A Novel Encryption Scheme over Group Codes using Vigenère Cipher

K Chittibabu[a], Jayanti Sravani[b], Prof. A Chandrasekhar[c]

[a]Research Scholar, Department of mathematics GITAM Deemed to be University, Visakhapatnam, India
[b]Research Scholar , Department of mathematics, GITAM Deemed to be University, Visakhapatnam, India
[b]Professor, Department of mathematics, GITAM Deemed to be University, Visakhapatnam, India
E mail: [a]121962101201@gitam.in, :[b]sjayanti@gitam.in, [c]cakkaped@gitam.edu

Article History: Received: 10 November 2020; Revised 12 January 2021 Accepted: 27 January 2021; Published online: 5 April 2021

Abstract: With the rapid growth in technology there is an imperishable need for maintaining security and integrity of data while communication. These needs can be met by the advancements in the field of cryptography which can be achieved using the concepts of Coding theory, Linear Algebra and Number Theory. In this paper we propose a mechanism where characters are assigned code words using an encoding function and encryption is performed using the concept of Vigenere cipher over group codes..

Keywords: Group codes, encoding functionVigenère cipher

1. Introduction

The security of data while communication is enhanced by developing a cipher which makes use of the concepts of basic mathematics [1][2][3][4]. In this paper we have proposed a method which has been worked out with the help of the concepts discussed below:

1.1 Group Code [4]

Group codes are the linear block codes which form a subgroup under a finite abelian group. Suppose that B^n consists of 'n' length code words consisting of either 0s or 1s. Then $|B^n| = 2^n$.An encoding function 'e' can be defined which maps code words of length 'm' to the code words of length 'n'. This can be defined as follows:

$$e : B^m \to B^n \text{ where } n > m \text{ by } e(x) = y \text{ for } x \epsilon B^m \text{ and } y \epsilon B^n$$

Here $x = x_1x_2x_3x_4 \dots\dots\dots x_m$ and $y = x_1x_2 \dots\dots x_mx_{m+1}x_{m+2}x_{m+3} \dots\dots\dots x_n$

where $x_i \text{ for } i = m+1, \dots\dots n$ can be represented as the sum of some or all of the $x_i s \text{ for } i = 1, \dots\dots m$ under XOR operation. ...(1)

Let 'P' be the $n-m \times m$matrix obtained from the above equations used to determine the values of the $x_i s for i = m+1, \dots\dots n.$

The entries in 'P' can be taken as '1'for the x_i bit present in the equation and '0' for the absent x_i bit where$i = 1,2, \dots, m.$

Then P^T(Transpose of P) is termed to be the parity matrix if the last $n-m \times n-m$submatrix of P^T forms an identity matrix. Then the codewords in B^n to which the code words in B^m are assigned can be put together in a set which will form a group under XOR operation. One can easily encode the code words in B^m to the codewords in B^n provided access to the parity matrix P^T.

1.2 Vigenere Cipher[2]

This cipher is named after Blaise de Vigenere who lived in the 16th century. This is a polyalphabetic cipher.

Let 'm' be any positive integer. Define P = C = K = $(Z_{26})^m$.

For a key $k = (k_1, k_2, \dots\dots k_m)$, we define

$$e_k = (x_1, x_2, \dots, x_m) = (x_1 + k_1, x_2 + k_2, \dots, x_m + k_m) and$$

$d_k = (y_1, y_2, \dots, y_m) = (y_1 - k_1, y_2 - k_2, \dots, y_m - k_m)$where all operations are performed in Z_{26}.

