Secure Data Storage On Web Using Matrix Based Cryptosystem.

N. P. Tiwary
Dept of CSE BIT, Mesra, Ranchi

G. Sahoo
Dept of IT BIT, Mesra, Ranchi

Abstract

The growing popularity of web applications in the last few years has led users to give the management of their data to online application provider, which will endanger the security and privacy of the users. This paper puts forward a safe mechanism of data transmission to tackle the security problem of data transmitted over internet. We propose a new technique for encrypting/ decrypting data to be transmitted using public network. The proposed technique is using the concept of coding theory and generating a number that can be expressed as a power of 2 to design a square matrix of plain text. The implementations and performance evolution demonstrate that the proposed technique is secure and efficient both in theory and practice.

1. Introduction

All manuscripts must be in English. These guidelines include complete descriptions of the fonts, spacing, and related information for producing your proceedings manuscripts.

With the rapid growth of Internet and Web 2.0 applications like Google, Gmail and Google Docs, people are moving their private data from their local storage to the online applications providers. While acquiring ease of use services users will have to give the control of their data privacy to the application providers. Although applications providers announce that these data are secure and will be handled without the involvement of administrators, these applications did not provide any mechanism to guarantee this promise. This paper discusses a new technique of encrypting private data before being moving them from their local storage to the online storage providers. It also gives a unique method of decrypting it back. Walsh Cryptosystem is completely based on coding theory. This concept is already being used to design different codes called chips in CDMA as one of the multiple access method with an objective to share a common channel among multiple stations on a multipoint link. Now we propose to employ this concept to secure user data to be stored in the premises of online application provider. Here we are using a Walsh Table, which is a two dimensional table of equal number of rows and columns. In a Walsh Table each row is an orthogonal sequence of elements and carries following properties:

1. Each sequence is made of N elements where N represents the number of rows in the Walsh Table and N must be expressed as a power of 2.
2. If we multiply a sequence by a number every element in the sequence is multiplied by that number.
3. The product of two equal sequences, element by element and there sum is equal to N.
4. The product of two unequal sequences element by element and their sum is zero.
5. The element by element addition of two sequences resulting a new sequence

According to Walsh if we know the table of N sequences denoted by W_N, we can design a table for 2N sequences as:

$$W_{2N} = \begin{bmatrix} W_N & W_N \\ W_N & \overline{W}_N \end{bmatrix}$$

2. Cryptography

Cryptography is the art of secret writing. More generally, people thing of cryptography as the art of mangling information into apparent unintelligibility in a manner allowing a secret method of unmangling. The basic service provided by cryptography is the ability to send information between participants in a way that prevents others from reading it. A message in its original form is known as plaintext or clear text The mangled information is known as ciphertext. The process for producing ciphertext from plaintext is known as encryption.The reverse process of encryption is called decryption. While cryptographers invent clever secret codes, cryptanalysts attempt to break these codes. These two disciplines constantly try to keep ahead of each other. Cryptographic systems tend to involve both an algorithm and a secret value known as the key. The reason for having a key in addition to an algorithm is that it is difficult to keep devising new algorithm that will allow reversible scrambling of information, and it is difficult to quickly explain a newly devised algorithm to the person with whom you would like to start communicating securely. It is important for cryptographic algorithms to be reasonably efficient for the good guys to compute. The good guys are the ones with knowledge of the keys. Cryptographic algorithms are not impossible to break without the key. A bad guy can simply try all possible keys until one works. The security of a cryptographic scheme depends on how much work it is for the bad guy to break it. A good cryptosystem is one that is computationally efficient and requires little storage space. Cryptographers are encouraged to develop systems that have a small key size, so that the keys are easy to share through stealthy channels; for example, short verbal communication, an encrypted email or disguised postal letter. The cryptosystem must also be secure against attack. Cryptanalysis is the art and science of decoding an encrypted message without knowing the keys. Cryptanalysis can be compared for finding the quickest and safest method of breaking into a house without the keys. Many view cryptanalysts as malicious and call them hackers, enemies, or adversaries. However, the tools, theory, and knowledge gained through cryptanalysis are essential. Law enforcement can use cryptanalysis to stop terrorism, child pornography, and fraud. If the government can decrypt emails, financial data, and other information stored on a terrorist's computer or website, then they may be able to save hundreds of lives by thwarting the terrorist's plans. Also, the information cryptanalysts discover while trying to break a cryptosystem can help cryptographers create stronger more secure systems. There are four basic types of attacks cryptanalysts use to break a cryptosystem.

