Asian Journal of Information Technology 15 (15): 2699-2701, 2016
ISSN: 1682-3915

DNA Based Secret Sharing Algorithm for Multicast Group

K. Saravanan and T. Purusothaman
Department of CSE, Government College of Technology, Coimbatore, India

Abstract: The transmission of a highly secured message that is highly essential in the global scenario today is possible by employing the DNA cryptography. Based on modified Shamir's secret algorithm as well as DNA based encryption and decryption, the present algorithm is arrived. In between server and clients, a provision of a three level secured message transmission is implemented. In this method, only when all the clients in the group are involved in the decryption process, decoding of the secret message is possible. This ensures security of first level. The server authenticates the clients to whom the secrets can be shared. It is the security of second level. The decoding of the original DNA sequence related to the secret is done only by the primers. That completes the final security level. The simulation validates that the proposed algorithm is much better from the point of security.

Key words: DNA encoding, secured message, shamir's secret, multicasting, decryption

INTRODUCTION

DNA cryptography is popularly known for its storage capacity of a large size data. It can store information related to living organisms. Each living organism will have unique DNA information. It is handy for present work as it is characterized by data transmission that is highly secured. Cryptography when combined with molecular biology provides more secured data transmission and also data hiding. A plain text message is encoded using DNA sequences. In a typical cryptography, hackers can access the encrypted messages. Encoding of genetic information is done by a sequence of nucleotides Adenine-A, Cytosine-C, Guanine-G and Thymine-T. The stated nucleotides are termed as base pairs. To maintain a helical structure, these are held to a phosphate and a sugar.

The DNA structures are classified into a single strand and a double strand. In the latter, two strands form complementary to each one. For the third level of security where primers are used, the double strand DNA concept is employed. DNA's storage capacity is typically 1 g = 10^8 TB. In simple language, the whole world's information can be stored in a tiny quantity of DNA. The shortcomings of cryptography employing DNA are its large computational complexity and need of a sophisticated biomolecular laboratory. Today's cryptography employs latest biological technologies. Hybridization and PCR amplification are included in the referred technologies. The later technologies are expensive, unwieldy and time consuming. Since, biological technologies are avoided in the proposed algorithm, it is much simple.

Literature review: Chen (2003) put forth a DNA based cryptographic approach founded on message transformation utilizing carbon nanotube as well as cryptosystem using DNA. The former helps transfer of data among formal binary storage and DNA.

Sabari and Setua (2012) suggested a technique which is based on parallel cryptography which uses one time pad, DNA based digital coding, molecular structure based on DNA and technique founded upon DNA hybridization. The technique used for generating encryption key is one time pad. The advantage is utilizes parallel technique for decrypting the message and has a storage capacity that is extraordinary. The disadvantage is it requires a bio molecular laboratory of high tech and also long computing time.

Sherif T Amin introduced a YAEA encryption algorithm implementation which is DNA based. The positive aspect in the suggested algorithm will never transfer the real message as such over the network. In addition it can handle comfortably large products of digital information. The primary disadvantage is that plain text size escalates both the encryption as well as decryption time.

Mandge and Choudhary (2013) offered a matrix manipulation based DNA encryption technique and scheme for generating secure key. The plus point of the

Corresponding Author: K. Saravanan, Department of CSE, Government College of Technology, Coimbatore, India

algorithm is that it provides a good security layer not giving any clue regarding the plaintext. The negative point is that the security depends only upon the key. Cui *et al*. (2008) suggested the encryption technique using public key, during the communication utilizing DNA digital coding, DNA synthesis and PCR amplification thus ensuring the security. In addition asymmetric keys are utilized. Cryptography computing along with the complexity associated with the biological scheme doubly ensures the scheme's security. The encryption cost for the scheme is small. But, the security is depending solely upon the decryption key. Li *et al*. (2008) utilized a key expansion matrix which is DNA based generation key scheme that is a novel algorithm proposed by them. The computational speed is hastened by a scheme that employs random key generation. The key totally holds the security. Kumar and Singh (2011) utilized DNA sequences and arrived at a technique for secret data writing. It easily detects any change in the cipher text. The drawback of this scheme is the key totally holds the security.

