A cryptosystem based on Vigenère cipher with varying key

Quist-Aphetsi Kester

Abstract—Privacy is one of the key issues information security addresses. Through encryption one can prevent a third party from understanding raw data during signal transmission. The encryption methods for enhancing the security of digital contents has gained high significance in the current era of breach of security and misuse of the confidential information intercepted and misused by the unauthorized parties.

This paper sets out to contribute to the general body of knowledge in the area of cryptography application and by developing a new way of implementing Vigenère cipher encryption algorithm by automatically changing the cipher key after each encryption step. The new method will use successive keys that will be dependent on the initial key value during the encryption process. The algorithm ultimately makes it possible for encryption and decryption of the text. The ciphertext will have different encryption key pattern and the Vigenère cryptosystem will be more difficult to decipher using frequency attack. The implementation will be done using Java programming.

Index Terms—Cryptography, Encryption, Vigenère, key.

I. INTRODUCTION

The broaden utilization of digital media for information transmission through secure and unsecured channels exposes messages sent via networks to intruders or third parties. Encryption of messages in this modern age of technology becomes necessary for ensuring that data sent via communications channels become protected and made difficult for deciphering. Enormous number of transfer of data and information takes place through internet, which is considered to be most efficient though it’s definitely a public access medium. Therefore to counterpart this weakness, many researchers have come up with efficient algorithms to encrypt this information from plain text into ciphers [1].

In cryptography, encryption is the process of transforming information using an algorithm to make it unreadable to anyone except those possessing special knowledge, usually referred to as a key. The result of the process is encrypted information. The reverse process is referred to as decryption [2]. There two main algorithmic approaches to encryption, these are symmetric and asymmetric. Symmetric-key algorithms [3] are a class of algorithms for cryptography that use the same cryptographic keys for both encryption of plaintext and decryption of cipher text. The keys may be identical or there may be a simple transformation to go between the two keys. The keys, in practice, represent a shared secret between two or more parties that can be used to maintain a private information link [4]. This requirement that both parties have access to the secret key is one of the main drawbacks of symmetric key encryption, in comparison to public-key encryption. Typical examples symmetric algorithms are Advanced Encryption Standard (AES), Blowfish, Tripple Data Encryption Standard (3DES) and Serpent [5].

Asymmetric or Public key encryption on the other hand is an encryption method where a message encrypted with a recipient's public key cannot be decrypted by anyone except a possessor of the matching private key, presumably, this will be the owner of that key and the person associated with the public key used. This is used for confidentiality. [6]. Typical examples of asymmetric encryption algorithms are Rivest Shamir Adleman (RSA), Diffie–Hellman key exchange protocol and Digital Signature Standard(DSS), which incorporates the Digital Signature Algorithm (DSA).

Modern day cryptography entails complex and advance mathematical algorithm are applied to encryption of text and cryptographic techniques for image encryption based on the RGB pixel displacement where pixel of images are shuffled to obtained a cipher image [7].

This research is aimed at contributing to the general body of knowledge in the area of the application of cryptography by developing a new way of implementing Vigenère cipher encryption algorithm by automatically changing the cipher key after each encryption step. The successive keys will be dependent on the initial key value during the encryption process. The algorithm ultimately makes it possible for encryption and decryption of the text and also makes the Vigenère cryptosystem more difficult against frequency attack using varying keys. The paper has the following structure: section II consist of related works, section III of the methodology, section IV The algorithm, section V Implementation, section VI Results and Analysis and section VII concluded the paper.

II. RELATED WORKS

Caesar cipher, also known as the shift cipher, is one of the simplest and most widely known classical encryption techniques. It is a type of substitution cipher in which each letter in the plaintext is replaced by a letter some fixed number of positions down the alphabet. For example, with a shift of 3, A would be replaced by D, B would become E, and so on. The encryption step performed by a Caesar cipher is often incorporated as part of more complex schemes, such as the Vigenère cipher, and still has modern application in the
The Vigenère cipher is a method of encrypting alphabetic text by using a series of different Caesar ciphers based on the letters of a keyword. It is a simple form of polyalphabetic substitution [8][9]. The Cipher spoils the statistics of a simple Caesar cipher by using multiple Caesar ciphers. The technique is named for its inventor, Blaise de Vigenère from the court of Henry III of France in the sixteenth century, and was considered unbreakable for some 300 years [13]. Vigenère can also be viewed algebraically. If the letters A–Z are taken to be the numbers 0–25, and addition is performed modulo 26, then Vigenère encryption \( E \) using the key \( K \) can be written,\[ E_n(x) = (x + n) \mod 26 \]

whereas \( M = M_0 \ldots M_n \) is the message, \( C = C_0 \ldots C_n \) is the ciphertext and \( K = K_0 \ldots K_m \) is the used key.

