A Hybrid Cryptosystem using DNA, OTP and RSA

Mahbuba Begum
Department of CSE
Mawlana Bhashani Science and Technology University, Tangail-1902, Bangladesh.

Jannatul Ferdush
Department of CSE
Jessore University of Science and Technology, Jessore, Bangladesh.

Md. Golam Moazzam
Department of CSE
Jahangirnagar University, Savar, Dhaka-1342, Bangladesh.

ABSTRACT
Cryptography is one of the major elements in data security and communication security. It is a technique of securing the communication by preventing third parties. A **hybrid cryptosystem** is the special set of rules using multiple ciphers or a combination of a series of well-defined steps that can be followed as a procedure of different types together by taking best advantage of each cipher. In this cryptosystem, a random secret key is generated used as a symmetric cipher which is required for all parties. Then the system encrypts this key via an asymmetric cipher using the recipient's public key. DNA cryptography is a new optimistic field in cryptography which hides the data in terms of DNA sequence to make it secured. The OTP **(one-time pad)** is an encryption technique in which a plaintext is paired with a one-time pre-shared key which is equal to or longer than the size of the message being sent. In RSA, a public key is used for encryption which specifies the transformation of plaintext into ciphertext and this encryption key differs from the decryption key which is kept private to secure data transmission. In this paper, a hybrid cryptosystem is proposed using DNA, the generic OTP technique and RSA to ensure high security in three levels. This method is very efficient for encrypting data, hiding text and preventing attacks.

General Terms
Cryptosystem, Hybridization.

Keywords
DNA, One Time Pad, RSA.

1. INTRODUCTION
Development and rapid growth of the internet and the computer, the exchange of information has become more desired and a big challenging task. Hence, it is necessary to concentrate on the security of the information so that we can protect the data from unauthorized access, use, replication or destruction. In order to get security requirements, different methods and systems have been developed in the mathematical cryptography for encoding and decoding the plaintext.

Encryption is one of the most effective data security methods. These methods are implemented using DNA cryptography. The DNA cryptography is a rising field in the area of DNA computing research. This cryptography plays a major role in next generation security. Simply this technique secures the data by using biological structure of DNA. It works on the concept of DNA computing.

The OTP **(one-time pad)** is an encryption technique in which a plaintext is paired with a one-time pre-shared key which is equal to or longer than the size of the message being sent. In this case, each bit or character of the original message is encrypted by integrating it with the exact bit or character from the pad using modular addition. If the key is truly random, is at least as long as the plaintext, is never reused in whole or in part, and is kept completely secret, then the resulting ciphertext will be impossible to decrypt or break [1][2][3].

DNA cryptosystem is used for concealing any kinds of message. OTP is known as the only theoretically unbreakable crypto- system when the random key (pad) is truly random, never reused, and kept secret [4][5]. True random numbers are generated using a physical process such as nuclear decay or thermal noise of electronic circuits, because mathematical algorithms can provide pseudo-random numbers only. In DNA-based cryptography, we can directly use the random hybridization process of DNA motifs for the random key generation [6][7][8][9]. Although several ideas of OTP using DNA strands have been presented [9][10][11].

OTP can be easily integrated with DNA. But they use the same key to encrypt and decrypt the data for the safe and secured communication over an unsafe channel. But, symmetric cryptosystem is not so secured and its security depends on key. So, in this cryptosystem, key sharing is a big disadvantage. On the other hand, RSA is an asymmetric cryptography where different keys are used to encrypt and decrypt the data. It is also known as public key cryptography. There are two (02) different keys in RSA where one of them can be given to everyone. The other key must be kept private.

But, the plain RSA cannot be used in real life implementation because of its limitation. So, development an authentic, secured and an impartial cryptosystem which perfectly balanced between security and efficiency is so tough and challenging.

Hybrid cryptosystem combines the best advantages of multiple ciphers. It incorporates more security because of combination of asymmetric and symmetric encryption. Also, it takes the higher speed of symmetric cipher. By this, it ensures the efficiency of the system. The main objective of this research work is to develop a hybrid cryptosystem using DNA, OTP and RSA to ensure high security in three levels by taking best advantage of each cipher.

