An Improved Cryptographic Technique to Encrypt Text using Double Encryption

Yashpalsingh Rajput
Research Scholar,
Government College of Engineering, Aurangabad

Dnyaneshwar Naik
Research Scholar,
Government College of Engineering, Aurangabad

Charudatt Mane
Research Scholar,
Government College of Engineering, Aurangabad

ABSTRACT
This work proposes an improved scheme to encrypt the plain text message for its security. All the conventional encryption techniques are very weak and brute force attack and traditional cryptanalysis can be used to easily determine the plain text from encrypted text. In this work of encryption technique, a new concept of conventional ceaser cipher algorithm with hill cipher algorithm is used to make encryption technique more secure and stronger than the earlier concept. The plain text is encrypted in such a way that it becomes difficult to decrypt it. The proposed system is divided into two phases. In first phase, the plain text message is converted to first encrypted text using a new substitution approach which uses poly-alphabetic cipher technique. The encryption is done using variable length key which depends on the string length. In the second phase, the hill cipher technique is applied on the first encrypted text to produce new encrypted text or cipher text. At the receiver end, if the receiver has appropriate decryption key, he can generate the text message similar to the original message.

This paper is organized into following sections. Section 1 contains a general introduction to the cryptography, types of cryptography and hill cipher. Section 2 contains literature review on some classical and modern text encryption techniques. Section 3 contains description of the proposed system. Finally paper is concluded in the Section 4.

General Terms
Information Security

Keywords
Cryptography, Plaintext, Ciphertext, Hill Cipher

1. INTRODUCTION
In today's world, it is impossible to imagine without web or internet. This modern era is dominated by paperless transactions in business, private or government offices by means of use of E-mail messages, E-cash transactions, etc. Due to this there is a great need of transmission of data through internet. In various business sectors, there may be sensitive and confidential information like banking transactions, credit information, government information, sensitive information is transferred over web using E-mails, etc. The confidentiality, authentication and integrity of such important information should be maintained and protected [1]. To protect this type of sensitive information from unauthorized access, there is a great need of security. To protect sensitive text information from unauthorized access various encryption techniques are used. Encryption technique is first used by Julius Caesar. When Julius Caesar sent messages to his generals, he didn't trust his messengers. So he replaced every A in his messages with a D, every B with an E, and so on through the alphabet. Only someone who knew the "shift by 3" rule could decipher his messages. Ex. For a message "secret", if shift by 3 rule (+3) is applied for number of times, then encrypted message will change as follows:

secret $\xrightarrow{+3}$ vhfufw $\xrightarrow{+3}$ ykixiz $\xrightarrow{+3}$ and so on.

To decrypt it again start from last encrypted message and reverse shift the characters by 3 (-3) and finally original message can be obtained.

Encryption plays a main role in information security. The encryption techniques methods are used to convert our text information in a non-readable form at sender side and convert that information in readable form again at receiver side.

Cryptology is the study of cryptosystems. It can be divided into two competing skills – concealment and solution.

The concealment portion of cryptology is called cryptography. The aim of cryptography is to render a message incomprehensible to the unauthorized reader. Cryptography concerns with the design of cryptosystems. The process of creating non-readable text information or cipher so that only intended person is only able to read the information is called Cryptography. It uses mathematical algorithms to encrypt and decrypt data. It enables you to store sensitive information or transmit it across insecure networks so that it cannot be read by anyone except the intended recipient. Cryptography is often called "code making".

The solution portion of cryptology is called cryptanalysis. While cryptography is the science of securing data, cryptanalysis is the science of analyzing and breaking secure communication i.e. breaking of cryptosystems. Cryptanalysis is often called "code breaking". Figure 1 shows the encryption process and Figure 2 shows the decryption process [2].

