ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Indian Journal of Science and Technology, *Vol 8(15), 33437, July 2015*

Universal Encryption Algorithm using Logical Operations and Bits Shuffling for Unicode

Abhas Tandon*, Rahul Sharma, Sankalp Sodhiya and P. M. Durai Raj Vincent

SSITE, VIT University, Vellore - 632014, Tamil Nadu, India;
abhastandon007@gmail.com, rahulsharma3493@gmail.com,
sankalp28011992@gmail.com, pmvincent@vit.ac.in

Abstract

Cryptography is the process of converting any information using some specific encryption algorithm to a secured form. Such secured form of information can be interpreted only by the intended recipient who possess equivalent decrypting algorithm. In this paper we have devised a new cryptographic primitive that concentrates on encryption and decryption of plain text string of any Unicode supported language. The ciphered text would be a combination of characters from several other languages. The algorithm defined here is an enhancement of one-time pad cipher and stream cipher using pseudo random numbers as public key and bits shuffling of the data.

Keywords: Bits Shuffling, Cryptography, Multi-Language, One Time Pad, Pseudo Random Key, Unicode

1. Introduction

Due to rapid growth of internet based services and applications such as Google Docs and Microsoft live which are available on cloud, large amount of data has become more open and accessible[1]. Recent incidents of data theft in IIM A[2] and Gurgaon city[3] have described the need for a secure and reliable data encryption algorithm which could not be easily deciphered by intruders. Confidentiality of data is of great significance for every individual from a normal PC user to a database administrator working in some large organization. In order to prevent data theft it is required to design complex algorithms that can efficiently encrypt the data. In any network environment, data must be transferred from source to destination via some secured channel complying the security protocols and encryption standards. Therefore cryptography is of great significance in area of network and information security. Most of the current encryption algorithms are found to increase the physical size of data by around 10% to 25%[4]. This could be a major hitch for industries and companies storing large amount of sensitive data.

Cryptographic ciphers are classified in two major types: symmetric ciphers and asymmetric ciphers[5]. Symmetric ciphers are conventional ciphers that uses single key (public key) for encryption as well as decryption purpose6. The algorithms under this approach are AES, TDES, RC5. Asymmetric ciphers use two types of keys (public key and private key), one for encryption and other for decryption[6]. Cryptosystem which belong to this category are RSA, PGP etc. While developing any new security algorithm, it is required to look at the mechanism of the system from intruder's point of view to determine possible vulnerabilities of the algorithm. Most of the text based ciphers are restricted to ASCII system or few well known languages. For public key based ciphers, intruders can use brute force attack[7] to predict the keys by analyzing the message pattern and checking for frequency of symbols.

Our algorithm tries to eliminate the above mentioned problems in existing system of text based ciphers. The

**Author for correspondence*

encryption technique suggested by us is an enhancement of stream cipher. We have used pseudo random numbers in our algorithm as public keys. Our algorithm can be used to encrypt any language supported by Unicode, thus making it a universal ciphering technique.

2. Motivation and Existing System

In[8] a new cipher has been demonstrated that is enhancement of one-time pad using a scheme involving arithmetic and logical operations. In same work new key generation algorithm has been suggested but method is restricted to ASCII characters only. A. Joshi et al.[9] have suggested a new randomized approach for cryptography in their work which is an enhancement of Caesar Cipher with a protocol using public key that is generated using actual global timestamp of the sending of message. However, algorithm is not using any secret key which is a major setback. In[10,11] methods have been suggested by authors that can be used to encrypt data in any language supported by Unicode. However both the algorithms are not using any secret key. The mapping constant chosen by both of these systems is restricted hence the encrypted message contains repetition of symbols in definite patterns which makes the system vulnerable to attacks. Also direct mapping technique is used which is not very efficient. Recently lot of innovative methods have been suggested for secured transmission of data over specific channels. In[12,13] methods have been suggested for text based steganography that is specifically designed for secured SMS delivery. Papers suggest how SMS language and commonly used informal SMS abbreviations or so called SMS language can be used for cryptography. However the suggestions made are limited to English language only and it is easy to decipher the message by person having general understanding of the language.

3. Proposed System

3.1 Encryption Algorithm

1. Input the plain text of length 'n'.
2. Convert each Plain text character into equivalent Unicode Value.
3. Take 9's Complement of each value obtained in Step 2.
4. Generate the public key values by providing proper seed value to pseudo random number generation algorithm.
5. Each group of value from Step 2 is XORed with key values from Step 4.

