

## Modifying Playfair Cipher by Using DNA and Amino Acids

AYUSHI KANSAL, SHRUTI SNEHA, MANISH KUMAR PATEL

Department of Information Technology

SRM University, NCR Campus

### ABSTRACT:

Cryptography in the field of technology is the result of promising technique involving the encoding of information into highly unreadable form, thus making it highly secured. After the invention of highly secured cryptographic techniques like, AES, DES, RSA, SHA etc, there is a requirement of quite modification in these techniques to bring out more superior cryptographic technique ever. In this paper, we have not only proposed, but also implemented a modified version of “symmetric key cryptographic, Playfair ” and making it superior and more compatible as the security provided by DES(Data Encryption Standard). With the making use of the properties of “DNA strands”, we have made our technique more secured and hence giving it the new name as “DNA modified playfair cipher using java technology”.

Keywords: Playfair cipher, DNA, Amino Acids, DES, Key.

### INTRODUCTION:

With the risk of broking of many cryptographic techniques, like DES and MD5, a new way of protecting the data is thought to implement. The new way behind this is the concept of DNA computing that could proved to be successful in bringing a new hope of making the cryptographic technique more powerful or really unbreakable. DNA can be defined as a nucleic acid that contains genetic instructions that are used in the development and functioning of all known living organisms and some viruses. The character form of a message or any form of an image can be easily transformed to the form of bits. The four bases found in DNA are Adenine(A), Cytosine(C), Guanine(G) and thymine(T). these four bases are attached to the phosphate/sugar to form the complete nucleotide. The genetic code can be expressed as either DNA codons or RNA codons. Both the codons are same except the nucleotide

Thymine(T) is found in the place of Uridine(U). So, in DNA, we have TCAG and UCGT in RNA. In our work, we applied the conversion of character form or binary form of data to the DNA form and then to amino acid form. Then the resulting form goes through the encryption algorithm which we choose for example; the classical Playfair cipher. Playfair Cipher is a popular poly-alphabetic cipher, in which the alphabets are arranged in a  $5 \times 5$  key matrix based on secret key allowing the use of 25-alphabets out of the total 26-alphabets of English language.

Thus, our proposed work focuses on the efficient modification of the playfair algorithm to eliminate its limitations and enrich its security parameters. The proposed method is implemented and also compared with other popular ciphers on the basis of certain parameters, like Avalanche Effect, Time Complexity, and Space Requirement.

### METHODOLOGY:

Playfair cipher is one of the old and simple symmetric key cryptographic technique, hence it is believed to be easily breakable. For making it more powerful, some modifications were required. This has been done by introducing the concepts of confusion and diffusion to the core of the encryption process. Additionally for the security feature, the plaintext message is restricted to be in all Upper Case without the letter ‘J’ and other numerical or symbolic values. These were the older problems in the playfair cipher that has been resolved by our new algorithm. All the alphabetical values of the entered plaintext is first converted in the form of

bits(according to their respective ASCII value). And this binary form is transformed to DNA form by the existing standardized table of DNA nucleotides. The further implementation is done by the JAVA technology and the existing problem is resolved giving it the enhanced look and providing high security to it. Playfair is based only on the English alphabetical letters, so for preserving this concept, we have used the English letters indirectly. DNA contains four bases that can be abbreviated as Adenine(A), Cytosine(C), Guanine(G), Thymine(T). And we have 20 amino acids abbreviated by single English alphabet. So, we can stretch these 20 characters to 26 characters. Then the DNA form is converted into amino acids which is passed through playfair cipher. The only problem that proved to be the nightmaring was the “ambiguity problem” that is now resolved and removed through our algorithm.

## PROPOSED MODIFIED PLAYFAIR ALGORITHM:

### ENCRYPTION ALGORITHM OF DNA BASED PLAYFAIR CIPHER:

Earlier, the cryptographic algorithm used four-square to the English alphabet characters of plaintext entered. It was unable to encrypt any special characters that caused many problems for the sender, as every letter was mandatory to be in the form of English alphabet. With this contrast, our proposed algorithm allowed any kind of character ( i.e. special character , a space or numerical value) to encrypt. The algorithm starts by converting the entered “plaintext” into the “binary format (in the form of ‘0’ and ‘1’)” according to the given ASCII value. Further, these values are represented into the DNA form according to the standardized table of DNA(fig. table1) which is considered to be the universal table of amino acids and their codons representation in the form of DNA. This formatted symbol is further used in the algorithm.

