DATA SECURITY IN CLOUD COMPUTING ENVIRONMENT USING CRYPTOGRAPHY

A Project Report

Submitted for Minor Project 6th Semester

Bachelor of Technology

In

Computer Science and Engineering

Submitted By

GAURAV (1806146)

MRINAL (1806149)

DEVANSH JAIN(1806174)

Under the supervision of

Dr. Suyel Namasudra

Asst. Prof. CSE Department
NIT PATNA

Department of Computer Science and Engineering

NATIONAL INSTITUTE OF TECHNOLOGY PATNA

PATNA - 800005

JAN - JUN 2020

राष्ट्रीय प्रौद्योगिकी संस्थान पटना
NATIONAL INSTITUTE OF TECHNOLOGY PATNA

Certificate

This is to certify that **Devansh Jain(1806174), Gaurav Antil(1806146), Mrinal(1806149)** have carried out the project entitled **"DATA SECURITY IN CLOUD COMPUTING ENVIRONMENT USING CRYPTOGRAPHY"** as their 6th Semester Minor Project (6CS191) under the supervision of **Prof Dr. Suyel Namasudra.**

DR. SUYEL NAMASUDRA
ASSISTANT PROFESSOR
CSE DEPARTMENT
NIT PATNA

DR. J.P. SINGH
HEAD OF DEPARTMENT
CSE DEPARTMENT
NIT PATNA

राष्ट्रीय प्रौद्योगिकी संस्थान पटना
NATIONAL INSTITUTE OF TECHNOLOGY PATNA

Declaration

I hereby declare that this project work entitled **"DATA SECURITY IN THE CLOUD COMPUTING ENVIRONMENT USING CRYPTOGRAPHY"** has been carried out by me as the requirement for the industrial training, under the supervision of **Dr. Suyel Namasudra**, Assistant Professor, Department of Computer Science and Engineering, NIT Patna. No part of this project has been submitted for the award degree or diploma to any other Institute.

NAME	**SIGNATURE**
DEVANSH JAIN	----------
GAURAV ANTIL	----------
MRINAL	----------

PLACE	**DATE:**
NIT PATNA	----------

Contents of Dissertation

राष्ट्रीय प्रौद्योगिकी संस्थान पटना
NATIONAL INSTITUTE OF TECHNOLOGY PATNA

Acknowledgement

I hereby take the privilege to express my gratitude to all the people who were directly or indirectly involved in the execution of this work, without whom this project would not have been a success.

I extend my deep gratitude, respect and obligation to the project supervisor, **Dr. Suyel Namasudra**, Asst Professor, Department of Computer Science and Engineering.

Devansh Jain 1806174

Mrinal 1806149

Gaurav 1806146

ABSTRACT

Cloud computing has been envisioned as the next-generation architecture of IT enterprise. In contrast to traditional solutions, where the IT services are under proper physical, logical and personnel controls, cloud computing moves the application software and databases to the large data centers, where the management of the data and services may not be fully trustworthy. And as we know the data production and storing of data in cloud environments has been increasing exponentially.

Huge amount of data is shared through devices. Therefore to avoid leakage of data it is necessary to secure it. Hence, Cryptography is the solution to it. It assures data security. The Cryptosystem consists of a key-pair generator(KPG), sender and receiver, where the key will encrypt and decrypt the data. AES Algorithm is used to encrypt and decrypt the shared data. RSA Algorithm is being used to encrypt the key when shared with the receiver to encrypt the shared data. Using this architecture would give a lot of protection to the data which is being shared against various security attacks.

Our results, as well as theoretical analyses, show the efficiency and effectiveness of the proposed scheme over existing schemes.

1. INTRODUCTION

Computers are progressively advancing day by day. Earlier we used to have room like big storage devices to store data. Now, we have such compact devices to store as much data as possible. Cloud computing is such an advanced on-demand service which stores vast amounts of data with comparatively faster computational speed without direct active involvement of the user. This term is used to describe data centres available on the internet to a large number of users.

Data Security has been the biggest concern since ages and still an ongoing issue. Cryptography is worldwide considered as the most efficient and prominent solution to ensure data security till date. The main motive of cryptography is to share data from the sender to the receiver such that the attacker is unable to decode and read the actual original data. Now, Cryptography is

classified into two types and that is s**ymmetric key encryption** and **asymmetric key encryption**.

In **symmetric key encryption** also known as **secret key encryption,** the same key is used to encrypt and decrypt data. If a secure algorithm is used, then symmetric key encryption can be really secure and also relatively fast. Although if a secure key-exchange mechanism is not established then there might be loss of secret information. **Figure 1** shows the symmetric key encryption process.

