Strong Key Mechanism Generated by LFSR based Vigenère Cipher
Abdul Razzaq, Yasir Mahmood, Farooq Ahmed, Ali Hur
School of Electrical Engineering & Computer Science (SEECS) National University of Sciences & Technology (NUST) Sector H-12, Islamabad 44000, Pakistan
{abdul.razzaq, yasir.mahmood, farooq.ahmad, ali.hur}@seecs.edu.pk

Abstract: Communication has extirpated distances transforming world into global village. But this also wreaked in security threats for transmission of information across the global village. Encryption is an effective solution for secure transmission of data. An efficient encryption methodology not only secures information but also conceal the attacking clues. This paper has proposed an effective cipher based on Kerckhoffs' principle. LFSR based stream cipher produce key, eliminating period longer than plaintext. Plain text is encrypted using classical Vigenere Cipher. A novel approach is presented by combing the LFSR key with Vigenere cipher key. Short key will be shared between sender and receiver but information will be encrypted with longer key. Paper provides confident analysis for the system will be more secure as compared to conventional Vigenere cipher and LFSR based Stream ciphers.

Keywords: Vigenere cipher, LFSR based stream cipher, frequency analysis

1. INTRODUCTION
Exponential increase in communication on the internet as well as within intranet has exposed the data for intruders. All organizations whether dealing with business data, secret information in the form of credit card or medical data are concerned with the security of their data. Different security and cryptographic mechanisms has been deployed but no system proven to be a perfect solution. Some solution may be considered secure but less efficient due to complex mechanism of encryption and decryption like AEC. Some cipher systems are simple but not secure enough as Vigenere and LFSR based stream ciphers.

LFSRs based ciphers inherently are weak ciphers, although they are stream ciphers but its application are used in various technologies. For example, LFSR based A5/1 and A5/2 are used in GSM cell phones whests E0, used in Bluetooth, and the shrinking generator. The A5/2 cipher has been broken and both A5/1 and E0 have serious weaknesses[17,19].

Vigenere cipher was considered secure for centuries but later its weakness was identified. Friedrich Kasiski discovered a method to identify the period and hence key and plaintext [1]. Stream ciphers (especially LFSR based) are an important class of symmetric ciphers used widely in encryption for hardware-based cryptographic systems. They are simple, efficient without compromising performance [2]. Key generation is the main problem during designing a stream cipher. It generates a key which is as long as the plain message [3]. Security of Vigenere cipher can be enhanced if period of key is greater than plaintext. To increase size of period, length of Vigenere key needs to be enlarged.

Longer keys are hard to remember and their transmission on secure channel is expensive as well. Paper introduces the cipher system which exploits the advantages of Vigenere cipher with LFSR based stream ciphers and mitigates the weaknesses of these individual ciphers.

The basic theme of Vigenère cipher is to disguise plaintext letter frequencies by defeating simple frequency analysis. This cipher encrypt plaintext message with different letters at different points. But the primary weakness of the Vigenère cipher is the repeating nature of its key. If a cryptanalyst correctly guesses the key's length, then the cipher text can be treated as can be easily broken. Various method like Kasiski and Friedman tests can help to determine the key length.

Proposed system generates large key from short keyword using LFSR concept. Plaintext is encrypted with large and pseudorandom letter generated key which help to flatten the letter frequencies of cipher text.

The basic theme of Vigenère cipher is to disguise plaintext letter frequencies by defeating simple frequency analysis. This cipher encrypt plaintext message with different letters at different points. But the primary weakness of the Vigenère cipher is the repeating nature of its key. If a cryptanalyst correctly guesses the key's length, then the cipher text can be treated as can be easily broken. Various method like Kasiski and Friedman tests can help to determine the key length.

