ISSN No:-2456-2165

# An Improved Method of Vigenere Cipher to Securely Compress the Text by using Relative Frequency

Nazia Shoukat, Dr.Muhammad Azam,Dr Imran Khan
Department CS & IT Superior University
Lahore, Pakistan

**Abstract:- Cryptography is an ancient technique of data security that is used to protect data from unauthorized access. It is used in many applications on commercial organizational or industrial scales to secure data. Data encryption is a process of converting data into cipher text. This is the most effective way to achieve data security. In this paper, we have imposed a new method to enhance the vigenere cipher for more efficiency and security. Though, the classical Vigenere cipher is extremely easier to crack and also vulnerable to attacks. To improve the data confidentiality of cipher text security features, we use the Relative Frequency of Alphabetic Letters. We arranged the sequence of the letter according to the relative frequency and arranged them in increasing order and then we performed some modifications in the Vigenere cipher by using the relative letter frequency technique and then compressed the text data. The proposed procedure improves the confidentiality of a text message by introducing multiple layers of security processes such as encryption and decryption, and compression, respectively. The purpose of this article is to improve the security of classical Vigenere cipher based on using relative frequency analysis for the encryption to secure the data and then the lossless Huffman technique compresses the resultant generated text. To sum up, the ensuing proposed technique produces a more secure and short length of data code than the existing process.**

***Keywords***:- *vigenere cipher; Huffman; relative letter frequency; and data compression.*

## I. INTRODUCTION

Cryptography is a field of data security technology. It seeks to protect and encrypt so that any third party who has access to the encrypted data can reconstruct and retrieve the required information. In practice, encryption methods use a specific function, algorithm, or formula in detail so that it cannot be accessed as before. Accurate data can be obtained with the correct key and data associated with an encrypted algorithm, function, or system that pre-determined data encryption. Cryptography is a trade of experts to demonstrate encoding algorithms specifically for organizational or martial purposes. Nowadays, it is used in many applications on commercial organizational, or industrial scales to secure data or information from stealers. All of this is likely due to the increasing use of computers and increasing processor power. Data encryption is a process of interpreting data into ciphertext. This is the most effective way to achieve data security. Unauthorized persons can easily understand the conversion of data into a form of technology. Encryption means converting encrypted data to its original state for understanding. Encryption makes the information useless for someone who does not have a decryption key [8]. In this paper, we have made use of the vigenere cipher. Vigenere cipher is an ancient cryptography technique and it can be used in data security. We have proposed a method to enhance the vigenere cipher for more efficiency and security. Though, the classical Vigenere cipher is extremely easier to crack and also vulnerable to attacks. To improve the data of ciphertext security features, we use the Relative Frequency of Alphabetic Letters. We arrange the sequence of the letter according to the relative frequency in increasing order. And then we made use of a Modified Vigenere ciphertext by using the relative letter frequency technique. The proposed method is to maintain and improves the security and confidentiality of a text message by introducing multiple layers of security processes such as encryption and decryption, and compression, respectively. Data compression is a common technique that can be used to compress data in most computerized applications. There are several algorithms based on data compression, that can be dedicated to compressing the various type of single data. This article scrutinizes lossless data compression algorithms and implemented them to assess the performance of text data compression. The purpose of this article is to enhance the classical Vigenere cipher based on using relative frequency analysis for the encryption to secure the data and then the lossless Huffman technique compresses the resultant generated text as shown in the flowchart of figure 1 given below. To sum up, the ensuing proposed method produces a more secure, short, much less foreseeable, and cost-effective ciphertext code length than the prevailing procedure.

## II. RELATED WORK

With the development of communication and network technology in modern life, there is a great demand for privacy in encryption technology. However, the matrix-based encoding methods available in the literature have limitations for their exploitation. Motivated by a large number of cryptographic techniques available in the literature for securing data proposed an approach based on the dimensionless vector space of Vigenere encryption relies on random decomposition to select the key in the key space [1].