| BIT 1 | BIT 2 | DNA |
|-------|-------|-----|
| 0     | 0     | A   |
| 0     | 1     | C   |
| 1     | 0     | G   |
| 1     | 1     | U   |

**Table 1. Standardized table for DNA representation of bits**

### FORMATION OF ALPHABETICAL FORMATTED TEXT:

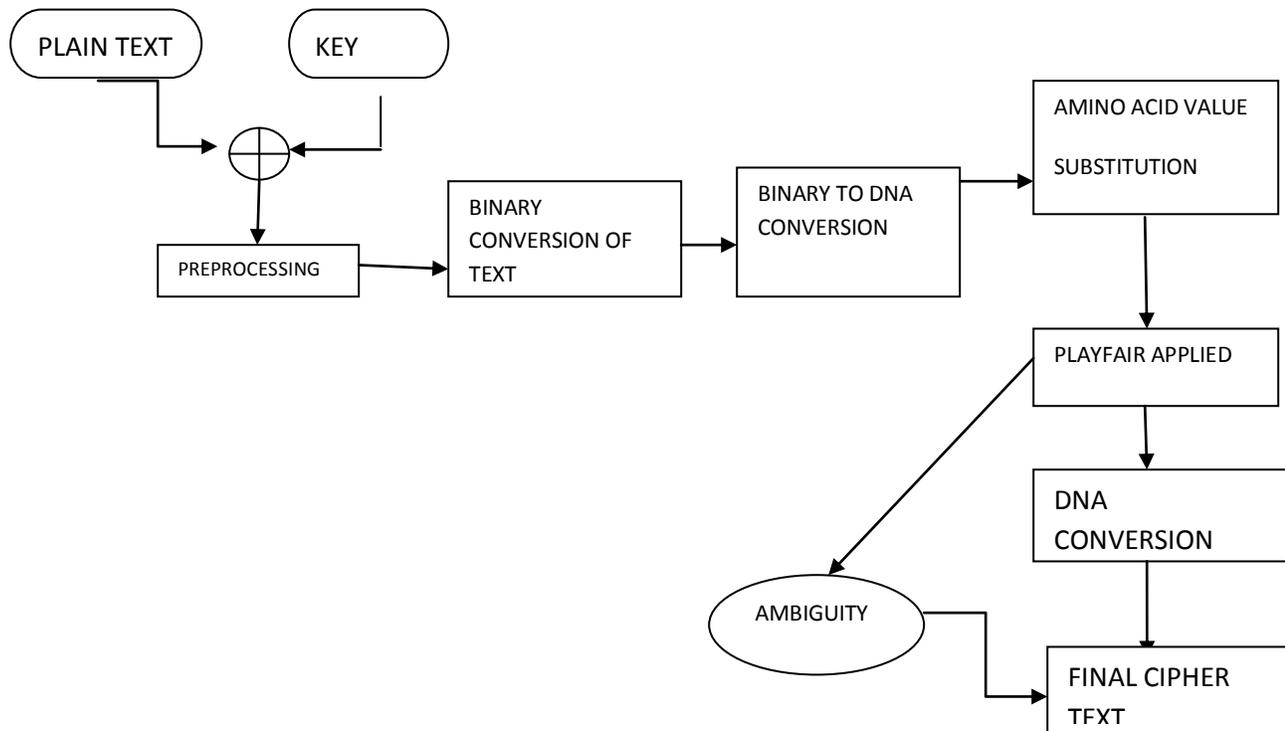
As far as the Amino acid codon is considered about, there are 20 amino acids in addition to 1 start and 1 stop. But, we need 25 letters for 5\*5 playfair matrix formation (ignoring ‘J’ due to its lowest frequency). We already have the standard alphabetical table for the amino acid codons(fig. table2). When the coupling of “three” of the generated codon is formed, after the formation of DNA representation of bits. The corresponding alphabet of its respective codon is noted down. The resultant alphabets are further taken for performing the playfair cipher action. The one thing to be noticed in this table is the number of ways of formation of codons in several alphabets. This number of terms is noted in the correspondence column of that particular alphabet. This number is termed as “ambiguity”. The previous existing algorithm could not be able to resolve this problem of ambiguity. But in our algorithm, it has been resolved by taking this “ambiguity number” and implementing it by “JAVA technology” prior use of performing playfair cipher.

