18CSE381T – Cryptography

MINOR PROJECT REPORT

Submitted by

Anant Aggarwal [RA2011030010129]

Mayank Mudgal [RA2011030010140]

Manubhav Sharma [RA2011030010159]

Under the Guidance of

Dr. Mary Subaja Christo

Professor Department of Networking and Communications

In partial satisfaction of the requirements for the degree of

BACHELOR OF TECHNOLOGY
in
COMPUTER SCIENCE AND ENGINEERING

with specialization in Cyber Security

SCHOOL OF COMPUTING

COLLEGE OF ENGINEERING AND TECHNOLOGY SRM INSTITUTE OF SCIENCE AND TECHNOLOGY KATTANKULATHUR – 603203

November 2022

A Dynamic DNA Cryptography Using RSA Algorithm and Stronger OTP

Abstract

A Dynamic DNA Cryptography Using RSA Algorithm and OTP is a combination of algorithm like RSA and One Time Pad key generation which is used for Encryption and Decryption of plain text, image, audio and video that are converted from binary numbers into a DNA sequence. The device MAC Address is a unique identification which helps to connect one or more nodes and the information is transmitted from one node to another in a secured fashion. The method of cryptography is used to protect the transmission of confidential information over wireless networks. If the key used in the OTP is randomly generated and not used more than once, then the algorithm is considered to be completely unbreakable. This technique does not provide complete unauthorized data access, only is a stronger method than most.**[1]**

Keywords—*DNA Cryptography, Encryption, Decryption, Cryptography, One Time Pad.*

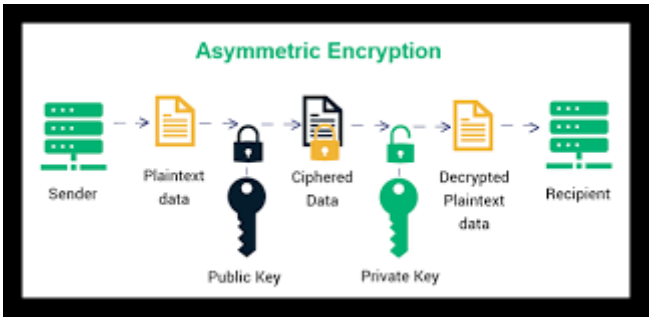

Figure 1: Asymmetric Encryption in Cryptography

Introduction

Data is any type of stored digital information. Security is about the protection of assets. Data security or Information Security refers to protective digital privacy measures that are applied to prevent unauthorized access to computers, personal databases and websites. To hide any data two techniques are mainly used one is Cryptography and the other is Steganography. In this paper we use Cryptography. Cryptography is the science of protecting data, which provides methods of converting data into unreadable form, so that Valid User can access Information at the Destination. Cryptography is the science of using mathematics to encrypt and decrypt data. **[4]**

In the digital world, cryptography plays an important role in day-to-day life that information security exists since historic times and it is existing in our modern life. The art of cryptography is considered to be born along with the art of writing. The roots of cryptography are found in Roman and Egyptian civilizations. Cryptography offers variety of aspects of data security. The main factors in cryptography are confidentiality, information integrity, authentication & non-repudiation.

The principal objectives of this paper are to more concentrate on the confidentiality phase & it help to discover the approach (ciphers) to make certain privacy via the use of DNA. DNA Cryptography is one of the most essential & promising disciplines in information security. Cryptography is considered as a technique of accomplishing confidentiality or privacy in data or message transmission. Additionally, it is a procedure of transferring the sender's data to a secret layout that is known as the ciphertext that the receiver will get recognize the secret message. The RSA algorithm is regarded as a robust asymmetric encryption algorithm. Asymmetric cryptography, also known as public-key cryptography, is a process that uses a pair of related keys -- one public key and one private key -- to encrypt and decrypt a message and protect it from unauthorized access or use. RSA has the more security; thus, the algorithm is proved to be cryptographically secure and it is suitable for applications where more than one layer of security is required. OTP is used for key generation purpose. The transferring information from one node to another node by using the MAC address it is very safe and secure as every device has its unique address in its DNA Cryptography Encryption.**[1] [2]**

Objectives

To create an Encryption Algorithm based on the RSA Encryption and then Encode that Cipher Text into DNA Sequences. Then, using an OTP based approach we shall Authenticate the user (receiver) to add even more security to the already secure transaction. However, the OTP generated shall be stronger than the one proposed in **[1]**.