Thus Given \( m \), a positive integer, \( P = C = (Z_{26})^n \), and \( K = (k_1, k_2 \ldots km) \) a key, we define:

Encryption: \[ Ek(p_1, p_2 \ldots pm) = (p_1+k_1, p_2+k_2 \ldots pm+km) \mod 26 \]

Decryption: \[ dk(c_1, c_2 \ldots cm) = (c_1-k_1, c_2-k_2 \ldots cm- km) \mod 26 \]

Example:

Plain text: CRYPTOGRAPHY

Key: LUCKLUCLKUC

Ciphertext: NLAZEBLJ

The algebraic description of the extended version is similar to that of the original cipher. It uses modulo 92 instead of modulo and cipher text \( C \) is derived using a sequence different from plain text sequence \( P \).

\[ C_i = E_k(C_i) \] and decryption \( D 

\[ C_i = E_k(C_i) = (C_i + K_i) \mod 92 \]

Friedrich Kasiski was the first to publish a successful general attack on the Vigenère cipher. Earlier attacks relied on knowledge of the plaintext, or use of a recognizable word as a key. Kasiski's method had no such dependencies. He published an account of the attack, but it's clear that there were others who were aware of it. Babbage was goaded into breaking the Vigenère cipher when John Hall Brock Thwaites submitted a "new" cipher to the Journal of the Society of the Arts. Thwaites challenged Babbage to break his cipher encoded twice, with keys of different length. Babbage succeeded in decrypting a sample, "The Vision of Sin", by Alfred Tennyson, encrypted according to the keyword "Emily", the first name of Tennyson's wife. Studies of Babbage's notes reveal that he had used the method later published by Kasiski [11][16].

III. Methodology

The new method employs the Vigenère square and key in its encryption process but the successive keys used will be dependent on the initial key value during the encryption process. That is the key varies as it is used in the encryption process. The first step key will be different from the second step key but the second step key will be as a result of a function that operated on the first step and so forth.

The algorithm ultimately makes it possible for encryption and decryption of the text and also makes the Vigenère cryptosystem more difficult against frequency attack. This is as a result of varying keys employed for each encryption process. A software program was written to demonstrate the effectiveness of the algorithm using java programming language and cryptanalysis performed on the ciphertext.
IV. THE MATHEMATICAL ALGORITHM

If the letters A–Z are taken to be the numbers 0–25, and addition is performed modulo 26, then Vigenère encryption E using the key K can be written,

\[ C_i = E_K(M_i) = (M_i + K_{ni}) \mod 26 \]

and decryption D using the key K,

\[ M_i = D_K(C_i) = (C_i - K_{ni}) \mod 26, \]

Thus Given m, a positive integer, \( P = C = (Z_{26})^n \), and \( K = (k_1, k_2 \ldots k_m) \) a key, we define:

Encryption:

\[
ek(p_1, p_2 \ldots p_m) = (p_1 + k_{n0}, p_2 + k_{n1} \ldots p_m + k_{mn}) \mod 26
\]

where \( k_{ni+1} = f(k_{ni}, si) \mod 26 \)

\( i = 0, 1, 2 \ldots n \)

\( s = x : [a, b] = \{ x \in \mathbb{I} : a \leq x \geq b, a=0, \text{ and } b=25 \} \)

Decryption:

\[
dk(c_1, c_2 \ldots c_m) = (c_1 - k_{n0}, c_2 - k_{n1} \ldots c_m - k_{mn}) \mod 26
\]

V. IMPLEMENTATION

The algorithm was implemented using java programming. A table of Vigenère cipher was generated and a chosen key was used to encrypt the plaintext. The key varies after each successive encryption in order to eliminate the repetition of the same key and plaintext encryption results within the cipher text.

With the encryption process, a table of Vigenère cipher was created and the plaintext is accepted from the user interface as an input as well as the key. The key is then used to operate on the message which is the plaintext to produce a ciphertext.

