DOI: https://doi.org/10.24113/ijoscience.v3i10.14

Hybrid DNA Cryptography for Secure Cloud Storage

Ms. Mona Vishwakarma
M.Tech Scholar
Department of CSE
NRI
Bhopal, M.P., India
shrma.meenu65@gmail.com

Mr. Umesh Lilhore
Professor
Department of CSE
NRI
Bhopal, M.P., India
umeshlilhore@gmail.com

***Abstract* - Now a day's cloud computing is being used in several areas like industry, medical, science and research, colleges etc for storage of huge amount of data. User can retrieve data files from cloud data center on request. While storing data files on cloud server several security issues may arise. To overcome from these security issues there are a number of techniques. Out of several security techniques Cryptography is more popular now a day's for data security. Use of a traditional cryptography algorithm is not effective or sufficient for high level security to data in cloud computing. In order to make sure security of the information at cloud data storage end, a design and implementation of an algorithm to enhance cloud security is proposed. With a concept, where the proposed algorithm combines features of DNA sequencing. Proposed methodology or system is provides security to data files by using hybrid encryption/decryption technique. So, the enhanced approach maintains confidentiality and authenticity of cloud data. As a result the proposed algorithm provides enhanced security as well as reduces time complexity during encryption and decryption process of data file.**

Keywords - Cloud Computing, Security, Confidentiality, Authentication, DNA Cryptography.

I. INTRODUCTION

Cloud computing depends on the idea of abstraction and virtualization and convey administrations to the customer. Essentially, the Cloud computing [1] has acquired these idea from Grid and Cluster Computing. All computing assets are amassed into a bundle and conveyed as a service over the internet to all cloud users. Cloud computing with a target conveys computing assets and administrations as a utility (like power, water supply, gas, phone, and so on) over the system to the client and thus acquires such pay-as-you-go property from utility.

Cloud computing is a subscription based service where we can obtain networked storage space and computing resources. It makes workable for us to get to our information from anyplace and at whatever time. It renders the expense of buying computational assets and capacity as everything is accessible on cloud and conveyed according to request of client and we need to pay for that only for which we need to utilize. Just necessity of Cloud computing is that we need internet accessible to make the most of their computational resources. Various organizations and associations are applying Cloud computing ideas to their business including GOOGLE, AMAZON and AZURE[2].

Cloud can be broadly classified into two parts: front end and the back end. Client part of the cloud computing system is referred as front end which consists of the applications and interfaces that are needed for accessing cloud computing services, e.g., Web Browser, mobile apps. Back end part provides different types of services to the user, e.g., Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a service (SaaS) [2,3] Deployment Models of Cloud Computing [4] are: Private Cloud is owned and used for particular organization that controls the virtualized resources. Public Cloud is owned and delivered for general public use by a particular organization or company to offer access to computing resources at minimal cost. Community Cloud is shared through various organizations or company. Hybrid cloud mean more than two cloud form a single cloud. In Cloud computing environment, there are set issues [5] such as privacy, security, performance, load balancing and reliability. The most important of these issues is the data security [6,7]. Secure cloud architectures [8] are proposed to enhance the data security at cloud end. Most effective technique to protect our data is cryptography. Different encryption schemes [9] for protection of data have been in use for many decades.

Thus in this presented work the cloud security and channel security is the main to improve. Therefore various security techniques are studied and found that the cryptographic techniques are providing the security for the communicated data.

In the recent times the biologically phenomenon have provided solutions to many computational problems. This area of research is called as Biologically Inspired Computing(BIC). Apart from the standard cryptographic algorithms, there has been research in the area of biologically inspired algorithms which are applied into the area of

cryptography. Biological systems are complex systems which are capable of processing complex information and are capable of providing solutions to many problems in engineering and technology. Modern artificial systems are mostly based on the biological phenomenon and systems like neural networks, immune systems, genetic behaviors etc [5]. BIC paradigms which are applied in the field of cryptography are categorized into: Genetic Algorithm(GA), Artificial Immune System(AIS), DNA, Cellular Automata (CA), Artificial Neural Networks(ANN), and Ant Colony Optimization(ACO).

II. RELATED WORK

In this section a detailed description of existing technique have described like that what is role of cryptography over cloud, how it's working, advantage and disadvantage and more important thing security of information over cloud environment. Moreover it described the performance parameter of information security in public network cloud environment and how can be improving to them.