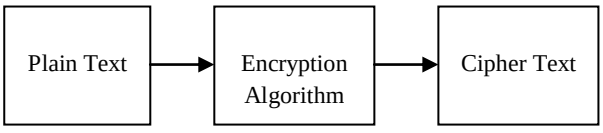

Fig. 1: Encryption Process

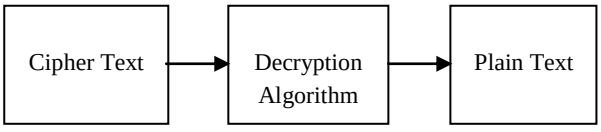

Fig. 2: Decryption Process

As shown in figure 1 and figure 2, Cryptography consists of two main steps: Encryption & Decryption. Using encryption process text information is converted non-readable form. Decryption is reverse of encryption process. Plaintext information is the intended original message. Cipher text

information is the coded message. There are two techniques of plain text encryption: Substitution Technique and Transposition Technique.

In substitution technique, the letters of plain text are replaced by other letters or any number or by symbols. Ex. Caesar cipher, Hill cipher, etc. In transposition technique, some sort of permutation is performed on plaintext. Ex. Rail Fence method, Columnar method, etc.

1.1 Terminologies in Cryptography

An encryption technique has five ingredients [2]:

1. Plaintext: This is the original intelligible message or data that is fed into the algorithm as input.

2. Encryption Algorithm: The encryption algorithm performs various substitutions and transformations on the plaintext.

3. Key: The key is also input to encryption algorithm. The key is a value independent of the plaintext and of the algorithm. The algorithm will produce a different output depending on the specific key being used at the time.

4. Ciphertext: This is the scrambled message produced as output. It depends on the plaintext and secret key.

5. Decryption Algorithm: This is essentially the encryption algorithm run in reverse. It takes the ciphertext and the secret key and produces the original plaintext.

1.2 Types of Cryptography

Cryptography is a technique in which secret messages are transferred in the encrypted form from sender to receiver over the communication line.

Cryptographic techniques are very useful to protect secret information. They protect the secret or confidential information by converting the information to some unintelligible form using a key. To retrieve the information, the encrypted information should be converted back to original information using some keys. Based on the key, the cryptography can be classified into two categories [1]:

1. Shared key cryptography

2. Public key cryptography

Shared key cryptography also called symmetric key cryptography or private key cryptography or secret key cryptography in which, same key is used for encryption and decryption i.e. both the sender and the receiver know the same key. Ex. DES, 3DES, AES, etc. Messages are encrypted by the sender using the key and decrypted by the receiver using the same key.

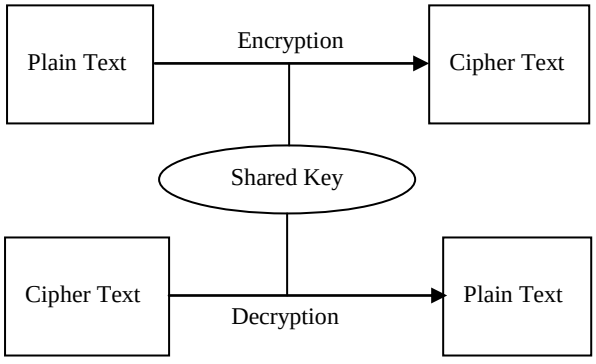

Fig. 3: Shared Key Cryptography

Figure 3 shows process of secret key cryptography [2]. Here same key is shared by both sender and receiver for encryption and decryption.

Public key cryptography also called asymmetric key cryptography which uses different keys for encryption and decryption. Ex. RSA, Digital signature scheme, etc. The public key is known to all the receivers, is used for encrypting the plaintext message. The private key is known only to the user of that key. With public key cryptography, keys work in pairs of matched public and private keys. Figure 4 shows the process of public key cryptography where public key used by sender for encryption and all the receivers use their private keys for decryption. Messages encrypted using the public key cannot be decrypted using the public key. Public key encrypted messages can only be decrypted using corresponding private key which is kept secure.

Asymmetric key cryptography is very slower and has very higher computational costs which are most of the time prohibitive for multimedia data. Symmetric key cryptography is fast, comparatively lower cost and may be used for multimedia data.