2. DNA STRUCTURE AND RSA ALGORITHM

2.1 DNA Structure
Deoxyribonucleic acid is a molecule called nucleotides. Each nucleotide is made up of one of four nitrogen-containing nucleobases — cytosine (C), guanine (G), adenine (A), or thymine (T) — a sugar called deoxyribose, and a phosphate group. DNA contains double hydrogen bond with (A, T) and (C, G) which are complement to each other. There are two strands of DNA sequence like stranded DNA (ssDNA) and double stranded DNA (dsDNA). An ssDNA can form double stranded DNA (dsDNA) with other ssDNA. SsDNA and

DsDNA are complementary with each other. This process is called Hybridization [12].

DNA Hybridization is a process by which two ssDNA segments are combined to form a DNA sequence. In hybridization process it is necessary that a hydrogen bond is formed only between two Watson-Crick pairs. If a Watson-Crick pair is not found at the corresponding position, then the hybridization of that pair fails. In DNA hybridization, it is necessary that the length of both ssDNA segments should be same. If not, fragmentation occurs and fragment assembly has to be done in order to repair the DNA molecule [13].

The figure 1 shows the DNA hybridization.

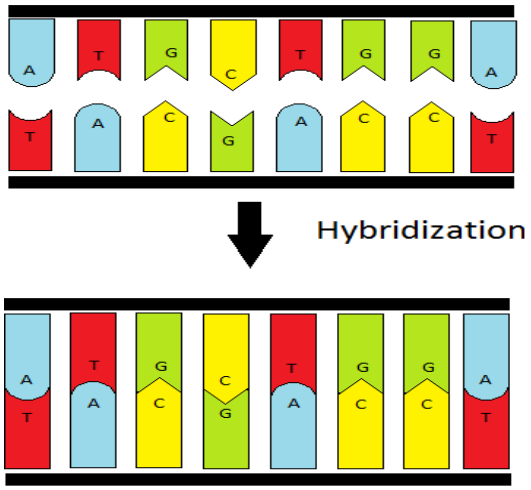


Fig 1: Hybridization [10]

DNA cryptography is known as concealing data in terms of DNA sequence. In this cryptosystem, each letter of the alphabet is transformed into a various combination of the four bases which make up the human deoxyribonucleic acid (DNA).

In this algorithm, Symmetric key exchange technique is used to calculate one time pad (OTP) DNA sequence to facilitate secure communication. Randomly generated OTP DNA sequence is used only once for encryption and decryption process. It gives unique result for a particular statistical calculation [9]. In this paper we used Symmetric key exchange technique to calculate symmetric key k.

2.2 RSA Algorithm

RSA algorithm is an asymmetric cryptographic algorithm. Asymmetric actually means that it works on two different keys i.e. **Public Key** and **Private Key.** In this algorithm, the public Key is given to everyone and the private key is kept private. This cryptosystem encrypts the message using the public key which is known to everyone and this key can only be decrypted with the private key.

Example:

The **RSA algorithm** involves four steps: key generation, key distribution, encryption and decryption. The public key can be known by everyone and is used for encrypting messages. The intention is that messages encrypted with the public key can only be decrypted in a reasonable amount of time using the private key. Here, plaintext is encrypted in blocks. This type of cryptosystem ensures that only the right people who knows the key can read the information. Here, the sender and receiver can confirm each other's identity and origin/destination of the information. In this case, the size of resulting encrypted text is more than the original plaintext size. But, key generation and decryption rate is slow in this cryptosystem.

3. HYBRID CRYPTOSYSTEM USING DNA, OTP AND RSA

3.1 Hybrid Cryptosystem

A hybrid cryptosystem is a set of rules using multiple different types of ciphers together by taking best advantage of each cipher. The common approach is to generate a random secret key for a symmetric cipher, and then encrypt this key via an asymmetric cipher using the recipient's public key.

In this paper, DNA, OTP and RSA are combined together by taking best advantage of each cipher to ensure speed and security.

3.2 Cryptographic Algorithm

In this paper, encryption of plaintext is done by combination of DNA cryptography, binary One Time Pad (OTP) and RSA scheme. This algorithm uses three keys for three stages of encryption. These keys are used for encryption on sender side and for decryption on the receiver side. Hence, this algorithm uses both symmetric and asymmetric keys. We have to find out RSA cipher. The length of binary cipher is equal to eight times of the length of RSA cipher.

3.2.1 Sender's Algorithm

Suppose, the Sender has plaintext, binary key and RSA public key. Now the steps of sender's algorithm are as follows:

1. Convert the plaintext message to its corresponding ASCII values. Then apply RSA algorithm to each ASCII values by RSA key and find out RSA cipher.
2. Convert each digit of RSA cipher to binary plaintext of 8 digit.
   Length of binary cipher = 8*length of RSA cipher.
3. Perform XOR operation between binary cipher and random Binary key. It is called XOR cipher.
4. Convert the result into DNA sequence by table 1 and Found DNA cipher.