Example:

Suppose that the alphabetic letters are mapped to the residues modulo 26 as A,B,…,Z to 0,1,…,25 respectively. Suppose that the keyword = M A N ⇒ 12 0 13 i.e. m=3

- *Encryption*

Plaintext	C R Y P TO S Y S T E M
Convert to residues modulo under 26	2 17 24 15 19 14 18 24 18 19 4 12
Divide them into groups of length m=3	2 17 24 | 15 19 14 | 18 24 18 | 19 4 12
Add keyword module 26 to each group	2 17 24 | 15 19 14 | 18 24 18 |19 4 12
	12 0 13 | 12 0 13 | 12 0 13| 12 0 13
	14 17 11 | 1 19 1 | 4 24 5 | 1 4 25
Cipher text	O R L B T B E Y F B E Z

- **Decryption**

While decrypting, we use the same keyword but we would subtract it module 26 from the cipher text instead of adding it.

2. Proposed Method

In this paper we proposed a crypto-system over group codes and a classical cipher Vigenere cipher. The process of generating key and carrying out the encryption and decryption is listed below:

1.1 Key generation:

- Let K = key matrix and A = invertible matrix compatible with K.

We consider Fibonacci Q_λ-matrix [4] for this invertible matrix 'A' whose order is shared to the receiver secretly where $Q_\lambda$is a matrix of order$* \lambda$.

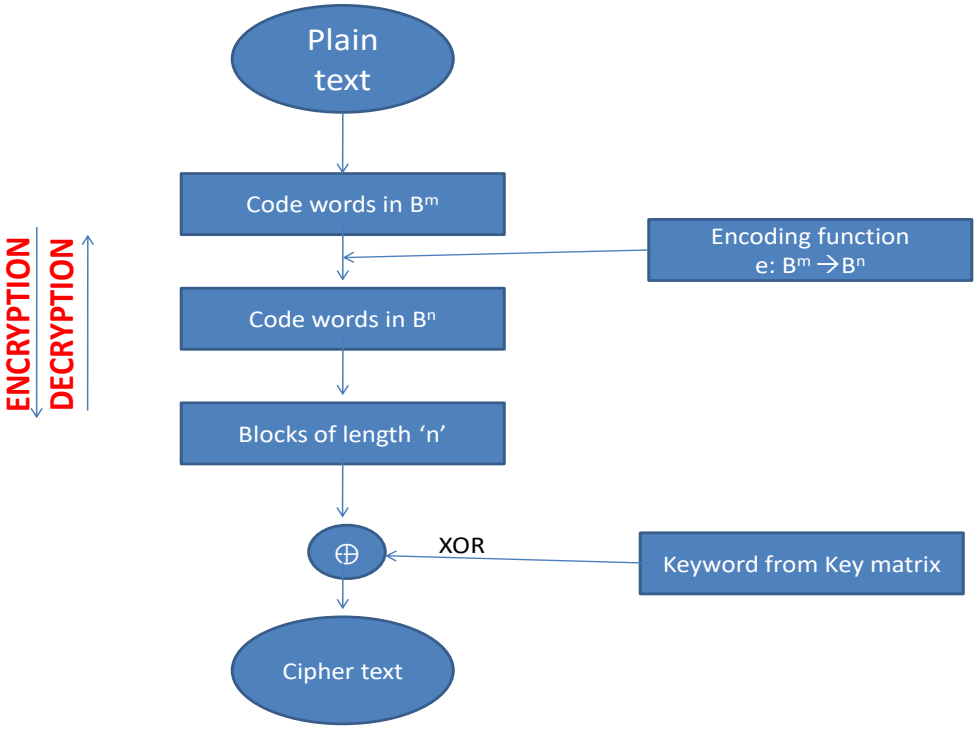
$$Q_\lambda = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

$$Q_\lambda{}^{-1} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & -1 & -1 & \dots & -1 \end{bmatrix}$$

- Compute $P = A * K$.
- 'P' is sent by sender to the receiver via a secure channel.
- The receiver on receiving 'P' retrieves 'K' by computing $A^{-1} * P = A^{-1} * (A * K) = K$