Zhang *et al*. (2012) suggested a DNA fragment assembly based cryptography. A symmetric key is employed and also the cipher text generated is secured and also it is short. Since, the length of fragment of DNA is small, any hacker can easily find out the secret.

MATERIALS AND METHODS

The proposed algorithm: DNA based secret sharing algorithm for multicast group

Algorithm for the new clients to join the group:

- Step 1: The Key Server (KS) authenticates each client
- Step 2: An ID number is specifically assigned to every client who newly joins the group. This is decided by the KS

Algorithm for generating a DNA sequence

Step 1: A primer value r (an integer chosen at random) sent by the KS to serve as key to the multicast group clients M_1, M_2, ..., M_j, M_{j+1}. Clients' individual Ids (I_1, I_2, ..., I_{j+1}) are utilized by the server to generate a secret Cusing the following equation:

$$C = \left(1_1 \times 1_2 \times \ldots, \times 1_{j+1}\right) + r \quad (1)$$

Step 2: The computed C as a message broadcasted by the KS to all the clients in the group.

Step 3: The primer r which is a remainder found by each member by dividing C by the respective ID.

Step 4: The KS converts the secret text message into the respective ASCII code. The latter in turn is converted into DNA form (a denotes 00, C denotes 01, G denotes 10, T denotes 11) leading to a single strand of DNA. So, a original DNA sequence is arrived by converting the completer text message.

Step 5: In the same way the primer value can also be finally ended up as a sequence of DNA. The DNA sequence of the corresponding primer gets divided into two equal parts (left half and Right half) by the KS. At the start of the original sequence of DNA, left half gets appended while at the end the complement of right half finds its place.

Step 6: On both the extremities of the DNA sequence which is modified recently, a duplicate DNA sequence is attached.

Generating the modified shamir's secret

Step 1: A positive secret integer of k digit is obtained from the complete sequence of DNA.

Step 2: The length of secret integer got through step 1 is divided among the number of clients. A polynomial equation of degree (n-1) is generated (Shamir, 1979) where n stands for multicast group clients:

$$F(x) = a_0 + a_1x + a_2x^2 + \ldots + a_{n-1}x_2^{n-1}$$

where, Q and R stand for the quotient and the remainder values, respectively. From the first digit of the secret integer, upto the (Q+R)th digit is alloted to a_0, the polynomial coefficient. The rest of the digits are split into equal parts (Q digits each from left to right). These are assigned to the remaining coefficients.

Step 3: Every client is communicated their respective (xi, yi = f(x)) value by the KS. The values is calculated with the help of the formula stated in Eq. 2.

Algorithm for decoding the secret:

Step 1: The (xi,f(xi)) value belonging to each client will be made known to the other members in the group.

Step 2: n simultaneous equations are generated by every client using the above mentioned values which includes the client's own value.

Step 3: The equations stated above are solved by Crammer's rule. The coefficients a_0, a_1, a_2, ..., a_n are obtained by each client. Then finally the secret integer value is arrived by placing a_0, a_1, a_2, ..., a_n successively,

Step 4: The latter will be converted by each client into its binary equivalent and finally a sequence of DNA is arrived. The LH, complement of RH of primers and also the duplicate sequence of DNA are eliminated.

Step 5: Every client next recovers the original DNA secret and in turn changes it into the form of binary and latter its respective ASCII.

Step 6: The final secret text is decoded by every client replacing each 8 bit by ASCII.

