Designing and Implementing Cloud Security Using Multi-layer DNA Cryptography in Python

Md. Irfan Alam and Satya Narayan Singh

Abstract Cloud computing is the latest technology. Provides various on-demand services and online for network services, platform services, data storage, etc. Many organizations are not thrilled with using cloud services due to data security concerns, as the data resides on the cloud service provider's servers. To address this problem, various researchers around the world have applied various approaches to strengthen the security of data stored in cloud computing. The latest development in the field of cryptography is DNA encryption. It arose after the disclosure of the computational ability of deoxyribonucleic acid (DNA). DNA encryption uses DNA as a computational tool along with various molecular techniques to manipulate it. Due to the large storage capacity of DNA, this field is becoming very promising. This paper used a layered DNA encryption method for the data encryption and decryption process. Using the four DNA bases (A, C, G, T), we generate dynamic DNA tables to replace the message characters with a dynamic DNA sequence. The implementation of the proposed approach is performed in Python and the experimental results are verified. The resulting encrypted text contains information that will provide greater security against intruder attacks.

Keywords Cloud computing · Cloud security · DNA computing · DNA sequencing · Encryption · Decryption · Cryptography techniques

Md. Irfan Alam (✉)
Jharkhand Rai University, Ranchi, Jharkhand, India
e-mail: irfan.alam1@ymail.com

S. N. Singh
Department of IT, Xavier Institute of Social Service, Ranchi, Jharkhand, India
e-mail: snsinghxiss@rediffmail.com

M. Chakraborty et al. (eds.), *Trends in Wireless Communication and Information Security*,
Lecture Notes in Electrical Engineering 740,
https://doi.org/10.1007/978-981-33-6393-9_38

1 Introduction

Cloud computing is the field of computing in which an organization or individual stores data such as text, images, or video on remotely hosted servers, rather than keeping it all on their local storage machine or computer. The concept of cloud computing has existed since the late 1970s in the form of distributed computing, but was popularized by Amazon.com in 2006. Today, cloud technology is one of the most widely used computing technologies.

1.1 Motivation

Cloud storage service offers tremendous benefits to customers. Despite these benefits, the concern over security and privacy associated with cloud model seems to be the biggest hurdle in adopting cloud by many individuals and organizations. First problem is, since the data resides on third party's premises, data owners lose control over their outsourced data. Second problem is data owners need to take high risk in trusting cloud service provider on all circumstances. Finally, the multi-tenancy nature of cloud brings in several malicious internal and external attacks. Lack of data security in cloud environment poses major challenges to data owners. The motivation of this research work is to seek cloud data security concerns and proposes secure and efficient protocols for multilevel security in cloud environment to preserve confidentiality, authorized access to stored data, authenticity, and integrity of data from the perspective of data owners.

1.2 Contributions of the Paper

The main contribution of this paper is to design and implement multi-layer DNA cryptography security system for the data outsourced to cloud data storage that preserves data confidentiality, authenticity, and integrity from the perspective of data owners. The main objective of this research work is to provide model for data protection. This model is designed in such a way to resist vulnerabilities and threats that jeopardize the data being transferred through an open communication medium. This could be possible with the strong cryptographic schemes with strong key generation mechanism.

1.3 Fundamental Concepts of Clouds

According to NIST, there are five core concepts in the cloud [1, 2] such as cloud functionality, service models, hosting, deployment models, and roles which are elaborated broadly as follows:

1.3.1 Cloud Features

The cloud contains five main features as follows:

(i) Services on demand: No human interaction with the resource provider is required for the provision of IT services such as storage, server time.
(ii) Access to the ubiquitous network: IT services are available on the network and can be accessed with the aid of standard methods using heterogeneous consumer platforms (e.g., mobile phones, laptops).
(iii) Location independent resource pool (multi-tenant): Resources are collected to serve multiple customers with the help of the multi-tenant paradigm where resources are dynamically allocated and reallocated on demand. Customers have no idea where the services are.
(iv) Quick elasticity: Resources are supplied quickly with good enough elasticity and similarly released to scale.
(v) Measured services: The use of resources is controlled by providing a measurement capability. Customers pay the bill based on the measured usage of the resources provided for a specific session.

