key permutations would take an impractically long time to exhaust, providing enhanced security compared to the original 95x95 Vigenère Cipher.

The researchers suggest further studies focusing on encryption time and memory usage of the Vigenère cipher. The researchers also recommend looking into the application of key management to further improve key secrecy and finally, the application of the BMP of Unicode to other poly or mono-alphabetic ciphers is encouraged for future improvements.

REFERENCES

[1] Cloudflare. (n.d). "What is encryption?" Retrieved from: https://www.cloudflare.com/learning/ssl/what-is-encryption/

[2] G. J. Simmons, "Vigenère cipher," Encyclopedia Britannica, April 2023. https://www.britannica.com/topic/Vigenere-cipher

[3] A. L. Hananto, A. Solehudin, A. Susilo, Y. Irawan, and B. Priyatna, "Analyzing the Kasiski Method Against Vigenere Cipher," International Journal of Computer Techniques, vol. 6, no. 6, pp. N/A, 2019. doi:10.29126/23942231/IJCT-V6I6P2

[4] J. De Groot, "What Is Data Encryption? (Definition, Best Practices & More)," Digital Guardian, 2023. https://www.digitalguardian.com/blog/what-data-encryption

[5] K. Brush, L. Rosencrance, and M. Cobb, "Asymmetric cryptography (public key cryptography)," TechTarget, Sep. 2021 https://www.techtarget.com/searchsecurity/definition/asymmetric-cryptography#:~:text=Asymmetric%20cryptography%2C%20also%20known%20as,from%20unauthorized%20access%20or%20use

[6] K. Nahar and P. Chakraborty, "A Modified Version of Vigenere Cipher using 95×95 Table," International Journal of Engineering and Advanced Technology, vol. 9, no. 5, pp. 1144-1148, 2020. doi:10.35940/ijeat.E9941.069520

[7] J. P. G. Perez, S. K. P. Sigua, D. M. A. Cortez, K. E. Mata, R. C. Regala, A. J. Alipio, M. C. R. Blanco, and A. M. Sison, "A Modified Key Generation Scheme of Vigenère Cipher Algorithm using Pseudo-Random Number and Alphabet Extension," 2021 7th International Conference on Computer and Communications, pp. 565–569, 2021. doi:10.1109/ICCC54389.2021.9674565

[8] A. Kaur, H. K. Verma, and R. K. Singh, "3D(4 x 4 x 4) - playfair cipher," International Journal of Computer Applications, vol. 51, no. 2, pp. 36–38, 2012. doi:10.5120/8017-1286

[9] T. H. Hameed and H. T. Sadeeq, "Modified vigenère cipher algorithm based on New Key Generation Method," Indonesian Journal of Electrical Engineering and Computer Science, vol. 28, no. 2, pp. 954–961, Nov. 2022. doi:10.11591/ijeecs.v28.i2.pp954-961

[10] A. P. Sidik, "Improve The Security of The Vigenère Cypher Algorithm by Modifying the Encoding Table and Key," International Journal of Basic and Applied Science, vol. 10, no. 2, pp. 42-50, 2021.

[11] Unicode Consortium, "What is Unicode?," Unicode, July 2017. https://unicode.org/standard/WhatIsUnicode.html

[12] The Unicode Consortium, "Unicode® Standard, Version 15.0.0: Chapter 2 - General Structure," Unicode.org, pp. 43, 2022. https://www.unicode.org/versions/Unicode15.0.0/ch02.pdf.

[13] C. R. Ryan, M. Kshirsagar, G. Vaidya, A. Cunningham, and R. Sivaraman, "Design of a cryptographically secure pseudo-random number generator with grammatical evolution," Scientific Reports, vol. 12, no. 1, 2022. doi:10.1038/s41598-022-11613-x

[14] L. Grigas, "The Ins and Outs of a Brute Force Attack," NordPass, 2022. https://nordpass.com/blog/brute-force-attack/#:~:text=Speed%20depending%20on%20password%20strength,symbols%20on%20a%20standard%20keyboard

[15] "Kasiski Examination," Michigan Technological University, [Online]. https://pages.mtu.edu/~shene/NSF-4/Tutorial/VIG/Vig-Kasiski.html. [Accessed: 9 June 2023].

[16] "Index of Coincidence (IOC)," Michigan Technological University, [Online]. https://pages.mtu.edu/~shene/NSF-4/Tutorial/VIG/Vig-IOC.html. [Accessed: 9 June 2023].

[17] A. Rukhin et al., "A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications," NIST, 2010. https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=906762