ARUNAI PUBLICATIONS

Asian Journal of Multi-Disciplinary Research

DOI: http://dx.doi.org/10.20468/ajmdr.2017.12.01

Research Article

Data Security using DNA Cryptography

P. V. Kumaraguru[1], V. J. Chakravarthy[2]

[1,2]PG Department of Computer Applications, Guru Nanak College, Chennai, Tamil Nadu, India

Address for correspondence:
V. J. Chakravarthy, PG
Department of Computer
Applications, Guru Nanak
College, Chennai,
Tamil Nadu, India. E-mail:
chakku_vjc@yahoo.co.in

ABSTRACT

Throughout history, there has ne'er been a modification a lot of dramatic than the one led to by the flexibility of humans to speak. Communication that at first materialized between folks at a selected place step-by-step augmented to communication across continents. With the appearance of email that revolutionized the communication field, communication between folks across massive distances became potential that once appeared a frightening task. This undoubtedly had its share of cons as the email may well be intercepted by anyone, whereas it is being transmitted from one place to different. To avert this interception, cryptography was evolved. Since its beginning, cryptography has been used for varied functions starting from storing knowledge of a selected system in encrypted kind, encrypting knowledge before causing it through a communication medium, etc. As this method evolved the quantity of attacks and alternative ways of attacks additionally evolved. Cryptography has returned a protracted method since it absolutely was initial used and currently an over plus of science techniques area unit out there so as to encode knowledge. During this paper, we have a tendency to gift a fancy cryptography algorithmic program that uses DNA sequence so as to encode the information. This method can modify to encode the information terribly very advanced kind and therefore encourage be a very economical algorithmic program with high accuracy. DNA has been wide used in concert of the advanced forms to represent data. Associate economical algorithmic program that uses DNA sequencing for cryptography has been mentioned during this paper.

Keywords:
Cryptography,
DNA cryptography,
DNA sequencing

Received: 15th September 2017
Accepted: 25th November 2017
Published: 05th December 2017

INTRODUCTION

When statesman sent messages to his generals, he did not trust his messengers. Therefore, he replaced each A in his messages with a D, each B with associate E, and then on through the alphabet.[1] One among the foremost vital wants of individual is to speak and share data by selection. This gave rise to the art and science of secret writing the messages and data specified solely the approved or meant folks may access this information. Notwithstanding unauthorized folks got hold of this data, it ought to be specified they might not extract any data from it.[2] In earlier time techniques to write in code, the data were employed by government, military folks, and through the war in order that their data does not get leaked to their enemy. However, with ever-changing times, we've entered into the age of technology. Therefore, they want and the use of cryptography has additionally become distinguished. Cryptography has diode to the expansion of e-commerce and net.

Cryptography is associate art and science of activity data. It is the flexibility to send data between sender and receiver specified it prevents others from reading it. A message in its original kind is understood as plaintext. The encoded data are understood as cipher text. The method of generating cipher text from plain text is termed cryptography. This includes associate algorithmic program and a secret worth known as key. The range of key depends on the extent of security concerned.[4] Privately key cryptography, one secret is used for each encrypting and decrypting. Whereas publicly key cryptography, there are a unit two keys - a public and a nonpublic key. Every user has each key and whereas the personal key should be unbroken secret the general public secret is in public proverbial.[5] Notable biracial key algorithms area unit AES, DES, FEAL, IDEA, BLOWFISH, and notable uneven key algorithms area unit RSA, Diffie-Hellman.[6]

In this paper, we've used computer memory unit rotation cryptography algorithmic program (BREA). The primary step of this algorithmic program is to interrupt the plaintext into blocks of 16 bytes every. Every block is drawn as the second array. The second step is to use the computer memory unit rotation on rows and columns to encode the text.[7] We have used a biracial key and that we would encode it victimization DNA sequence. During this, every letter of the secret is regenerate into totally different combos of the four bases that conjure the DNA. DNA strands contain long polymers of legion joined nucleotides. These nucleotides carry with it one among four chemical element bases, a 5-carbon sugar and a phosphate cluster. The nucleotides that conjure these polymers area unit named once the chemical element base that it consists of purine (A), pyrimidine (C), purine (G), and pyrimidine (T). One among the foremost rising techniques within the world of cryptography is DNA cryptography that works on the ideas of DNA computing. Victimization, the biological structure of DNA a brand new technique for securing knowledge, was introduced known as DNA computing/biological computing.

