DIGITAL FILE SECURITY

A PROJECT REPORT

Submitted by

Anshika Verma	(19BCY10030)
Soumya Singh	(19BCY10070)
Trisha Sinha	(19BCY10072)
Yukti Sharma	(19BCY10157)

in partial fulfillment for the award of the degree
of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING
(specialization in cyber security and digital forensics)

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING

VIT BHOPAL UNIVERSITY

KOTHRIKALAN, SEHORE
MADHYA PRADESH – 466114

OCTOBER, 2020

VIT BHOPAL UNIVERSITY,KOTHRIKALAN, SEHORE MADHYA PRADESH – 466114

BONAFIDE CERTIFICATE

Certified that this project report titled **“DIGITAL FILE SCURITY”** is the bonafide work of “**Anshika Verma (19BCY10030), Soumya Singh (19BCY10070), Trisha Sinha (19BCY10072), Yukti Sharma (19BCY10157)”** who carried out the project work under my supervision. Certified further that to the best of my knowledge the work reported here does not form part of any other project / research work on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

PROGRAM CHAIR
Dr. Shishir K. Shandilya,
Division Head - Cyber Security & Digital Forensics,
School of Computer Science and Engineering
VIT BHOPAL UNIVERSITY

PROJECT GUIDE
Prof. Muneeswaran V,
Assistant Professor,
School of Computer Science and Engineering
VIT BHOPAL UNIVERSITY

The Project Exhibition I Examination is held on October 28,2020.

ACKNOWLEDGEMENT

First and foremost we would like to thank the Lord Almighty for His presence and immense blessings throughout the project work.

We wish to express our heartfelt gratitude to **Dr Manas Kumar** Mishra head of the Department, School of Computer Science for much of his valuable support encouragement in carrying out this work.

We would like to thank our internal guide **Prof. Muneeswaran V**, for continually guiding and actively participating in our project, giving valuable suggestions to complete the project work.

We would like to thank all the technical and teaching staff of the School of Computer Science, who extended directly or indirectly all support.

Last, but not the least, we are deeply indebted to our parents who have been the greatest support while we worked day and night for the project to make it a success.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The project title mentioned “Digital File Security” focuses on providing security to files and images by encrypting it followed by 2-step authentication. As nowadays people need to store many important documents in their system, so they need to ensure their security.

There are several software/applications available which can encrypt and decrypt files. But, the main difference between what we did and what already exists is in our software we are working on original files and not creating any separate copies for encrypted and decrypted files, also checking the originality of file by comparing hash values of any file before encrypting and after decrypting it.

INTRODUCTION

1.1. INTRODUCTION

With an increasing involvement of digital world in humans lives the need of security has increased and people are concerned for the security of their files. This increased concern had lead to the development of various security applications for users to secure their personal images, official documents, etc. Today, there are around __ security apps available in the market but many of these applications while striving to provide security to the files end up risking the privacy of the users.
We, the students of VIT Bhopal present to you ‘Digital File Security’ which will ensure the security of your files without compromising your privacy. We aim to provide security to the files present on the user's desktop without asking for any personal information of the user such as email-id or phone number.

File Security software provides assurance to users as they consider a high-level security as an important parameter for accepting any security software. It has major significance in day-to-day operation of file systems. Files need to be secured thoroughly to enhance its capability to handle abnormal conditions. File security is performed by concentrating on points or situations where files may behave abnormally in case of compromising security and can result in failure later on.

The basic principle of securing a file is to ensure the user’s file security. File Security process is performed to incorporate security features i.e. authentication, confidentiality, integrity, availability etc. to the user. Security software helps files secure itself from unwanted actions and does not allow other entities/intruders to vanish the system’s integrity. The unwanted action may comprise the system's unauthorized access to suspiciously altered file(s). So, security software is an act of making a system defendable from attacks.