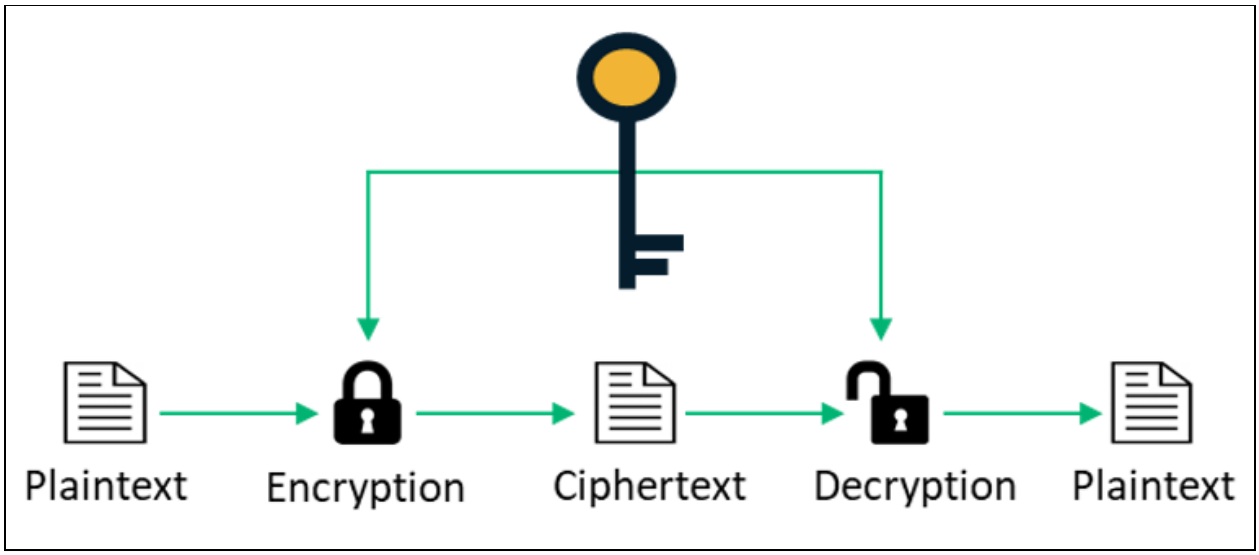

FIGURE 1 : Symmetric key Encryption Process

In **Asymmetric key encryption** also known as **public key encryption,** two different keys are used for encryption and decryption of data. The two keys are known as public key and private key. The private key is kept a secret whereas the public key is made available to all. Increased security is the prime advantage of asymmetric key encryption. **Figure 2** shows the asymmetric key encryption process.

