DNA Cryptography using randomly generated DNA sequence table

[1]Nagaraj S M, [2]Mr. S Lokesh

[1]Student of Information Technology, [2]Associate Professor
Computer Science and Engineering,
The National Institute of Engineering, Mysuru, India

***Abstract*— DNA cryptography is another branch of data security. It encodes the data as DNA nucleotides (A, T, G and C). It makes utilization of the endless stockpiling limit of DNA and its natural properties like exceptionally stable atom, strength, practical and effectively accessible. DNA cryptography is a blend of organic and software engineering area. The specialist of this field must have the learning of DNA, software engineering and Information security techniques. This paper gives the system to DNA cryptography and it likewise gives the abridged data of calculations utilized for giving security to the information put away in DNA. In this paper, creators propose another technique for giving security to the information as DNA succession. The proposed technique gives security at the three level utilizing binary conversion, spiral transposition and DNA arrangement word reference table.**

***Index Terms*—DNA, Cryptography, nucleotides, sequence table, Binary, Spiral transposition.**

I. INTRODUCTION

Deoxyribo nucleic corrosive develops as another capacity gadget for putting away information. A solitary gram of DNA can store around 106TB of information. Storing data in DNA originates from the way that DNA is a characteristic transporter of data as nucleotides. DNA can store any sort of information whether it is a picture or a sound or a video or any content document. The information put away in the DNA might bank points of interest, passwords, messages or some other mystery data. As each capacity gadget requires security, DNA likewise requires security. Security of information put away in DNA is given in an indistinguishable route from security is given to other stockpiling gadget. DNA cryptography develops as a promising recorded in the region of security. It gives security to the information put away in the DNA as nucleotides. It encodes the data in such a way, to the point that it can be gone through the open system. DNA cryptography has an endless and element scope in howdy tech world. The measure of information and need of security develops at an extremely quick rate. Security necessities are not satisfied by the customary cryptography and steganography techniques. Information stockpiling in DNA and cryptography satisfies the security and capacity prerequisites.

In this paper, the creators proposed a structure for DNA cryptography and give a concise writing study about the security calculations in the zone of information stockpiling in DNA. In the wake of concentrate the past work, the creators display another strategy for encoding data in DNA. The proposed strategy will gives the security at three levels.

This paper is isolated into six areas. To begin with area gives the presentation, second segment talks about the examination system utilized by the specialists for this paper, third segment examines the DNA cryptography, fourth segment abridge the work in the field of security in information stockpiling in DNA, in the fifth segment creators clarifies the strategy proposed by them took after by outline and conclusion.

DNA cryptography is another recorded in the territory of cryptography. It is a mix of two space i.e. software engineering and organic science. For understanding DNA cryptography, the specialist must have the learning of DNA.

DNA is a shortened form of De-oxyribo nucleic acid. It is the fundamental natural unit of each living life form. DNA was found by Swiss doctor Friedrich Miescher in 1869 [1]. DNA comprise of four nucleotides to be specific Adenine (A), Thymine (T), Cytosine (C) and Guanine (G). It is a twofold stranded reciprocal structure. Watson Crick proposed the corresponding principle that adenine will make a couple with thymine (A-T) and cytosine dependably make a couple with guanine (G-C) . DNA cryptography is the investigation of giving security to the information put away as DNA succession. It scrambles the information as DNA nucleotides. The information might be picture, sound, video or content document.

The second level of DNA cryptography is paired change. In this progression the information is changed over into paired information. The analysts may pick whatever other number framework moreover. It differs from creator to creator. The third period of DNA cryptography is paired encryption. Paired encryption scrambles the parallel information. The scientists may utilize the conventional technique to encode the twofold information. In the event that the information is not all that significant, a few scientists skirt the twofold encryption. The fourth period of DNA cryptography is the DNA change. In this stage the encoded paired information or the straightforward twofold information is changed over into DNA nucleotides. The tenets for DNA change are distinctive for various scientists.

The last phase of DNA cryptography is the DNA information encryption. In this stage the information as DNA nucleotides is scrambled. The encryption is finished by utilizing organic operations like grafting, interpretation, polymerase chain response and so on. The DNA encryption should likewise be possible by numerical or legitimate operations like expansion, OR, XOR, supplement, substitution and so on.