1.2 Encryption and Decryption process:

- Consider a group code B^m to which the characters are assigned. Define an encoding function $e : B^m \rightarrow B^n \; where \; n > m$ as discussed in (1).
- The matrix 'P' of order $n \times n$ is determined using the codewords in B^n to which the codewords in B^m are assigned. This 'P' is exchanged between the sender and the receiver as discussed in key generation.
- Each row in 'P' is a codeword of length 'n' which is operated on the encoded message after grouping it into blocks of 'n' length each. Therefore 'n' number of keywords each of length 'n' is operated under XOR on the encoded message. Therefore, length of the keyword used in this encryption process which uses the basic idea of Vigenere cipher is n^2. The cipher text can be sent as blocks of 'n' bits each.
- While performing decryption each row in the matrix P is operated under XOR with the cipher text to retrieve the codeword in B^n.
- The codeword in B^m can be received by dropping the extra bits added in B^n.Thus, the character equivalent to the codeword can be retrieved at the receiver's end.
- The flowchart explains the process of Encryption and Decryption:



3. Example

We understand the proposed method through an example where we limit the number of characters to 26 which are the English Alphabets.

Consider a groupcode of bit length 5. Then the total number of codewords will be $2^5 = 32$.

We assign the 26 alphabets to these codewords as listed below:

Code word in B^5	**Character equivalent**
00000	-
00001	A
00010	B
00011	C
00100	D
00101	E
00110	F
00111	G
01000	H
01001	I
01010	J
01011	K
01100	L
01101	M
01110	N
01111	O
10000	P
10001	Q
10010	R
10011	S
10100	T
10101	U
10110	V

10111	W
11000	X
11001	Y
11010	Z
11011	-
11100	-
11101	-
11110	-
11111	-

Now let us define an encoding function $e: B^5 \rightarrow B^6 bye(x_1x_2x_3x_4x_5) = x_1x_2x_3x_4x_5x_6\, where x_6 = x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus x_5$. Then the equivalent code words in B^6 for the code words in B^5 can be given as below:

Code words in B^5	**Code words in B^6**
00000	000000
00001	000011
00010	000101
00011	000110
00100	001001
00101	001010
00110	001100
00111	001111
01000	010001
01001	010010
01010	010100
01011	010111
01100	011000
01101	011011
01110	011101
01111	011110
10000	100001
10001	100010
10010	100100
10011	100111
10100	101000
10101	101011
10110	101101
10111	101110
11000	110000
11001	110011
11010	110101
11011	110110
11100	111001
11101	111010
11110	111100
11111	111111

Suppose that the plaintext is "CRYPTOGRAPHYENSURESSECURITY" and the key matrix is $K=\begin{pmatrix}0&0&0&0&0&1\\0&0&0&0&1&0\\0&0&0&1&1&1\\1&0&0&0&1&1\\1&1&0&0&0&1\\1&1&1&0&1&1\end{pmatrix}$.

3.1 Key exchange mechanism:

- Suppose that the invertible matrix used to exchange key is Fibonacci $Q_\lambda$matrix [4].
- We consider Q_6 matrix for exchanging key matrix since the code words in B^6 is considered.

Then $Q_6=\begin{pmatrix}1&1&1&1&1&1\\1&0&0&0&0&0\\0&1&0&0&0&0\\0&0&1&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\end{pmatrix}$ and $Q_6{}^{-1}=\begin{pmatrix}0&1&0&0&0&0\\0&0&1&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\\1&-1&-1&-1&-1&-1\end{pmatrix}$

- Sender sends the matrix P over a secure channel to Receiver[6] where $P = A * K=\begin{pmatrix}3&2&1&1&4&5\\0&0&0&0&0&1\\0&0&0&0&1&0\\0&0&0&1&1&1\\1&0&0&0&1&1\\1&1&0&0&0&1\end{pmatrix}$
- Receiver on receiving P retrieves K as follows: $K = A^{-1} * P = \begin{pmatrix}0&0&0&0&0&1\\0&0&0&0&1&0\\0&0&0&1&1&1\\1&0&0&0&1&1\\1&1&0&0&0&1\\1&1&1&0&1&1\end{pmatrix}$