Vigenere is an example of a permutation cipher with a variety of restrictions. In this paper, the authors propose an advanced cryptographic algorithm that improves the security of the Vigenere scheme by combining it with a modern cryptography method that uses plaintext, ciphertext, and binary format where the key is a string of bits (rather than characters). It has been noticed that repeated parts of the plaintext are always encrypted with different parts of the keyword or binary key. This is because odd-placed characters are encoded with a stream cipher and even-placed characters

ISSN No:-2456-2165

with a Vigenere cipher, resulting in different ciphertext segments. This means that the proposed algorithmic technique makes ciphertext more complex to analyze [2]. By applying an algorithm, Goldbach uses vigenere to encrypt the results of the data protection process. Since the number of characters in the original ciphertext is different from the value of the code as a result of compression is difficult to predict the plaintext [3]. The author proposes a new form of key symmetric cipher called the Rectangular Generalized Vigenere Cipher, where the encryption and decryption tables are not necessarily square matrices, but rectangular. The main advantage of using square ciphers is to reduce the memory requirement for storing encryption and decryption tables. Using a rectangular cipher table can reduce the memory space required to develop a cryptographic program but is hard to implement [14]. Researchers expanded the original vigenere table to 95x95. It introduced all the possible letters, mathematical symbols, numbers, and punctuation marks found in common keyboard layouts. They can be easily encoded with this technique and are case-sensitive. The researchers in this paper seek to further expand this set of characters by processing additional 95 x 95 characters in a vigenere table. Due to the large character set, this algorithmic method does not suffer from frequency attacks. [4].The writer proposed Vigenere's advanced encryption to perform obfuscation that preserves users' data privacy. It combines Vigenere's advanced encrypted algorithm with smart rules to perform encryption with different rule sets. It primarily focuses on data privacy and minimizes the complexity ( encryption and decryption) of time. By comparing with the previous research, the proposed system suffers from performance degradation based on user-side execution time and attacks [18]. The authors are expected to cover the weaknesses of the vigenere cipher and develop new methods to mitigate the weaknesses of the vigenere cipher by combining Caesar's and Hill's cipher techniques in the cryptographic key generation process. The simulation results show that the Vigenere encoding algorithm may repeat words and predict the final key information. However, Vigenere's modified cryptographic algorithm does not have it, so the information cannot be predicted [5]. Many types of encryption, such as single-letter and multi-letter ciphers, have been developed to protect information. The computation of the encryption process involves only additive encryption, which makes this algorithm vulnerable to attackers based on character frequency analysis. The proposed method in this research complicates the visioner code by combining monoalphabetic code and vigenere code. Combining the Vigenere cipher with the Affine cipher creates a new method that is a more complex algorithm. In this paper, we propose to combine affine ciphers and Bigenelle ciphers to create more complex algorithms. This is proven by the process of cryptanalysis using a vigenere analyzer and monoalphabetic permutation, cryptanalysis cannot decode plain text and this method is more secure [6].More recently, work has been done to address the duplicate key nature of the algorithm by matching the key length to the plaintext. Results based on mono-bit frequency research and analysis show that there is no key duplication and generating keys can be used to encrypt larger character sets. The results show that the modified version does not work in the proposed algorithm due to its large character set but a modified version is more complex to understand [7].The encryption approach presented by the author avoids the key duplication problem. The writer suggests a new cryptographic strategy to exchange the data securely. It has been noticed that the proposed encryption approach avoids the key duplication problem by using a new key generation process. A variety of modifications were performed in previous research based on various techniques of transposing the key for the Vigenère cipher with various other cipher techniques to secure the data. The results demonstrated the suggested method's superior performance as compared to the other techniques used in this paper [12]. As a result, various approaches have been described in the above literature but this article proposed a new improved method of cryptographic strategy to maintain the integrity and securely compress the data by combining a modified version of the vigenere table with a modified relative letter frequency table.