II. LITRATURE REVIEW

DNA encryption technique is based on mathematical matrix manipulation where they used a secure generation algorithm to generate new key for encryption process. The benefit of this key generation scheme is that they always get a new cipher data for same plaintext and same key. It provides a good security layer which does not give any hint about plaintext. DNA cryptography can be combined with traditional cryptography to provide hybrid security. It is still taking its initial steps, so there is a lot of scope to work in this area of cryptography and need more works and researches to reach the realization and to enhance the technical issues.[1]

A hybrid approach of cryptography by using traditional symmetric cipher to encrypt part of message while hiding other part inside a DNA microdot to ensure the attacker can't obtain the message without compromising both parts. This will in itself be made untenable for the attacker by hiding data-holding DNA behind a myriad of similar DNA strands which would be acting like camouflage agents. So, provided the secret keys are transferred securely we can expect that the algorithm will retard any type of attacker to compromise the communication by raising his level of efforts and cost of successful attack much beyond the traditional computer based computations. [2]

Several attempts have been made to remove the deficiencies in the scheme of DNA steganography and cryptography. A data hiding algorithm has been designed by using DNA sequences concept and traditional steganography technique. Using steganography they hide the data into the DNA sequences and send encrypted DNA sequences along with a key to the receiver side. Using this key value and encrypted text the receiver easily recovers the plain text. Using this technique they send and receive the data without any deficiency. If any attempt is made to make a fake data then the receiver is able to know after applying the algorithm into fake data because it cannot gives the same results when we apply key on it. This algorithm is very efficient and easy to use. DNA computing has brighter development possibilities in field of Steganography and authentication, which have a more layer protection than a single encryption. [3]

Some methods stress on utilizing DNA cryptography for providing security in communication, especially in data transmission in wireless sensor networks. SSL (Secure Socket Layer) is used to resolve the problem of sharing keys in a WSN. Asymmetric key encryption is followed, for which the keys are generated using RSA algorithm. When the sensor nodes are deployed, each node is assigned a key pair and digital certificate, owing to its tiny storage and low power. Public keys are exchanged through SSL. Security here is achieved in 3 stages - information, computation and DNA representation. The advantage of this method lies in the achievement of access control, availability and signature. The only drawback noticed, is the assignment of key pairs and thesignature to each and every node in the WSN, prior to deployment. [4]

The current paper give the security at two levels for any sort of record. To begin with, at the twofold level encryption and second, at the DNA change. The parallel encryption is finished by doing the transposition of 8x8 grids spirally by gathering the 4 sections at an opportunity to stir up the double values. The following stride is transformation of the main segment into ASCII design and scrambling the information as indicated by the static DNA word reference table. This outcomes in the scrambled record.[5].

III. EXISTING SYSTEM

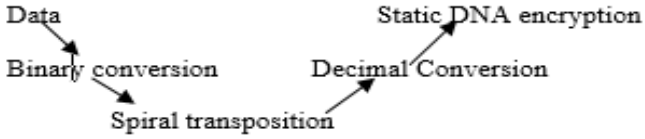

Figure1. Existing method

IV. PROPOSED APPROACH

In this paper we have defined a new method through DNA cryptography. The input is given as the file which consists the data as text, image, binary file etc. The data in the file is converted into ASCII data which is then converted into binary data.

We chose first 64 bits of binary data and arrange them into 8x8 matrix. We perform the same steps for the next 64 bit blocks of data and if the bits are less than 64 in a matrix, the rest of the place in the matrix is padded as zero. For the first 64 binary bits we will perform the spiral transformation which may be row wise or column wise. After getting the resultant spiral matrix we retrieve the decimal value for the first 8 bits of the first row and similarly for the next 8 rows. The decimal values obtained are substituted with the corresponding DNA sequence using the generated table. The same will be performed for the rest of the blocks. This results in the encrypted file for the given input file.

In the next phase, we concentrate on randomizing the table. For this purpose we take 32 character primary key i.e. 256bit key. Then we divide the key into four parts.

First 96 bits as first part, next 32 bit as second part, next 96 bits as third part, last 32 bits as fourth part. We perform operations like reversing first 96 bits of first part and swapping with 96 bits of third part. Similarly reversing last 32 bits of fourth part and swapping with 32 bits of second part. We combine all the four parts which results into a 256 bit key. The resultant key is XOR with original 256 bit key given by the user. Using the similar procedure we obtain eight more such keys. Next we take eight digit secondary key from the user and minimize that key into one to four digits. On getting the digits we use the first digit and the corresponding keys to alter the DNA sequence table. Using the randomly generated DNA sequence table we will encrypt the data and store into a encrypted file and sent to the receiver using a secure algorithm. This acts as one time pad for providing more security.