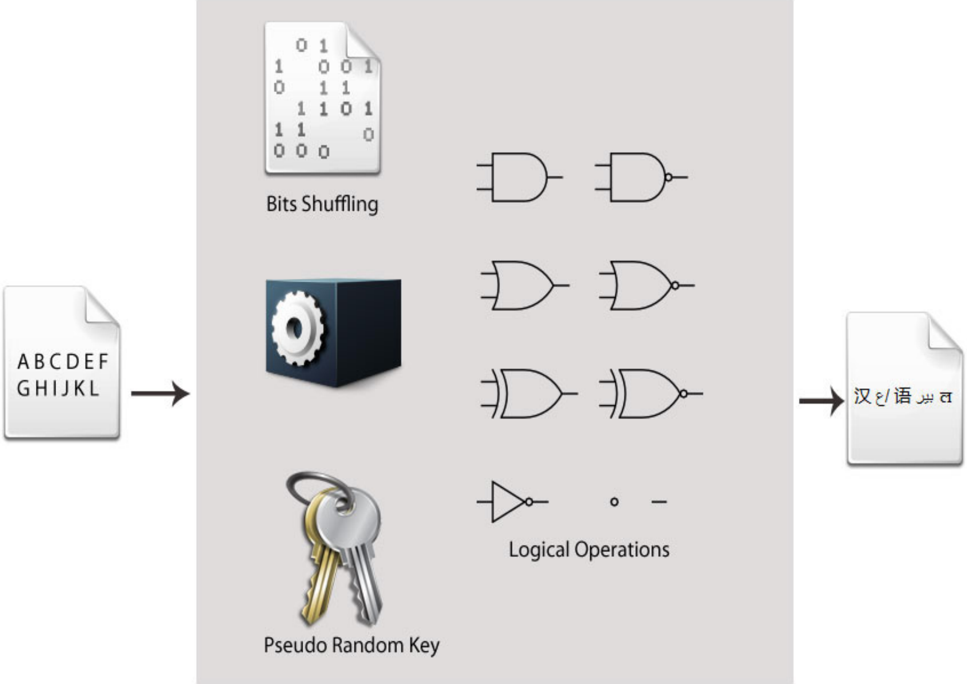

Figure 1. Scales of performance, P (y-axis) and time stamp, ts (x-axis).

6. Arrange all values from Step 4 in vertically with corresponding binary value aside (pad with zeros before binary value if length not equal to 16) and then list them in sequence by taking column bits longitudinally from MSB to LSB.
7. Group the value from Step 5 for 16 bits and convert each group into equivalent hexadecimal value.
8. Now hexadecimal values are shuffled by taking first values from each group and then second values and so on until all values are listed then again grouped for 16 bit hexadecimal value.
9. Equivalent binary value for each group in Step 8 is obtained and aligned together horizontally to get value of length 16n.
10. Value from Step 9 is divided into two portions in equal halves at 8nth position as First Section and Second Section.
11. Now binary bits are shuffled and grouped by 4 bits taking first two bits from first section and first two bits from second section and this step is repeated until all bits from both sections are grouped.
12. Equivalent Hexadecimal value for each group obtained from Step 11 is calculated.
13. Group the hexadecimal values from Step 12 of 4 digits each and find the Unicode Symbol corresponding to each new group.
14. The resultant String of Symbols obtained from Step 13 is the Ciphered Text.

3.2 Decryption Algorithm

1. Input the Encrypted(Ciphered) Text.
2. Apply the above Steps in reverse direction from Step 12 to Step 3.
3. The Resulting Unicode value after applying above steps will retrieve back the plain text.

4. Encryption Process

Plain Text	G	o	D
Unicode Value	71	111	100
9's Complement	28	888	899
Pseudo Random keys	1001	1532	36578
XOR operation	1013	1668	36193
Binary value	0000001111110101	0000011010000100	1000110101100001
Longitudinal alignment	0010000000000010	1111010111010110	1100000110000101
Hexadecimal Value	0x2002	0xF5D6	0xC185
Bits Shuffling	0x2FC0	0x510D	0x8265
Binary value	0010111111000000	0101000100001101	1000001001100101
Bits Shuffling and Hex Unicode Value	0000 1000 1111 1101 0 8 F D	1110 0000 0000 0010 E 0 0 2	0101 0110 0001 0101 5 6 1 5
Ciphered Text[14]			

5. Merits

Any encryption algorithm can be regarded efficient if it satisfies one or both of these mentioned criteria6:

- The cost of breaking the cipher exceeds the value of the encrypted information.
- The time required to break the cipher exceeds the useful lifetime of the information.