## III. PROPOSED METHODS

### *A. Vigenere Cipher*

Vigenere encryption is the process of data encryption by using the Ceaser cipher technique embroidered based on key letters. It used a kind of polyalphabetic substitution that uses the Caesar cipher formula. Its algorithm is an older cryptographic method that is much safer than Caesar's cipher. Its algorithm provides a more secure ciphertext. Vigenereciphertext is a way to encode text into ciphertext based on key characters. The real ciphertext encryption is generated by using a letter frequency analysis based on the Vigenère square table. Vigenere cipher is a polyalphabetic substitution cipher that consists of horizontal and vertical blocks with a 26 x 26 alphabetic letter matrix [9]. Its algorithm uses a Relative Letter frequency analysis table in vigenere to encrypt the plaintext. Each letter moves forward to the left proceeding letters, consistent with all possible alphabetic cipher letters. Each row of the square table represents keywords column of the square table shows plaintext. In this paper, we have to use the vigenere cipher technique to secure text or data based on relative letter frequency analysis. As we know, Relative letter frequency analysis is the calculation of the probability of whole alphabetic characters. Frequency analysis is evaluated based on the fact that a given data or text contains a collection of characters with specific characters and different frequencies. It is based on the fact that any part of the written language has a collection of characters of different frequencies. Furthermore, the characters have a specific distribution that is almost identical to almost all patterns in that linguistic. According to relative letter frequency analysis, letters O, T, A, and E have the highest frequency because the usage of these letters is most common but the Z, Q, X, and J have the least frequency because they are rarely used. As we know, simple letter frequency analysis or vigenere cipher is providing no more secrecy of data. However, In this paper, we are going to represent a more secure alternative method of vigenere cipher by using a letter frequency analysis table. Here, an alternate modified square table of vigenere cipher based on increasing order of relative letter frequency is given in "*Fig.1"*.

ISSN No:-2456-2165

| | Z | J | Q | X | K | V | B | P | G | W | Y | F | M | C | U | L | D | H | R | S | N | I | O | A | T | E |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Z | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 |
| J | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 |
| Q | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 |
| X | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 |
| K | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 |
| V | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 |
| B | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 |
| P | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| G | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| W | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Y | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| F | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| M | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| C | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| U | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| L | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| D | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| H | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| R | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| S | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
| N | 21 | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| I | 22 | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| O | 23 | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
| A | 24 | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| T | 25 | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| E | 26 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |

Fig. 1: Modified Vigenere Cipher(using Relative Letter Frequency analysis)

In this paper, we will use the key for plain text to secure our encrypted ciphertext by using the relative letter frequency analysis table to arrange the sequence of the letter according to the order of the relative letter frequency cipher table. Each letter of the plaintext is matched with each character of key letters according to the increasing order of relative letter frequency analysis [10] as shown in *"Fig. 2"*, and "*Table 1"* given below:

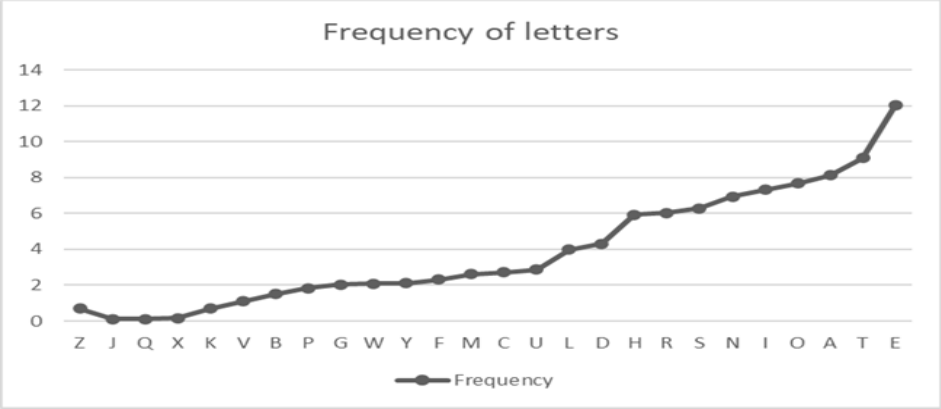


Fig. 2: Graphical representation

ISSN No:-2456-2165

Now, we have sorted the sequence of whole alphabetic letters in increasing order according to the relative frequency analysis in *"Table1"* given below

| Arranged alphabetic letters in an increasing order | | |
|---|---|---|
| **Order of sequences** | **Arranged letters according to their frequency (increasing order)** | **An Increasing order letter frequency** |
| 1 | Z | O.O7 |
| 2 | J | 0.10 |
| 3 | Q | 0.11 |
| 4 | X | 0.17 |
| 5 | K | 0.69 |
| 6 | V | 1.11 |
| 7 | B | 1.49 |
| 8 | P | 1.82 |
| 9 | G | 2.03 |
| 10 | W | 2.09 |
| 11 | Y | 2.11 |
| 12 | F | 2.30 |
| 13 | M | 2.61 |
| 14 | C | 2.71 |
| 15 | U | 2.88 |
| 16 | L | 3.98 |
| 17 | D | 4.32 |
| 18 | H | 5.92 |
| 19 | R | 6.02 |
| 20 | S | 6.28 |
| 21 | N | 6.95 |
| 22 | I | 7.31 |
| 23 | O | 7.68 |
| 24 | A | 8.12 |
| 25 | T | 9.1 |
| 26 | E | 12.0 |