Section 2 describes the existing and related work to the proposed technique. Section 3 explains the proposed cipher technique. Section 4 provide the analysis for proposed technique along with data and graphs. Conclusions has been provided in section 5.
2. RELATED WORK

In a latest scenario of information security, the need of securing the secret information is multifold. Multiple applications are running on web concerning the important public information like credit card information, national security number, business or defense secrets and medical data of person. In April 23, 2009, Interhack [4] has classifying data loss incidents with public information and examined publicized data breaches by type and industry and found significant results for Finance, Education, Public Administration, and Health Care. The classical and modern cryptographic systems along with their theoretical and computational properties are presented by Whitefield Diffie and Martin E. Hellman [5], along with cryptanalytic examination of various important systems and provide a base for cryptographic literature. Various Classical Cryptographic Transposition ciphers as Substitution ciphers, One-time pad or Codebook has been used initially but easily breakable through the cryptanalysis, using the divide and concourse policy and frequency analysis of alphabet of particular language. Poly-Alphabetic Substitution cipher was very successful to counter and flatten the frequency analysis technique. Vigenere cipher using the Poly-alphabet algorithms to counter the frequency analysis was thought to be impregnable “uncrackable” till late 1917. But later it had been broken [6] and now it is no more secure. Index of co-incident [7] of two letters in a language play a revolutionary role in the cryptanalysis of the classical ciphers especially the myth of Vigenere cipher break by a Prussian scientist, Friedrich Kasiski [8].

Lots of contributions have been made by IBM to meet the growing needs for information security in a cryptographic research effort in the late 1960s. They have produced several important cryptosystems that contribute to the quality of the unclassified cryptography [9, 10, 11,12]. In January 1977, IBM introduce the DES [15] as nation stranded based on Shannon’s brilliant research on information theory[18] that suggested using product ciphers to build a strong system out of simple, individually weak, components. One time Pad Stream ciphers considered to be an unbreakable cipher.

Stream ciphers normally based on linear feedback shift registers (LFSRs) because they can be easily implemented in hardware and can be readily analysis mathematically. But inherently the LFSRs based stream ciphers are weak and are insufficient to provide good security. Various schemes have been proposed to increase the security of LFSRs. Multi Linear Shift Registers (MLSR) are often used generate the pseudorandom bit sequences in cryptosystems which make the system much secure as one time pad. MLSRS with m-stage will have very long period. If m = 100 then the sequence will not repeat in 1016 years on a 1-Mbit/s data link [13,14].

3. PROPOSED TECHNIQUE

Proposed technique uses the concepts and methodologies of classical Vigenere cipher plus modern stream cipher. Vigenere cipher is an old but a practical encryption method that can be efficiently used for many applications. Technique generates the long key for Vigenere Cipher. Poly alphabetic is one of the classical cipher techniques that make the frequency count flatter with few clues [5]. This is the result of encrypting same letter with multiple letters. This flatness can be more lowered if key is random instead of a phrase or plain English sentence. Basic concept of our technique is based on stream cipher. General model of Stream Cipher has been shown in Figure. 1. Key setup function H takes a short key K to initialize the key stream K. Function F updates the states of Ki. G generates the key stream; this will be used in poly alphabetic cipher for encryption and decryption.

Proposed technique exploits the concept and methodologies of LFSR based stream ciphers. Vigenere Cipher is an old but a practical encryption method that can be effectively and efficiently used for many applications. Difference between general stream cipher and proposed key stream generator is alphabetic letters and bits. A long key enhances the security of the Vigenere cipher. A key that is as long as the plaintext message would result in having no period [16]. Randomness of the letters in the key also increases protection of key. Kasiski test and Index of Coincidence determine the period of the Vigenere Cipher leading to identify key and plaintext [9].

Proposed key stream is pseudorandom list of alphabetic letters. These alphabetic letters are used to encrypt the plain text using the classical methodology of Vigenere Cipher [5]. Proposed technique generates the period less (period is larger than plaintext) and pseudorandom letter key stream to be used for encryption/ decryption. Period less and randomness of key lowers the flatness of letter frequency. Previously it was easy to identify the period and hence key for Vigenere Cipher. Since proposed technique is period less, it is more difficult to identify key and hence confidentiality integrity of plain text. Sender “S” and Receiver “R” have proposed system of encryption with desired polynomial for LFSR based Stream Cipher.“S” encrypts message with desired keyword “K” using proposed technique. Proposed technique generates as long as plaintext P, period less and randomized key stream “ki” from K. An example has been shown in equation 1.