3.2 Encryption and Decryption process:

- *Encryption*

1. Plaintext ⇒ CRYPTOGRAPHYENSURESSECURITY
2. Equivalent codewords in B^5

000111001011001100001010001111001111001000001100000000110000010001100111011001010111010011101011001000101100111001100101000111010110010010011010011001

3. Equivalent codewords in B^6

000110100100110011100001101000011110001111100100000011100001000011100001
010001110011111011000101001110110011111010111001000010101001111001110010100001101010111001000100101010000110011

4. Convert them into blocks each of length 6 and perform XOR operation with the keywords listed as rows in the key matrix K. This is listed below:

Character equivalent codeword in B^6	XOR with Key	Cipher text
000110	000001	000111
100100	000010	100110
110011	000111	110100
100001	100011	000010
101000	110001	011001
011110	111011	100101
001111	000001	001110
100100	000010	100110
000011	000111	000100
100001	100011	000010
010001	110001	100000
110011	111011	001000

001010	000001	001011
011101	000010	011111
100111	000111	100000
101011	100011	001000
100100	110001	010101
001010	111011	110001
100111	000001	100110
100111	000010	100101
001010	000111	001101
000110	100011	100101
101011	110001	011010
100100	111011	011111
010010	000001	010011
101000	000010	101010
110011	000111	110100

5. Final Cipher text is sent in the form of blocks of length 6 bits each as 000111,100110,110100,... .

- *Decryption*

1. Convert the received Ciphertext to the equivalent codeword by operating XOR with the keywords as follows:

Cipher text	XOR with Key	Codeword in B^6
000111	000001	000110
100110	000010	100100
110100	000111	110011
000010	100011	100001
011001	110001	101000
100101	111011	011110
001110	000001	001111
100110	000010	100100
000100	000111	000011
000010	100011	100001
100000	110001	010001
001000	111011	110011
001011	000001	001010
011111	000010	011101
100000	000111	100111
001000	100011	101011
010101	110001	100100
110001	111011	001010
100110	000001	100111
100101	000010	100111
001101	000111	001010
100101	100011	000110
011010	110001	101011

011111	111011	100100
010011	000001	010010
101010	000010	101000
110100	000111	110011

2. After receiving the codewords convert them back to their equivalents in $B^5$by dropping the last bit in each codeword of length 6. On doing so we receive,
000111001011001100001010001111001111001000011000000001100000100011001110110010101110100111
010110010001011001110011001010001110101100100100110100110001

3. Decrypt them back to their original character equivalent by grouping them to a block of bit length 5. Finally we obtain the plaintext as :

00011|10010|11001|10000|10100|01111|00111|10010|00001|10000|00001|10000|01000|11001|11011|00101|01110|10011|10101|10010|00101|10011|10011|00101|00011|10101|10010|01001|10100|11001

⇒CRYPTOGRAPHYENSURESSECURITY

4. Cryptanalysis

Cryptanalysis is the art of tracing information communicated between the sender and the receiver in a disguised format without having access to the key. Different attacks are possible for stealing the message communicated between the sender and the receiver which are Known Plain text Attack, Chosen Plain text Attack, Cipher Text Only Attack, Man-In-The- Middle Attack and Brute force attack.

The confidentiality of the method used in the Cryptosystem designed is mainly dependent on:

- the encoding function,
- the Key matrix and
- the position where the parity bits are added to the codeword in B^m to convert it into a codeword in B^n.

3.1 Known plain text attack: In this type of attack some part of the plain text-cipher text combinations are known to the attacker. This attack can be easily carried out but is the least prone.