1.2. MOTIVATION FOR WORK

Security is a vital task or property. Providing security to a system is very complex in comparison to a simple software testing process which involves black box and white box testing. For securing a system, we need to check the system’s two important things: First, validity of implemented security measures that provide functionality and security to the system. Security measures also include features like cryptography, strong authentication, and access control measures. And second, the system’s behavior when it gets attacked by attackers, resulting in destruction by accessing secured and confidential information. File’s security failure may cause million-dollar business loss all over the world which is not generally acceptable to any organization at any cost. As a result, file securing process has been accrued a lot of importance and is given almost 40% time of the total time required in the SDLC process. As with any business, this effort is difficult without a “secured file” type of application.
Attackers can attack systems with their most powerful and exotic skill set to create room for himself in it. Developer and user need to understand the attacker’s mindset so that they can restrict the exploitable activities for the hacking system.

1.3. OVERVIEW OF THE PROJECT

Software comprises the entire set of programs, procedures, and routines associated with the operation of a computer system. The term was coined to differentiate these instructions from hardware—*i.e.,* the physical components of a computer system. A set of instructions that directs a computer's hardware to perform a task is called a program, or software program.

The project title mentioned 'Digital File Security' focuses on securing files, which includes images, pdfs and text documents, on the user's system through AES encryption. The software has a 2-step authentication service enabled, the first is the password and second, answer of the security question. Whenever a new user registers himself to use the software they are provided with 3 security questions; 1. 2. 3.

The user needs to select any one of these questions and then provide an answer to it. In case the user forgets his password, we have featured an option of 'forget password' but to ensure that the user is a genuine one the user must select the security question he chose while registering and then give the answer to that question. If both the question and answers are matched the user can create a new password.

1.4. SUMMARY

The overall objective of the project titled 'Digital File Security' is to secure the system files of the user from any kind of intrusion.

LITERATURE SURVEY

As per our title ‘digital file security’, our software helps to secure a user's file by encrypting the file that the user wants to encrypt and later decrypting it on user’s demand.

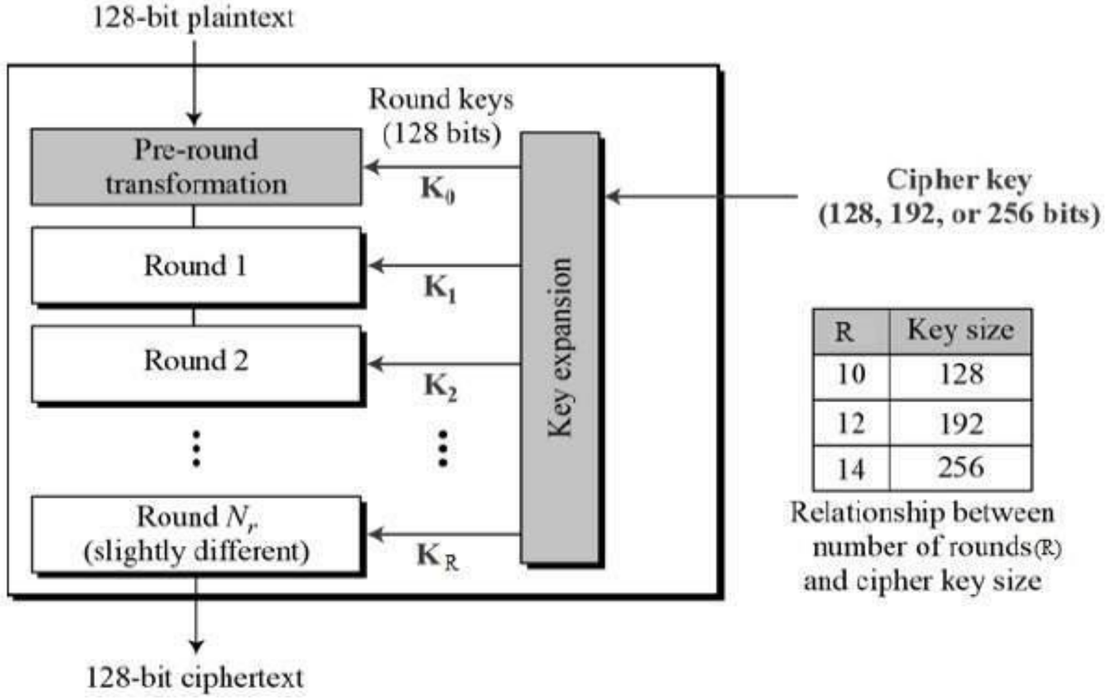
This software uses AES 256 encryption algorithm. AES means Advanced Encryption Standard . It is a symmetric block cipher chosen by the U.S government to protect classified information. when we say it is a symmetric block cipher, it means it uses the same secret key for encryption and decryption.