\[ K_{i+1} = (K_i + K_{i+1}) \mod 26 \]
Where \( i \) is length of keyword \( K \). \( K_i \) encrypts each letter \( P_i \) of \( P \) and generates cipher text letter \( C_i \) for cipher text \( C \) as shown in equation 2. “S” sends \( C \) to \( R \) using unsecured channel “\( U_C \)” but sends \( K \) using secure channel “\( S_C \)”. 

\[
C_i = (K_i + P_i) \mod 26
\]  
\( (2) \)

It is supposed \( S_C \) is very expensive hence sending \( P \) on \( S_C \) is relatively costly. Expenses on sending of short keyword \( K \) is not as costly as \( P \). Receiver “\( R \)” receives \( K \) and \( C \). \( R \) using the proposed technique generates key stream \( K_i \) from \( K \) as shown in equation 1. Key stream is used for decryption of \( C \) using equation 3.

\[
P_i = (K_i + C_i) \mod 26
\]  
\( (3) \)

**4. ANALYSIS OF PROPOSED DESIGN**

For the analysis purpose, proposed system has been implemented in JAVA programming language for demonstration intention. Proposed technique has been analyzed for identification of keyword, identification of keyword period, identification of polynomial and identification of key stream.

This section also highlights weakness of proposed technique and applications for which the system provides enhanced security. Period for small letter keyword is long enough to encrypt the entire plain text letters. For example period for four-lettered keyword “nust” has been identified 10980 letters, whereas period for five-lettered keyword “label” could not be known even running the Java based program for 3+ hours (8006280+ cycles) on Intel(R) Core2 Duo CPU T7250 @ 2.00GHz 2.00 GHz with 2 GB RAM.

Further frequency analysis for ciphered text is more difficult than traditional classical cipher. This has been stated in the table I with English letter frequency, Vigenere cipher frequency and proposed cipher frequency.

<table>
<thead>
<tr>
<th>English Alphabet</th>
<th>Frequency of English letters</th>
<th>Vigenere Cipher</th>
<th>Proposed Cipher</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>8.17%</td>
<td>4.60%</td>
<td>4.20%</td>
</tr>
<tr>
<td>b</td>
<td>1.49%</td>
<td>3.24%</td>
<td>3.92%</td>
</tr>
<tr>
<td>c</td>
<td>2.78%</td>
<td>5.63%</td>
<td>4.60%</td>
</tr>
<tr>
<td>d</td>
<td>4.25%</td>
<td>10.91%</td>
<td>4.80%</td>
</tr>
<tr>
<td>e</td>
<td>12.70%</td>
<td>1.63%</td>
<td>5.40%</td>
</tr>
<tr>
<td>f</td>
<td>2.23%</td>
<td>4.50%</td>
<td>3.20%</td>
</tr>
<tr>
<td>g</td>
<td>2.02%</td>
<td>3.40%</td>
<td>4.20%</td>
</tr>
<tr>
<td>h</td>
<td>6.09%</td>
<td>5.70%</td>
<td>4.20%</td>
</tr>
<tr>
<td>i</td>
<td>6.97%</td>
<td>2.66%</td>
<td>4.20%</td>
</tr>
<tr>
<td>j</td>
<td>0.15%</td>
<td>4.55%</td>
<td>4.72%</td>
</tr>
<tr>
<td>k</td>
<td>0.77%</td>
<td>4.22%</td>
<td>3.92%</td>
</tr>
<tr>
<td>l</td>
<td>4.03%</td>
<td>3.40%</td>
<td>4.62%</td>
</tr>
<tr>
<td>m</td>
<td>2.41%</td>
<td>8.63%</td>
<td>3.32%</td>
</tr>
<tr>
<td>n</td>
<td>6.75%</td>
<td>8.16%</td>
<td>3.42%</td>
</tr>
<tr>
<td>o</td>
<td>7.51%</td>
<td>2.57%</td>
<td>3.95%</td>
</tr>
<tr>
<td>p</td>
<td>1.93%</td>
<td>1.00%</td>
<td>4.18%</td>
</tr>
<tr>
<td>q</td>
<td>0.10%</td>
<td>5.42%</td>
<td>3.52%</td>
</tr>
<tr>
<td>r</td>
<td>5.99%</td>
<td>5.00%</td>
<td>3.82%</td>
</tr>
<tr>
<td>s</td>
<td>6.33%</td>
<td>6.79%</td>
<td>4.13%</td>
</tr>
<tr>
<td>t</td>
<td>9.06%</td>
<td>4.22%</td>
<td>3.12%</td>
</tr>
<tr>
<td>u</td>
<td>2.76%</td>
<td>0.85%</td>
<td>3.82%</td>
</tr>
<tr>
<td>v</td>
<td>0.98%</td>
<td>1.52%</td>
<td>2.90%</td>
</tr>
<tr>
<td>w</td>
<td>2.36%</td>
<td>0.08%</td>
<td>3.52%</td>
</tr>
<tr>
<td>x</td>
<td>0.15%</td>
<td>0.04%</td>
<td>2.80%</td>
</tr>
<tr>
<td>y</td>
<td>1.97%</td>
<td>0.06%</td>
<td>2.72%</td>
</tr>
<tr>
<td>z</td>
<td>0.07%</td>
<td>1.22%</td>
<td>2.80%</td>
</tr>
</tbody>
</table>
It can be observed for small keyword there are very large and/or unknown periods. It is hard to break proposed cipher compared to traditional Vigenere cipher. Possible attack on Vigenere cipher is identification of period leading to discovery of key stream (i.e keyword in traditional Vigenere Cipher). Since period of proposed Cipher is large enough, further key stream is pseudorandom letters leading to decrease the effectiveness of Kasiski and Index of Coincidence (IC) attacks as shown in Table 1 and Figure 2.