In [4] researchers proposed a model which is a combination of dynamic hashing fragmented component for data confidentiality and neural data security model for encryption and decryption of data. In [5] a security framework is proposed to enhance security for cloud computing using ANN. The model proposed mainly consists of two phases. One is authentication phase using symmetric key process and second phase to ensure only the authenticate user can send or receive data. The author proposes the model inspired by counter propagation neural network. In [6] a data security strategy based on Artificial Immune Algorithm is designed. According to the model, negative selection was selected to mature antibodies by then files stored in data node could be detected and arranged in optimized node. New files were created and manage into a data node based on the clone selection algorithm. In [8] author had proposed an asymmetric key encryption algorithm which was designed using the bio-inspired genetic algorithm. Genetic algorithm is generically used for solving the optimization problem as it provides robustness. Here the genetic algorithm is applied for the cryptographic domain. The basic advantage of genetic algorithm is that it does not break with any change in the inputs. In this paper, the authors have utilized the crossover and mutation functions of the genetic algorithm for the generation of the key pair for the encryption and the decryption algorithms. The length of the secret key is determined by the number of crossover points, mutation points along wit the random byte and permutation factor. So as to maintain uniformity, each parameter is represented using 4 bits. The size of the key generated is 36 bits.The strength of the algorithm is the permutation of the asymmetric key, and the number of permutations is agreed upon by both the sender and the receiver. Authors of the paper claim that, this randomness makes the algorithm robust. At first in [7] genetic algorithm in crypt-analysis is proposed to analyze simple substitution ciphers. In his work he even specified that genetic algorithm can also be used to cryptanalyse knapsack ciphers. He derived his fitness function using single character and diagram frequency distribution.

In [10] a genetic algorithm based security mechanism is proposed. The proposed model utilizes the properties of attribute-based encryption for the key generation and management. The proposed methodology is more suitable for the cloud scenario, as it takes less execution time thereby decreasing the latency and increasing the performance of cloud. In [1] researchers proposed a data hiding mechanism to hide data in DNA sequences, there buy increasing the confidentiality and complexity in cloud environments. There exists many more authors who have contributed in the field of DNA cryptography and continuous research is going on in the particular domain. In [11] a work is presented based on DNA cryptographic algorithm for security. The DNA based cryptographic technique is basically developed using the substitution and other basic operator's implementation. Thus this technique is less computational cost effective and efficient.

III. PROPOSED METHODOLOGY

The main motive of the proposed work is security concern in respect of confidentiality of the data file at the Cloud End while data is getting stored in cloud disk. In order to keep securities at cloud storage following skeleton of the proposed work which is hybrid in nature containing three stages is given.

The idea for the algorithm comes with the basis on bio inspired model i.e. DNA Cryptography where the encryption and decryption strategy is based on the following steps.

1. Generate Secret Key used for encryption & decryption
2. Encrypt the File using modified DNA cryptography.
3. Stores the encrypted data file at Data Centre.

The based algorithm is quit efficient to work with cloud with little time complex. Figure 1 represents the proposed architecture of the system.

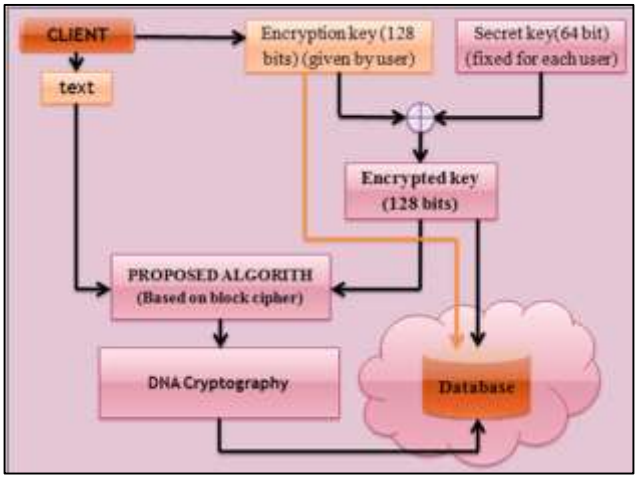

Figure 1: Proposed Encryption Architecture

In the first level of proposed work deals with new designed encryption algorithm which is based on the symmetric cryptographic concept (block based). This work uses 128-bit block size for the encryption purpose and this 128-bits block size provides more security level and at the same time this 128-bits block size is encrypted with the help of the encrypted key which size is also 128-bits. Encrypted key is generated by applying XOR operation over private key of user and secret key of the cloud. So in this way through the new designed encryption algorithm provides the double level of data security. The proposed algorithm steps are shown in figure 1. This algorithm is iterated for 10 rounds. After 10 rounds, the encrypted data is send to DNA algorithm for re-encryption. Detailed steps of encryption process is described below:

Note: PT = Plain Text, K = Key, L = Left part of PT, R = Right part of PT, ${}^1K_{64}$ & ${}^2K_{64}$ = Sub-Keys of K , CT_1 = Cipher Text generated after encryption.

