International Journal of Research In Science & Engineering e-ISSN: 2394-8299
Volume: 3 Issue: 2 March-April 2017 p-ISSN: 2394-8280

EFFICIENT TRIE BASED BOOK CIPHER ALGORITHM IMPLEMENTATION FOR DATA ENCRYPTION AND DECRYPTION

Ramanathan Arun[1]

[1]*Student, Computer Technology, Madras Institute of Technology,* ***ramanathanarun14@gmail.com***

ABSTRACT

The last quarter of a century has seen a tremendous change and improvement to our lifestyle. This includes shopping with our fingertips in our mobile to transferring important business details over email and personal messages with our colleagues in social media. Everything involves transfer of data over the network. These data are so vulnerable, such that these information's you pass onto the network may be accessed by another person who was not intended to receive those details. This restricts us to utilize these features suspecting that our information might be stolen. Here comes the need for a concept called Encryption and Decryption. We have introduced a novel approach called as Book Cipher Method. This is a Symmetric Cryptographic method as same key is used for both Encryption and Decryption. We have utilized Trie Data Structure to make this cryptographic method more efficient in terms of speed. The Book Cipher method was started to be utilized as early as World War time. But it wasn't a hit at that time because the books couldn't be digitized and stored at that time due to the lack of digitization methods. Because of this they used actual books to encrypt information. This was a tedious process and also since books were easily accessible one can easily decrypt that information, so this technique was not carried out further.But now since the development of World Wide Web we have a large repository of files over the internet (more than that we can think of). So it's practically impossible to find which document we are using to mask the data and unmask the data. This is one of the major advantages of the Book Cipher Method because the fact of uncertainty is very large.

Keywords*: Algorithm design and analysis, Book Cipher, Ciphers, Encryption and Decryption, Network Security.*

---

1. INTRODUCTION

Encryption is nothing but transforming our information to something like a having a mask such that information stealers can't find what data we send over the network. Decryption is the removal of this mask. This field involves various algorithms.

There are two main types of keys in this field namely private key and public key. Private Key as the name indicates it's specific to a particular user and Public key is available to any number of users. But recent research show that these keys can be decrypted within a single day.

Book Cipher is chosen over traditional One Time Pad (OTP) here because One Time Pad method requires the key to be as long as the length of data we transmit. The huge drawback of One Time Pad method is that it is absolutely secure only if a key is used only once i.e. each message must have its own key, this leads to having a key as long as the length of all the messages we expect to transmit in the future in order for it to be absolutely secure. We shouldn't use the same key for multiple messages because the information hacker has a very high probability of obtaining the messages transmitted having the same key by XORing the bits and get combined information about those messages. The infinite key space which was involved in One Time Pad method isn't a issue in Book Cipher method because all the documents have a definite size.

2. PROPOSED ANALOGY

The Book Cipher starts off by taking up a piece of text which can be any document, pdf's etc. This piece of text is used as a reference key. Before the data is transmitted over the network the data is mapped with this reference key to create a ciphertext which is to be transmitted over the network. After the data is mapped with the reference key and before sending it over the network we hash the data using secure hashing algorithm (SHA). The cipher text is first passed through secure hashing algorithm to obtain the original ciphertext and is then

again mapped with the reference key to produce to the actual data. This reference key may also be changed from time to time in order to increase further uncertainty.

Trie is being implemented by us in order for faster lookups for a word. Example: suppose the word's first occurrence to be encrypted is present in the 700th page, then each time the word occurs the algorithm must scan from 1st page until 700th page unnecessarily. On using trie we can get the word and its associated information in order of n in which n denotes the number of characters of the string to be passed. This provides a substantial improvement in lookup process.

Secure Hashing Algorithm provides us a extra layer of security. The SHA includes various operations like mod 2^32, AND, XOR, etc. The Secure Hash Algorithm was developed for use with the Digital Signature Standard (DSS) is specified within the Secure Hash Standard. SHA is a cryptographic message digest algorithm similar to the MD4 hash functions developed. It differs in that it adds an additional expansion operation, an extra round and the whole transformation was designed to accomodate the DSS block size for efficiency. The Secure Hash Algorithm takes a message of less than 2^{64} bits in length and produces a 160-bit message digest which is designed so that it should be computationally expensive to find a text which matches a given hash. I.e. if you have a hash for text t1, H (A), it is difficult to find a text t2 which has the same hash.

3. LITERATURE SURVEY

3.1 AES (Advanced Encryption Standard)

This algorithm falls under symmetric encryption category. It is able to produce blocks of length 128, 192, and 256 bits. It is based upon both substitution and permutation techniques. Same key is used for both encryption and decryption purposes but the ciphering and deciphering process are different to each other. There are ten rounds for 128-bit, twelve for 192-bit, and 14 for 256-bit. A round is made up of several processing steps such as substitution, transposition and mixing of the plaintext and transforms it into cipher text.

It is vulnerable under attacks but it takes long time to decrypt under the brute force technique which is performing trials for each possible key in the sequence.

