International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 03 | Mar-2016 www.irjet.net p-ISSN: 2395-0072

A Double Layer Encryption Algorithm based on DNA and RSA for Security on Cloud

Karandeep Kaur

Assistant Professor, Dept. of Computer Science, Guru Nanak Dev University, Amritsar

---

Abstract - *Cloud computing nowadays is being comprehended as the potential solution to the problem of ever rising demand for storage, resources and computing power. Its pay-as-you-use policy is turning out lucrative especially for small and medium sized organizations. The pros of cloud service are well exhibited by the fact that it is one of the most worked upon area in the field of information technology research, with a large number of new service providers coming in the market with improvised plans. Conversely, its weaknesses are deferring cloud users from adopting it for confidential works. Out of these, security is the most important aspect. Research work is signifying new techniques to make the cloud model more secure and trust worthy and a lot of work is being done regarding this. This paper aims to suggest an approach which is a double layer encryption method to ensure security in cloud. It is based on a popular cryptography algorithm RSA and Deoxyribonucleic Acid (DNA) sequences, which is a relatively novel technique.*

Key Words: Cloud computing, Cryptography, RSA algorithm, DNA sequences, Cloud

1. INTRODUCTION

1.1 Cloud Computing

Cloud computing has emerged as a resource provider; both data and applications are stored on cloud servers and can be used when required with the help of internet. The computing power of remote systems can also be exploited using cloud service with minimal charges (mostly pay-per-use policy).

Cloud services can be offered by the means of various deployment models. Software as a service (SaaS) model offers application software to its users. Infrastructure as a service (IaaS) model offers various resources like data storage space, computing power etc. to its users. Platform as a service (PaaS) model offers platform and several other tools and services to the users.

There are different types of clouds like Public cloud which is open to all the users online. The users who have subscribed to the public cloud service provider can access its services based on their rental plans. Private cloud on the other hand, belongs to a particular organization and has everything reserved for its own users only. It can't be accessed by the common public. Community cloud is a special-purpose cloud which is built by a group of organizations which share a common concern or purpose. Hybrid cloud which is a latest development combines the features of both public and private cloud.

1.2 Security in Cloud

Cloud computing is undoubtedly a budding technology which offers numerous benefits to its users. Scalability, flexibility, low cost and pay-per-use policy forms a major part of its appeal. Easy availability of services through internet adds to its fascination. Also, it can be easily integrated with the existing facilities. In short, cloud can empower its users with great power through comparatively easy and less complicated interfaces and minimal resources.

As every new work of research has its downside, similar is the case with cloud. The cloud storage servers are owned by some third party and located in different locations rather than the users or organizations themselves, which makes them quite apprehensive of the security part of it. Cloud is susceptible to various threats like disclosure of information, Denial of Service (DOS) attacks, malicious users, encryption failures etc. Security on cloud, therefore, is a major area of research as well as concern to the users.

1.3 Cryptography Basics

Cryptography is an ancient art of hiding information to protect it from malicious users. The information to be sent over the network called plaintext is converted into another form called cipher text by some Encryption algorithm. On the receiver's side, the original information can be recovered from the cipher text by applying a Decryption algorithm on it. There are various algorithms to encrypt and decrypt the secret information. These are broadly classified as symmetric and asymmetric methods. Symmetric methods use the same key for encryption and decryption while asymmetric methods use two keys (private key and public key).

A symmetric algorithm uses the same key for both encryption and decryption. In this method, the decryption algorithm is the exact opposite of the encryption algorithm.

The sender uses the common key, sometimes called the secret key, along with the encryption algorithm to generate cipher text. The receiver uses the common key along with the decryption algorithm to generate plain text back. Popular examples of such algorithms are DES (Data Encryption Standard), AES (Advanced Encryption Standard) etc.

An asymmetric method uses two separate keys: public and private. The private key is available only to the receiver while the public key is announced to everyone. The sender encrypts the plain text using the public key and the receiver decrypts it using its private key.

The decryption algorithm is designed such that it is not the exact opposite of the encryption algorithm. So anyone can encrypt the information using public key but only the authorized receiver can decrypt it using its private key. Popular examples of asymmetric algorithms are RSA, Diffie Hellman etc.