The first encryption step will yield a length ciphertext, \( L_{ci} = L_{k_i} \)

The second to the nth encryption step will yield a length ciphertext, \( L_{ci} = L_{k_i-1} + L_{k_i} \)

Where \( i = 1, 2, 3 \ldots n \) and \( L_k = \text{length of key} \).

The will vary by a shift during each encryption process.

Below is the java codes written to implement the algorithm. The decryption process follows the similar trend but with an inverse operation in order to decipher the ciphertext.

A. Encryption

```java
int tableRowSize = 26;
int tableColumnSize = 26;

int vignereTable[][] = new int[26][26];
Scanner input = new Scanner(System.in);
for (int rows = 0; rows < tableRowSize; rows++) {
    for (int columns = 0; columns < tableColumnSize; columns++) {
        vignereTable[rows][columns] = (rows + columns) % 26;
    }
}
System.out.println("Enter the plaintext");
String plainText = input.nextLine();
System.out.println("Enter the key: ");
String key = input.nextLine();
for (int ptextIndex = 0; ptextIndex < plainText.length(); ptextIndex++) {
    char pChar = plainText.charAt(ptextIndex);
    int asciiVal = (int) pChar;
    if (pChar == ' ')
        continue;
    if (asciiVal < 65 || asciiVal > 90)
        continue;
    int basicPlainTextValue = ((int) pChar) - 65;
    char kChar = key.charAt(keyIndex);
    kChar += 1;
    int basicKeyValue = ((int) kChar) - 65;
    int tableEntry = vignereTable[basicPlainTextValue][basicKeyValue];
    char cChar = (char) (tableEntry + 65);
    plainText += cChar;
    keyIndex++;
    if (keyIndex == key.length())
        keyIndex = 0;
}
System.out.println(" cipher text is "+cipherText);
```

B. Decryption

```java
int tableRowSize = 26;
int tableColumnSize = 26;
int vignereTable[][] = new int[26][26];
Scanner input = new Scanner(System.in);
for (int rows = 0; rows < tableRowSize; rows++) {
    for (int columns = 0; columns < tableColumnSize; columns++) {
        vignereTable[rows][columns] = (rows + columns) % 26;
    }
}
System.out.println("Enter the cipher text");
String cipherText = input.nextLine();
System.out.println("Enter the key: ");
String key = input.nextLine();
int keyIndex = 0;
for (int ctextIndex = 0; ctextIndex < cipherText.length(); ctextIndex++) {
    char cChar = cipherText.charAt(ctextIndex);
    int asciiVal = (int) cChar;
    if (cChar == ' ')
        continue;
    System.out.println(" Enter the key: ");
    String key = input.nextLine();
    key = key.toUpperCase();
    String cipherUpperCase = ";
    int keyIndex = 0;
    for (int ctextIndex = 0; ctextIndex < cipherText.length(); ctextIndex++) {
        char cChar = cipherText.charAt(ctextIndex);
        int asciiVal = (int) cChar;
        if (cChar == ' ')
            continue;
        System.out.println(" Enter the key: ");
        String key = input.nextLine();
        key = key.toUpperCase();
```
if (asciiVal < 65 || asciiVal > 90){
    plainText += cChar;
    continue;
}
int basiccipherTextValue = ((int)cChar) - 65;
char kChar = key.charAt(keyIndex);
kChar+=1;
int basicKeyValue = ((int)kChar ) - 65;
for (int pIndex = 0; pIndex < tableColumnSize;
    pIndex++){
    if (vignereTable[basicKeyValue][pIndex] ==
        basiccipherTextValue){
        char potChar = (char)
            vignereTable[basicKeyValue][pIndex] + 65);
        char potpChar = (char) (pIndex + 65);
        plainText += potpChar;
    }
}
keyIndex++;
if (keyIndex == key.length())
    keyIndex = 0;
}
System.out.println(" plain text is "+plainText);

VI. RESULTS AND ANALYSIS

The program written was used to encrypt a message and the result was analyzed with the existing Vigenère cipher. The Plaintext was encrypted by the software and analyzed below by applying the Kasiski attack.