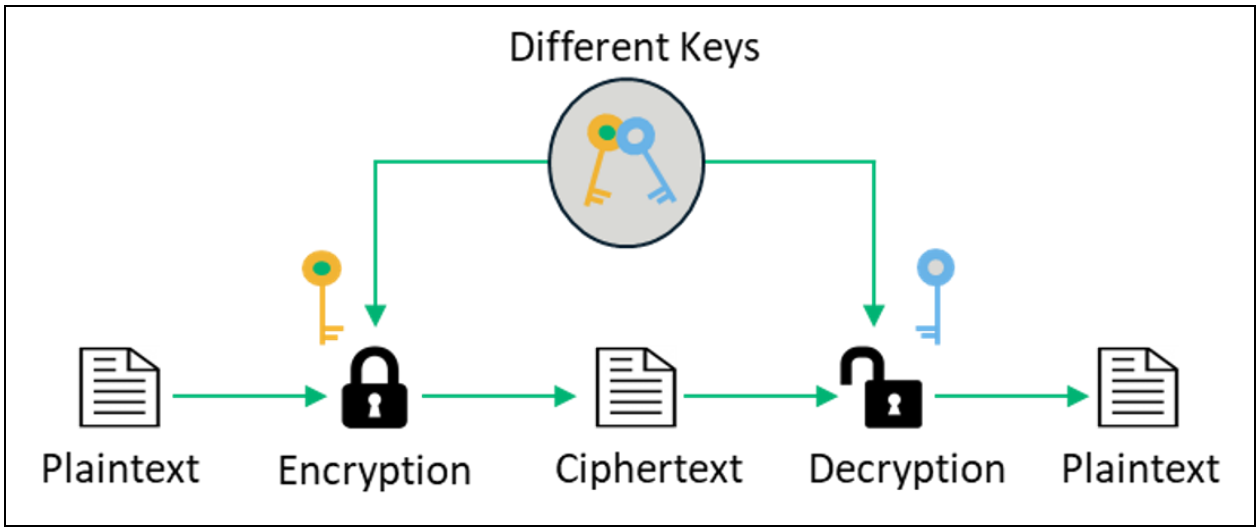

FIGURE 2 : Asymmetric key Encryption Process

1.1 WHAT HAVE WE APPLIED IN OUR PROJECT?

Our project consists of designing a new approach that combines Cloud Computing and cryptography. We have applied the Advanced Encryption Standard (AES) symmetric encryption algorithm, also known by its original name Rijndael, to encrypt and decrypt the data text file and Rivest–Shamir–Adleman (RSA) asymmetric key encryption algorithm to exchange the keys. The architecture of our model is shown in **Figure 4** below. This model secures the data and protects it against attacks.

1.2 WHAT IS RSA?

RSA Algorithm is an asymmetric cryptography algorithm used extensively for secure data transmission. It works on two different keys as discussed earlier i.e. private key and public key.

Steps to work on RSA Algorithm which is used in key generation:

Step 1: Generate the RSA modulus

The initial procedure begins with selection of two prime numbers namely p and q, and then calculating their product N, as shown−

N=p*q Here, let N be the specified large number.

Step 2: Derived Number (e)

Consider number e as a derived number which should be greater than 1 and less than (p-1) and (q-1). The primary condition will be that there should be no common factor of (p-1) and (q-1) except 1.

Step 3: Public key

The specified pair of numbers n and e forms the RSA public key and it is made public.

Step 4: Private Key

Private Key d is calculated from the numbers p, q and e. The mathematical relationship between the numbers is as follows −

ed = 1 mod (p-1) (q-1)

Encryption Formula: C = (P^e) mod n

Decryption Formula: Plaintext = (C^d) mod n

PROS AND CONS OF RSA ALGORITHM:

PROS:

RSA algorithm is a complex and super secure cryptography algorithm. It has its fair share of advantages. Firstly, the RSA algorithm is safe and secure for its users through the use of complex mathematics. Secondly, the RSA algorithm is hard to crack since it involves factorization of prime numbers which are difficult to factorize. And moreover, the RSA algorithm uses the public key to encrypt data and the key is known to everyone, therefore, it is easy to share the public key.

CONS:

RSA algorithm is a complex and super secure cryptography algorithm. But it has its fair share of disadvantages. Firstly, the RSA algorithm can be very slow in cases where large data needs to be encrypted by the same computer. Secondly, it requires a third party to verify the reliability of public keys. And finally, data transferred through the RSA algorithm could be compromised through middlemen who might temper with the public key system.

1.3 WHAT IS AES?

AES stands for Advanced Encryption Standard and is a majorly used symmetric encryption algorithm. It is mainly used for encryption and protection of electronic data. It was used as the replacement of DES(Data encryption standard) as it is much faster and better than DES. AES consists of three block ciphers and these ciphers are used to provide encryption of data.

PROS AND CONS OF AES ALGORITHM:

PROS:

AES algorithm is a complex and super secure cryptography algorithm. It has its fair share of advantages. Firstly, it can be implemented on both hardware and software, as it is super flexible. Secondly, it provides high security to the users as the algorithm is highly complex and secure. Thirdly, it provides one of the best open source solutions for encryption. And lastly, it is a very robust algorithm.

CONS:

RSA algorithm is a complex and super secure cryptography algorithm. It has its fair share of disadvantages. Firstly, it requires many rounds for encryption. Secondly, it is hard to implement software. Thirdly, it needs much processing at different stages. Lastly, it is difficult to implement when performance has to be considered.

1.4 WHAT ARE THE ADVANTAGES?

The advantages of Data security in cloud computing environments using cryptography are adverse and are discussed in this section. Firstly, it reduces the huge capital cost of buying hardware and software, which is a huge advantage as it makes it affordable and economical for users. Secondly, resources can be accessed very quickly as it is fast and efficient. Thirdly, productivity is high as we have to put less operational effort. Also, backup and recovery of data are less expensive and very fast. Also, we can increase or decrease the requirement of resources according to our budget or needs. It also improves redundancy and disaster recovery. And also, it prevents or curtails viruses and malware infection.

1.5 WHAT ARE THE DISADVANTAGES?

The disadvantages of Data security in cloud computing environments using cryptography exist and are discussed in this section. Firstly, you can’t do anything with the cloud as long as your internet access is out. Secondly, from an angle security is also an issue with cloud computing. Last but not the least, there is a possibility that your cloud service provider runs out of money and closes their doors forever.