Step 1: Input PT & K
Step 2: Divide PT = L & R
Step 3: Divide K = ${}^1K_{64}$ & ${}^2K_{64}$
Step 4: L>>r -----> L (2-bit Right Circular shift)
Step 5: L ⊕ R -----> L (XOR operation)
Step 6: R>>r -----> R (2-bit Right Circular shift)
Step 7: Swap L & R
Step 8: L ⊕ ${}^1K_{64}$ -----> L
Step 9: L << l -----> L (2-bit Left Circular shift)
Step 10: L ⊕ R -----> R
Step 11: R << l -----> R (2-bit Right Circular shift)
Step 12: Swap L & R
Step 13: L ⊕ ${}^2K_{64}$-----> L
Step 14: Repeat step 4 to 13 up to 10 rounds
Step 15: CL + CR = CT_1

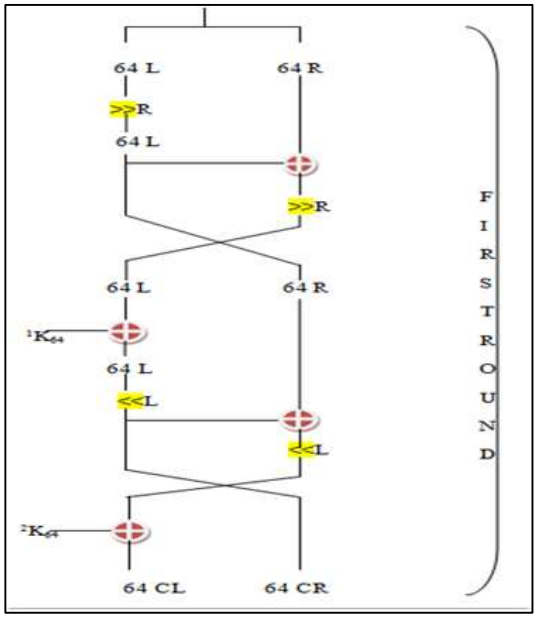

Figure 2: Proposed Algorithm

Further DNA Cryptography is used to re-encrypt the encrypted data file in first level. In information science, the binary digital coding encoded by two state 0 or 1 and a combination of 0 and 1. But DNA digital coding can be encoded by four kind of base that is ADENINE (A) and THYMINE (T) or CYTOSINE (C) and GUANINE (G). There are possibly 4!=24 pattern by encoding format. Here in this work, ATGC is being used as a key. Every bit have 2 bits like A=00, T=01, G=10, and C=11 and by using ATGC, key combinations is generated and give numbering respectively. To understand the scenario of proposed DNA cryptography flow chart is illustrated in figure 3.

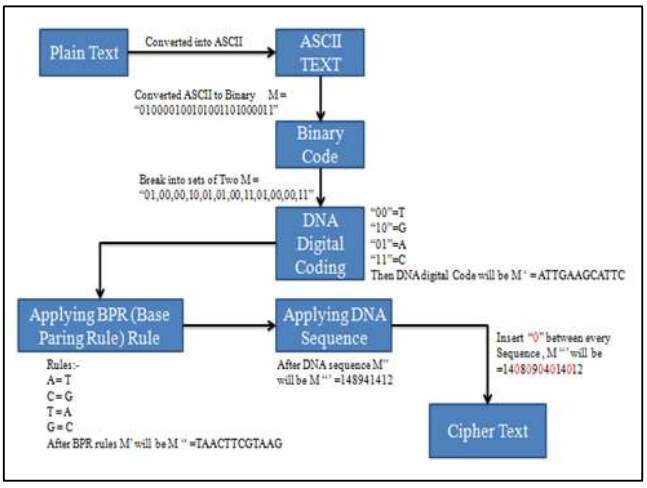

Figure 3: Proposed DNA Cryptography

File retrieval process is also termed as decryption of data file from cloud data center. When a user wants to access his data file that is saved at data center, he sends his password to the cloud service provider. Cloud service provider then re-decrypts the data file and retrieves the original data file using hybrid DNA decoding process as shown in Figure 4