1.3.2 Service Models

Cloud service models can be grouped into three classifications as follows:

i. SAAS: Software-as-a-Service is a product conveyance model in which applications are facilitated by a supplier or specialist co-op and made accessible to clients over an organization, and as a rule, the Internet. SaaS is additionally regularly connected with a pay-more only as costs arise membership authorizing model.
ii. PAAS: Platform-as-a-Service (PaaS) is a lot of improvement devices and programming facilitated on the supplier's workers. It is a layer on head of IaaS in the stack and edited compositions everything down to the working framework. This offers an incorporated arrangement of advancement condition that a designer can use to manufacture their own applications without understanding what is happening under the administration.

iii. IAAS: Model that incorporates its arrangement administrations, for example, stockpiling, network limit, preparing components to permit clients to run their applications [3].

1.3.3 Deployment Model

Cloud can be conveyed in four models:

i. Private cloud: These are executed distinctly inside an endeavor or association. Venture or outsider claims it. Private mists are worked inside an endeavor firewalls and on location worker run them. They offer types of assistance, for example, virtualization, multi-occupancy, consistent arrangement, security, and access control [4].
ii. Public cloud: Public cloud portrays distributed computing in the customary standard sense, whereby assets are powerfully provisioned on a fine-grained, self-administration premise over the Internet, by means of Web applications/Web administrations, from an off-Webpage outsider supplier who shares assets and bills on a fine-grained utility figuring premise. It is ordinarily founded on a compensation for each utilization model, like a prepaid power metering framework which is adaptable enough to cook for spikes popular for cloud advancement [5]. Public mists are less secure than the other cloud models since it puts an extra weight of guaranteeing all applications and information got to on the public cloud.
iii. Hybrid cloud: It is a collection of private, public, and network mists. Public and private mists both are worked by cross-breed cloud at the same time. Half breed cloud is a private cloud connected to at least one outside cloud administrations, midway oversaw, provisioned as a solitary unit, and delineated by a safe organization [6].
iv. Community cloud: A mutual foundation is characterized in this sort of cloud and a few associations uphold it.

1.4 Cryptography

Cryptography is the study of techniques or methodology to encode the plain text into ciphered text and vice versa. Cryptography consists of basically two complementary sub-techniques:

(1) Encryption
(2) Decryption.

Encryption is the technique to convert a plain text (understandable form) into ciphered/encrypted text (not understandable form). This process is known as ciphering or encrypting. Decryption is the technique to convert the ciphered/encrypted text (not understandable form) to plain text (understandable

form). This is process is also known as deciphering or decrypting. Implementing cryptography makes the medium of communication secure and thus channel becomes a more reliable medium to send some secret data.

Cryptography can be broadly divided into two categories:

(i) Symmetric cryptography
(ii) Asymmetric cryptography [5].

Symmetric cryptography consists of same secret key at side, sender and receiver. This means that in symmetric cryptography, the Encryption Key and Decryption Key are same. Asymmetric cryptography uses two types of secret keys:

i. Private key
ii. Public key.

Cryptography is time taking and requires intensively complex processing but yet maintaining the security as maximum as possible. To make DNA-based cryptography, a more reliable but yet a fast medium to implement security we will use the symmetric cryptography technique. To increase security further, we will use various existing cloud-based system security techniques. There are various cryptography implementing techniques but the core thing among all is that the degree of uncertainty and randomness in the process of generation of secret key. Higher the degree of uncertainty and randomness, the more secure and stable medium we have but eventually increasing the processing requirement.