Table 1. Binary codes [10]

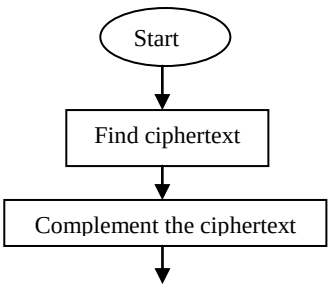
Nucleotide	**A**	**C**	**G**	**T**
Code	00	01	10	11

5. Complement the DNA cipher and this is the final ciphertext.
6. Sent this ciphertext to the receiver.

3.2.2 Receiver's Algorithm

Receiver's side decryption process is shown below:

.



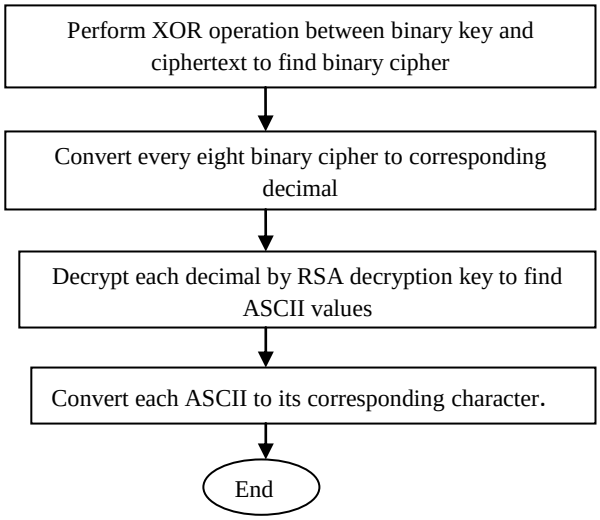


Fig 2: Receiver side

4. SYSTEM MODEL

Fig 3 shows the total system model of proposed cryptosystem. There are two parties: sender and receiver. Sender wants to share securely some information with receiver. To securely share information, sender needs to encrypt his message with some keys. Receiver also needs keys to decrypt message. So the proposed algorithm has total seven (07) steps that have been shown in fig 3.

From the receiver and sender algorithm it has been clear that we use only two keys for our encryption and decryption process. One is random binary key and other is RSA key. Conversation of DNA sequences and its complement is not part of key. It is just an encoding process for proposed algorithm.

So when sender wants to share something with receiver, he generates a random binary key. The size of binary key is equal to the eight times multiple of plaintext size. Then the sender sends it to receiver. At the same time, receiver generates RSA key pair that means RSA public and private key. Receiver only sends the public key to sender. Then, sender encrypts the plaintext according to sender's algorithm. After encryption he sends it to receiver. Receiver then decrypts it by receiver's algorithm.

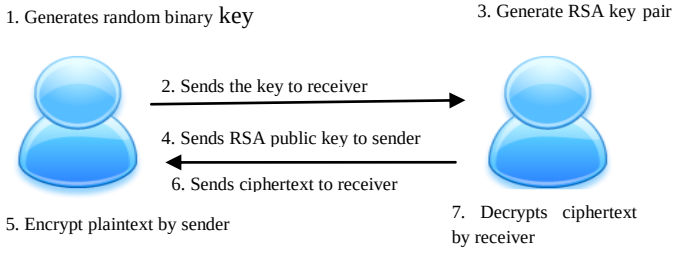


Fig 3: System Model

Here, steps 1, 2, 5 6 are performed by sender and 3, 4, 7 are performed by receiver.

5. EXPERIMENTAL ENVIRONMENT

The prototype of this algorithm is developed by Matlab 2013 on Intel(R) Pentium(R) Dual CPU T3400 @2.16 GHZ 32 bit processor with 2 GB RAM running under Windows 7.

6. EXPERIMENTAL RESULT

Suppose sender wants to send 'I love my country' to receiver. So size of plaintext=17

Step 1: Sender generates a random binary key of size=8*size of plaintext. So for this plaintext it becomes =136. Let it is:

1 1 0 1 0 1 1 0 0 0 0 0 1 1 1 1 0 1 0 1 0 1 0 1 1
1 1 1 0 0 1 1 1 0 1 1 0 1 1 0 1 1 0 1 1 1 1 1 0 0 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0
0 0 0 0 0 0 0 1 1 1 0 0 0 1 0 0 0 0 0 1 0 1 1 0 1
1 1 1 0 1 0 0 1 1 0 0 1 1 1 1 0 1 0 1 0 1 1 0 0 0
0 0 1 0 1 0 0 0 0 1 1

Step2: Sender now sends this key to receiver.