Cipher text Only Attack

In a cipher text only attack, the adversary only has access to strings of cipher texts. For example, suppose we have a substitution cipher, one in which we create a mapping from the English letters to a permutation of the letters. Then, one can use statistical properties of the English language to figure out what the message says. A cryptosystem is considered extremely weak if it is susceptible to this type of attack because this is the most difficult and time consuming way to discover the key used in the cryptosystem. A ciphertext only attack also usually requires extensive computational power.

Known Plaintext Attack

In a known plaintext attack, the adversary has access to a set of plaintexts and their corresponding cipher texts. For example, if it is known that the cryptosystem is a specific linear transformation, then one can try to use mathematical properties of the linear transformation and the pairs of text to compute the key. A cryptosystem is considered weak if it is susceptible to this type of attack using current computational technologies.

Chosen Plaintext Attack

In a chosen plaintext attack, the adversary obtains temporary access to the encryption machine. He can input any message he wants and see the cipher text that it creates.

Chosen Cipher text Attack

In a chosen cipher text attack, the adversary obtains temporary access to the decryption machine. He can input any cipher text he chooses and see the message that produced it.

These last two attacks are more difficult to implement than the known plaintext attack or cipher text only attack because gaining access to the encryption or decryption machine is usually more difficult than gaining a list of cipher texts and/or plaintexts. Also, the chosen plaintext and chosen cipher text attacks are more successful when the adversary's chosen texts are within a narrow range, so more strategic planning is usually necessary. If one can prove that a cryptosystem is secure against these four types of attacks, even if the attacker has infinite computational resources, then it is highly probable that the cryptosystem is not susceptible to attack. However, there are always other methods of attack being developed.

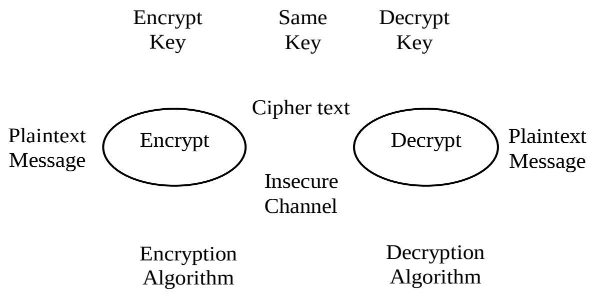


Fig-1

3. Web Application Security

Web application security focuses on examining application code for vulnerabilities in order to protect against exploits or exposing sensitive user data. In recent years a number of evaluation techniques have appeared trying to protect against attacks such as cross-site scripting and SQL injection. An example of analyzing application security includes examining and modifying a popular web application to ensure it is free from any bugs that might be exploitable. For a concrete example, consider a JavaScript based application, such as a document editor or web email, that relies heavily on the web browser for execution. Researchers have developed taint-tracking, static, and dynamic analysis techniques inside the browser that use programming languages techniques to analyze JavaScript based attacks. This involves developing custom tools that analyze the application provided, and for most web applications this is a combination of JavaScript and HTML and requires no reverse engineering. Then using the results of this analysis, tools automate rewriting the web application code to demonstrate safe transformation techniques. To evaluate the analysis and rewriting, the tools are demonstrated on a few popular web based applications. This scenario can pose a number of legal and ethical concerns depending on the terms of service and agreements between users and the application provider. If researchers were examining the document editor provided by Google, the relevant terms of service specially prohibit[You agree not to access or attempt to access any of the services by any means other than through the interface that is provided by Google. You specially agree not to access (or attempt to access) any of the services through any automated means (including use of scripts or web crawlers)."You agree that you will not reproduce, duplicate, copy, sell, trade or resell the services for any purpose. Going down the list, analysis tools written by researchers access the web application in a manner that is different than intended (i.e. not through a web browser), and accessing an application automatically is also a violation of the terms of service. Further, any downloading, copying or modifications of the web application being studied (even for non-commercial, proof of concept purposes) are also a violation of the terms, limiting researcher's ability to maintain temporary copies for examination. In addition to the terms of service limits on copying and duplication, other legal consequences could arise due to federal copyright law.