Suppose that the some part of the plain text-cipher text combinations along with the complete cipher text and the methodological approach of the cryptosystem are known to the attacker. Then in order to retrieve the key matrix in the method discussed above one must have access to the values of *'n'* and *'m'*.

The keywords used in the key matrix are selected from the 2^n codewords. Therefore, the number of possible key matrices is 2^nC_n. Thus it is highly infeasible to carry out an exhaustive key search for deducing a key by sufficiently increasing the value of 'n' even if the value of '*n'* is known.

On considering the worst case of values of *'n'* and *'m'* being unveiled, the security of the cryptosystem can be enhanced by hiding the position of the parity bits added to the codeword in B^m to convert it into a codeword in B^n.

In the example furnished above, the value of $m = 5\ and\ n = 6$. Therefore the number of parity bits added$= n - m = 1$. The position of the parity bit can easily be traced by placing the parity bit in the possible 6 positions. This problem can be overcome by sufficiently increasing the values of *'n'*, *'m'* and '*n-m'* and not disclosing the encoding function.

3.2 Chosen Plain text Attack: Suppose that the attacker has access to the cryptosystem along with the cipher text and the attacker tries to fetch some random plaintext-cipher text combinations. Then the situation is similar to that of Known Plain text Attack except for the point that some part of the original plain text is known in case of KPA to guess the original message unlike CPA. Thus the security of the proposed method can be enhanced as discussed in the section 4.1.

3.3 Cipher Text Only Attack: Suppose that the attacker has access to the Cipher text only. This attack is the most prone but a difficult one to crack. Let us assume a situation where the methodological approach for the proposed cryptosystem is known to the attacker. Then in order to carry out decryption one must have access to the key matrix to proceed further which is transmitted over a secure channel. Thus Cipher text only Attack is not possible.

3.4 Man-In-The-Middle-Attack: In this attack the key exchanged between the sender and the receiver is attacked by a third person. In order to overcome this problem the key matrix used for encrypting and decrypting information is exchanged by using an invertible Q_λ matrix.

Thus the designed cryptosystem is resistant to cryptanalytic attacks.

5. Conclusion

- In this paper group codes have been used for performing encryption and decryption. In general, the parity bits added to convert 'm' length codeword to 'n' length codeword is used for detecting and correcting errors but here they are used to create confusion to the attacker. Also the position to which the parity bits are added can be changed to even more complicate the cipher by sending the position where the parity bit has been placed as a key. This can be carried out by using a predefined function in Mathematica or any other language.
- By increasing the values of 'm' and 'n' sufficiently large, the security of the cipher can be enhanced which is discussed in section 4.
- Any invertible matrix can be used for exchanging key matrix but the usage of Q_λ matrix helps us to reduce memory while exchanging matrix since it is enough to send the order of the matrix to the receiver for performing decryption.

The novelty of the cryptosystem can be witnessed in the fact that the transmitted Cipher text consists of only 0s and 1s. The algorithm can be made even more complex by treating these 0s and 1s as binary digits and grouping them together to convert them to a simple decimal equivalent number that can be transmitted as Cipher text to the receiver..

References

César Polcino Mllies.(2019).Group algebras and coding theory: a short survey:Revista Integración, vol. 37, no. 1

D.R. Stinson & M B Paterson.Cryptography Theory and Practice: Fourth edition, CRC Press Taylor & Francis Group

Garcia Pillado C, Gonzalez S, Martinez C, Markov V &Nechaev A..(2013).Group codes over non-abelian groups: Journal of Algebra and its Applications ,Vol 12, No.7,1350037

J P Tremblay & R Manohar .Discrete Mathematical Structures with Applications to Computer Science:TATA McGraw-Hill Edition

Kalika Prasad & Hrishikesh Mahato.(2020).Cryptography using generalized Fibonacci matrices with Affine-Hill cipher: arXiv:2003.11936v1 [cs.CR] 25 Mar 2020

https://www.calculator.net/matrix-calculator.html