1.5 DNA Structure

DNA stands for deoxyribonucleic acid. DNA is molecule that is made up of two components. First, it has four types of bases that are Thymine (T), Guanine (G), Cytosine (C), and Adenine (A). Second, DNA contains sugar phosphate which makes its backbone as shown in Fig. 1. DNA bases are also known as nucleotides. DNA consists of two biopolymer strands and forms a double helix structure that is described in Fig. Between strands, lie base pairs that are bonded together with strong hydrogen bonding. This pair exists only in certain manner such that "Adenine (A) can only make Hydrogen Bond (double bond) with Thymine (T)" and "Cytosine (C) can only make Hydrogen Bond (triple bond) with Guanine (G)". The sequence of these base pair defines the rules for formation of cell, eventually whole body of an organism. DNA is found in every cell of every living being. Basically, it is found in every nucleus of the cell and also in mitochondria. It is a biological storage device that stores all the genetic information and instructions which helps in formation of cells. DNA molecule has two chemical polarities that are 5′ and 3′ at top and bottom as described in Fig. 1. These two polarities enable DNA to bind together and transform itself into a double strand helix structure from single strand structure. DNA molecule

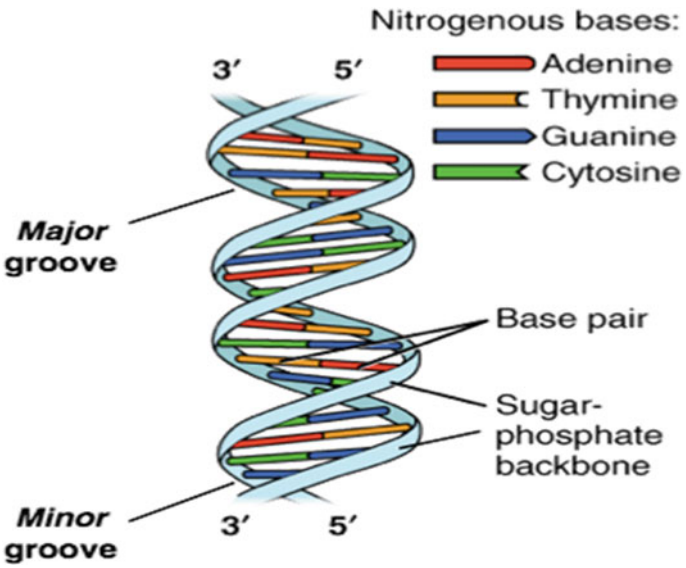


Fig. 1 DNA double helix (https://openstax.org/books/concepts-biology/pages/9-1-the-structure-of-dna)

is responsible for transmitting massages among the cells. DNA uses proteins to interact with its environment. DNA uses messenger ribonucleic acid (mRNA) to send information. There are two processes that are involve in transmitting a message: (i) Transcription (ii) Translation. In transcription, DNA passes the information to the mRNA. In translation, mRNA uses the information to interact with proteins and passes on the desired message. DNA has the phenomenal biological property where it replicates without losing the original DNA.

1.5.1 DNA Computing and DNA Cryptography

DNA computing is the field of computing bringing together the computer science, Biological Science and Molecular Science to understand and solve some primary NP problem [6, 7]. Earlier, it was introduced by Leonard Max Adleman but now it has evolved as one of the most fascinating platform to develop something new by teachings of Mother Nature. DNA computing is the best example of biomolecular computing. Biomolecular computers are those computers where all the computing components are made up of molecular compounds, i.e., all Input/Output and Software/Hardware are all in form of a molecular compound. DNA computing involves various steps such as melting, annealing, merging, amplification, and selection. DNA actually behaves like a turing machine that is why it can be used as a data storage