4. Related Work

Symmetric key encryption is the best approach for encryption. Nath et. Al. have developed an algorithm called MSA[1] for encryption and decryption of any file using a random key square matrix containing 256 elements. It is absorbed that if someone applies the brute force method then he has to give a trial for factorial 256 to find the actual key matrix. In modern world this number of trial runs may not be impossible for the hackers. To overcome the problem of MSA, Dripto et.Al.[1] has developed a better algorithm called DJSA. In DJSA the authors have considered the size of the key matrix to be 65,536 and in each cell two characters are placed in place of one characters as stored in MSA. If someone wants to apply brute force method to find actual key then one has to give a trial for factorial 65536 runs. To make the system more secure the authors have also introduced multiple encryptions. The authors have made all the effort to make DJSA more secure encryption algorithm but it is suitable for encryption of a file of small size normally less than or equal to 2MB.The best known types of symmetric encryption are the Data Encryption Standard (DES), triple DES (3DES), the Advanced Encryption Standard (AES), and Rivest Cipher (RC4). The former three are block ciphers while RC4 is a true stream cipher. The AES was developed as a replacement for DES and 3DES. It supports key lengths of 128, 192, and 256 bits and a variable block length. AES is based on the Rijndael encryption algorithm. Rijndael is a block cipher adopted as an encryption standard by the U.S. government, developed by Joan Daemen and Vincent Rijmen. It has been analyzed extensively and is now used widely worldwide as was the case with its predecessor, DES [3]. During the evaluation of candidates for the AES standard, Rijndael was analyzed by some of the world's best cryptanalysts. It has proven to be very effective against known attacks, very efficient, and simple to implement. Rijndael supports a larger range of block and key sizes; AES has a fixed block size of 128 bits and a key size of 128, 192 or 256 bits, where as Rijndael can be specified with key and block sizes in any multiple of 32 bits, with a minimum of 128 bits and a maximum of 256 bits [4]. Previously, there were several attempts to combine the RSA algorithm with the other security mechanism to provide a fast and secure implementation. For instance, number of researchers combined RSA algorithm with the Chinese remainder theorem (CRT) [5][6]. A number of other cryptographic techniques provide security through a number of mathematical transformations that can be proven to be mathematically secure provided some optimum conditions [10]. We however need to cognizant that cryptography on its own is insufficient to ensure a high level of security within an organization, that is to say that cryptography is not the silver bullet to solve all information security issues and should be used in conjunction with good security practices [11]. Cryptography, like the Information Security field itself, is an incredibly broad field involving many existing disciplines such as abstract algebra to provide mathematical proofs for the guaranteed correctness of an algorithm, statistics for analysis of cryptographic algorithms and quantum physics for quantum based random number generation for quantum cryptography.

5. Proposed Encryption Algorithm

The plaintext chosen is arranged into a Bi-directional circular queue data structure in a matrix of order N x N. Random () function is used to generate a random positive integer N which must be represented as a power of 2. The N so generated is used to arrange the plaintext character by character in a square matrix of order N x N. Then the characters stored in the matrix are replaced by their ASCII values. A Walsh table of order N x N is now used to encrypt the character stored in the matrix. Here we make introduction to our encryption algorithm. The entire process of encryption is summarized below:

1. Select a random number M and define N = 2^M where 2≤ M≤ 8
2. Let us consider a table W_1 = [+1]. Use this table to generate a matrix W_N of order N x N which is based on the concept of Walsh table.

$$W_2 = \begin{bmatrix} W_1 & W_1 \\ W_1 & \overline{W}_1 \end{bmatrix}$$

$$W_4 = \begin{bmatrix} W_2 & W_2 \\ W_2 & \overline{W}_2 \end{bmatrix}$$

For general

$$W_N = \begin{bmatrix} W_{N/2} & W_{N/2} \\ W_{N/2} & \overline{W}_{N/2} \end{bmatrix}$$

3. Read the text and arrange the characters in a matrix of order N x N. Let us assume this matrix is T_N.
4. Replace the character in the table so obtained by their corresponding ASCII value.
5. Multiply T_N with W_N and store the product in a matrix C_N of order N x N.
i.c. $C_N = T_N$ x W_N
6. Send the product matrix (C_N) for their online storage provider where the same will be stored in the encrypted form.