1.4 RSA Algorithm

This is an asymmetric cryptographic algorithm suggested by Rivest, Shamir and Adleman in 1977. The algorithm uses two keys, public (which is announced to all) and private (which is kept secretly with the receiver) .The public key 'e' is used to encrypt the plain text 'P' into cipher text 'C' and a private key 'd' converts cipher text back to plain text.[2]

Encryption Algorithm:

$C = P^e$ modulo N

This generates the cipher text 'C' from the plain text 'P' using public key 'e'.

Decryption Algorithm:

$P = C^d$ modulo N

This generates the plain text 'P' from the cipher text 'C' using private key 'd'.

Key Generation

- First two very large prime numbers 'p' and 'q' are chosen
- Calculate N = p x q and Ø = (p-1) x (q-1)
- Choose 'e' such that 1 < e < Ø and GCD(e, Ø) = 1
- Choose 'd' such that (d x e) modulo Ø = 1

The RSA algorithm is very secure because it is very difficult to find the values of p and q from N if p and q are very large prime numbers. Anyone can use the public key to encrypt the message but only an authorized user can decrypt it using the private key.

1.5 DNA Algorithm

DNA is the starting point of all life species. DNA molecules contain strands of nucleotides which are named after the nitrogen base they are made of: Adenine (A), Cytosine (C), Guanine (G) and Thymine (T). The DNA molecules have a double helix structure which is a combination of complementary strands A to T and G to C. These four bases can be represented by a combination of two bits as shown in Fig 1.

Using DNA features for cryptography is a relatively novel concept. However, inter-disciplinary research has always offered mammoth sized benefits to computer science. Similar is the DNA encryption technique. There are many ways we can exploit the DNA characteristics. The plain text can be converted into ASCII codes and further into binary code, which can then be replaced by DNA bases A, T, G and C as shown in Fig. 1. Once there is a sequence of A, T, G and C characters, they can be assigned some decimal numbers based upon a reference DNA sequence from a gene sequence database. The probability of finding the reference sequence used is very low as there are millions of sequences available to be used freely. The concept will be clarified with the help of an example in the later part of this paper.

Bits	Base
00	A
01	T
10	G
11	C

Fig -1: DNA encoding

2. PROPOSED ALGORITHM

The algorithm that has been proposed in this paper consists of a double layer security algorithm. DNA algorithm is used followed by RSA algorithm to ensure twofold protection in cloud environment. All the data stored in cloud servers can be encrypted using this algorithm. Only the authorized user, who possesses the private key as well as the DNA reference sequence, will be able to decrypt the data from the cloud.

Encryption Algorithm:

Step A: Convert the plain text into its binary format.

Step B: Assign DNA bases to the binary format as given in Fig. 1.

Step C: Replace DNA bases with the numbers given in the reference DNA sequence.

Step D: Using RSA algorithm, convert it into cipher text C using $C = P^e$ modulo N

Step E: Transmit the cipher text C.

Decryption Algorithm:

Step A: Convert the cipher text into plain text using $P = C^d$ modulo N

Step B: Replace the numbers with DNA bases given in the reference DNA sequence.

Step C: Replace DNA bases with the binary numbers given in Fig. 1.

Step D: Convert the numbers back into plain text.

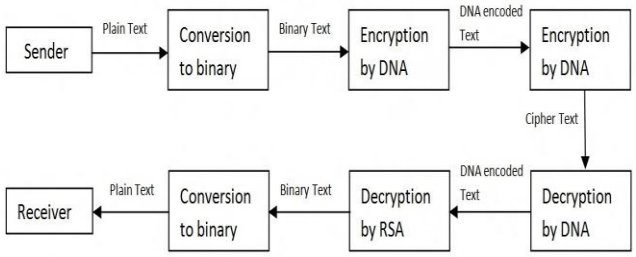

Fig -2: The proposed Algorithm

To show the working of our algorithm, we are citing an example. Let the plain text be '12'. After conversion to binary it becomes '1100'. Now we assign DNA codes to it using values in Fig. 1. The output is 'CA'. The next step is to use our DNA reference string to assign numbers to the output. The reference string used by us is shown below. We have taken one of the DNA strings available freely at BDGP (Berkeley Drosophila Genome Project) site. According to the reference string, 'CA' gets replaced by number '14'. Now we use RSA algorithm on this using the public and private keys calculated below. The output is '20'. The procedure is illustrated in Fig. 3 and 4.