| D   | E   | F   | G   | H   | I   | J   | K   | L   | M   | N   | O   | P   | Q   | R   | S   | T   | U   | V | W | X | Y   |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|---|---|---|-----|
| 2   | 2   | 2   | 4   | 2   | 3   | 2   | 6   | 1   | 2   |     | 4   | 2   | 6   | 6   | 4   |     | 4   |   |   |   |     |
| GAU | GAA | UUU | GGU | CAU | AUU | AAA | UUA | AUG | AUU | UUA | CCU | CAA | CGU | UCU | ACU | AGA | GUU |   |   |   |     |
| GAC | GAG | UUC | GGC | CAC | AUC | AAG | UUG |     | AAC | UUG | CCC | CAG | CGC | UCC | ACC | ADG | GUA |   |   |   |     |
|     |     |     | GGA |     | ACA |     | CUU |     |     |     | CCA |     | CGA | UCA | ACA |     |     |   |   |   | GUU |

**Fig. table 2. Standard alphabetical table in correspondence with its codons**

**WORKFLOW:**

The flowchart shown is the step-by-step procedure that is followed in our algorithm. The plaintext along with a sharing key is taken as input and is processed. The plaintext is further converted into its binary formatted text. The pairing of two binary text are taken together and according to the universal table of DNA representation of bits, as shown previously is converted into the DNA. The key matrix for further playfair cipher is constructed. The generated DNA is converted into its respective English alphabet from the universal standardized table of the amino acids. Now the ambiguity number is taken into account and then the playfair algorithm is further implemented. The resultant encoded letters are not the final cipher text. The ambiguity length of the text is appended at the end of the resultant cipher text to give its final cipher text.



**Fig 3.Flowchart of the proposed algorithm**

## EXPERIMENTS, RESULTS AND INTERPRETATION:

The proposed DNA modified Playfair cipher is implemented on java platform and number of tests are considered to observe the encryption efficiency in terms of certain parameters like, Avalanche effect, Key randomness and time. DES is considered for comparative encryption analysis with the proposed algorithm in respect of the above mentioned parameters. Some experimental results are given below:-

### AVALANCHE EFFECT:

It refers to a desirable property of cryptographic algorithm. The avalanche effect is idle, when an input is changed slightly (just by flipping a single bit), the output changes significantly. In our algorithm, the result of Avalanche Effect lies between 30% - 40%. The ideal case or strict avalanche criterion (SAC) is the probability of 50% bits change. So the obtained result proves the efficiency of the algorithm strongly. Though in this algorithm the type or variety of rotation is fully dependent on “key” and on the “content of the plaintext file”, the result may vary for different files. A details comparison of Avalanche effect between different algorithms is given in the chart below.

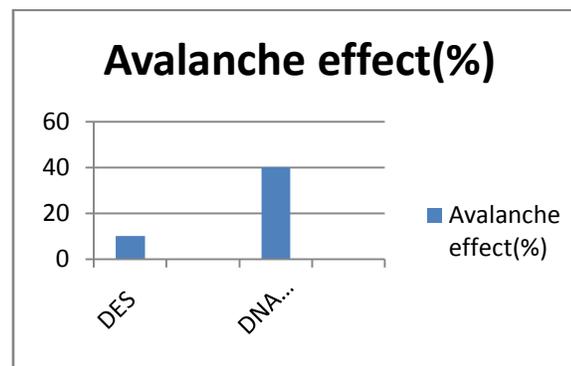


Fig 4 Avalanche effect

It can be observed that the DNA playfair gives the avalanche effect of almost 40% change, which satisfies nearly the strict avalanche criteria of 50%. Thus it can be inferred that DNA playfair gives comparatively better than the DES.

### LOW SPACE REQUIREMENT:

The space requirement of DNA modified Playfair is very low in comparison with the DES. Since, it requires several rounds of both encryption and decryption. While, DNA modified Playfair requires only 5\*5 matrix and two more bytes to read and write.

### BRUTE FORCE ATTACK:

The proposed algorithm uses key matrix of 5\*5 and in addition, it uses DNA codons for further encapsulation. So, the attacker won't be able to crack the key. So, the proposed algorithm is safe from the Brute Force Attack.

### MAINTENANCE COST:

Since, the proposed algorithm, DNA modified Playfair algorithm is light- weighted. So, its complexity level is lower in comparison to DES. Hence, less maintenance cost will be required in case of DNA modified playfair algorithm.

## PLAYCOLOR CIPHER:

The playcolor cipher indicates the problem in completely different way by generating Block Cipher using color substitution (ARGB color substitution is used).