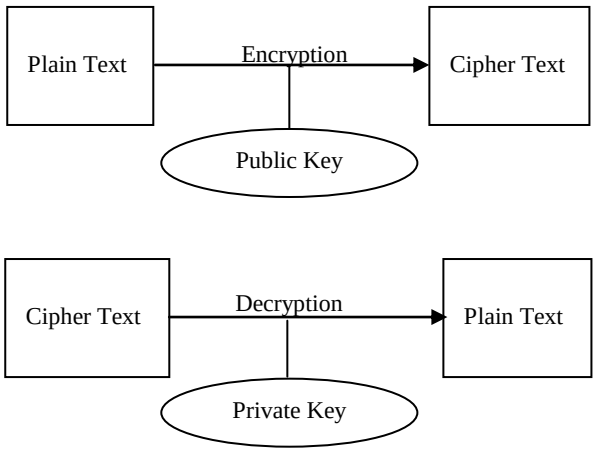

Fig. 4: Public Key Cryptography

Hash function is another type of cryptography which makes use of some mathematical transformation.

1.3 Hill Cipher

The Hill cipher (HC) algorithm is one of the famous and known symmetric algorithms in the field of cryptography. It is a poly-alphabetic cipher proposed by the mathematician Lester Hill in 1929 in the journal of mathematics. Hill cipher requires a matrix based polygraphic system [5]. The Hill cipher takes m successive plaintext characters and substitutes for them m ciphertext characters.

The core of Hill cipher is matrix manipulations. For encryption, algorithm takes m successive plaintext characters and instead of that substitutes m cipher characters.

For example m=2; {abcdef…}=ab cd ef… or abc def… and so on.

In Hill cipher, each character is assigned a numerical value like a = 0, b = 1, ... , z = 25. The substitution of ciphertext characters in the place of plaintext characters leads to m linear equation. For m = 3, the system can be described as follows:

$$C=KP$$

where C and P are column vectors of length 3, representing the plaintext and ciphertext, respectively and K is a 3 x 3 matrix, which acts as key for encryption. All operations performed with modulus of 26. In Hill cipher key is an invertible m x m matrix, where m is block length. Decryption process uses inverse of matrix K. The inverse matrix K^{-1} of a matrix K is defined by following equation.

$$KK^{-1} = K^{-1}K=I$$

where, I is the Identity matrix. But the inverse of matrix does not always exist and when it exists, it satisfies above equation. The inverse matrix K^{-1} is used to decrypt the ciphertext. In general it can be written as follows:

Encryption Process:

$$C=E_k(P)=Kp$$

Decryption Process:

$$P=D_k(C)=K^{-1}C= K^{-1}Kp=P$$

If the block length considered as m, there are 26^m different m characters blocks are possible.

2. LITERATURE SURVEY

This section consists of brief description of few classical and modern plain text encryption techniques.

2.1 Caesar Cipher

The Caesar cipher is one of the earliest known and simplest ciphers. The method is named after Julius Caesar, who apparently used it to communicate with his generals. It is a type of substitution cipher in which each letter in the plaintext is 'shifted' a certain number of places down the alphabet. For example, with a shift of 1, A would be replaced by B, B would become C, and so on. The Caesar cipher offers essentially no communication security, and it can be easily broken even by hand.

2.2 Playfair Cipher

The Playfair cipher is a polygraphic cipher which enciphers two letters at a time. The Playfair cipher was the first practical digraph substitution cipher. The scheme was invented in 1854 by Charles Wheatstone, but was named after Lord Playfair who promoted the use of the cipher [4]. The technique encrypts pairs of letters, instead of single letters as in the simple substitution cipher. The Playfair is significantly harder to break since the frequency analysis used for simple substitution ciphers does not work with it. Frequency analysis can still be undertaken, but on the 25*25=625 possible digraphs rather than the 25 possible monographs. Frequency analysis thus requires much more ciphertext in order to work.

It was used for tactical purposes by British forces in the Second Boer War and in World War I and for the same purpose by the Australians during World War II. This was because Playfair is reasonably fast to use and requires no special equipment. A typical scenario for Playfair use would be to protect important but non-critical secrets during actual combat. By the time the enemy cryptanalysts could break the message the information was useless to them.

2.3 Affine Cipher

The Affine cipher is a special case of the more general monoalphabetic substitution cipher. The cipher is less secure than a substitution cipher as it is vulnerable to all of the attacks that work against substitution ciphers, in addition to other attacks. The cipher's primary weakness comes from the fact that if the cryptanalyst can discover (by means of frequency analysis, brute force, guessing or otherwise) the plaintext of two ciphertext characters, then the key can be obtained by solving a simultaneous equation.