device. Adleman has showed that DNA computing can be used as an effective tool to solve the NP problems like Hamiltonian graph problem or travelling salesman problem (TSP) [9]. He showed that DNA computing can be used to solve complex combinatorial problems like TSP and finite state problem. Here, the basic idea is that all the operations are performed over DNA (more precisely using DNA Bases or Nucleotide) not in DNA. DNA computing can be classified as intermolecular DNA computing, intramolecular DNA computing and supramolecular DNA computing. DNA Computers form a self-replicating system. DNA cryptography uses DNA nucleotides only to generate a set of symmetric cryptographic key. For, DNA cryptography, many techniques has already been established in many researches [8, 9] but here, I aimed to develop a technique to make the existing cloud-based data storage security systems more accurate and giving the encrypting and decrypting capability directly to the authorized client on its own machine. In this technique, first, we have to define three types of information. First of all, we need certain standard library named as DNA reference sequence that includes the 4-bit base sequence uniquely defined for all 256 ASCII characters in random order as shown in Table 1.This DNA reference sequence encodes the plain text message into DNA bases sequence text. Second table will replace the existing genome DNA base sequence to other DNA base sequence. Third component we need is the base-binary library that store the information about the equivalent conversion of DNA base sequence message to a long binary string, i.e., Table 3. This base-binary library is also not standardized as it can be defined by the user itself (Table 2).

2 Proposed Algorithm

Aim of this proposed algorithm is to provided client end cryptography using DNA computing which when.

integrated with existing cloud system storage server security methods can increase reliability and secrecy of the data. Important feature of this algorithm is that the data getting stored or data under transmission if even get.

hacked or intruded, that data will be of no use for the middle man even to that data administrator of the cloud storage facility.

Table 1 DNA reference sequence

TTTT	NUL	TTTC	Space	TTTG	@	TTTA	`	TTCT	Ç	TTCC	á
TTCG	└	TTCA	Ó	AAAG	■	AAAA	nbsp				
TTGT	SOH	TTGC	!	TTGG	A	TTGA	a	TTAT	ü	TTAC	í
TTAG	┴	TTAA	ß	TCTT	STX	TCTC	"	TCTG	B	TCTA	b
TCCT	é	TCCC	ó	TCCG	┬	TCCA	Ô	TCGT	ETX	TCGC	#
TCGG	C	TCGA	c	TCAT	â	TCAC	ú	TCAG	├	TCAA	Ò
TGTT	EOT	TGTC	$	TGTG	D	TGTA	DEL	TGCT	ä	TGCC	ñ
TGCG	─	TGCA	õ	TGGT	ENQ	TGGC	%	TGGG	E	TGGA	e
TGAT	à	TGAC	Ñ	TGAG	┼	TGAA	Õ	TATT	ACK	TATC	&
TATG	F	TATA	f	TACT	å	TACC	ª	TACG	ã	TACA	µ
TAGT	BEL	TAGC	'	TAGG	G	TAGA	g	TAAT	ç	TAAC	º
TAAG	Ã	TAAA	þ	CTTT	BS	CTTC	(	CTTG	H	CTTA	h
CTCT	ê	CTCC	¿	CTCG	╚	CTCA	Þ	CTGT	TAB	CTGC	)
CTGG	I	CTGA	i	CTAT	ë	CTAC	®	CTAG	╔	CTAA	Ú
CCTT	LF	CCTC	*	CCTG	J	CCTA	j	CCCT	è	CCCC	¬
CCCG	╩	CCCA	Û	CCGT	VT	CCGC	+	CCGG	K	CCGA	k
CCAT	ï	CCAC	½	CCAG	╦	CCAA	Ù	CGTT	FF	CGTC	,
CGTG	L	CGTA	l	CGCT	î	CGCC	¼	CGCG	╠	CGCA	ý
CGGT	CR	CGGC	-	CGGG	M	CGGA	m	CGAT	ì	CGAC	¡
CGAG	═	CGAA	Ý	CATT	SO	CATC	.	CATG	N	CATA	n
CACT	Ä	CACC	«	CACG	╬	CACA	¯	CAGT	SI	CAGC	/
CAGG	O	CAGA	o	CAAT	Ä	CAAC	»	CAAG	¤	CAAA	´
GTTT	DLE	GTTC	0	GTTG	P	GTTA	p	GTCT	É	GTCC	░
GTCG	ð	GTCA		GTGT	DC1	GTGC	1	GTGG	Q	GTGA	q
GTAT	æ	GTAC	▒	GTAG	Ð	GTAA	±	GCTT	DC2	GCTC	2
GCTG	R	GCTA	r	GCCT	Æ	GCCC	▓	GCCG	Ê	GCCA	‗
GCGT	DC3	GCGC	3	GCGG	S	GCGA	s	GCAT	ô	GCAC	│
GCAG	Ë	GCAA	¾	GGTT	DC4	GGTC	4	GGTG	T	GGTA	t
GGCT	ö	GGCC	┤	GGCG	È	GGCA	¶	GGGT	NAK	GGGC	5
GGGG	U	GGGA	u	GGAT	ò	GGAC	Á	GGAG	ı	GGAA	§
GATT	SYN	GATC	6	GATG	V	GATA	v	GACT	û	GACC	Ã
GACG	Í	GACA	÷	GAGT	ETB	GAGC	7	GAGG	W	GAGA	w
GAAT	ù	GAAC	À	GAAG	Î	GAAA	¸	ATTT	CAN	ATTC	8
ATTG	X	ATTA	x	ATCT	ÿ	ATCC	©	ATCG	Ï	ATCA	°
ATGT	EM	ATGC	9	ATGG	Y	ATGA	y	ATAT	Ö	ATAC	╣
ATAG	┘	ATAA	¨	ACTT	SUB	ACTC	:	ACTG	Z	ACTA	z
ACCT	Ü	ACCC	‖	ACCG	┌	ACCA	·	ACGT	ESC	ACGC	;
ACGG	[	ACGA	{	ACAT	ø	ACAC	╗	ACAG	▉	ACAA	¹
AGTT	FS	AGTC	<	AGTG	\	AGTA	|	AGCT	£	AGCC	╝
AGCG	▄	AGCA	³	AGGT	GS	AGGC	=	AGGG	]	AGGA	}
AGAT	Ø	AGAC	¢	AGAG	¦	AGAA	²	AATT	RS	AATC	>
AATG	^	AATA	~	AACT	×	AACC	¥	AACG	Ì	AACA	▪
AAGT	US	AAGC	?	AAGG	_	AAGA	DEL	AAAT	*f*	AAAC	┐