Table 1: Arranged alphabetic letters in an increasing order

Then, we have created a flow chart of the proposed method of vigenere cipher as shown in *"Fig. 3"* below:

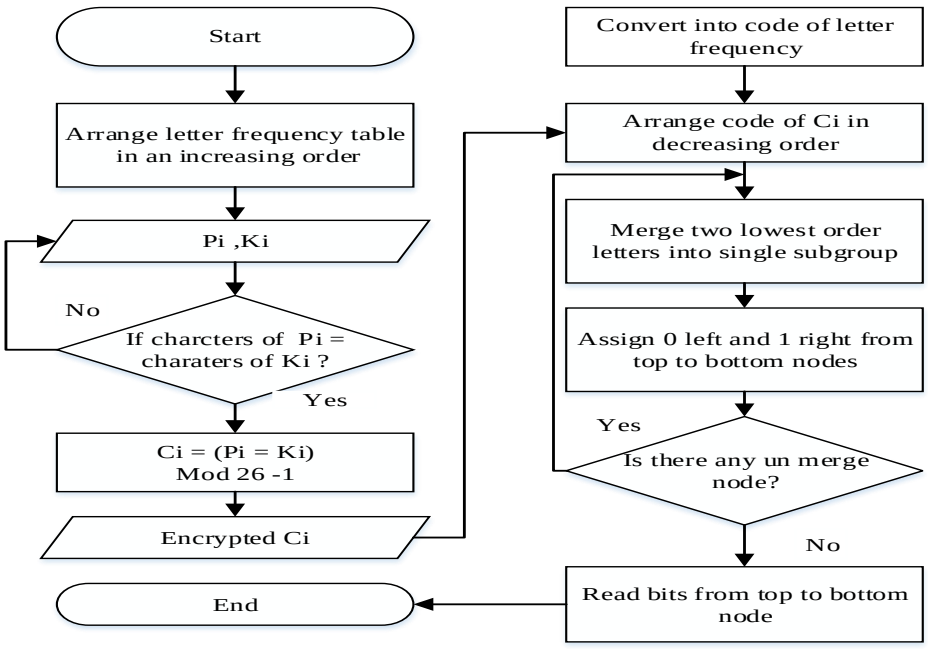


Fig. 3: Modified Vigenere Cipher(using Relative Letter Frequency analysis)

Where : Pi = "Plain text" , Ki = "Key", Ci = "Cipher text"

ISSN No:-2456-2165

In this paper, The encryption of a modified Vigenere Cipher by using the relative letter frequency technique is the same as the classical Vigenere cipher. Let's suppose the plaintext NOMORE is encrypted by using the secret keywords TRY. Then, the use of a key is given below in *"Table II"*.

| **Supposed plaintext with the key** | | | | | | |
|---|---|---|---|---|---|---|
| **Plaintext** | N | O | M | O | R | E |
| **Keyword** | T | R | Y | T | R | Y |

Table 2: supposed plaintext with the key

To get the ciphertext from the plaintext and the keyword above, Based on the Modified Vigenere square matrix, The Initial plaintext letter is N(21) and the keyword letter is T(25) so by following the Letter frequency table the intersection of the ciphertext of the first letter would be R(19). The same way is to follow on the next plaintext O(23) and the keyword letter is R(19) so by following the Letter frequency table the intersection of the ciphertext of the first letter would be U(15). Similarly, the next plaintext is M(13) and the keyword letter is Y(11) so by following the Letter frequency table the intersection of the ciphertext of the first letter would be O(23). The same way has been done on all the next characters of plaintext and letters of keywords. The result of the encrypted ciphertext is as follows:

| **Encrypted cipher text** | | | | | | |
|---|---|---|---|---|---|---|
| **Cipher text** | R | U | O | N | Y | W |

Table 3: Encrypted cipher text

On the classical vigenere cipher the mathematical formula is:

Ci=(Pi+Ki) mod26

Where : Ci = "Cipher text", Pi = "Plaintext" and Ki = "Keyword"

Uses the decimal value of the character starting from 0....25, that is A=0 …..to…..Z=25.While in the modified vigenere cipher, we use a letter frequency table .so in this paper the mathematical formula uses alphabetic letters would start from 1...26, but apply to letters randomly according to the increasing order of alphabetic letters by using Letter Frequency and will also obtain a slightly different mathematical formula for encryption [10].