6. Decryption Algorithm

The proposed algorithm is a kind of symmetric encryption algorithm, with decryption process is done by reversing the operation done in the encryption process. The cipher text arranged in a matrix of order N x N denoted by C_N. The following are the steps in decryption:

1. Read the data stored at the online storage provider and arrange their character in a matrix C_N of order N x N.
2. Multiply C_N with W_N and store the product in T_N.
3. Divide each element of T_N by N to get the ASCII value of the character.
4. Replace the ASCII value stored in the matrix by their corresponding characters. Use the data

7. Simulation & Experimental Results

Example for Encryption:

Let us consider

T_4 =

1	3	4	6
0	2	12	10
8	9	11	14
6	2	4	0

W_4 =

+1	+1	+1	+1
+1	-1	+1	-1
+1	+1	-1	-1
+1	-1	-1	+1

Hence the product matrix C_4 = T_4 x W_4

C_4 =

1	3	4	6
0	12	12	10
8	9	11	14
6	2	4	0

X

+1	+1	+1	+1
+1	-1	+1	-1
+1	+1	-1	-1
+1	-1	-1	+1

Encrypted Matrix:

i.c. C_4 =

14	-4	-6	0
24	0	-20	-4
42	-4	-8	2
12	8	4	0

The above matrix C_4 is the encrypted matrix which can further be decrypted by multiplying the same with W_4 and dividing the resultant matrix by 4.

Example for decryption:

T_4 =

14	-4	-6	0
24	0	-20	-4
42	-4	-8	2
12	8	4	0

x

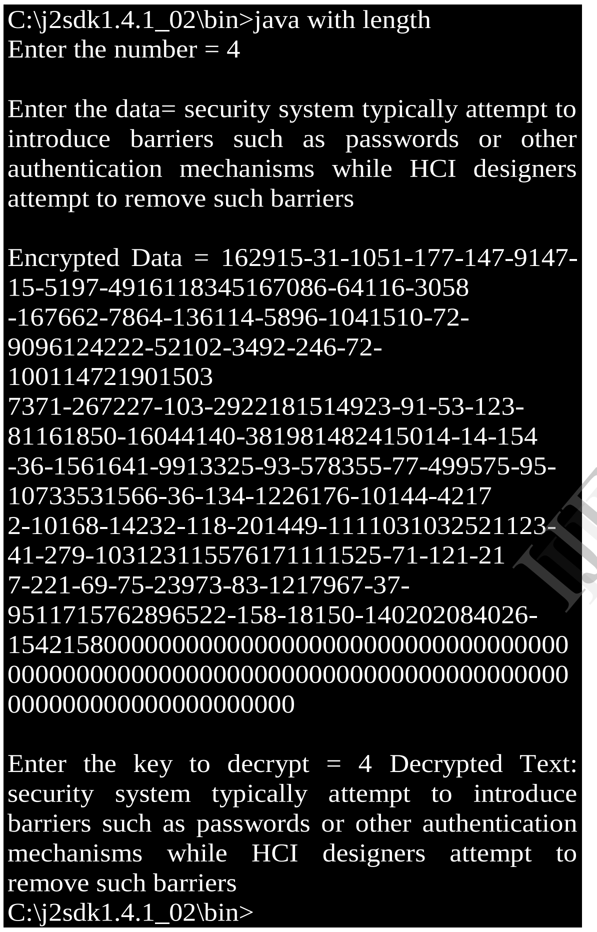
+1	+1	+1	+1
+1	-1	+1	-1
+1	+1	-1	-1
+1	-1	-1	+1

x $\frac{1}{4}$

$$\text{i.c. } T_4 = \begin{array}{|c|c|c|c|}\hline 4 & 12 & 16 & 24 \\ \hline 0 & 8 & 48 & 40 \\ \hline 32 & 36 & 44 & 56 \\ \hline 24 & 8 & 16 & 0 \\ \hline \end{array} \quad \text{X} \quad \frac{1}{4}$$