Proposed system allows to generate period less key stream of desired length to minimize the effectiveness of above mentioned attacks. Weakness of proposed system is first n-letters of key stream. Each of the letters can be identified form cipher text with exhaustive attack on each letter. Once n-lettered keyword is known, it can be helpful for identification of polynomial and hence key stream can be identified. Such kind of attacks can be stopped if first n-letter from key stream are eliminated and rest of key stream is used for encryption. Keyword is exchanged through secure channel to generate key stream. As far Vigenere Cipher is concerned proposed system enhanced its security.

Proposed system can provide security for email, small to medium sized text files and communication systems where amount of data transmitted is not very large enough. For above mentioned systems, proposed system has period larger than plaintext.

The following are the strong point of the proposed cipher design:

- Proposed system generates key stream with very large period, larger than plaintext and hence provide more encryption security than conventional Vigenere cipher and LFSR based stream ciphers.
- Unlike Vigenere cipher Key length is hard to guess.
- It defeat the frequency analysis of cipher text so Kasiski method, Expected Values or Friedman test or Index of Coincidence are failed to guess the Key value.
- The speed/cost ratio of the design almost same as Vigenere cipher with negligibale overhead
- The design also perform reasonably well in restricted resource environments.
- It embedding the strong point of LFSRs and Vigenere Cipher and minimized the limitation of both ciphers
- It follow the Kerckhoffs' principle in true sense
- Proposed system can provide security for email, small to medium sized text files and communication systems where amount of data transmitted is not very large enough. It also helpful to prevent algebraic, chosen-ciphertext and dictionary attacks
- Technique has been analyzed with intentions of cryptanalysis and has been found more secure that traditional techniques.

![Figure 2: Graphical representation of Frequency Analysis Proposed Cipher with traditional Vigenere Cipher](image-url)
CONCLUSION
Paper presented a novel idea combining the classical encryption technique and LFSR based stream cipher technique for improved security of cipher text. Proposed system generates key stream with very large period size having possibility of key size larger than plaintext. This larger period size provide more encryption security than conventional Vigenere cipher and LFSR based stream ciphers. The proposed solution has been found secure for frequency analysis attack since the key period is larger than plain text. Technique has been analyzed with intentions of cryptanalysis and has been found more secure than traditional techniques.

In future the work may be extended for hardware based secure communication solution. The idea will help in securing communication by adding new hardware with existing communication devices.

Acknowledgements
This research is funded by National ICT R&D Fund Pakistan under the ICT-Related Development and Research Grant on Project “Semantics based Web Application Security: Concept, Design and Implementation”.

REFERENCES
[8]. Easy on Cryptology by Oliver Pell http://www.ridex.co.uk/cryptology/