Ci=(Pi+Ki) mod 26 -1

In this formula, Ci = "Cipher text", Pi = "Plaintext", and Ki = "Keyword". Cipher text would be generated by putting the value of characters of plaintext and Character of Key separately by using the value table of the Relative Letter Frequency table. For example, when our plaintext key works safely then the encrypted value of cipher text according to an increasing order Relative to Letter Frequency would be

Cipher text: RUONYW

*B. Data Compression*

Text is a set of letters or a single-character unit. It has limited storage devices and a lot of characters that can always cause problems with data transfer speeds from time to time. Although storage can be replaced with larger ones, it is not a good solution if there is another solution. And everyone is thinking of finding a way to use it to shorten the text. Pressing converts real data into code form to save storage requirements and data transfer time [17]. Compression is a technique of data representation in a condensed form rather than its real form. It is a process that is used to reduce the size of the data and remove extra information, or redundancy. Data compression performs a significant role in the area of distributive schemes and file storage in interactive programs and software, text credentials, and record catalogs. The size of data, files, or text can be reduced by using data compression. There are several ways to classify compression. Lossless data compression techniques recreate the real text or data without any loss of data from the compression file. Therefore, the data does not change the processes during the compression and decompression. Lossless compression is also known as revocable or reversible compression. Lossless compression techniques are used to compress text, picture images, data files so on [11]. The lossy data compression technique is generally used when a perfect uniformity with the real data is not essential after decompressing. An example of lossy compression is frequently used to compress video or picture data [16]. In our article, the compression process is performed by using the Huffman coding algorithm. Each letter is encoded with lots of bits to give a good result. The purpose of this article is to examine the security and shortcut compression of Huffman's algorithm on ciphertext and to explain text compression methods using the vigenere ciphertext based on the letter frequency analysis [17].

*C. Huffman code*

David A. Huffman developed an appropriate source code used to compress lossless data known as Huffman coding. The algorithm generates code words with variable lengths instead of a character, based on a table based on the letter frequency of characters from plaintext [10]. Huffman code is a famous lossless compression algorithm to minimize code redundancy as compared to other algorithms. The algorithm of Huffmanis efficiently used in data, text, and audio-video compression [19]. It is a simple compression algorithm compiled by David Huffman in 1952. This algorithm is included in the type of lossless data compression, which means that data does not eliminate or change the number of bytes and is stored according to the

ISSN No:-2456-2165

original data [13]. The Huffman technique is compatible with specific algorithms, prefix codes created from a set of options by the Huffman coding algorithm. It compressed the data based on ASCII alphabetic characters. It compresses and reduces the resultant binary code by constructing a binary tree from the topmost to the bottom node. Data is properly compressed because it reduces the data storage capacity, size, or volume.

The compression is frequently stated as binary codes such as input characters orbits are issued by a specific source of information and must be coded before being sent to their destination [15]. A Huffman is the most efficient and useful technique for lossless data compression. In this technique, all characters or letters of text convert into binary digits, and the least -frequent characters have long binary bits but most-frequent used letters have short binary codes [16]. Let's suppose the above vigenere cipher produces a cipher text RUONYW by using relative letter frequency analysis. The newly generated characters and their corresponding code words based on relative frequency are shown in *"Table IV"* given below:

| **Encrypted cipher text with code** | | | | | | |
|---|---|---|---|---|---|---|
| **Ciphertext** | R | U | O | N | Y | W |
| **Frequency** | 19 | 15 | 23 | 13 | 11 | 10 |

Table 4: encrypted cipher text with code

We have characters of ciphertext which produce secure vigenere ciphertext by using the relative letter frequency table as shown above in "Table IV". To construct Huffman code, First, arrange all characters in non-increasing order from higher to lower frequency as given below in *"Table V"*.