Table 1: Simulation results

Description	Results
Number of new clients joined the group	3
Random number generated by KS (r)	33
The cipher text generated (C)	6699363
The secret plain text created by KS	Hello
The DNA sequence corresponding to the secret plain text generated by KS	TCGA TCTT TCGA TCGA TCGG
ASCII equivalent	01101000011001010110110001101100011011111
Binary equivalent of the random number 33	00100001
The DNA equivalent of the random number	ACAT
LH of random number	AC
Complement of RH (AT)	TA
The DNA sequence after appending LH and complement of RH	ACTCCATCTTTCGATCGGTA
The DNA sequence after adding the duplicate DNA sequence	GATGAAT ACTCCATCTTTCGATCGGTA TTGT
The DNA sequence converted into binary equivalent and then into positive secret integer	9178084191017463362653
The coefficients a_0, a_1 and a_2 of f(x)	$a_0 = 9178084$, $a_1 = 1901746$, $a_2 = 3362653$
The KS sends the (x_i, y_i) to the clients joining the group	$(x_1, y_1) = (1,14442483)$, $(x_2, y_2) = (1,26432188)$, $(x_3, y_3) =)$ 1,451447199)
The clients receiving the (x_i, y_i) of all other clients, compute the final secret after removing the duplicate sequence from original DNA sequence	Hello

RESULTS AND DISCUSSION

The validation by simulation is done for three clients. It can be repeated for more clients. The presence of many steps ensures better security. So, any unauthorized person who receives the intermediate message will never decode the original message which is intended for the receiving clients. The DNA sequence chosen is just random and the length of the sequence can be increased to a desired safe level (Table 1).

Unless the primer is known, decoding the secret is impossible. Obviously whenever a client leaves or a new client enters the group, the primer, secret and also the (x_i, y_i) values must be changed. The proposed algorithm is so secured that no hacker can access the encrypted message.

CONCLUSION

The proposed algorithm employs the DNA encoding which leads to a tightly secured message transmission for multicast applications. It must be mentioned that multiple securities are possible in the DNA based secret sharing algorithm as there are three levels of security. In particular where high security is essential like that of military applications this is more suitable. The proposed protocols are validated using Python and Java tools. The outputs reliably justify the efforts. The algorithm may be implemented for image as a future extension in place of text message.

REFERENCES

Chen, J., 2003. A DNA-based, biomolecular cryptography design. Proceedings of the IEEE International Symposium on Circuits and Systems, December 27-30, 2003, Cairo, Egypt, pp: 822-825.

Cui, G., L. Qin, Y. Wang and X. Zhang, 2008. An encryption scheme using DNA technology. Proceedings of the 3rd International Conference on Bio-Inspired Computing: Theories and Applications, September 28-October 1, 2008, Adelaide, Australia, pp: 37-42.

Kumar, D. and S. Singh, 2011. Secret data writing using DNA sequences. Proceedings of the International Conference on Emerging Trends in Networks and Computer Communications (ETNCC), April 22-24, 2011, IEEE, Udaipur, India, ISBN: 978-1-4577-0239-6, pp: 402-405.

Li, X.S., L. Zhang and Y.P. Hu, 2008. A novel generation key scheme based on DNA. Proceedings of the International Conference on Computational Intelligence and Security, December 13-17, 2008, IEEE, Suzhou, China, ISBN: 978-0-7695-3508-1, pp: 264-266.

Mandge, T. and V. Choudhary, 2013. A DNA encryption technique based on matrix manipulation and secure key generation scheme. Proceedings of the International Conference on Information Communication and Embedded Systems, February 21-22, 2013, Chennai, Tamilnadu, India, pp: 47-52.

Sabari, P. and S.K. Setua, 2012. DNA cryptography. Int. Conf. Elect.Comput. Eng., 7: 551-554.

Shamir, A., 1979. How to share a secret. Commun. ACM, 22: 612-613.

Zhang, Y., B. Fu and X. Zhang, 2012. DNA cryptography based on DNA Fragment assembly. Proceedings of the 8th International Conference on Information Science and Digital Content Technology (ICIDT), June 26-28, 2012, IEEE, Jeju, ISBN: 978-1-4673-1288-2, pp: 179-182.