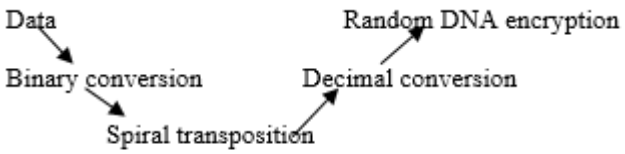

Figure2. Proposed method

V. DNA SEQUENCE TABLE

Decimal Number	DNA Sequence	Decimal Number	DNA Sequence	Decimal Number	DNA Sequence	Decimal Number	DNA Sequence
1	AAAA	65	TAAA	129	GAAA	193	CAAA
2	AAAT	66	TAAT	130	GAAT	194	CAAT
3	AAAG	67	TAAG	131	GAAG	195	CAAG
4	AAAC	68	TAAC	132	GAAC	196	CAAC
5	AATA	69	TATA	133	GATA	197	CATA
6	AATT	70	TATT	134	GATT	198	CATT
7	AATG	71	TATG	135	GATG	199	CATG
8	AATC	72	TATC	136	GATC	200	CATC
9	AAGA	73	TAGA	137	GAGA	201	CAGA
10	AAGT	74	TAGT	138	GAGT	202	CAGT
11	AAGG	75	TAGG	139	GAGG	203	CAGG
12	AAGC	76	TAGC	140	GAGC	204	CAGC
13	AACA	77	TACA	141	GACA	205	CACA
14	AACT	78	TACT	142	GACT	206	CACT
15	AACG	79	TACG	143	GACG	207	CACG
16	AACC	80	TACC	144	GACC	208	CACC
17	ATAA	81	TTAA	145	GTAA	209	CTAA
18	ATAT	82	TTAT	146	GTAT	210	CTAT
19	ATAG	83	TTAG	147	GTAG	211	CTAG
20	ATAC	84	TTAC	148	GTAC	212	CTAC
21	ATTA	85	TTTA	149	GTTA	213	CTTA
22	ATTT	86	TTTT	150	GTTT	214	CTTT
23	ATTG	87	TTTG	151	GTTG	215	CTTG
24	ATTC	88	TTTC	152	GTTC	216	CTTC
25	ATGA	89	TTGA	153	GTGA	217	CTGA
26	ATGT	90	TTGT	154	GTGT	218	CTGT
27	ATGG	91	TTGG	155	GTGG	219	CTGG
28	ATGC	92	TTGC	156	GTGC	220	CTGC
29	ATCA	93	TTCA	157	GTCA	221	CTCA
30	ATCT	94	TTCT	158	GTCT	222	CTCT
31	ATCG	95	TTCG	159	GTCG	223	CTCG
32	ATCC	96	TTCC	160	GTCC	224	CTCC
33	AGAA	97	TGAA	161	GGAA	225	CGAA
34	AGAT	98	TGAT	162	GGAT	226	CGAT
35	AGAG	99	TGAG	163	GGAG	227	CGAG
36	AGAC	100	TGAC	164	GGAC	228	CGAC
37	AGTA	101	TGTA	165	GGTA	229	CGTA
38	AGTT	102	TGTT	166	GGTT	230	CGTT
39	AGTG	103	TGTG	167	GGTG	231	CGTG
40	AGTC	104	TGTC	168	GGTC	232	CGTC
41	AGGA	105	TGGA	169	GGGA	233	CGGA
42	AGGT	106	TGGT	170	GGGT	234	CGGT
43	AGGG	107	TGGG	171	GGGG	235	CGGG
44	AGGC	108	TGGC	172	GGGC	236	CGGC
45	AGCA	109	TGCA	173	GGCA	237	CGCA
46	AGCT	110	TGCT	174	GGCT	238	CGCT
47	AGCG	111	TGCG	175	GGCG	239	CGCG
48	AGCC	112	TGCC	176	GGCC	240	CGCC
49	ACAA	113	TCAA	177	GCAA	241	CCAA
50	ACAT	114	TCAT	178	GCAT	242	CCAT
51	ACAG	115	TCAG	179	GCAG	243	CCAG
52	ACAC	116	TCAC	180	GCAC	244	CCAC
53	ACTA	117	TCTA	181	GCTA	245	CCTA

Table1. Randomly generated DNA table

VI. CONCLUSION

In this paper, the authors have proposed a new method for encrypting type of data. The proposed method provides three level security which provides more security than the previous existed approach first at binary level and second at DNA encryption and third using DNA random sequence table which is similar to one time pad.

REFERENCES

[1] Mandge, Tushar, and Vikas Choudhary. "A DNA encryption technique based on matrix manipulation and secure key generation scheme". Information Communication and Embedded systems (ICICES), 2013. International Conference on. IEEE,2013.

[2] Chaudhary, Himanshu, and Vishal Bhatnagar. " Hybrid approach for secure communication of data using chemical DNA." Confluence The Next Generation Information Technology Summit (Confluence),2014 5th International Conference. IEEE, 2014.

[3] Kumar, dinesh, and Sushil Sing. "Secret data writing using DNA Sequence ." Emerging Trends in Network and computer Communications (ETNCC), 2011 International conference on. IEEE, 2011.

[4] Sundaram, G. Shanmuga, et al. " Cellular automata based DNA cryptography algorithm." Intelligent System and Control (ISCO), 2015 IEEE 9th International Conference on. IEEE, 2015.

[5]Shipra Jain, Dr.Vishal Bhatnagar. "A novel DNA sequence Dictionary method for securing data in DNA using Spiral Approach and Framework for DNA Cryptography". IEEE International conference on advances in Engineering & Technology Research.,2014.