| **Arranged cipher text code in decreasing order** | | | | | | |
|---|---|---|---|---|---|---|
| **Ciphertext** | O | R | U | N | Y | W |
| **Frequency** | 23 | 19 | 15 | 13 | 11 | 10 |

Table 5: Arranged cipher text code in decreasing order

Then to construct a Huffman tree, we choose the smallest-frequent letters and rationally assembled them, and also add the frequencies of these characters. To compress and generates binary code bits, we use the Huffman code binary tree based on the vigenere cipher by using the letter frequency table. For instance, letters X and X have the same least frequency so join them together. When the least frequency of both letters is added then a new child node frequency 8 is produced. Now, we again select the two least-frequent letters and also added their frequencies to get a new node frequency. Again, follow the same methods select two low-frequency characters, joined them, and also add the frequencies of these letters. Up until to grasp letter also remains a single root node. In the Huffman tree, the edges connecting to nodes are labeled with the binary digits 0 and 1. Here, each edge of the child node is connected to the right side of the parent node linked with binary bit 0, and each edge of the left side of the parent node shows with binary bit 1. The binary digits for each source character are arranged along the path from a parent node to a child node. Based on the Letter frequency table in vigenere cipher text, the creation of new codes by using a tree is shown in *"Fig.4"* as given below:

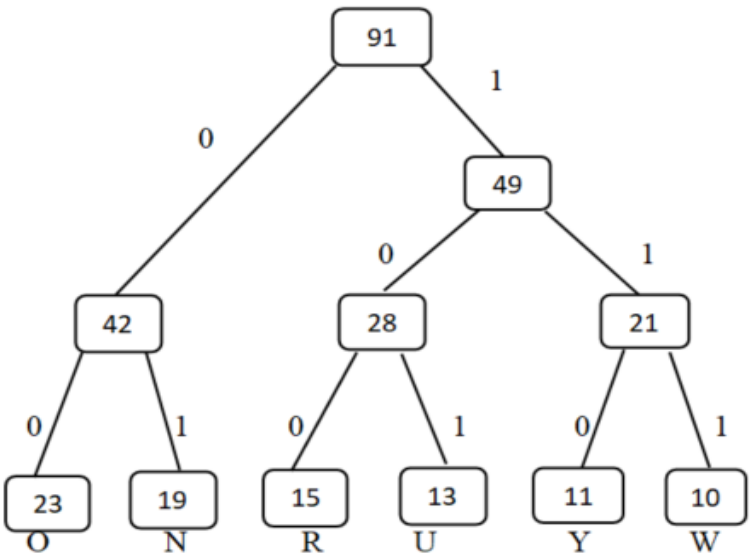


Fig. 4: Cipher text tree generated by the Huffman technique

ISSN No:-2456-2165

By drawing downward, the Huffman tree generated the shortest unique binary code also allocated to the highest frequency character ciphertext, based on the relative letter frequency analysis table as shown in *"Table VI"* given below. All ciphertext letters are replaced by Huffman binary codes.

| **Arranged cipher text code after using the Huffman technique** | | | | | | |
|---|---|---|---|---|---|---|
| **ciphertext** | O | N | R | U | Y | W |
| **code** | 00 | 01 | 100 | 101 | 110 | 111 |

Table 6: Arranged cipher text code after using the Huffman technique

Finally

, The compressed binary bits sequence, obtained as a result of replacing the symbols in the example used with the Huffman codes, is arranged in *"Table VI"* At the stage of reconstructing a compressed text or data, the Huffman tree is redeveloped from the frequencies of the ciphertext and transformed into real characters. The binary bits are used to crisscross the Huffman tree. Each iteration is monitored from the parent(root) node until it reaches the node of that letter. This procedure starts all over again for the next node.

Cipher text: R U O N Y W

Compressed code: 1001010001110111

## IV. RESULT AND DISCUSSIONS

In this given section, we have used CrypTool 2.1 to provide a comparative result of both methods. Initially, we have to create a framework according to our proposed method then we applied the same plain text message with the same keywords on both (classical vigenere cipher and modified vigenere cipher) methods shown in *"Fig.5"*.Then it is notified that the compressed cipher text data of classical vigenere cipher is not secure and easy to break. This is due to the sequential alphabetic order repetition of keywords, so there is a chance to easily predict the text message by applying various cryptographic methods. As shown in *"Fig. 5"*, by applying the Huffman technique before modification in the vigenere cipher it has been noticed that the text data will be un securely compressed and there is also a chance to predict the message.