The National Institute of Standards and Technology (NIST) started development of AES in 1997 when it announced the need for an alternative to the Data Encryption Standard (DES), which was starting to become vulnerable to brute-force attacks.

NIST stated that the newer, advanced encryption algorithm would be unclassified and must be "capable of protecting sensitive government information well into the [21st] century." It was intended to be easy to implement in hardware and software, as well as in restricted environments -- such as a smart card -- and offer decent defenses against various attack techniques.
AES is based on ‘substitute- permutation network’. It comprises a series of linked operations, some of which involve replacing inputs by specific outputs (substitutions) and others involve shuffling bits around (permutations).

Interestingly, AES performs all its computations on bytes rather than bits. Hence, AES treats the 128 bits of a plaintext block as 16 bytes. These 16 bytes are arranged in four columns and four rows for processing as a matrix −

Unlike DES, the number of rounds in AES is variable and depends on the length of the key. AES uses 10 rounds for 128-bit keys, 12 rounds for 192-bit keys and 14 rounds for 256-bit keys. Each of these rounds uses a different 128-bit round key, which is calculated from the original AES key.



R	Key size
10	128
12	192
14	256

Relationship between number of rounds(R) and cipher key size

Figure 1

How does AES 256 work?

The basic principle of all encryption is that each unit of data is replaced by a different one according to the security key.

First, the initial key is added to the block using an XOR ("exclusive or") cipher, which is an operation built into processor hardware. Then each byte of data is substituted with another, following a predetermined table. Next, the rows of the 4x4 array are shifted: bytes in the second row are moved one space to the left, bytes in the third row are moved two spaces, and bytes in the fourth are moved three. The columns are then mixed—a mathematical operation combines the four bytes in each column. Finally, the round key is added to the block (much like the initial key was), and the process is repeated for each round. This yields ciphertext that is radically different from the plaintext. For AES decryption, the same process is carried out in reverse.

Each stage of the AES encryption algorithm serves an important function. Using a different key for each round provides a much more complex result.

How secure AES 256 is?

The National Institute of Standards and Technology selected three "flavors" of AES: 128-bit, 192-bit, and 256-bit. The difference lies in the length of the key. As the longest, the 256-bit key provides the strongest level of encryption. With a 256-bit key, a hacker would need to try 2256 different combinations to ensure the right one is included. This number is astronomically large, landing at 78 digits total.

The three AES varieties are also distinguished by the number of rounds of encryption. AES 128 uses 10 rounds, AES 192 uses 12 rounds, and AES 256 uses 14 rounds. The more rounds, the more complex the encryption, making AES 256 the most secure AES implementation. It should be noted that with a longer key and more rounds comes higher performance requirements. AES 256 uses 40% more system resources than AES 192, and is therefore best suited to high sensitivity environments where security is more important than speed.

PROJECT PROCEDURE

Nowadays, people are increasingly dependent on electronic data and computer networks to conduct their daily operations, growing pools of personal and financial information are being transferred and stored online as well as offline (on their PC and laptops). This can leave individuals exposed to privacy violations, and financial institutions and other businesses exposed to potentially enormous liability, if and when a data security breach occurs.

According to a survey:
Identity theft is currently a gold mine for cybercriminals—one that reached an all-time high in 2016, with up to $16 billion worth of losses caused by fraud and identity theft. Most people are already aware that theft can happen due to high visibility cases that occurred during the past couple of years, like the attack on Yahoo during latter half of 2016. While identity theft should be concerning in itself, the real, tangible damage usually comes after, when an attacker uses the stolen information for malicious purposes. Device loss or theft was also one of the primary cause of stolen information.

1. Stolen information can lead to following damages:
2. Financial information
3. Healthcare information
4. Payment card information
5. Credentials
6. Education information

Therefore, it has become necessary to store your data safely. But wherever you store your data, it should never be unencrypted. An access **password** to your computer is not sufficient protection to protect your confidential data from unauthorized access. If your computer is stolen or disposes of an old device, it is no problem for technical experts to read all your stored data. For this reason, it always makes sense to store all your data encrypted on your own computer.