2.4 Vigenere Cipher

Vigenere cipher was proposed by Blaise de Vigenere in the 16th century. Vigenere cipher is poly-alphabetic substitution cipher in which a single plain text letter can be converted into multiple cipher text letters. This conversion depends on the position of the letter in the plain text e.g. c may be converted into g because it is at position 3 in the plain text but c can be changed into z because its position in the plain text is 5. Vigenere cipher makes use of Vigenere table of size 26 X 26.

2.5 Blow-Fish Cipher

Blowfish cipher was developed by Bruce Schneier. Blowfish is a variable-length key, 64-bit block cipher. The algorithm consists of two parts: a key-expansion part and a data-encryption part. Key expansion converts a variable-length key of at most 56 bytes into several sub key arrays totaling 4168 bytes. Data encryption occurs via a 16-round Feistel network. Each round consists of a key-dependent permutation, and a key- and data-dependent substitution.

2.6 Rail Fencing Cipher

Rail fencing technique involves: Writing plain text message as a sequence of diagonal and reading it as a sequence of row to produce cipher text. In a Rail Fence cipher, after removing the spaces from the original message, write the characters in the message in the zig-zag pattern. The key for the Rail Fence cipher is just the number of rails.

2.7 Modern Ciphers

Modern ciphers use both substitution and transposition to encrypt the message that increases the security of data. The data is encrypted in blocks instead of single characters at time. The well-known example of block cipher is Data Encryption Standard (DES). DES uses 56-bits key and encrypt 64-bits of data as a single block.

AES uses the same key for both encryption and decryption. The AES standard states that the algorithm can only accept a block size of 128 bits and a choice of three keys - 128, 192, 256 bits. Depending on which version is used, the name of the standard is modified to AES-128, AES-192 or AES- 256 respectively.

The Rijndael Algorithm is the new Advanced Encryption Standard (AES) recommended by the US National Institute of Standards and Technology (NIST) for protecting sensitive, unclassified government information. Since Rijndael is an iterated block cipher, the encryption or decryption of a block of data is accomplished by the iteration of a specific transformation. As input, Rijndael accepts one-dimensional 8-bit byte arrays that create data blocks. The plaintext is input and then mapped onto state bytes. The cipher key is also a one-dimensional 8-bit byte array. With an iterated block cipher, the different transformations operate in sequence on intermediate cipher results.

The columnar transposition cipher is a fairly simple, easy to implement cipher. It is a transposition cipher that follows a simple rule for mixing up the characters in the plaintext to form the ciphertext. Although weak on its own, it can be combined with other ciphers, such as a substitution cipher, the combination of which can be more difficult to break than either cipher on its own.

2.8 Cryptanalysis

In Cryptanalysis, the attacker uses various methods to get the plain text from the cipher text. They try to find out the way in which plain text is converted into cipher text and the encryption key used. Various methods were used for identifying ciphers. Identification of permutation, substitution and Vigenère ciphers was done using frequency analysis. An attempt was made to identify block ciphers like DES and Blowfish using pattern recognition method. Other ciphers like stream cipher SEAL and Enhanced RC6 have been identified using neural networks.

3. PROPOSED SYSTEM

In this section, the main idea used for the proposed system is described. The proposed system, which is used to encrypt the plain text message, is divided into the following 2 main phases:

Phase-1: Improved substitution cipher**.**
Phase-2: Hill Cipher technique.
In phase 1, some improvements are applied on classical substitution cipher by applying dynamic key for each letter in a string. The dynamic key depends on the length of the string to be encrypted and for each letter in the string to be encrypted key changes as encryption proceeds through the length of the string. The use of variable and dynamic key for each letter makes the system more secure and unbreakable.

In phase 2, classical hill cipher is applied which uses 3 X 3 key matrix for encryption. Use of Hill cipher makes the string unstructured due to which it becomes difficult to get the original text string.

3.1 Algorithm

Following is the algorithm for proposed plain text encryption system. Algorithm is explained with the example using a string "Encryption" as an input.