3.2 DES (Data Encryption Standard):

This algorithm also falls under symmetric encryption category and it's a Block Encryption Algorithm, and hence same key is used for both the users using this algorithm. Key size is 56 bits and block size is 64 bits. Main operations include Permutation and Substitution; six different permutation operations are done. This algorithm is said to be scalable because of the difference in the key and block size.. Since it has only one key size the number of rounds is also constant owing to sixteen rounds. This algorithm is more efficient to hardware when compared to software. The ciphering and deciphering process are different to each other as in Advanced Encryption Standard.

It is not as secure as the Advanced Encryption Standard, it is susceptible to attacks from brute force as well as linear and differential cryptanalysis because of the short key length. It was not designed for software, hence if used for software it'll be slow.

3.3 Blowfish:

Blowfish was developed as a replacement to Data Encryption Standard (DES). This is a Symmetric Cryptographic Technique. The key length of this algorithm is a variable one, it varies from 32 to 448 bits. It has a maximum of 14 rounds. The block size of Blowfish is 64 bits.

It might or can be cracked by birthday attack, especially in secured HTTP (HTTPS), and the starting four rounds of BLOWFISH are susceptible under 2nd order differential attack.

3.4 RSA (Rivest, Shamir and Adelman):

It primarily uses two keys: public and private keys for encryption process. The public key is visible to all whereas the private key is specific to the user. It is also known as Asymmetric Cryptographic algorithm. Signature verification is mostly done using RSA algorithm. RSA's security is based on factoring the product of two large prime numbers.

It has a possibility of being cracked by using Hastad's attack and CopperSmith's attack.

Features	Book Cipher	Blowfish	AES	DES
Key Size	Size of the document used	32-448 bits	128/192/256 bits	56
Block Size	Varies according to each word	64	128	64
Rounds	2(can be extended to any number)	14	10/12/14	16
Mathematical operations involved	Only for non-alphabetic characters	For all characters (XOR)	For all characters (XOR)	For all characters (XOR, Left shift)

Table-1: Comparison of some Key Features between Book Cipher and other Cryptographic Methods

4. ALGORITHMS INVOLVED

4.1 Random PDF Selector

A Pdf is formed by putting up random words from available dictionary with no more than 50 words with a same starting letter. The document goes on until all the alphabets are used as starting characters. And it is seen to that no word duplication is encountered. This document is uploaded to be used as a Reference Key. This function is be called again after a certain period of time repeatedly so that the hackers cannot crack the message because the combinations of words in the document changes continuously. This function is not available to the user of the application.

4.2 ENCRYPTION

This function is called before the user places over the data onto the network. First a Trie is constructed with all the words from that reference key with each word having a position attribute mentioning the offset of the word from the start of the document. After the user approves for the message or data to be sent, the sentence is broken up into individual words. Each word starting from the first one is being searched in the trie. Searching the word in the trie first checks the complete word, if found the position of the String and the length of the String is added to the Encrypted String. If the entire word is not found, then the prefixes of the word is searched in the trie until found. In the worst case maximum length of prefix is one because we have words starting with all twenty-six characters. Here the position of the word whose prefix is matched and the length of the prefix is attached to the Encrypted String. Special Characters are handled by giving them unique numbers in the prefix length field.

This Encrypted String is being given another layer of security by using Secure Hashing Algorithm (SHA). Then the SHA encrypted message is passed over the network.

4.3 DECRYPTION

This function is utilized after the receiver receives the SHA hashed message. This message is first passed to SHA algorithm and then the output of this algorithm is the String consisting of positions of the word and its corresponding prefix length.

This Encrypted String is passed to another function which performs the actual Book Cipher Decryption. This function searches the word/prefix in the trie such that the trie returns the word which has the position and offset value and adds to the Decrypted String. Special Characters are handled by giving them unique numbers in the prefix length field. This happens until the entire Encrypted String is parsed through. After this the message or data is visible to the user.