Benefits of DNA computing include - speed, tokenism power necessities, and tokenism storage necessities.[8] A gram of DNA contains concerning 1021 DNA bases or concerning 108 terabytes. Hence, some grams of DNA could have the potential of storing all the information keep within the world.[9] Activity knowledge within the type of DNA is termed DNA cryptography. It is a theme of study concerning the way to use DNA as an associate data carrier, and it uses fashionable biotechnology as alive to transfer ciphertext into plaintext.[10] Multiple DNA algorithms that have been studied and researched area unit biradial and uneven key cryptosystem victimization DNA, DNA steganography systems, and triple stage DNA cryptography.[8] Therefore, we've used a mix of BREA and DNA sequence in our algorithmic program to form the cryptography assured.

CONNECTED WORK

After DNA was cited because the most advanced type of data illustration, several new algorithms were developed and projected by the researchers so as to make sure knowledge security. This section highlights a number of the algorithms that used ester sequence so as to encode the digital knowledge.

One of the analysis recommended victimization biserial DNA cryptography algorithmic program within which the text message is regenerate into positional representation system code and computer code. Currently, this message is split into two components out of that one is employed as key and also the different is employed as message and adding to that, XOR operation is additionally performed so as to extend the compression issue. DNA primarily based coded message is received once applying DNA digital secret writing, so the polymerase chain reaction (PCR) amplification is enforced by victimization two prime try as key and compression are performed for the variable length of knowledge.[11] This algorithmic program undoubtedly will increase the safety of the cryptography technique; however, it additionally will increase the process complexness because it uses two prime numbers victimization PCR amplification.

Another algorithmic program that was projected by Shreyas Chavan was the cryptography of plaintext victimization DNA cryptography and binary only once pad (OTP) theme. This algorithmic program primarily uses two keys that area unit used for the cryptography on the sender aspect and cryptography on the receiver aspect. One among the keys could be a random string of nucleotides forming a DNA sequence and also the length of this key depends on the length of the plaintext. The second secret is binary sequence that is used for OTP. The length of the binary secret is doubly the length of the DNA sequence key.[12]

Kang Ning implies the thought of securing knowledge by away known as pseudo-DNA cryptography technique. The safety approach adopted during this technique is of changing the information or text into super molecule consistent with the ordination table and also the key is send to the receiver through a secure channel. Even, if the theoretical analysis of this technique is also powerful, the length of ciphertext is far beyond the plaintext and also the partial data that exists once cryptography will be compromised simply.[13]

Another algorithmic program projected on similar lines was that of Kritika Gupta and Shailendra Singh. The algorithmic program that they projected comprised changing the plaintext into its code that was then regenerate into binary kind. Currently, these binary values are encoded in DNA sequences. Following that, a DNA sequence is chosen as a key and classified in blocks of eight characters every. Supported the character positions within the key, a table is made, and with the assistance of table and key, knowledge gets regenerate within the encrypted kind.[14]

PROJECTED METHODOLOGY

The algorithmic program that is largely used here is biradial Key Block Cipher algorithmic program. The key points of this algorithmic program area unit mentioned below:

1. Every block size during this algorithmic program is taken as 16 bytes.
2. The scale of the key matrix is additionally a similar, that is, 16 bytes.
3. The values of the key matrix area unit at random generated and these values vary from 0 to 127.
4. During this algorithmic program, the construct of polyalphabetic substitution is followed.
5. This algorithmic program additionally makes use of the byte-rotation technique.[7]
6. The cryptography of the secret is finished the assistance of DNA sequencing.