Literature Survey

We started with paper **[6]** to introduce ourselves with the idea and need for Cyber Security. After which, papers **[3]** and **[4]** were followed to give a basic understanding of Cryptographic Techniques, how they work and how is a good and robust Encryption takes place, along with the basics of Ciphers, Symmetric and Asymmetric Algorithms and finally AES, DES and RSA. With the knowledge gained, Papers **[1]**,**[2]** and **[5]** were studied with the sole purpose to find a robust and modern Encryption Technique to provide maximum security. Paper **[1]** forms the main base for this project since there was a vulnerability found in the sense that although the Encryption is extremely secure with RSA and DNA Encryption, however, the Technique does not allow for complete User Authentication and the OTP system used uses a random OTP of **4** digit only which can be easily guessed ad also only a file size of **100 MB** is allowed, leading to limited applications. In this project we shall attempt to rectify these errors by generating truly strong random OTPs and hypothesizing a way to provide larger file support. Also, we are authenticating the receiver with the OTP rather than generating the keys.

The table below shows the Literature Survey summary of all journals referenced in this project along with their results and limitations.

Table 1:Literature Survey of all the journals read by the team as part of original inspiration for this project

S. No.	Title of Paper	Author and Journal Name	Objectives of the Paper	Methodologies Used	Results Obtained	Limitations
1.	"A Dynamic DNA Cryptography Using RSA Algorithm and OTP"	Poojashree Kamble,Firoz Nagarchi,Akshata Akkole,Vanishree Khanapur,Bahubali Akiwate, International Journal of Scientific Research in Computer	In this approach, the encryption and the decryption take place through RSA Algorithm and the OTP. The use of a combination of DNA with RSA ensures twofold protections in a	RSA Algorithm DNA Encoding Algorithm One Time Pad Key Generation	A robust Cryptography Technique using DNA Encryption and OTP System	In future the network can be connected with a greater number of nodes with both IP Address and MAC address with

		Science and Engineering	cloud environment where there are greater probabilities of breaches.			improvement in performance. Proposed File size only up to 100MB. Only 4-digit OTP, can be guessed easily by a fast-computing machine.
2.	"A Research Paper on Cryptography"	Gurdeep Singh, Prateek Kumar, International Journal for Technological Research in Engineering	Abstract Data Types, Data Encryption, Data Compression, Asymmetric Key Cryptography	Symmetric and Asymmetric Encryption and Decryption	Cryptography is used to ensure that the contents of a message are confidentiality transmitted and would not be altered.	If the attacker will get the key and sender and receiver is not aware about it, it may harm the CIA triad. So, proper method should be used to Encrypt and Decrypt the data.
3.	"Research Paper on Cyber Security"	Mrs. Ashwini Sheth, Mr. Sachin Bhosale & Mr. Farish Kurupkar, Contemporary Research in India	In the current world of technology, it is crucial to know what Cyber Security is and how to use it effectively. How	1)Symmetric and Asymmetric Encryption 2) Types of Phishing: - Ransomware, Malware. 3)Goals of	Awareness on Online Attacks, Types of Viruses, Goals of Cyber Security, Advantages and	Increasing Threats targeting user devices, devices used by employees who are working from

			to Secure Systems, important files, data and other important virtual things.	Cyber Security: Confidentiality, Integrity, Availability. 4)Attacks on IOT	Disadvantages of Cyber Security	home aren't protected well enough from attacks and preventing hackers.
4.	"A Study on Cryptographic Techniques"	Anjali Krishna A, Dr. L C Manikandan, International Journal of Scientific Research in Computer Science, Engineering and Information Technology	Data security using Symmetric and Asymmetric Key Encryption.	Symmetric and Asymmetric Key Encryption	There are different techniques and algorithms researched, and various types of work have been performed. In this paper briefly discussed cryptography and its form of symmetric key cryptography and algorithms for asymmetric key cryptography.	Key leakage, software bugs, holes in operating systems, side-channel attacks, phishing attacks, and social engineering ttacks. So, it is important to understand and acknowledge that cryptography ≠ security.