Step 3: Receiver generate RSA key pair. Let it is (7,209) and (103,209)

Step 4: Receiver sends just (7,209) public key to sender.

Step 5: Now sender's algorithm works.

Step 5.1: Convert every plaintext character to corresponding ASCII character. So for this example it is:

73 32 108 111 118 101 32 109 121 32 99 111
117 110 116 114 121

Step 5.2: Now every ASCII value is encrypted by RSA public key. So now RSA cipher is:

17 10 48 188 101 161 10 98 121 10 44 188
116 165 52 38 121

Step 5.3: Now every RSA cipher is converted to eight digit binary value. It can be called binary cipher. It becomes:

0 0 0 1 0 0 0 1 0 0 0 0 1 0 1 0 0 0 1 1 0 0 0 0 1
0 1 1 1 1 0 0 0 1 1 0 0 1 0 1 1 0 1 0 0 0 0 1 0 0
0 0 1 0 1 0 0 1 1 0 0 0 1 0 0 1 1 1 1 0 0 1 0 0 0
0 1 0 1 0 0 0 1 0 1 1 0 0 1 0 1 1 1 1 0 0 0 1 1 1
0 1 0 0 1 0 1 0 0 1 0 1 0 0 1 1 0 1 0 0 0 0 1 0 0
1 1 0 0 1 1 1 1 0 0 1

Step 5.4: Now XOR operation is performed between binary cipher and binary key: Then the result becomes:

1 1 0 0 0 1 1 1 0 0 0 0 0 1 0 1 0 1 1 0 0 1 0 1 0
1 0 0 1 1 1 1 1 1 0 1 0 0 1 1 0 1 1 1 1 1 1 1 0 1
1 0 1 0 1 0 1 1 1 0 0 0 1 0 0 1 1 1 1 1 0 0 0 0 0
0 1 0 1 0 0 0 0 1 0 1 0 0 0 0 1 1 1 1 1 0 1 0 1 0
1 0 1 0 0 0 1 1 1 1 0 0 1 1 0 1 1 1 1 0 1 1 1 0 0
1 1 1 0 0 1 1 1 0 1 0

Step 5.5: Now this XOR cipher is converted into DNA cipher:

TACTAACCCGCCCATTTCATCTTTCGGGTGAGCTTAAAGGACCA
ATTGGGGGATTATCTGTGCTATGG

Step 5.6: Complement it to find complement cipher.

ATGATTGGGCGGGTAAAGTAGAAAGCCCACTCGAATTTCCTGG
TTAACCCCCTAATAGACACGATACC

Step6: Sender sends this complement cipher as the ciphertext to receiver.

Step 7: Receiver receives the ciphertext and runs receiver's side algorithm. Actually receiver's side algorithm is just opposite to sender's side algorithm. Here just RSA decryption

key means here RSA private key is used. After performing RSA decryption when receiver converts it to character it becomes:

"I love my country".

7. PERFORMANCE ANALYSIS

Fig 4 represents the encryption and decryption time for different plaintext size. From graph it has been clear that decryption time is greater than encryption time.