1. Take the input string.

   Ex. Encryption

2. Find the length of the input string and use it as keylength.

   Ex. Length of "Encryption" = 10

3. Find the odd and even positioned character from the given string & divide them into two separate groups.

4. Treat the group of odd positioned & even positioned characters as two separate strings. Store odd positioned characters into "OddGroup" and even positioned characters into "EvenGroup".

   For above example:

   OddGroup= Ecyto

   EvenGroup= nrpin

5. Reverse both the strings i.e. OddGroup and EvenGroup and store them in ReverseOddGroup and ReverseEvenGroup, respectively.

   ReverseOddGroup= otycE

   ReverseEvenGroup= niprn

6. Now apply Caesar cipher on the reversed odd positioned character group string for which the key depends on the length of the original input string. i.e. apply Caesar cipher on the string from right end to left end i.e. from last character to first character with the key value equal to length of the original input string for last character & will be decreased by 1 for each character till reached to first character.

   The ReverseOddGroup string "otycE" becomers "uaglO".

7. For next string i.e. reversed even positioned characters group, apply Caesar cipher for which the key depends on the length of the original input string. i.e. apply Caesar cipher on the string from left end to right end i.e. from first letter to last character with the key value equal to length of the original input string for first character & will be decreased by 1 for each character till reached to last character.

   The ReverseEvenGroup string "niprn" becomers "xrxyt".

8. Apply steps 1 to 7 on all the other strings.

9. Replace the character at the start position, middle position & end position of the string, with a new character whose ASCII value is equal to the ASCII value of the original character at that place + length of the original string - 26 (for alphabets) or 10 (for numbers).

10. Apply hill cipher technique.

4. CONCLUSION

The proposed system is a poly-alphabetic substitution cipher whose substitution key depends on the string length, which makes the key dynamic and variable for each string. The proposed system is an improvement over traditional substitution cipher encryption methods. In addition hill cipher on new string makes it more secure and unbreakable. Due to use of hill cipher the information becomes unstructured. The text encrypted using proposed method, can't be decrypted using traditional crypto-analysis tools. The brute force attack technique also fails to decrypt the text which is encrypted by using proposed technique.

5. ACKNOWLEDGMENT

The authors express their gratitude to all the faculty members of the Computer Science & Engineering Department of Government College of Engineering, [Autonomous], Aurangabad for their support and enthusiasm.

6. REFERENCES

[1] Gary C. Kessler, "An Overview of Cryptography", 2013.

[2] Ramandeep Sharma, Richa Sharma, Harmanjit Singh, "Classical Encryption Techniques", International Journal of Computer & Technology, Vol. 3, No. 1, August 2012, pp. 84 – 90.

[3] "CRYPTOGRAPHY", https://en//.wikipedia.org/wiki/cryptography

[4] Henk C. A. van Tilborg, "FUNDAMENTALS OF CRYPTOLOGY", pp. 9 - 21.

[5] M. Nordin A. Rahman, A.F.A. Abidin, Mohd Kamir Yusof, N.S.M. Usop, "Cryptography: A New Approach of Classical Hill Cipher", IJSA, Vol. 7, No. 2, March 2013.

[6] Kashish Goyal, Supriya Kinger "Modified Caesar Cipher for Better Security Enhancement", International Journal of Computer Applications, Vol. 73, No. 3, July 2013.

[7] Brian Worthingotn, "An Introduction to Hill Ciphers Using Linear Algebra" 2010.

[8] Jawahar Thakur, Nagesh Kumar, "DES, AES and Blowfish: Symmetric Key Cryptography Algorithms Simulation Based Performance Analysis", International Journal of Emerging Technology and Advanced Engineering, Vol. 1, Issue 2, 2011.

[9] Ajit Singh, Aarti Nandal and Swati Malik, "Implementation of Caesar Cipher with Rail Fence for Enhancing Data Security", International Journal of Advanced Research in Computer Science and Software Engineering. Vol. 2, Issue 12, December 2012.

[10] Amogh Mahapatra, Rajballav Dash "Data Encryption and Decryption by Using Hill Cipher Technique and Self Repetitive Matrix", in 2007.