Thus, our project focuses on how can we store files and images safely on users PC and laptop. So we have come up with an idea of making software which can encrypt and decrypt user's confidential files and images.

For accessing the software, the user needs to login first using a password and a security question answer. If both of them are correct then only the user will be able to encrypt and decrypt their documents. In case the user forgets his password, they can reset it as well.
For encrypting and decrypting the documents, we are using AES-256 (Advanced Encryption Standards) encryption algorithm. According to our research, AES-256 algorithm is the strongest algorithm for this purpose.

AES is a symmetric key cipher. This means the same secret key is used for both encryption and decryption. The advantage of symmetric systems like AES is their speed. Because a symmetric key algorithm requires less computational power than an asymmetric one, it's faster and more efficient to run.

AES is also characterized as a block cipher. In this type of cipher, the information to be encrypted (known as plaintext) is divided into sections called blocks. The basic principle of all encryption is that each unit of data is replaced by a different one according to the security key. More specifically, AES was designed as a substitution-permutation network. AES brings additional security because it uses a key expansion process in which the initial key is used to come up with a series of new keys called round keys. These round keys are generated over multiple rounds of modification, each of which makes it harder to break the encryption.

Each stage of the AES encryption algorithm serves an important function. Using a different key for each round provides a much more complex result. Byte substitution modifies the data in a nonlinear manner, obscuring the relationship between the original and encrypted content. Shifting the rows and mixing the columns diffuses the data, transposing bytes to further complicate the encryption. Shifting diffuses the data horizontally, while mixing does so vertically. The result is a tremendously sophisticated form of encryption. AES 256 uses 14 rounds. The more rounds, the more complex the encryption, and making AES 256 the most secure AES implementation. AES-256, which has a key length of 256 bits and 1.1 x 10^{77} possible key combinations, supports the largest bit size and is practically unbreakable by brute force based on current computing power.

Hash value of the file is stored before it gets encrypted. When the user decrypts the file, the program compares the hash value that is generated after the file has been decrypted with the one that is stored earlier of that particular file. If the hash values are same it will give the user a message that no amendments have been done with the data and the file integrity is secured. If a slightest change has also been done with the file it will give an error message, that file has been altered.

WORK DONE

4.1 Module 1:

The first window which would be visible to the user once the code runs is the welcome window. The user is displayed login and exit option.

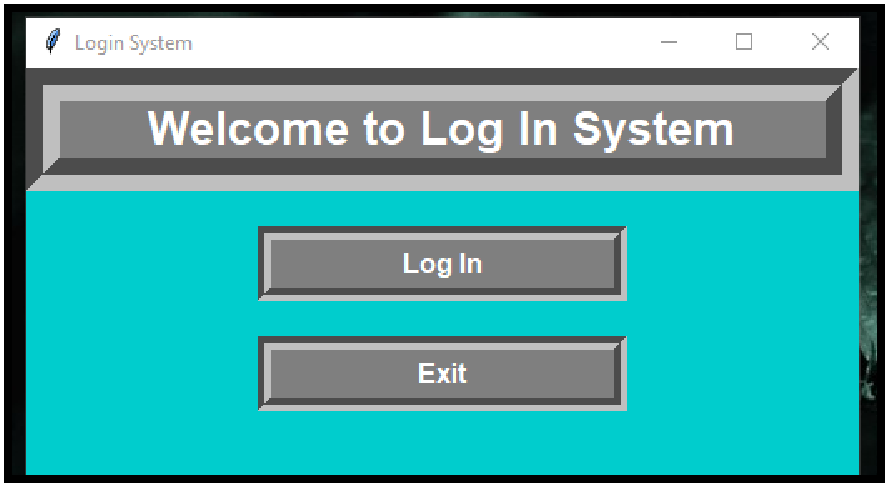
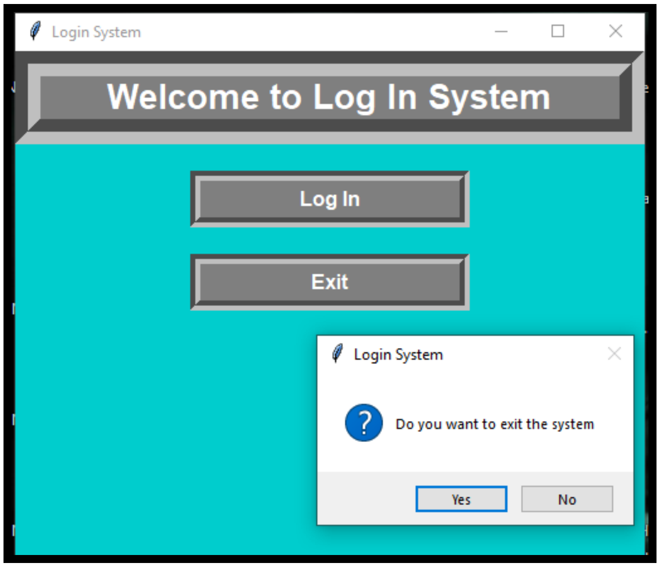