Architecture Diagram

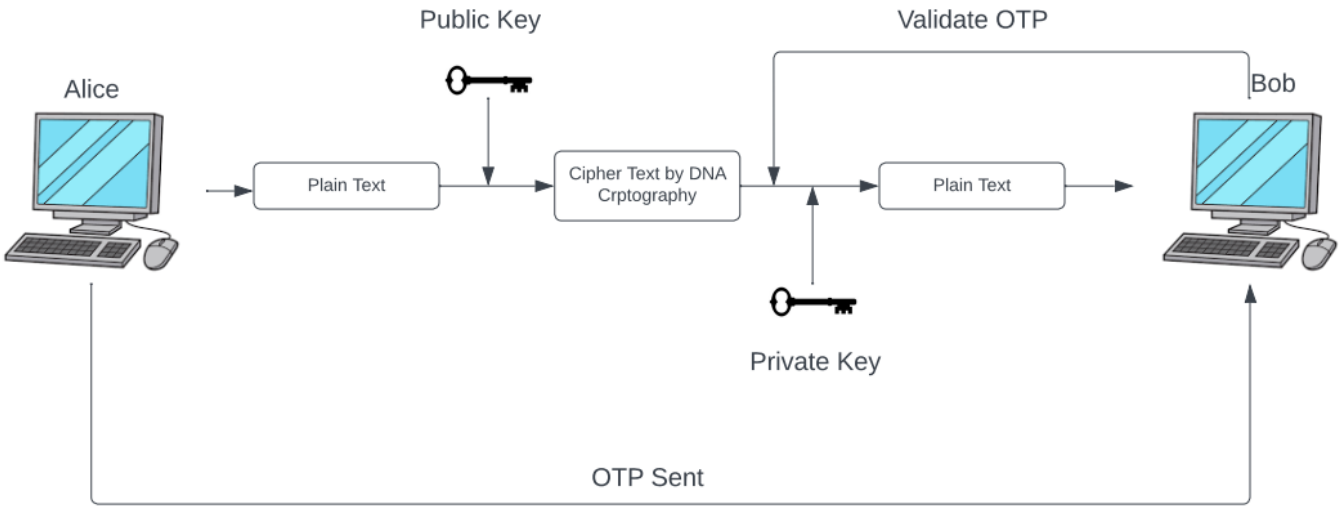

Figure 2: Architecture Diagram for the Encryption and Decryption Process between Communicators

Binary Digits	DNA Base
00	A
01	G
10	C
11	T

Figure 3: DNA Encoding as per binary digits

space → CCAG	! → CACT	" → TCGA	# → GTAC
(→ GCTG	) → CGTG	* → ATGG	+ → TGTA
, → AAAT	- → GGCC	. → TGGG	/ → TCCT
0 → CGCT	1 → TCAC	2 → GAGG	3 → CTAC
4 → CCTC	5 → CCTT	6 → AAAG	7 → GGGT
8 → TTGT	9 → TAAT	: → AGGG	; → GTTT
¡ → GTGT	= → CAAG	¿ → AACA	? → CTTG
@ → CAAA	A → TGTT	B → CAAC	C → TTAA
D → GAAA	E → CCTG	F → TGAG	G → ACCC
H → CCCC	I → GGAT	J → TGGT	K → CAGA
L → CTTC	M → ATAC	N → CCAA	O → GGCA
P → TGAA	Q → CTGG	R → GGGC	S → GCTA
T → CCCG	U → GGAA	V → AGAC	W → ACTG
X → GCAT	Y → ACCT	Z → TCTT	[→ CGTT
\ → TGGC	] → CTAT	^ → AGGA	_ → AGAA
' → ACGG	a → CTCT	b → GGTG	c → GGAG
d → TAAA	e → GCCA	f → GACC	g → GTGA
h → TGCT	i → ATAT	j → GAGA	k → CAGT
l → AATT	m → TTGG	n → GTAG	o → TCTC
p → TTTG	q → TTCC	r → GTCT	s → AGTT
t → ACAC	u → GCAA	v → TTCT	w → TCAA
x → GGTC	y → TCTG	z → AAGA	{ → GCTT
| → GTGC	} → CCCA	~ → ATGT	