Decrypted matrix:

$$\text{i.c. } T_4 = \begin{array}{|c|c|c|c|}\hline 1 & 3 & 4 & 6 \\ \hline 0 & 2 & 12 & 10 \\ \hline 8 & 9 & 11 & 14 \\ \hline 6 & 2 & 4 & 0 \\ \hline \end{array}$$



```
C:\j2sdk1.4.1_02\bin>java with length
Enter the number = 4

Enter the data= security system typically attempt to
introduce barriers such as passwords or other
authentication mechanisms while HCI designers
attempt to remove such barriers

Encrypted Data = 162915-31-1051-177-147-9147-
15-5197-4916118345167086-64116-3058
-167662-7864-136114-5896-1041510-72-
9096124222-52102-3492-246-72-
100114721901503
7371-267227-103-2922181514923-91-53-123-
81161850-16044140-381981482415014-14-154
-36-1561641-9913325-93-578355-77-499575-95-
10733531566-36-134-1226176-10144-4217
2-10168-14232-118-201449-1111031032521123-
41-279-1031231155761711111525-71-121-21
7-221-69-75-23973-83-1217967-37-
9511715762896522-158-18150-140202084026-
1542158000000000000000000000000000000
0000000000000000000000000000000000000
0000000000000000000

Enter the key to decrypt = 4 Decrypted Text:
security system typically attempt to introduce
barriers such as passwords or other authentication
mechanisms while HCI designers attempt to
remove such barriers
C:\j2sdk1.4.1_02\bin>
```

Conclusion

In this paper, we first analyze the security aspect of online data storage. We make a thorough explanation on the security problem in data stored at online storage provider. Then we propose a new security mechanism with Walsh Cryptosystem which is based on based encryption / decryption algorithm. The proposed concept is secure and easily applicable.

REFERENCES

1. Dripto Chatterjee, Joyshree Nath, Suvadeep Dasgupta and Asoke Nath, 2011, International Conference on Communication Systems and Network Technologies "A new Symmetric key Cryptography Algorithm using extended MSA method: DJSA symmetric key algorithm"

2. Neeraj Khanna, Joel James and Amlan Chakrabarti, 2011 International Conference on Communication Systems and Network Technologies, **"**New Symmetric key Cryptographic algorithm using combined bit manipulation and MSA encryption algorithm: NJJSAA symmetric key algorithm"

3. Risley, J. Roberts, P. LaDow, "Electronic security of real-time protection and SCADA communications", Schweitzer Engineering Laboratories, SEL 2003 Inc. Pullman WA USA.

4. J. Nechvatal, E. Barker, L. Bassham, W. Burr, M. Dworkin, J. Foti, and E. Roback, "Report on the Development of the Advanced Encryption Standard (AES)", October 2, 2000

5. J. Blömer, M. Otto, J. Seifert. " A new CRT-RSA algorithm secure against bellcore attacks." Proceedings of the 10th ACM Conference on Computer and Communications Security, pp.310 –320, Washington D.C., USA, October 2003.

6. D. Wagner, "Cryptanalysis of a provably secure CRT-RSA algorithm." Proceedings of the 11th ACM conference on Computer and communications security, pp. 92 – 97, Washington D.C., USA, 2004.

7. K. Yumbul and E. Savas. "Efficient, secure, and isolated execution of cryptographic algorithms on a cryptographic unit." Proceedings of the 2nd international conference on Security of information and networks, pp. 143 – 151, Famagusta, North Cyprus, 2009

8. Ronald Monzillo, Chris Kaler, Anthony Nadalin, and Phillip Hallem-Baker. Web services security: Saml token profile 1.1. Technicalreport, OASIS, 2010.

9. Lou, D.C and Liu J. L.2002. "Steganography Method for Secure Communications." Elsevier Science on computers & Security, 21,5:449-460.

10. Schneier. Applied Cryptography. Wiley and Sons, 1996.

11. B. Schneier and N. Ferguson. Practical Cryptography. Wiley Publishing, 2003.

12. OWASP Top Ten Most Critical Web Application Security Vulnerabilities, http://www.owasp.org

13. Cryptography and Network, William Stallings, Prectice Hall of India