Table 2 DNA base to other DNA base sequence

DNA code	Corresponding DNA
A	T
T	G
G	C
C	A

Table 3 DNA base to binary library

DNA code	Binary value
A	00
T	01
G	10
C	11

A. ENCRYPTION:

Step1: Let Message be: M = "A"

Step 2: M can be encode as follows Using Table 1 as follows:
M1 = "TTGG".

Step 3: Message M1 can be encoded using Table 2 as follows:
M2 = "GGCC".

Step 4: Using Base-Binary Library Message M2 can be encoded as follows using Table 3:
M3 = 10101111.

Step 5: A Random number is generated, this random number is XORed with the message M3
Let the random number generated=5
M4=10101111 XOR 00000101=10101010

Step 6: Converting Binary String M4 into equivalent Decimal Value.
M5 = 170.

Step 7: Convert message M5 to its equivalent ASCII character
M6=ASCII character of (M5)
Send this data, M6 to the Cloud Server.

Step 8: Generate the decryption key as:
[Corresponding DNA Base Value for A][Base-Binary for A][Corresponding DNA Base Value for C][Base-Binary for C][Corresponding DNA Base Value for G][Base-Binary for G][Corresponding DNA Base Value for T][Base-Binary for T][Random Number]
Example: Decryption Key: - "T00A11C10G0100000101".

B. DECRYPTION:

Step1: Download the file from the Cloud Server i.e.
M6

Step 2: Retrieve the decryption key i.e.
Decryption Key: - "T00A11C10G01".