Figure 2



Do you want to exit the system

Figure 3

4.2 Module 2:

If the user clicks on ‘LOGIN’, a window appears which asks the user to enter the password and security question answer. If the both of them are correct, then user can move ahead. For password and security answer the database is accessed.

Database table Record

Table 1

Password	question	sans

In the above table there are three columns: password, question, sans . Password saves the password through which the user enters into the application. Question saves the security question of the user and sans saves the answer of the security ques.

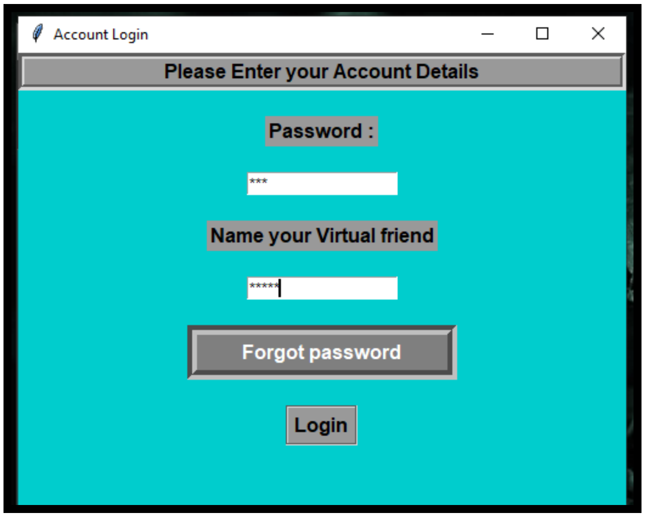
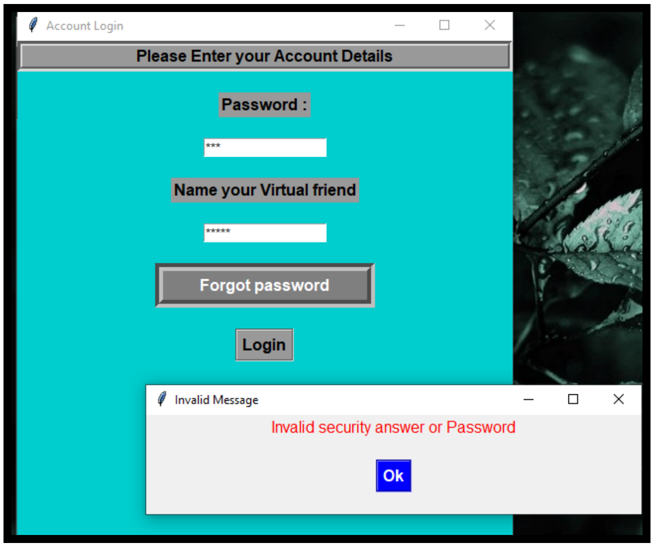