## TIME COMPLEXITY:

The proposed algorithm has been applied randomly on a set of 100 files in order to calculate the average time requirement for encryption / decryption. Average Encryption / Decryption Time In Figure 5 the time taken to encrypt 24 files of size ranging from 100 KB to 150 KB has been depicted. It can be observed that on an average 2 sec 832 millisecond time is required to encrypt a file of size 122 KB. As a matter of fact it can be well inferred that average time required to encrypt more number of large size files will be more.

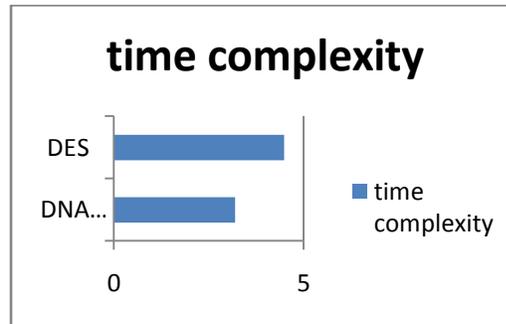


Fig 5. Time complexity

It can be observed clearly that the proposed DNA modified Playfair algorithm is taking the least time of 3.2 seconds to encrypt the file. Whereas, other algorithm, DES is taking more time for encrypting the same file, which is around 4.5 seconds.

## 1. APPENDIX

### CODE SNIPPET:

Some of the snap shots of the main encryption / decryption interface and also some testing interfaces, which are used to test efficiency of the modified algorithm, are given below.



Fig.6. Main interface for encryption/decryption



**Fig 7. Modified input text interface**



**Fig 8. Key matrix formed according to the entered “key”**



**Fig 9. Input Form for Cipher Demo**



Fig 10. Binary to DNA conversion



fig 11. DNA to Amino Acid codon encryption



Fig 12. Amino to Playfair for encryption interface



Fig 13. Cipher text conversion with ambiguity



Fig 14. Ambiguity to DNA encryption



Fig 15. Final cipher text formation



**Fig 16. Interface for Decryption**

## CONCLUSION:

The algorithm initially succeeded in overcoming some main problems in "Playfair cipher" like restriction of plaintext to "English Alphabet". As in our algorithm the plaintext is to be converted to its binary value before encryption, it is now clear that the plaintext message can be written in upper or lower case, with any punctuation, and numerical values. Other papers conducted the idea of amino acids way of representation from the point of view of the central dogma design. But they were unable to clearly handle the problem of ambiguity as performed by our algorithm. Our algorithm made few preprocessing steps to handle this problem and the result was quite accurate (same input message obtained after decryption). This feature is very important when regarding an encryption algorithm in order to verify the concept of data integrity or in other words, to assure that data after decryption to be the same input data before encryption. The project is very versatile as many amendments will be possible at any time of computing because of the support for a number of intermediary processes during encryption. The project has good current market value as it is important from the view of research and has an extremely bright future as regards the importance of DNA Cryptography in an era where the DES and MD5 have been broken.

## REFERENCES:

2. Sherif T. Amin, Magdy Saeb, Salah El-Gindi, "A DNA-based Implementation of YAEA Encryption Algorithm," IASTED International Conference Computational and Intelligence <http://www.actapress.com/PaperInfo.aspx?PaperID=29058>
3. Ayan Lahiri, "DESIGN AND IMPLEMENTATION OF A ENHANCED BINARY PLAYFAIR ALGORITHM", Independent Research Study Project, October 2011.
4. TAYLOR Clelland Catherine, Viviana Risca, Carter Bancroft, 1999, "Hiding Messages in DNA Microdots". Nature Magazine Vol.. 399, June 10, 1999
5. Mona Sabri, "A DNA AND AMINO ACID-BASED IMPLEMENTATION OF PLAYFAIR CIPHER", Independent Research Study Project for CS5231, October 2004.
6. Leonard Adleman. "Molecular Computation of Solutions to Combinatorial Problems". Science, 266:1021-1024, November 1994.
7. D. Kumar, Anand Pandey, and S. K. Singh, "'Reliability' An Emerging Problem in Network Security," Int. J. Appl. or Innov. Eng. Manag., vol. 2, no. 12, pp. 199-202, 2013.
8. Anand Pandey, D. Kumar, and S. K. Singh, "Performance Evaluation of TORA Protocol with Reference to Varying Number of Mobile Nodes," Int. J. Appl. or Innov. Eng. Manag., vol. 2, no. 12, pp. 203-209, 2013.
9. Varying Number of Mobile Nodes," Int. J. Appl. or Innov. Eng. Manag., vol. 2, no. 12, pp. 203-209, 2013.