Encryption Steps	INPUT	OUTPUT
A: Decimal to binary conversion	12	1100
B: DNA encoding	1100	CA
C: DNA reference string coding	CA	14
D: RSA using $C = 14^7$ modulo 33	14	20

Fig -3: Encryption process

Decryption Steps	INPUT	OUTPUT
A: RSA using $P = 20^3$ modulo 33	20	14
B: DNA reference string decoding	14	CA
C: DNA decoding	CA	1100
D: Binary to decimal conversion	1100	12

Fig -4: Decryption process

DNA Reference String: [TA1], [GC2], [TG3], [AG4], [CT5], [CT6], [TT7], [TG8], [AC9], [TC10], [TC11], [TA12], [AT13], [CA14], [CC15], [CC16], [TC17], [CG18], [TG19], [CT20] [3]

RSA Keys:

1. Choose two primes, p = 3 and q = 11
2. Find n = p x q = 3 x 11 = 33
3. Compute φ = (p - 1) x (q - 1) = 2 * 10 = 20
4. Choose 'e' such that 'e' and 20 are co-prime and 1 < e < 20. Lets take e = 7
5. Calculate the value of 'd' such that (d x e) modulo φ = 1. One such value is d = 3 as {(3 x 7) modulo 20 = 1}
6. Our Public key 'e' = 7
7. Our Private key 'd' = 3

3. RESULTS

There are approximately 163 million DNA sequences which are open for public access [8]. It is very tough for anyone to find out the particular DNA reference sequence used in the proposed algorithm. It definitely incorporates higher level of security than traditional cryptography algorithms.
RSA algorithm has been proven over the years, as one of the well-built algorithms to hack. Its security can be comprehended from the fact that it is very complicated to find out the values of 'p' and 'q' from the value of 'N' if both are taken as huge prime numbers (containing about 100 digits).
Using a combination of DNA with RSA ensures twofold protection in cloud environment where there are more chances of breaches.

4. CONCLUSION

With the escalating boom in cloud services, as users move into the cloud environment, they have relatively lesser control over their data. This is what makes the issue of security on cloud a crucial concern to the users. In this paper, we have discussed about the cloud computing scenario in context of the encryption techniques which can be used in it. A new cryptographic algorithm has been proposed based on the DNA based algorithm followed by the popular RSA algorithm to guarantee double security on cloud. Exploiting the power of DNA sequences and their very large number available to access, DNA algorithm provides a novel concept of cryptographic algorithms which is very hard to break (almost impossible as millions of DNA sequences are available). When followed by RSA algorithm, our algorithm makes sure there is no way to crack the cipher text.

REFERENCES

[1] Subashini, S., Kavitha, V.: A survey on security issues in service delivery models of cloud computing. J. Netw. Comput. Appl. 34, 1–11 (2011)

[2] Forouzan, Behrouz A., and Sophia Chung. Fegan. *Data Communications and Networking*. New York: McGraw-Hill Higher Education, 2007. Print.

[3] Gugnani, G., Ghrera, S. P., Gupta, P. K., Malekian, R., & Maharaj, B. T. (2015). : Implementing DNA Encryption Technique in Web Services to Embed Confidentiality in Cloud. Advances in Intelligent Systems and Computing Proceedings of the Second International Conference on Computer and Communication Technologies, 407-415.

[4] Thakur, A.S., Gupta, P.K., Gupta, P.: Handling data integrity issue in SaaS cloud. In: Satapathy, S.C., Biswal, B.N., Udgata, S.K., Mandal, J.K. (eds.) FICTA 2014. LNCS, vol. 328, pp. 127–134. Springer International Publishing (2015)

[5] Terec, R., Vaida, M.F., Alboaie, L., Chiorean, L.: DNA security using symmetric and asymmetric cryptography. Int. J. New Comput. Archit. Their. Appl. 1, 34–51 (2011)

[6] Ramesh, B., Bhavani, S. A., & Muralidhar, P. (2015). Enhancement of Stream Ciphers Security Using DNA. *Advances in Intelligent Systems and Computing Proceedings of the Second International Conference on Computer and Communication Technologies,* 637-643.

[7] Liu, H., Lin, D., Kadir, A.: A novel data hiding method based on deoxyribonucleic acid coding. Comput. Electr. Eng. 39, 1164–1173 (2013)

[8] Shiu, H.J., Ng, K.L., Fang, J.F., Lee, R.C., Huang, C.H.: Data hiding methods based upon DNA sequences. Inf. Sci. 180, 2196–2208 (2010)