Figure 4

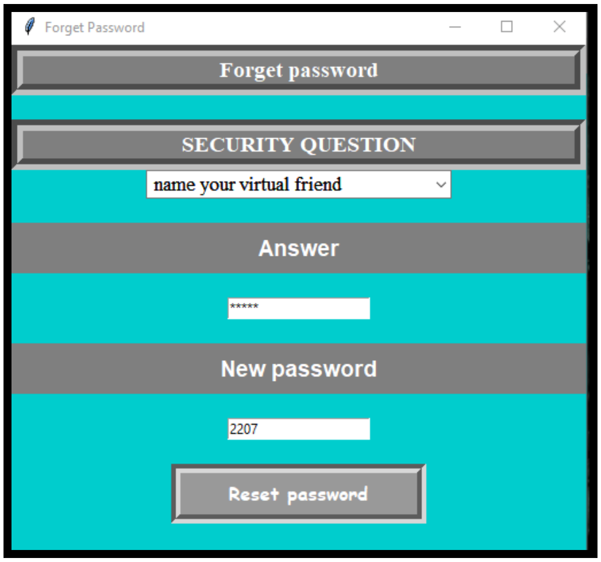
If the login password or security answer is incorrect the screen stating "Invalid us security answer or password".



Invalid security answer or Password

Figure 5

If the user forgets the password, a 'forget password' option is also given to the user. When the user clicks on it, a new window opens and the user is asked to select their security question and enter its corresponding answer. If the security question and its answer is correct, then the password will get reset.



SECURITY QUESTION: name your virtual friend

Answer: *****

New password: 2207

Figure 6

4.3 Module 3:

If login is successful, it shows a tab where encryption , decryption and exit option is available.

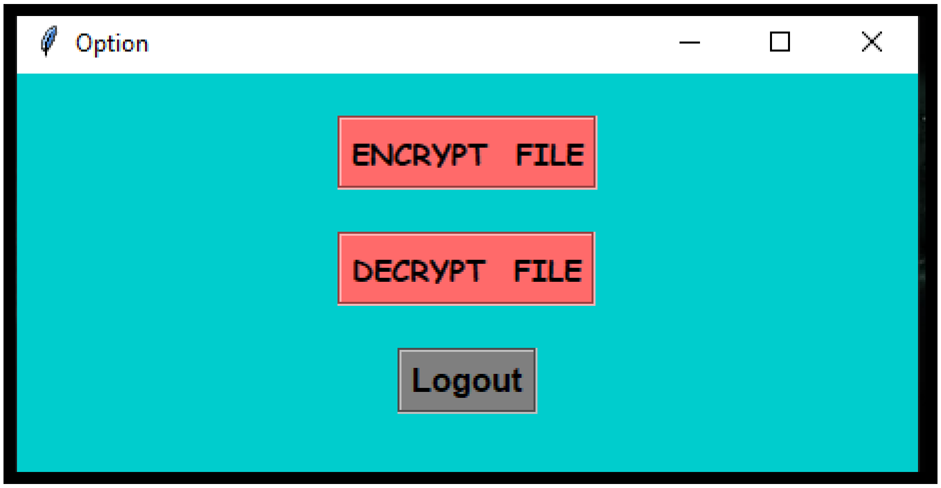
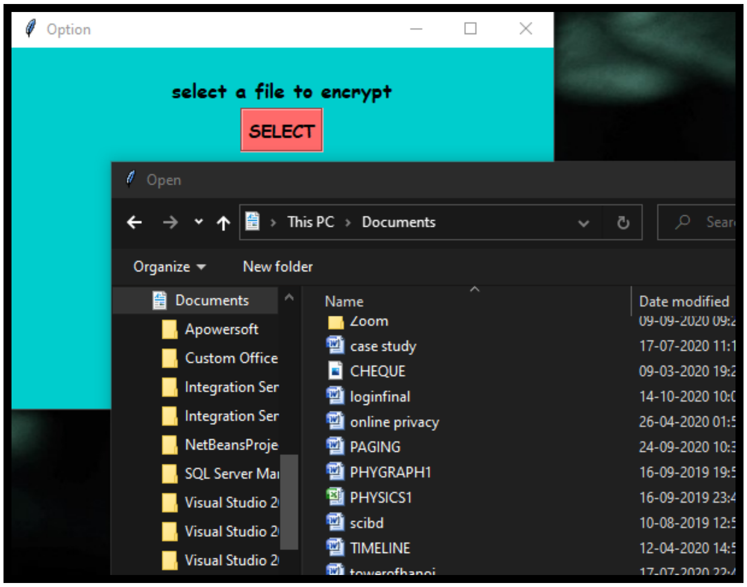


Figure 7

4.4 Module 4:

When ‘encrypt file’ is clicked, a file dialog is open from where user can select files or images that are present on their system. After the file/image is selected, it gets encrypted.