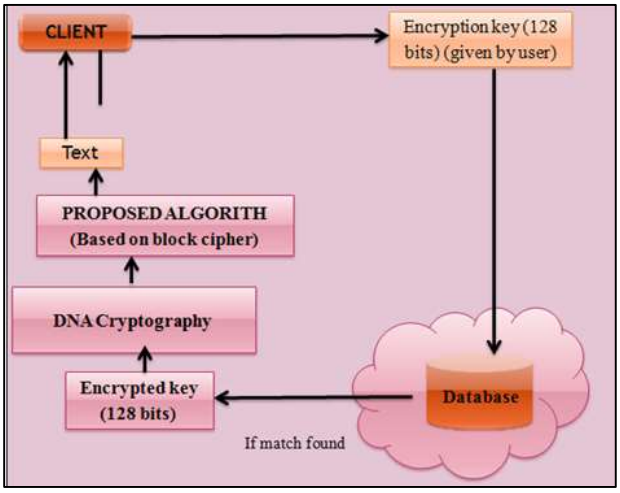

Figure 4: Proposed Decryption Architecture

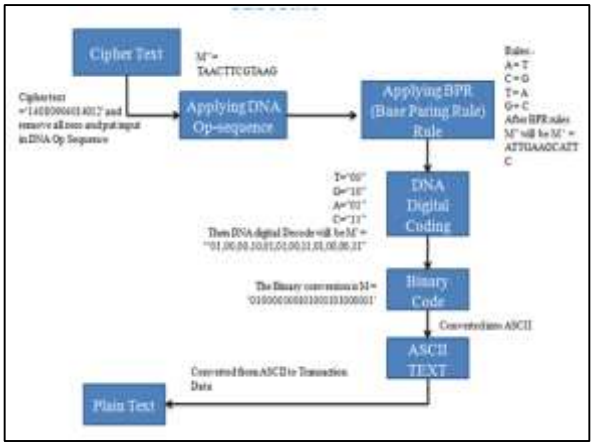

Figure 5: DNA Decryption Architecture

After decryption through DNA cryptography, proposed decryption module re-decrypt the data file using given below algorithm as illustrated in figure 6.

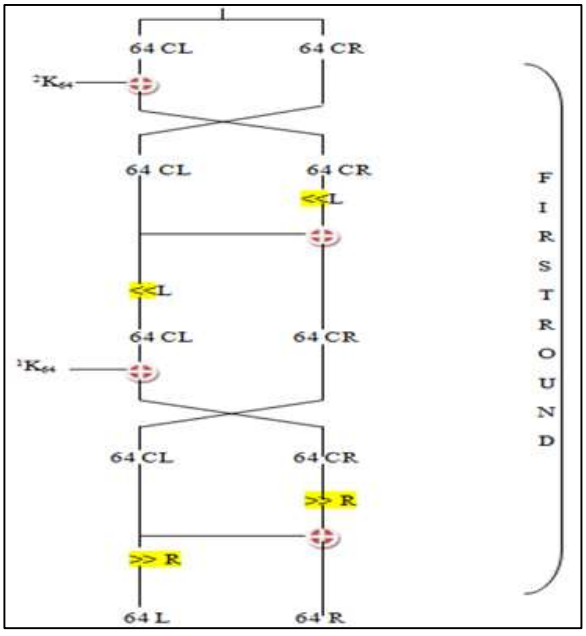

Figure 6: Proposed Decryption Architecture

Note: CT_1 = Cipher Text retrieved from database, K = Encrypted Key, CL = Left part of CT_1, CR = Right part of CT_1, $^1K_{64}$ & $^2K_{64}$ = Sub-Keys of K , PT = Plain Text generated after completion of decryption process of proposed algorithm.

Step 1: Input CT_1 & K
Step 2: Divide CT_1 = CL & CR
Step 3: Divide K = $^1K_{64}$ & $^2K_{64}$
Step 4: CL $\oplus$ $^2K_{64}$ -----> CL (XOR operation)
Step 5: Swap CL & CR
Step 6: CR >> R -----> CR (2-bit Right circular shift)
Step 7: CL $\oplus$ CR -----> CR
Step 8: CL >> R -----> CL (2-bit Right circular shift)
Step 9: CL $\oplus$ $^1K_{64}$ -----> CL
Step 10: Swap CL & CR
Step 11: CR <<L -----> CR (2-bit Left circular shift)
Step 12: CL $\oplus$ CR -----> CR
Step 13: CL << L -----> CL (2-bit Left circular shift)
Step 14: Repeat step 4 to 13 up to 10 rounds
Step 15: CL + CR -----> PT

IV. PERFORMANCE EVALUATION

For evaluation of performance of proposed algorithm the parameters or criteria is to be determined to analyze or test its efficiency. Here the execution time is preferred factors to analyze the performance of the proposed algorithm to encrypt/decrypt data blocks of various sizes.