Step 3: Convert the ASCII character of message to equivalent decimal number, thus we get message M5.
M5=170

Step 4: Convert the decimal of message M5 to its binary equivalent .message M4 obtained
M4=10101010

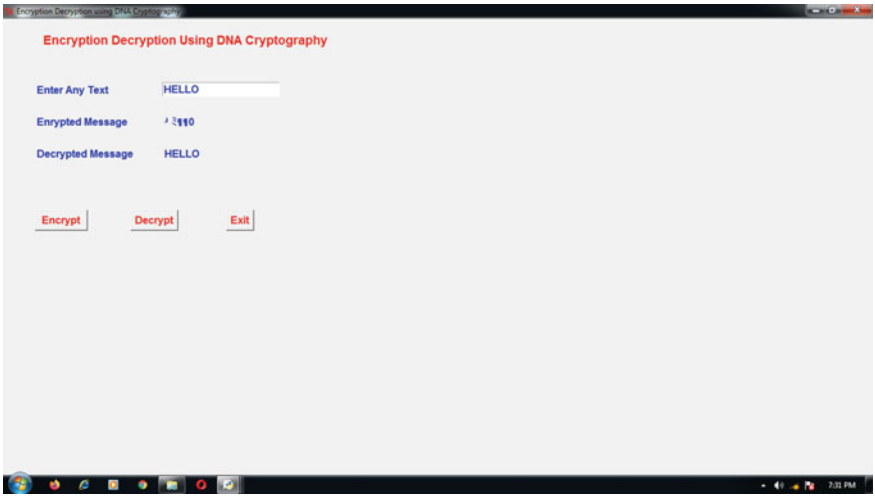
Step 5: XORED the message M4 with random number, message M3 obtained
M3=10101111

Step 6: Convert the binary to its equivalent DNA base
M2=GGCC

Step 7: Convert the message M2 to its corresponding DNA base
M1=TTGG

Step 8: Convert the message M1 to its equivalent character
M=A

3 Implementation of Above Algorithm in Python



4 End and Future Work

In this paper, a DNA-based multi-layer encryption technique is proposed for storing data in the cloud mainly in the public cloud and for SaaS users where security is a major concern. The technique will provide improved security as it includes the

computational complexity by using biocomputing techniques in addition to cryptography. User can check the integrity of the data without relying on the third party. The proposed DNA cryptography is a novel encryption technique for secure storage of data in the cloud environment, using DNA cryptography for cloud has great scope considering the importance of cloud storage in the industries and day-to-day life. Everywhere data is bombarding in the form of video, image, and other digital forms. So, a platform for storage is very important and DNA encryption is a trending new concept which will dominate the security world in the future. The proposed DNA cryptographic method utilizing dynamic character-DNA succession table, DNA base to its equivalent DNA base table and DNA base-binary table to builds the degree of information security. The above algorithm is implemented in Python. Many cryptanalyst has just said that the future of cryptography lies in the multidisciplinary studies of different aspects of science and Mathematics.

References

1. Global Netoptex Incorporated. Demystifying the cloud. Significant chances, pivotal choices. p 414. Available: https://www.gni.com. 13 Dec 2009
2. Brodkin J (2008, June) Gartner: seven distributed computing security hazards. Infoworld, Available: https://www.infoworld.com/d/security-focal/gartner-seven-loudcomputingsecurity-chances 853? Page=0, 1 13 Mar 2009
3. Kuyoro SO, Ibikunle F, Awodele O (2011) Distributed computing security issues and challenges. Global J Comput Networks (IJCN) 3(5)
4. Mell P, Grance T (2011) The NIST definition of cloud computing. IT Laboratory NIST, Gaithersburg, MD, Tech. Rep. 800-145, pp 1–3
5. Pallavi N, Singh A, Dwivedi SP (2016) A DNA based secure data hiding technique for cloud computing. Int J Current Eng Technol 6(4)
6. Adleman L (1994) Sub-atomic calculation of arrangements of combinatorial issues. Science 266:1021–1024
7. Nimje AR (2012) Cryptography in cloud-security using DNA (genetic) techniques. Int J Eng Res Appl (IJERA) 2(5):1358–1359 ISSN: 2248-9622. www.ijera.com
8. Jain A, Rajpal N, Adaptive key length based encryption algorithm utilizing DNA approach. In: International conference on machine intelligence research and advancement
9. Rahman NHU, Balamurugan C, Mariappan R, A novel DNA computing based encryption and decryption algorithm. In: International conference on information and communication technologies