Zoom 09-09-2020 09:2
case study 17-07-2020 11:1
CHEQUE 09-03-2020 19:2
loginfinal 14-10-2020 10:(
online privacy 26-04-2020 01:5
PAGING 24-09-2020 10:3
PHYGRAPH1 16-09-2019 19:5
PHYSICS1 16-09-2019 23:4
scibd 10-08-2019 12:5
TIMELINE 12-04-2020 14:5
towerofhanoi 17-07-2020 22:4

Figure 8

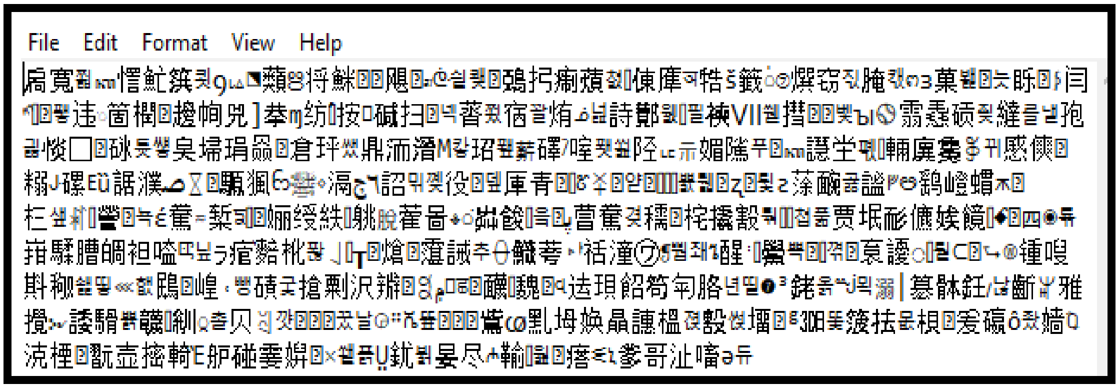
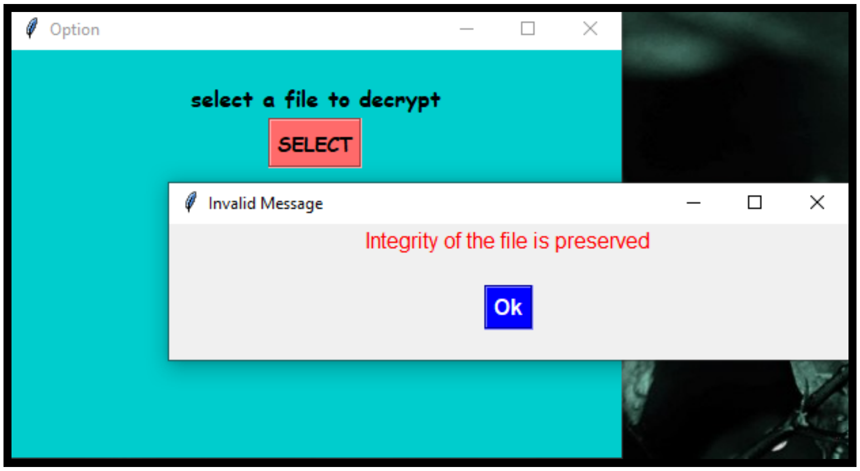


Figure 9

4.5 Module 5:

When ‘decrypt file’ is clicked, a file dialog is open from where user can select files or images that are present on their system. After the file/image is selected, it gets decrypted. When the user decrypts the file, the program compares the hash value that is generated after the file has been decrypted with the one that is stored before encryption of that particular file. If the hash values are same it will give the user a message that ‘file integrity is preserved’ that is no amendments have been done with the data.