The multidimensional language illustration of the cryptography method is given as Figure 1:

The steps to be followed for implementing this algorithm along with example are as follows:

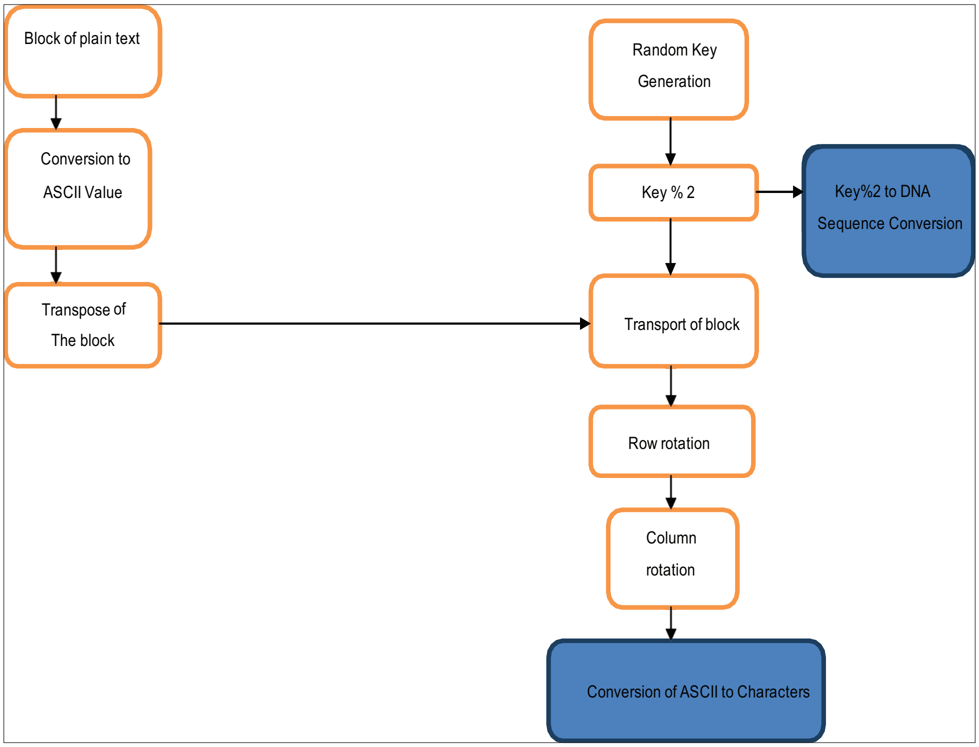
1. Plain text to cipher text conversion

1.1. The plaintext is partitioned into fixed-length blocks of size 16 bytes each. These blocks are represented by a matrix M. For example: Let the plaintext be "VANHUESENSUITING."

M=

V	A	N	H
U	E	S	E
N	S	U	I
T	I	N	G

1.2. The letters of alphabet are assigned numerical values based on their ASCII values. If the plaintext is smaller than the block size, then the empty spaces after the complete plain text are replaced by (ASCII value: 46) till the length of the plaintext becomes equal to block size.



Figurer 1: Encryption Process

M=

86	65	78	72
69	85	83	69
78	83	85	73
84	73	78	74

1.3. Calculate the transpose of the matrix M denoted by Mt.

M=

86	69	78	84
65	85	83	73
78	83	85	78
72	69	73	71

1.4. The random key is generated using some random function; the values in the key matrix are generated in the range similar to that of the range of the characters that can be encrypted by this algorithm. K=[k1,k2,......................, k16] K=

108	101	112	52
124	65	79	45
76	60	31	97
75	39	15	09

1.5. Find the matrix Kt using - Kt = K mod2.

Kt=

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

*(To enhance the security, this key Kt is converted into DNA sequence using DNA sequencing/ciphering technique before it is stored or transmitted. The conversion of key (Kt) to DNA sequence (Ke) is explained in section 2. Key conversion).

1.6. Add the matrices Mt and Kt and the resultant matrix is denoted by Mk.

Mk =

86	70	0	0
0	1	1	1
0	0	1	1
1	1	1	1

1.7. Rotate all the rows horizontally of Mk matrix to the right such that rotate 4 bytes from the first row, rotate 3 bytes from the second row, rotate 2 bytes from the third row, and rotate 1 byte from the fourth row. The resultant matrix is denoted by Mr. After rotating the rows the matrix will be:

Mr =

86	70	78	84
86	84	74	65
86	79	78	83
72	73	70	74

1.8. Rotate all the rows vertically downward of Mr matrix such that rotate 4 bytes from the first column, rotate 3 bytes from the second column, rotate 2 bytes from the third column, and rotate 1 byte from the fourth column. Denote the resultant matrix by Mc. After rotating the columns, the matrix will be-

Mc=

86	84	78	74
86	79	70	84
86	73	78	65
72	70	74	83

1.9. This matrix values containing the ASCII values are then converted back to their respective symbols or characters to be stored or used further. The resultant matrix is denoted by C.