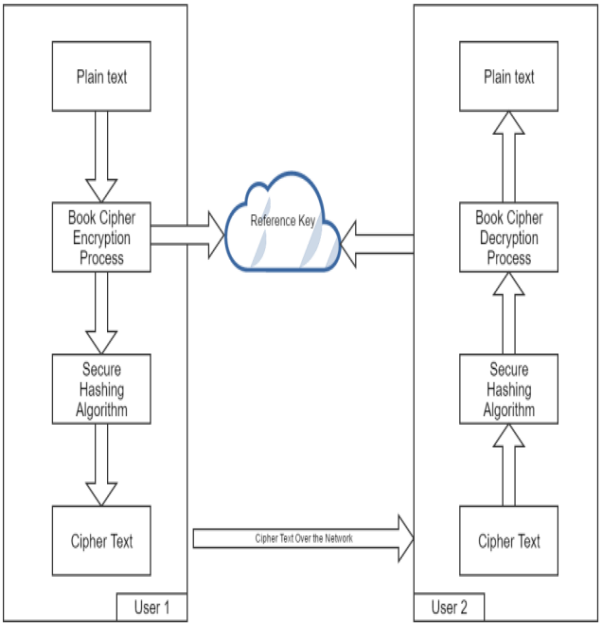

Fig-1: Overall view of the Cryptographic method

5. RESULTS

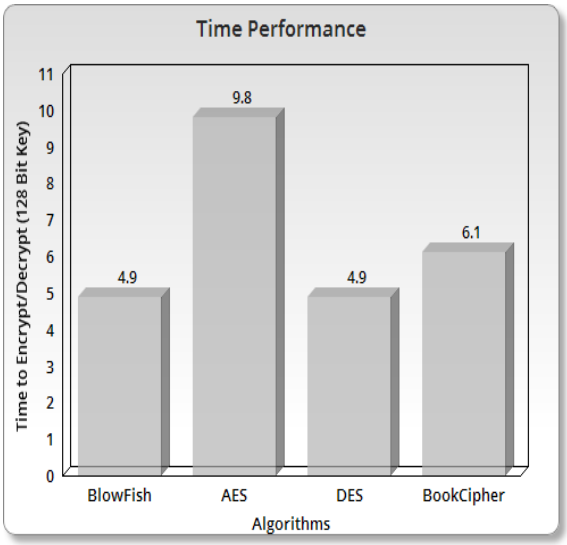

Fig-1: Time in seconds for encryption/decryption process of an input file of 9.8 mb

6. CONCLUSION

As every year passes by, there is an ever growing demand of utilize our smart phones and laptops instead of having to actually going to a particular place to finish that job so as to save our time and utilize it for another work in this highly competitive world. This has been evident from Digital Signatures which survey tells that it has saved one week per year on average if we had to get manual signatures from respective persons for business documents and other purposes. Increasing growth in Virtual Reality field has started to see that we can now get our own glasses by virtually wearing it over the Internet. Because of the vast improvement of Internet and Transfer of information over the network we need absolute secure algorithms in order to prevent our data from being misused by others.

Our approach may not be the best time efficient cryptographic technique but cryptographic techniques should rely more on the security the algorithm provides rather than the time it takes. We, the general people can wait half a minute or so for secure transactions. Even though the time factor is not important we still need to consider it to be manageable by the users and surveying our algorithm shows that: This is faster than AES and not too far behind DES and Blowfish.

This main motto before developing any cryptographic algorithm is to have the uncertainty level at the maximum to prevent our algorithm from being decrypted. We have thus served this purpose because our technique is capable of taking up any document from the World Wide Web as a reference key. Even a Supercomputer finds it tough to find out which document we use and this can be enhanced further by changing the documents (reference key) time and again to boost the Key from being accessed to steal the encrypted information.

ACKNOWLEDGEMENT

I have worked assiduously for this project. However, it would not have been possible without the guidance and support of my fellow professors. I would like to extend my sincere thanks to all of them. I am highly indebted to Mr. Gunasekaran for his guidance and constant supervision as well as for providing necessary information regarding the project and also for his support in completing the project. I would like to express my gratitude towards my parents for their kind co-operation and encouragement which help me in completion of this project.

REFERENCES

[1] P. Vadhera and B. Lall, "Review Paper on Secure Hashing Algorithm and Its Variants". 2014. International Journal of Science and Research (IJSR).ISSN (Online): 2319-7064

[2] V. Agrawal, S. Agrawal and R. Deshmukh, "Analysis and Review of Encryption and Decryption for Secure Communication", 2014. International Journal of Scientific Engineering and Research (IJSER).ISSN (Online): 2347-3878

[3] S. Nandan Kumar, "Review on Network Security and Cryptography", 2015. International Transaction of Electrical and Computer Engineers System. DOI: 10.12691/iteces-3-1-1.

[4] P. Mahajan and A. Sachdeva, "A Study of Encryption Algorithms AES, DES and RSA for Security", 2013. Global Journal of Computer Science and Technology, Network, Web & Security. ISSN (Online): 0975-4172.

[5] S. Dhull, Mamta and Reetu, "A Review Paper on Data Encryption and Decryption", 2016. Imperial Journal of Interdisciplinary Research (IJIR). ISSN: 2454-1362.

[6] A. Yassir and S. Nayak, "Cybercrime: A threat to Network Security", 2012. IJCSNS International Journal of Computer Science and Network Security.

[7] B. Eka Purnama and Zaneiah, "An Analysis of Encryption and Decryption Application by using One Time Pad Algorithm", 2015. (IJACSA) International Journal of Advanced Computer Science and Applications.

[8] K. Agarwal and S. Kumar Dubey, "Network Security: Attacks and Defence", 2014. International Journal of Advance Foundation and Research in Science & Engineering (IJAFRSE).

[9] R. Lele, R. Jainani, V. Mikhelkar, A. Nade and V. Meshram, "The Book Cipher Optimised Method To Implement Encryption And Decryption", 2014. International Journal Of Scientific & Technology Research. ISSN: 2277-8616.

[10] C. Wang and S. Ju, "A Novel Method to Implement Book Cipher", 2010. Journal of Computers.