Figure 4: DNA Encoding Sequences - Example

Methodologies Used

1. RSA Encryption and Decryption

The concept of RSA is based on the fact that big integers are challenging to factor. The public key is made up of two numbers, one of which is the product of two enormous prime numbers. The same two prime numbers are also used to create the private key. Therefore, the private key is compromised if someone is able to factorise the huge integer. As a result, the key size completely determines how strong an encryption is, and doubling or tripling the key size significantly boosts encryption strength. RSA keys can normally be 1024 or 2048 bits long, however experts think that keys with 1024 bits could be cracked soon. But as of right now, it appears to be an impossible feat. **[7]**

1.1. Encryption Using Public Key

- Select two prime numbers P and Q.
- Now, first part of the Public Key: n = P*Q.
- We also need a small exponent say e (integer) which is not a factor of n.
- $\Phi(n) = (P-1) * (Q-1)$
- $1 < e < \Phi(n)$
- Our Public Key is made of n and e.
- Encrypted Data c = (plain text) ^e mod n.

1.2. Decryption Using Private Key

- We know, $\Phi(n) = (P-1) * (Q-1)$
- Now calculate Private Key, $d = (k * \Phi(n) + 1) / e$ for some integer k
- Decrypted Data = c ^ d mod n.

Example:

- We shall consider trivial values so as to understand easily what is supposed to happen
- Let p=53 and q=59 => n=3127
- Take e=3 and k=2
- Φ(n)=3016
- We get d = (k * Φ(n) + 1) / e => d=2011
- Taking a=0, b=1……. z=25
- This means, h=8
- So, encrypting h will be: (8^3) mod 3016 which gives 512
- So, encrypting “hello” using the same encoding will give:
- h -> 512
- e -> 125
- l -> 1728
- l -> 1728
- o -> 248
- This will now lead onto Binary string Manipulation

2. Binary String Manipulation

With the Cipher Text obtained after RSA Encryption we can convert the digits (data) into Binary strings of equal length.

It is to be taken into careful consideration that the binary strings are of equal length so that the DNA Encryption can be successfully done over it so as to prevent any confusions and data loss at the time of Decryption.

At time of Decryption, the received data has to be first manipulated just in the reverse order of the above-mentioned steps to get to our digit Cipher Text, to finally get towards our original Plain Text.

Example

Plain Text: **hello**

RSA Cipher Digit Data: **512 125 1728 1728 248**

Binary Strings: **1000000000 1111101 11011000000 11011000000 11111000**

Extended Binary Strings (to accommodate all letters of the alphabet to equal lengths): **001000000000 000001111101 011011000000 011011000000 000011111000**

3. DNA Encoding

This is the part which mainly gives our Cipher Text and added advantage of extra security. The obtained Binary string after RSA Encryption shall be converted into DNA Sequences as per certain predefined values.**[1] [2]**

Example

If we have the Plain Text as “hi” and using the above RSA Encryption as well as Binary string Manipulation, we get the string as:

Plain Text: **hi**

Binary String: **001000000000001011011001**

DNA Sequence (as per Table 1): **ACAAAAACTGCG**

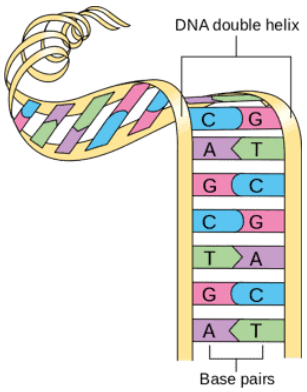

Figure 5: DNA Double Helix Structure and the Complementary pairs

4. Random OTP Generation

We can generate random OTP numbers using cryptographically secure pseudo random generators and such to make them more secure. Also, longer OTP numbers are obviously more secure.