Our system satisfies both of the above criteria due to level of complexity in the encryption algorithm. Even if the third party manages to access the secret seed key then too time required to break the algorithm would exceed the useful lifetime of information due to use of various logical operations and bits shuffling. Proposed system is Unicode based cipher which allows localization of cryptographic application. Due to use of Unicode system, it is possible to encrypt wide range of major languages which are supported by Unicode. Due to presence of characters from unfamiliar languages, the hacker would require aid of language expert. The encrypted string would contain symbols in different languages; this would make it difficult for attackers to even predict the actual language of the message. If the original message is transliterated to some other language (e.g. "Mera Bharat Mahaan"), the probability that the intruder is able to make out the original meaning of message is further minimized.

Another main advantage of algorithm suggested is no increase in size of data due to encryption. System proposed can be used for secure transmission of sensitive data through proper channel e.g. Military applications where data being transferred is often related to national security and reliable storage of sensitive data to minimize data theft from organizational database.

6. Conclusion and Future Work

The proposed system can be used to encrypt and decrypt any language in Unicode system hence it is named Universal Encryption Algorithm. Also by usage of logical operations and bits shuffling this makes algorithm efficient and immunes to fight against intruders. Every character encrypted by this algorithm maps to unique element in Unicode system and hence data length remains same. The basic notion must be taken care that the secret keys must be communicated via secure channel. Future Study would be required to compress data which needs intensive research over above system to maintain its data efficiently without any loss and unique schema.

7. References

1. Kumar GP, Kumar Murmu A, Parajuli B, Choudhury P. MULET: A multilanguage encryption technique. 2010 7th International Conference on Information Technology: New Generations (ITNG). 2010 Apr 12–14. p. 779–82.
2. Available from: http://articles.timesofindia.indiatimes.com/2012-07-13/ahmedabad/32663176_1_data-theft-iims-confidential-data
3. Available from: http://articles.timesofindia.indiatimes.com/2006-10-09/india/27805615_1_cyber-crime-data-theft-gurgaon-police
4. Available from: http://www.computerweekly.com/news/2240085449/A-guide-to-practical-encryption-across-the-business
5. Al Housani H, Baek J, Yeun CY. Survey on certificateless public key cryptography. 2011 International Conference on Internet Technology and Secured Transactions (ICITST). 2011 Dec 11–14. p. 53–8.
6. Stallings W. Cryptography and Network Security Principles and Practice. 5th ed. Prentice Hall; 2011.
7. Asokan N, Janson PA, Steiner M, Waidner M. The state of the art in electronic payment systems. IEEE Computer. 1997 Sep; 30(9):28–35.
8. Srikantaswamy SG, Phaneendra HD. Enhanced One Time pad cipher with more arithmetic and logical operations with flexible key generation algorithm. International Journal of Network Security and Its Applications (IJNSA). 2011 Nov; 3(6):243–8.
9. Joshi A, Joshi B. A randomized approach for cryptography. 2011 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC). 2011 Apr 22–24:293–6.
10. Praveen Kumar G, Arjun Kumar M. Parajuli B, Prasenjit C. MULET: A Multilanguage Encryption Technique. IEEE 7th International Conference on Information Technology; 2010. p. 779–82.
11. Rajendiran M, Ibrahim BS, Pratheesh R, Babu CNK. Multilanguage block ciphering using two dimensional substitution array. 2011 Third International Conference on Advanced Computing (ICoAC); 2011 Dec 14–16. p. 117–20.
12. Shirali-Shahreza M, Shirali-Shahreza MH. Text Steganography in SMS. 2007 International Conference on Convergence Information Technology. 2007 Nov 21-23. p. 2260–5.

13. Rafat KF. Enhanced text steganography in SMS. 2nd International Conference on Computer, Control and Communication (IC4'2009); 2009 Feb 17–18. p. 1–6.

14. Available from: http://www.fileformat.info/info/unicode/char/search.htm