Integrity of the file is preserved

Figure 10

PERFORMANCE ANALYSIS

5.1Observation:

- The file or image selected by the user is taken as input

OUTPUT:

- The file selected by user under encryption operation gets encrypted. The file does not open or if it gets open the text appears in encrypted form only i.e it will not be in readable format.
- The file selected by user under decryption operation, gets decrypted and a message appears regarding the integrity of the file. If the content of the files has not been altered it shows integrity of the file preserved otherwise an error occurs.

5.2 Performance measures:

Table 2

SIZE OF FILE (in KB)	TIME TAKEN FOR ENCRYPTION (in seconds)
1KB	1.00
11KB	1.23
50KB	2.40
100KB	7.10
200KB	8.70
500KB	13.30
1000KB	23.40

5.3 Performance analysis:

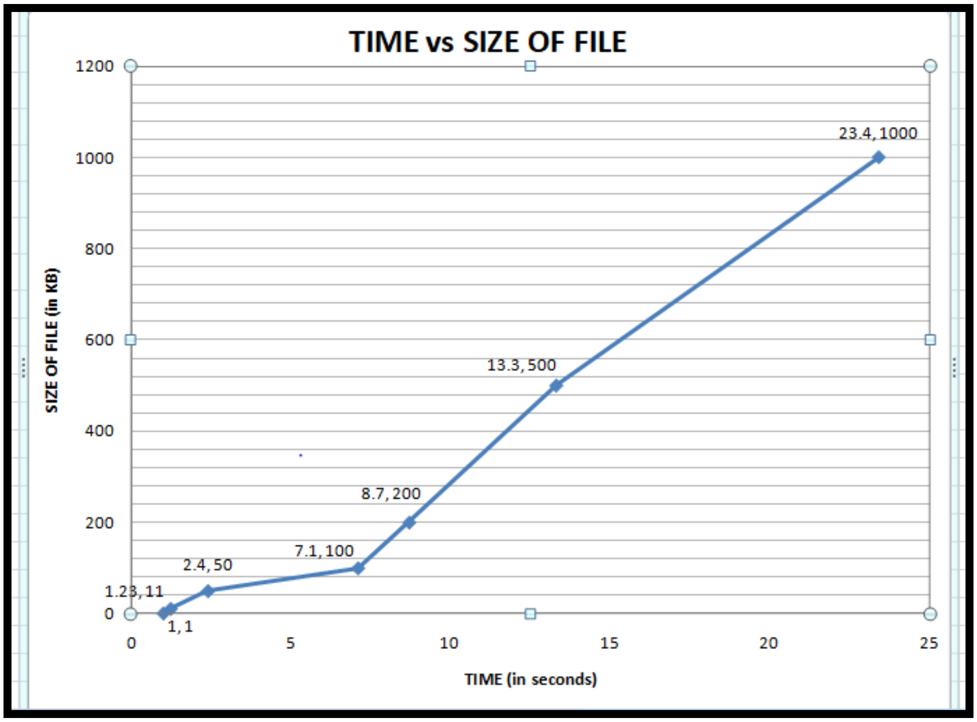


Figure 11

5.4 Summary:

As the size of file is increasing, there is a slight increase in the time taken for encryption.

RESULT AND CONCLUSION

The final result of the project is:

- User can safely store their file on the system by using the software.
- The code uses AES encryption method which is very powerful and nearly impossible to brute-force.
- File size is not affected after encryption and decryption.
- Users can check the integrity of their file with the help of the hash function which is enabled in the software.

RECCOMENDATION FOR FUTURE WORK

Since there is always a slight chance of improvement even in the most perfect work, so there are some flaws in the ‘Digital File Security’ on which we would like to work upon in near future. They are:

- The GUI of the software is not aesthetically designed, so we can work on this.

- Currently there is no provision of file recovery as the files are stored on the system itself, after gaining enough knowledge on the topic we will try introduce it in the software.

- For naive users ‘Digital file Security’ as an application would be easier and convenient to use so, we can work in the direction of moving from the software to application.

REFERENCES

1. www.youtube.com
2. www.geeksforgeeks.org
3. www.edureka.co
4. www.stackoverflow.com
5. www.pynative.com
6. www.coursera.org
7. www.tutorialspoint.com
8. www.w3schools.com