But after modification, we have used the same plain text message with the same key in an improved method of modified vigenere cipher as shown in *"Fig.5"*, where the implementation of the result of the modified vigenere cipher has been shown. It can be seen that an improved method of vigenere cipher is more secure and hard to break. This is due to the modification of the vigenere square table by applying an increasing order letter frequency technique mentioned in *"Fig. 3"*.So, the message is encrypted by a modified vigenere cipher and will be safely compressed, and cannot be predicted easily.

Hence, we can say that this technique provides confidentiality so by using this improved method we can easily and securely deliver compressed data because it is stronger than others and there is less chance to predict the message easily.

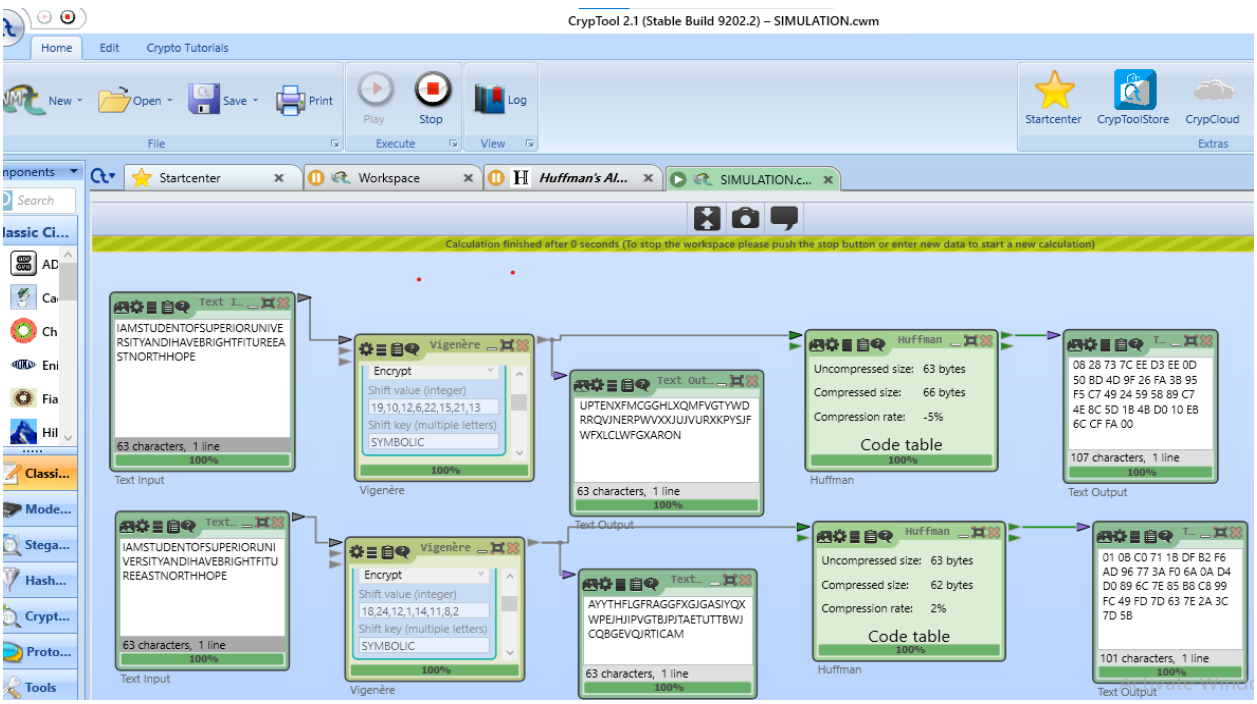


Fig. 5: Comparative result of classical vigenere cipher as well as modified vigenere cipher