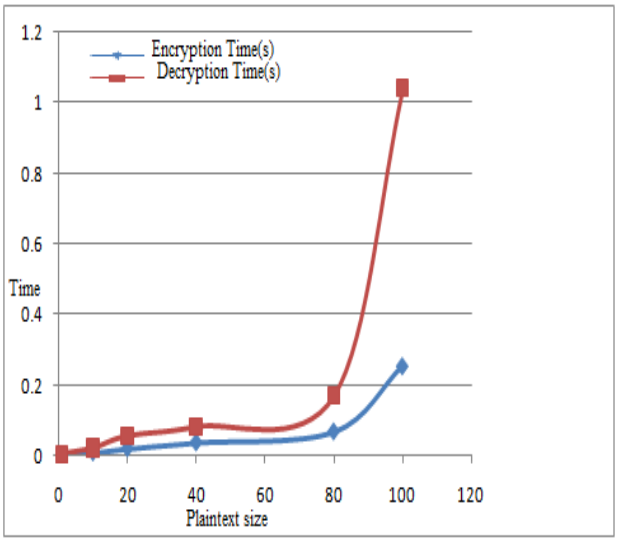


Fig 4: Time requirement for encryption and decryption

8. CONCLUSION AND FUTURE WORK

In this paper a new text encryption technique based on DNA, One Time Pad and RSA is implemented to ensure high security in three levels. Actually encryption of plaintext is done in three levels. It is a hybrid approach which combines the idea of symmetric and asymmetric cryptosystem. Thus, it ensures more security than those processes which used symmetric cryptosystem. But in this paper, when ASCII values are encrypted by RSA, we limit our ciphertext size is multiple of 8 bit. This is the limitation of the system. In future, this limitation will be tried to be solved. Also, the random binary key needs to be shared between sender and receiver. So, how securely the binary key shared between two parties is also a future concern.

9. REFERENCES

[1] ^a b c "Intro to Numbers Stations". Retrieved 13 September 2014.

[2] "The only unbreakable cryptosystem known—the Vernam cipher". Pro-technix.com. Retrieved 2014-03-17.

[3] "One-Time Pad (OTP)". Cryptomuseum.com. Retrieved 2014-03-17.

[4] Schneier, B.: Applied Cryptography: Protocols, Algorithms, and Source Code in C. John Wiley & Sons. Inc., Chichester (1996).

[5] Cameron, P.J.: Notes on cryptography (2003), http://www.maths.qmw.ac.uk/~pjc/notes/crypt.pdf3, Number 1, Nov 27, 2002.

[6] Adleman, L.M., Rothemund, P.W.K., Soweis, S., Winfree, E.: On Applying Molecular Computation to the data encryption standard. Journal of Computational Biology 6, 53–63 (1999).

[7] Leier, L., Richter, C., Banzhaf, W., Rauhe, H.: Cryptography with DNA binary strands. Biosystems 57, 13–22 (2000).

[8] Chen, J.: A DNA-Based, Biomolecular Cryptography Design. In: 2003 IEEE International Symposium on Circuits and Systems, vol. 3, pp. 822–825 (2003).

[9] Gehani, A., LaBean, T., Reif, J.: DNA-Based Cryptography. In: Jonoska, N., P˘aun, G., Rozenberg, G. (eds.) Aspects of Molecular Computing. LNCS, vol. 2950, pp. 167–188. Springer, Heidelberg (2003).

[10] Hirabayashi, M., Kojima, H., Oiwa, K.: Effective Algorithm to Encrypt Information Based on Self-Assembly of DNA Tiles. In: Nucleic Acids Symposium Series (53), pp. 79–80 (2009).

[11] Chen, Z., Xu, J.: One-Time-Pads Encryption in the Tile Assembly Model. In: Kearney, D. (ed.) Third International Conference on Bio-Inspired Computing: Theories and Applications. IEEE BICTA 2008, pp. 23–29 (2008).

[12] Tausif Anwer, Abhishek Kumar, Sanchita Paul: DNA Cryptography Based on Symmetric Key Exchange, International Journal of Engineering and Technology, Vol. 7, No. 3, 2015, pp. 938 – 950.

[13] Shreyas Chavan: DNA Cryptography based on DNA Hybridization and One Time Pad scheme, International Journal of Engineering Research & Technology, Vol. 2, Issue 10, October 2013, pp. 2679-2682.

10. AUTHOR PROFILE

Mahbuba Begum received her B.Sc. and M.S. degree in Computer Science and Engineering from Jahangirnagar University, Bangladesh, in the year 2007 and 2009, respectively. Currently, she is serving in the Department of Computer Science and Engineering, Mawlana Bhashani Science and Technology University, Tangail, Bangladesh as an Assistant Professor. Her research interest is Cryptography, Image Processing, Pattern Recognition, Face Recognition and Trademarks Recognition.

Jannatul Ferdush received her B.Sc. degree in Computer Science and Engineering from Khulna University of Engineering & Technology, Bangladesh, in the year 2016. Currently, she is serving in the Department of Computer Science and Engineering, Jessore University of Science and Technology, Jessore, Bangladesh as a Lecturer. Her research interest is Cryptography, Image Processing, Cloud Computing and Artificial Intelligence.

Md. Golam Moazzam received his B.Sc. and M.S. degree in Computer Science and Engineering from Jahangirnagar University, Bangladesh. Currently, he is serving in the Department of Computer Science and Engineering, Jahangirnagar University, Dhaka, Bangladesh as a Professor. His research interest is Cryptography, Image Processing, Pattern Recognition, Face Recognition and Trademarks Recognition.