Table 2:Sample Cipher Text for n=1000000016000000063

Alphabet	DNA Cipher
a	AAAAAAAAAAAAAAAAAAAAAAAAAG
b	AAAAAAAAAAAAAAAAAAAAACAAAAA
c	AAAAAAAAAAAAAAAAAACCTATTTCT
d	AAAAAAAAAAAAAAAGAAAAAAAAAAA
e	AAAAAAAAAAAAACTCCGAATCTGTG
f	AAAAAAAAAAAGGGCGTTTGCAAAAA
g	AAAAAAAAAAGTGGTGCTCGTACGGT
h	AAAAAAAAACAAAAAAAAAAAAAAAA
i	AAAAAAAAGTGATCGTGATTCAAGCG
j	AAAAAAAGGTGACAGTGCTCCAAAAA
k	AAAAAAGAACGCTGTCGCCGCGGGAT
l	AAAAAACCTATTTCTAAAAAAAAAAA
m	AAAAAGCCAGGAGGAGGGGCCTATGG
n	AAAAATCCTCTGTATCGACTCAAAAA
o	AAAAGTTGTGTCTGCCAAAGCGCCTT
p	AAAGAAAAAAAAAAAAAAAAAAAAAA
q	AAAGTTACCTCATAATGACAGTCTAG
r	AAATCCGTATCCGTTAAATACAAAAA
s	AAGCCGTTACCAAACAAAGGTGACCT
t	AACTCCGAATCTGTGAAAAAAAAAAA
u	AGATTCCGATGGGGTTTCAGTGAATG
v	ACAGATGCTTATGATACCCGCAAAAA
w	ATGCACCGATGATTACCACTATGAGT
x	GGGCGTTTGCAAAAAAAAAAAAAAAA
y	CAGTCAGCGTCAATACGCTCCCTACG
z	TGAACCACCACCCTGGCGCCCAAAAA

Results and Discussions

Assessing the Algorithm from a performance point of view, it does not take too long of a time to generate random OTPs of a certain predefined size, however, all the extra computation that takes place can prove to be a little less efficient as compared to simple RSA Cryptographic Encryptions and Decryptions. Another viewpoint is that we can potentially use this algorithm over any platform since the code used to implement a simple version of this was developed using Python 3 which has vast integration capabilities, so, any data size limitations can only be imposed by the platform and not the actual Algorithm. The data necessarily need not be in a character format to be manipulated. We can use digits directly to encrypt and encode but the uses do not even stop there. We can potentially encrypt any form of data we like since we already are using Binary strings to our use which is the form data travels in anyway.

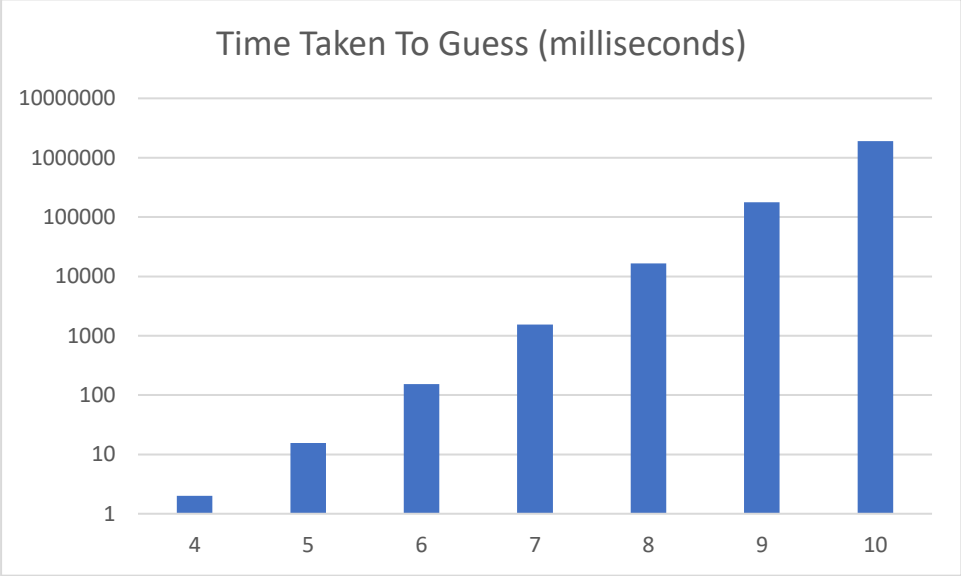

Figure 6: Time taken to Brute Force Guess OTPs of certain digits on a Logarithmic Scale

We can see from Fig 5. That longer OTPs take significantly longer to guess, approximately, we can see that the time to guess the key increases by an average of **9.981234702575582 (~ 9.98) times per 1 digit increase in size**.