C=

V	T	N	J
V	0	F	T
V	I	N	A
H	F	J	S

C= “VTNJVOFTVINAHFJS”

2. Key conversion

The key to be stored for decryption of the data is converted using 2 bytes at a time. The key matrix is Kt.

Kt=

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

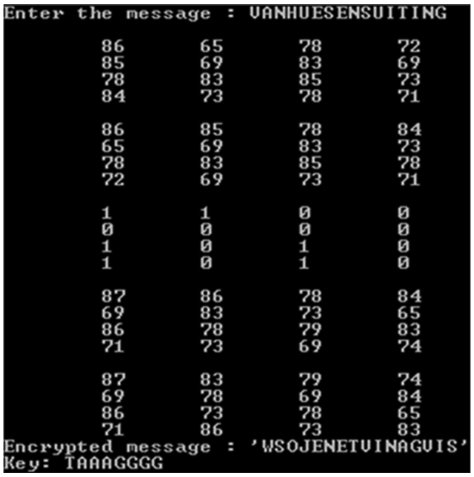
The generated key (Kt) is encrypted using the DNA sequences which will be further stored and used for decryption purpose. The conversion of the key is done by considering 2 bytes at a time and uses some basic permutation logic. As there are four DNA strands, namely, A (Adenine), C (Cytosine), G (Guanine), and T (Thymine), there are 256 combinations possible. Since the key is a sequence of pairs of either **00** or **01** or **10** or **11**, these pairs can be substituted by A or C or G or T as per the following table**.** The sequence of strands used is taken to be alphabetic and not as per their literal combination.

Sequence	Nucleotide used
00	A
01	C
10	G
11	T

Hence, our encrypted key would be Ke-CACTATTT.

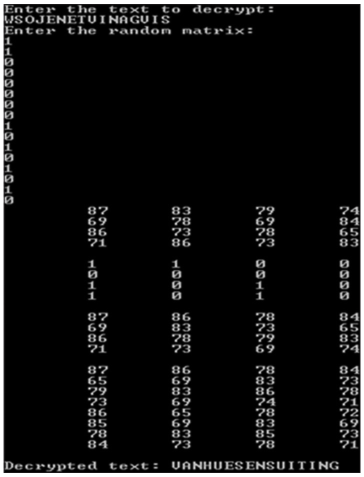
Decryption process

The cryptography is simply the reverse of the cryptography method. During this method, first of all, the DNA encrypted key should be obtained to convert it to the conventional key matrix that is to be used for cryptography purpose by subtracting it from the code values of the ciphertext. After that, the reverse column and row rotation are to be performed severally, and at last, the transpose of that matrix is to be taken. This matrix, therefore, offers America the first plain text Figures 2 and 3.



```
Enter the message : VANHUESENSUITING

        86        65        78        72
        85        69        83        69
        78        83        85        73
        84        73        78        71

        86        85        78        84
        65        69        83        73
        78        83        85        78
        72        69        73        71

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

        87        86        78        84
        69        83        73        65
        86        78        79        83
        71        73        69        74

        87        83        79        74
        69        78        69        84
        86        73        78        65
        71        86        73        83
Encrypted message : 'WSOJENETVINAGVIS'
Key: TAAAGGGG
```

Figure 2: Encryption



```
Enter the text to decrypt:
WSOJENETVINAGVIS
Enter the random matrix:
1
1
0
0
0
0
0
0
1
0
1
0
1
0
1
0
           87        83        79        74
           69        78        69        84
           86        73        78        65
           71        86        73        83

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

           87        86        78        84
           69        83        73        65
           86        78        79        83
           71        73        69        74

           87        86        78        84
           65        69        83        73
           79        83        86        78
           73        69        74        71
           86        65        78        72
           85        69        83        69
           78        83        85        73
           84        73        78        71

Decrypted text: VANHUESENSUITING
```

Figure 3: Decryption

RESULTS

In this section, the results that area unit obtained once applying the algorithmic program on a plaintext area unit displayed. The subsequent results area unit obtained:

In figure, a pair of, the plaintext that has been used is “VANHUESENSUITING.” Once applying the algorithmic program, we have a tendency to get the encrypted message as “WSOJENETVINAGVIS” and also the key generated is TTGGAGAC.