The plain text = "IN THE FOREST THERE ARE MANY TREES WITH THE SAME HEIGHT. FOR EXAMPLE MANY"

The keyword was "TREE"

A. Results from existing Vigenère Cipher

Key: TREE

Message: IN THE FOREST THERE ARE MANY TREES WITH THE SAME HEIGHT. FOR EXAMPLE MANY

Ciphertext: BE XAV IDHWM XLXII TII FRRC KVIIXJ
            ABKL MYI LRQI YIMZYX. WSV VBEFGPI DERR

B. Result from proposed algorithm

Key: TREE

Message: IN THE FOREST THERE ARE MANY TREES WITH THE SAME HEIGHT. FOR EXAMPLE MANY

Ciphertext: CF YMY XTWYKY YBWWJ UJJ RUFD
            YLWIX QAYM NZJ XUEJ MYALMN. XTW YPFRRJDJ RUFD

Fig 3 frequency graph of B

Fig 4 Percentage of alphabet frequency in A (a) and B (b) respectively

C. Index of Coincidence of A and B

The I.C. of a piece of text does not change if the text is enciphered with a substitution cipher. It is defined as:

\[ I.C. = \frac{\sum_{i=A}^{Z} f_i(f_i - 1)}{N(N - 1)} \]

Where \( f_i \) is the count of letter \( i \) (where \( i = A, B...Z \)) in the ciphertext, and \( N \) is the total number of letters in the ciphertext.
From the analysis the index of coincidence (IC) of A and B was calculated to be to be 0.0590 and 0.0631 respectively. For a normal English text of alphabet of A-Z, the Variance is normally 14.50603 and the standard deviation of 3.80868. The results indicated there is a larger deviation in B than in A.

The table below shows the N-Gram analysis of A which is the result of the ciphertext from existing Vigenère cipher and B which is the result of the ciphertext from the new proposed method of encryption using the modified Vigenère cipher. The N-Gram analysis indicates the frequency of monogram, bigram as well as trigram with their respective percentages of occurrence within the ciphertexts.

### Table 2: N-Gram Analysis of A

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Histogram</th>
<th>Bigram</th>
<th>Trigram</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>28.69%</td>
<td>5.06%</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>9.47%</td>
<td>3.39%</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>9.47%</td>
<td>3.39%</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>6.78%</td>
<td>3.39%</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5.02%</td>
<td>1.60%</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>5.02%</td>
<td>1.60%</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>5.02%</td>
<td>1.60%</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>5.02%</td>
<td>1.60%</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>5.02%</td>
<td>1.60%</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>3.37%</td>
<td>1.60%</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>3.37%</td>
<td>1.60%</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>3.37%</td>
<td>1.60%</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>3.37%</td>
<td>1.60%</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>3.37%</td>
<td>1.60%</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>21</td>
<td>21</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>22</td>
<td>22</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>23</td>
<td>23</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>24</td>
<td>24</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>25</td>
<td>25</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>26</td>
<td>26</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>27</td>
<td>27</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>28</td>
<td>28</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>29</td>
<td>29</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
<tr>
<td>30</td>
<td>30</td>
<td>1.66%</td>
<td>1.60%</td>
</tr>
</tbody>
</table>

### Table 3: N-Gram Analysis of B

From the analysis the index of coincidence (IC) between a string and that same string with its first few characters deleted (sometimes called a shift of the string) was obtained for A to be 0.0590 and B to be 0.0631 which indicates a stronger approach of the new algorithm. For a normal English text of alphabet of A-Z, the Variance is normally 14.50603 and the standard deviation of 3.80868. For stronger implementation of the new algorithm, the length of the key have to be long and the variation function of the key must be very difficult to be predicted.

### REFERENCES


Quist-Aphetsi Kester: is a global award winner 2010 (First place Winner with Gold), in Canada Toronto, of the NSBE’s Consulting Design Olympiad Awards and has been recognized as a Global Consulting Design Engineer. He is a PhD student in Computer Science. The PhD program is in collaboration between the AWBC/USFC Academics Without Borders/Universitaires Sans Frontieres (formerly AHED-Academics for Higher Education and Development) Canada and the Department of Computer Science and Information Technology (DCSIT), University of Cape Coast. He had a Master of Software Engineering degree from the OUM, Malaysia and BSC in Physics from the University of Cape Coast-UCC Ghana.

He has worked in various capacities as a peer reviewer for IEEE ICAST Conference, lecturer and Head of Computing Science Department. He is currently a lecturer and Head of Digital Forensic Laboratory Department at the Ghana Technology University.