A. Execution Time Evaluation

The total time taken by a process to encrypt data files i.e. convert plain text into cipher text is called encryption time or cipher text to plain text is called decryption time. The table I and II shows the execution time observation of encryption process and decryption process of proposed algorithm respectively. Execution process is evaluated using different file size.

Table I: Encryption Time Analysis of Proposed Method

File Size	Existing [1]	Proposed
200 Byte	7.1	6
400 Byte	13.3	12
600 Byte	20.4	19
800 Byte	27.7	27
1000 Byte	35.1	29

Table II: Decryption Time Analysis of Proposed Method

File Size	Existing [1]	Proposed
200 Byte	8.2	1
400 Byte	14.5	1
600 Byte	22.2	4
800 Byte	30.3	5
1000 Byte	35.6	8

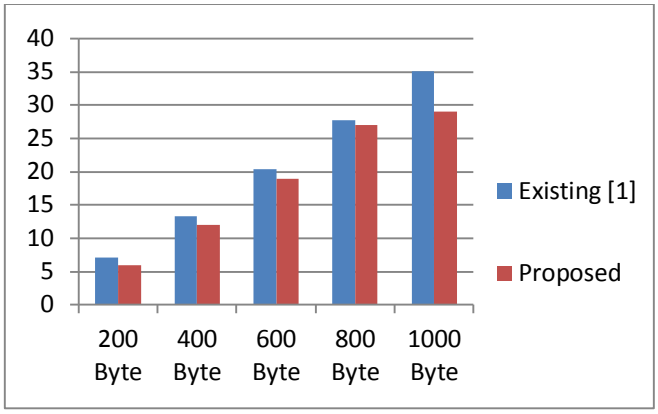

Figure 7: Encryption Time Analysis of Proposed Method

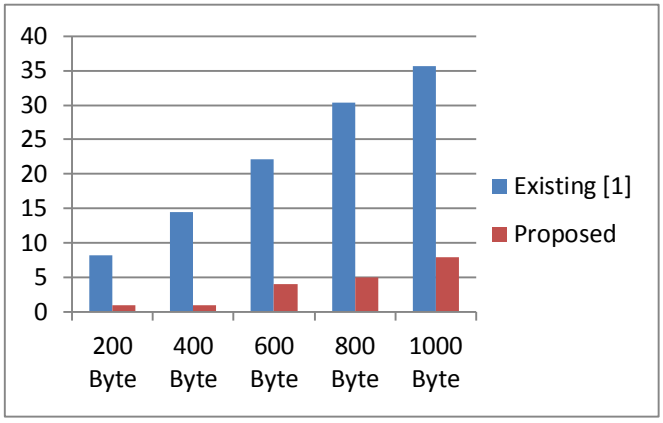

Figure 8: Decryption Time Analysis of Proposed Method

Figure 6 and 7 shows the execution time observation of encryption process and decryption process of existing and proposed algorithm. Execution process is evaluated using different file size such as 200 to 1000 bytes. After analyzing all data files on existing techniques it is concluded that as data file size increases, execution time increases. But it is also observed that as file size increases, the execution time in proposed algorithm is quite efficient than existing algorithm.

In cryptography, the avalanche effect is a most impressive property for block ciphering and hash function algorithms [12, 13]. Avalanche Effect Formula is given below as:

$$\text{Avalanche Effect} = \frac{\text{Number of change bit in cipher text}}{\text{Number of bit in cipher text}}$$

Table III: Comparative Analysis of Avalanche Effect

File Size	DES	AES	Proposed
200 Byte	35.03%	33.28%	48.94%
400 Byte	34.94%	33.70%	42.25%
600 Byte	34.96%	33.81%	57.00%
800 Byte	34.79%	32.95%	48.92%
1000 Byte	34.94%	33.50%	44.10%
Average	34.93%	33.45%	48.24%