In figure three, the ciphertext used is the same one that is received once cryptography in figure one that is “WSOJENETVINAGVIS.” After applying the decryption algorithm, we get the encrypted messages “VANHUESENSUITING” and also the key generated is TAAAGGGG, therefore, proving the correctness of the projected algorithmic program.

FUTURE WORK

This algorithmic program is often additional increased by implementing the DNA sequencing to the ciphertext additionally. This may modify the ciphertext to urge double encrypted. The area units within which this algorithmic program are often used are military functions as any country’s knowledge is incredibly vital and confidential. It is often additionally used for banking applications, wherever it is often accustomed encode the very important knowledge of the client like the account range or pin or countersign.

CONCLUSIONS

In this paper, we've devised a brand new science technique. This new science technique uses DNA sequencing to encode the information. The algorithmic program conferred during this paper has been tried and also the results are conferred. The results depict the effectiveness of the algorithmic program conferred and are a sign that it is often utilized in varied applications.

ACKNOWLEDGMENTS

The authors would like to acknowledge for the useful discussions on this topic and their inputs and feedback while writing this paper.

REFERENCES

1. Rodriguez-Henriguez F, Saqib NA, Diaz-Perez A, Koc CK. Cryptographic Algorithms on Reconfigurable Hardware; 2007.
2. Akdeniz Y. Cryptography and Encryption, Cyber-Rights and Cyber-Liberties (UK); 1996. Paid Access Only.
3. Noubir G. Fundamentals of Cryptography: Algorithms and Security Services. Boston: Northeastern University; 2012.
4. Bhati S, Bhati A, Sharma SK. A new approach towards encryption schemes: Byte-rotation encryption algorithm. In: Proceedings of the World Congress on Engineering and Computer Science. Vol. 2. San Francisco, USA: WCECS; 2012.
5. Marquis J. The Future of Data Security: DNA Cryptography and Cryptosystems. Boston: Northestern University; 2011.
6. Gehani A, La Bean T, Reif J. Department of Computer Science, DNA-Based Cryptography. Durham, NC: Duke University; 1999.
7. Zhang Y, Fu LH. Research on DNA Cryptography. Xi "an, China: College of Software and Microelectronics. North Western Polytechnical University; 2015.
8. Prabhu D, Adimoolam M. Bi-Serial DNA Encryption Algorithm; 2011. Available from: https://www.pdfs.semanticscholar.org/1754/f0eb5852500598a70af4002e186cd2f3c6ce.pdf.
9. Chavan S. DNA cryptography based on DNA hybridization and one time pad scheme. Int J Eng Res Technol 2013;2:194.
10. Ning K. A Pseudo DNA Cryptography Method, Cornell University Library, March; 2009. Paid Access Only.
11. Gupta K, Singh S. DNA based cryptographic techniques: A review. Int J Adv Res Comput Sci Softw Eng 2013;3:607-10.
12. Watson JD, Hopkins NH, Roberts JW, Steitz JA, Weiner AM. Molecular Biology of the Gene. 4th ed. Menlo Park, CA: The Benjamin/Cummings Publishing Co., Inc.; 1987.
13. Seeman N C. Nanotechnology and the double helix. Sci Am 2004;290:34-43.
14. Debao L, Ping X. Theory and Methods of Recombinant DNA. Hangzhou: Zhejiang Science and Technology Publishing Co.; 1994.

Cite this article: Kumaraguru PV, Chakravarthy VJ. Data security using DNA cryptography. Asian J Mult-Disciplinary Res. 2017;3(9):1-5.

Source of Support: Nil, **Conflict of Interest:** None declared.