ISSN No:-2456-2165

## V. CONCLUSION

In this article, we have performed some modifications to the vigenere cipher table because it's very simple and easily crackable. So we have to change the sequence of letters and apply modified a vigenere cipher square table according to the relative frequency of letters and then we have to compress the text data by using the lossless compression Huffman algorithm to create more complexity and maintain the data confidentiality. The above-described methods which are the combination of modified vigenere cipher and relative letter frequency technique provide a much more secure cipher text and the implementation of lossless Huffman algorithms produces effective compression by reducing the numbers of binary digits code. The combination of these classic techniques provides more confidentiality and strong cipher text. Therefore, It is concluded that the purpose of the above newly imposed method can prove that it is valuable to safely maintain and compress the data that has been achieved with the use of a modified vigenere cipher by using the relative letter frequency analysis table. Simulation results show that the cipher text generated by using the relative letter frequency is providing more secrecy with the concept of compression using the Huffman lossless technique.

## REFERENCES


[1.] N. Uniyal, G. Dobhal,A. Rawat, and A. Sikander, "A novel encryption approach based on vigenere cipher for secure data communication," India, March 2021.

[2.] Fairouz Mushtaq Sher Ali and Falah Hassan Sarhan, "Enhancing security of vigenere cipher by stream cipher," vol. 100. August 2014

[3.] Surya Darma Nasution, Guidio Leonarde Ginting, Muhammad Syahrizal, and Robbi Rahim, "Data security using vigenere cipher and Goldbach code algorithm," vol. 6. January 2017.

[4.] Khairun Nahar and Partha Chakraborty, "A modified version of vigenere cipher using 95*95 tale," vol. 9. June 2020.

[5.] Z Qowi1 and N Hudallah, "Combining caesar cipher and hill cipher in the generating encryption key on the vigenere cipher algorithm," Universitas Negeri Semarang, Indonesia, 2021.

[6.] Siti Agustini, Weny Mistarika Rahmawati, and Muchamad Kurniawan, "Modified vigenere cipher to enhance data security using Monoalphabetic cipher," vol. 01, pp. 26-32, 2019

[7.] John Paul G. Perez, Sean Kevin P. Sigua, Dan Michael A. Cortez, and Khatalyn E. Mata, "A modified key generation scheme of vigenere cipher algorithm using Pseudo-Random number and alphabetic extension," December 2021.

[8.] Dr. Mukesh Sharma and Smiley Gandhi, "Compression and encryption: An integrated approach," vol. 01. July 2012.

[9.] Jan Carlo T. Arroyo, Jenny A. Espadero, Marife A. Ganas, Randy F. Ardeña, and Ramcis N. Vilchez, "An efficient least significant bit image Steganography with secret writing and compression techniques," vol. 9, Philippines, June 2020.

[10.] Jan Carlo T. Arroyo and Allemar Jhone P. Delima, "Caesar cipher with Goldbach code compression for efficient cryptography," vol. 8. July 2020.

[11.] S.R. Kodituwakku and U. S.Amarasinghe, "Comparison of lossless data compression algorithms for text data," vol. 1. Sri Lanka.

[12.] Thamer Hassan Hameed and Haval Tariq Sadeeq, "Modified Vigenère cipher algorithm based on new key generation method," vol. 28. Iraq, November 2022.

[13.] Laurentinus, Harrizki Arie Pradana, and Dwi Yuny Sylfania, "Improving the SMS security and data capacity using advanced encryption standard and Huffman compression,"vol. 172. Indonesia, 2019,

[14.] YumnamKirani Singh, "Rectangular generalized vigenere cipher," vol. 9, pp75-81, India, 2021.

[15.] Md. Jayedul Haque and Mohammad Nurul Huda, "Study on data compression technique," vol. 159. Dhaka, July 2017.

[16.] Shrusti Porwal, Yashi Chaudhary, Jitendra Joshi, and, Manish Jain, "Data compression methodologies for lossless data and comparison between algorithms," vol. 02. India, March 2013.

[17.] Suherman and Andysah Putera Utama Siahaan, "Huffman text compression technique,"vol. 03. Indonesia , 2016.

[18.] A.K. Jaithunbi, S.Sabena, and L.Sairamesh, "Preservation of data integrity in the public cloud using enhanced vigenere cipher based obfuscation," Anna University Chennai, December 2021.

[19.] Rohit Kumar Gangwar, Mukesh Kumar, A.K.Jaiswal, and And Rohini Saxena, "Performance analysis of Image compression using a fuzzy logic algorithm,"vol. 5. India, April 2014.