One drawback could be that the extended cipher text could mean that we need to use some compression techniques to make the data bit more compact so as to be easily and safely transmittable.

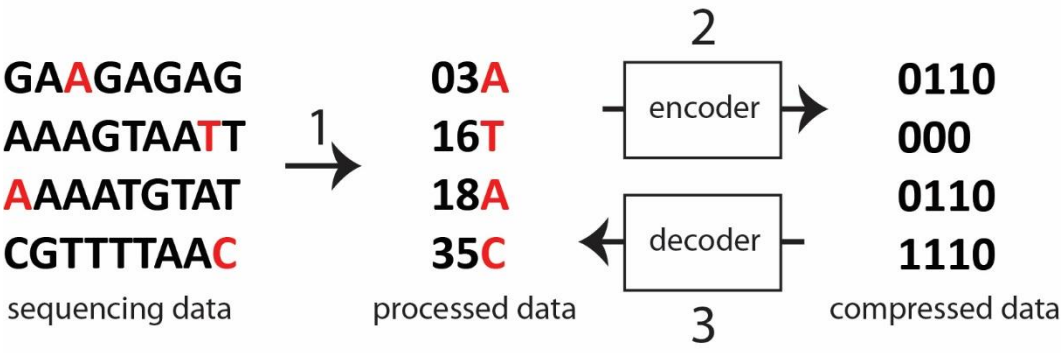

Figure 7: Compressing DNA Encoded data

Conclusion

In order to create a strong and impenetrable technique, the fundamental idea of this work was to use RSA encryption with DNA Encoding and random OTPs to authenticate the Receiver. The algorithm was discussed in detail with the help of Steps and Architecture Diagrams. The performance measurements pertaining to the general Brute Force Guessing of the Keys used for the underlying base of the RSA Encryption and Decryption was given and discussed. We also discussed the potential limitations of the Algorithm which was that it may consume more performance resources compared to regular Cryptographic Techniques but doing an ROI analysis will show that the extra security is worth it. Also, we have discussed the uses of this approach in many different areas as it just gives a way to encrypt and decrypt data without posing limitations to any file size or type. As was previously mentioned, the OTP gets increasingly more difficult to break, and with the additional security provided by the DNA Encoding, we may draw the conclusion that this algorithm offers significantly greater security than standard RSA cryptographic techniques.

References

[1] Poojashree Kamble,Firoz Nagarchi,Akshata Akkole,Vanishree Khanapur,Bahubali Akiwate, "A Dynamic DNA Cryptography Using RSA Algorithm and OTP", International Journal of Scientific Research in Computer Science and Engineering Vol.8, Issue.4, pp.78-82, August (2020).

[2] Mahbuba Begum, Jannatul Ferdush, Md. Golam Moazzam, "A Hybrid Cryptosystem using DNA, OTP and RSA," International Journal of Computer Applications (0975 – 8887), Volume 172 – No.8, August 2017. (As extra reference, not part of Literature Survey)

[3] Anjali Krishna A, Dr. L C Manikandan, "A Study on Cryptographic Techniques", International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Volume 6, Issue 4, pg. 321-327, July-August 2020.

[4] Gurdeep Singh, Prateek Kumar, "A Research paper on cryptography", International Journal for Technological Research in Engineering, Volume 7, Issue 4, December-2019.

[5] Hamza Hammami, Hanen Brahmi, Sadok Ben Yahia, "Secured Outsourcing Towards a Cloud Computing Environment Based on DNA Cryptography," IEEE, pp. 31-36, 2018. (As extra reference, not part of Literature Survey)

[6] Mrs. Ashwini Sheth, Mr. Sachin Bhosale & Mr. Farish Kurupkar, "Research Paper on Cyber Security", Contemporary Research in India, Special Issue: April, 2021.

[7] R.L. Rivest, A. Shamir, and L. Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems", Communications of the ACM (120-126), February 1978. (As extra reference, not part of Literature Survey)