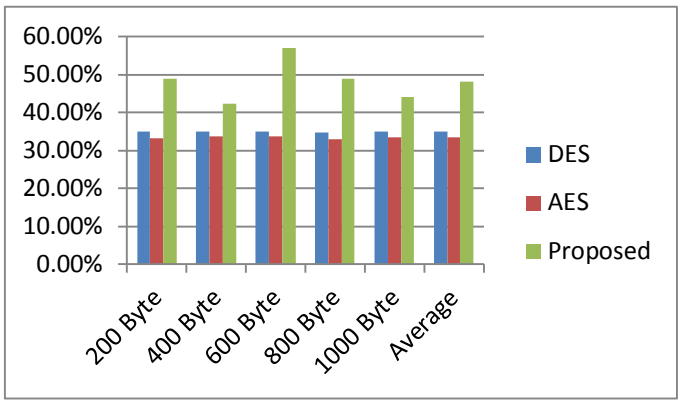

Figure 9: Comparative Analysis of Avalanche Effect

Avalanche effect allow small variations to propagate through iterations of the algorithm in a way that every bit of the cipher text should depend on every bit of plaintext before the termination of an algorithm. Table III and Figure 9 give a comparative analysis of some existing algorithms with proposed algorithms.

V. CONCLUSION

Presented research work focused on the cloud data protection or security at cloud end. To make sure data protection or security of cloud data storage at cloud end, security architecture is designed that secures the data using encryption/decryption algorithm where the proposed algorithm is a hybrid encryption algorithm. The proposed work maintains confidentiality in the cloud environment. For confidentiality maintenance hybrid encryption algorithm based on DNA cryptography and AES is proposed. Presented experiment results show that the proposed concept is reasonable, it enhancing efficiency in terms of execution time i.e. encryption as well as decryption time and security and providing confidentiality of cloud data and security and providing confidentiality of cloud data at could end. The security level is validated by the concept of Avalanche effect which is compared with some traditional algorithms i.e AES and DES algorithm. The proposed method provides a secure framework for confidentiality of text information at cloud storage data that can be useful in a number of applications at cloud. As part of future work data sharing between multiple users can also be performed. For future point of view it is required to add a third party auditor for invigilation of the entire communication.

REFERENCES

[1] Abbasy, M. R. and Shanmugam, B, Enabling data hiding for resource sharing in cloud computing environments based on dna sequences. In 2011 IEEE World Congress on Services, pages 385390. IEEE,2011.

[2] Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., and Brandic, I, Cloud com-puting and emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility. Future Generation computer systems, 25(6):599616, 2009.

[3] Clark, A. J. (1998). Optimisation heuristics for cryptology.60 Dimovski, A. and Gligoroski, D., Attacks on the transposition ciphers using optimization heuristics. In International Scientific Conference on Information, Communication & Energy Systems & Technologies ICEST, 2003.

[4] Jegadeeswari, S., Dinadayalan, P., and Gnanambigai, Enhanced data security using neural network in cloud environment. International Journal of Applied Engineering Research, 11(1):278285, 2016.

[5] Negi, A., Singh, M., and Kumar, S, An efficent security framework design for cloud computing using artificial neural networks. International Journal of Computer Applications, 129(4):1721, 2015.

[6] Phangal, S. and Kumar, M, A dual security scheme using dna key based dna cryptography. In Proceedings of the 2014 International Conference on Information and Communication Technology for Competitive Strategies, page 37. ACM, 2014.

[7] Spillman, R., Janssen, M., Nelson, B., and Kepner, M., Use of a genetic algo- rithm in the cryptanalysis of simple substitution ciphers. Cryptologia, 17(1):3144, 1993.

[8] Naik, P. G. and Naik, G. R, Symmetric key encryption using genetic algorithm. International Journal of Latest Trends in Engineering and Technology (IJLTET), 3(3):118128,2014.

[9] Ranalkar, R. and Phulpagar, B, DNA based cryptography in multi-cloud: Security strategy and analysis. pages 189192, 2014.

[10] N. Hitaswi and K. Chandrasekaran, "A Bio-Inspired Model to Provide Data Security in Cloud Storage", IEEE, 2016.

[11] Shruti Goyal, Sourabh Jain, "A secure cryptographic cloud communication using DNA cryptographic technique", ICICT, IEEE, 2016.

[12] Rajkumar Buyya, Karthik Sukumar, "Platforms for Building and Deploying Applications for Cloud Computing", CSI Communications, 2011.

[13] Ganesh Patidar, Nitin Agrawal, SitendraTarmakar, "A block based Encryption Model to improve Avalanche Effect for data Security", International Journal of Scientific and Research Publications, Volume 3, Issue 1, 2013.