1.6 WHAT ARE ITS APPLICATIONS?

We tried to develop an application based on the hybrid architecture (Architecture above) in order to encrypt the data passing through the Cloud. This hybrid model secures data and protects it against attacks. It uses, on the one hand, the AES symmetric encryption algorithm in order to benefit from its advantages in terms of robustness and processing speed, and on the other hand, the asymmetric RSA encryption algorithm to exchange the keys.
The applications of Data security in cloud computing environments using cryptography are discussed in this section. It is used in **Authentication and identity**. Maintaining confidentiality, integrity and availability for data security is a function of the correct application and configuration of familiar network, system, and application security mechanisms at various levels in the cloud infrastructure. Authentication of users takes several forms, but all are based on a combination of authentication factors: something an individual knows(such as a password), something they possess (such as security token), or some measurable quality that is intrinsic to them(such as a fingerprint). It has been applied to a few places in the real world. Some of them are mentioned in this section. Firstly, Art applications like Moo, Adobe creative cloud , Vistaprint have been users of data security via cryptography. Secondly, Business applications like Salesforce , PayPal etc have also adopted data security using cryptography in cloud computing environments . Thirdly, Data storage and backup applications use it. Lastly, Educational Applications, Entertainment applications like online games, Management applications where data security is a huge concern have adopted cryptography for their data security .

1.7 WHAT WAS OUR MOTIVATION?

The motivation for selecting our topic data security in cloud computing environment using cryptography are discussed below in this section.
Cloud computing helps startups manage shifting computing requirements by providing flexibility in computer service they purchase. Also, Cloud has a strong impact on company budgets, it improves use of resources reducing IT spending by 20 and 80%. And, It allows for the possibilities to generate new income streams via the implementation of disruptive business models. Lastly, It allows a greater flexibility in the deployment of services, which permits a “granular” focus in completing or replacing capacities.

1.8 WHY CLOUD?

The problem with the local database system is that it is prone to physical damage. In such cases there is a high possibility of data loss. This could affect the whole program running on the database. So, we need the cloud database system: Data could be synced to any device. Multiple users could access the data. It provides security against data loss as data is stored at various data

centers, so it is easy to recover lost data. Users could share the data by inviting them to their server. Now, as the multiple users could access the cloud data, it needs to be designed in such a way that only desired users could access the given data. For the same purpose we could encrypt the data and share the key with the receivers who want to access the same data by sender.

1.9 AGENTS INVOLVED

There are various agents involved in the process of data sharing:

- Server: A cloud server which stores all the information regarding the keys. During the time of data sharing, it helps to validate the user using the key pair generated. If the user is valid, a secure connection is established for the data transfer.
- Data owner: Data owners or service providers, are those who are the owner of the data. They store either the normal data or the entire database and depend on the server to manage data.
- Client: Users are the authorized parties who could have access to the data if the server establishes connection between them.

2. LITERATURE REVIEWS

Literature review is the most important step in the development process. Before developing an application it is necessary to determine the time factor, economy and company strength. Once these things are satisfied, then the next step is to determine which operating system and language can be used for developing the application. Once the programmers start building the application the programmers need a lot of external support. This support can be obtained from senior programmers, from books or from websites. Before building the system the above consideration is taken into account for developing the proposed system.

Also a few algorithms that we studied before going forward with the algorithm used in our project are as follows:

1. **ELGAMAL ALGORITHM**: Elgamal encryption is a public-key cryptosystem. It uses asymmetric key encryption for communicating between two parties and encrypting the

message. This cryptosystem is based on the difficulty of finding discrete logarithms in a cyclic group that is even if we know ga and gk, it is extremely difficult to compute gak.

.

2. **KNAPSACK ALGORITHM**: Knapsack Encryption Algorithm is the first general public key cryptography algorithm. It was developed by Ralph Merkle and Martin Hellman in 1978. As it is a Public key cryptography, it needs two different keys. One is the Public key which is used for the Encryption process and the other one is the Private key which is used for the Decryption process. In this algorithm we will have two different knapsack problems in which one is easy and the other one is hard. The easy knapsack is used as the private key and the hard knapsack is used as the public key. The easy knapsack is used to derive the hard knapsack.
   For the easy knapsack, we will choose a Super Increasing knapsack problem. Super increasing knapsack is a sequence in which every next term is greater than the sum of all preceding terms.

OUR PROJECT:

Our project, inspired from asymmetric cryptography, uses RSA algorithm for key encryption-decryption and AES algorithm for message encryption-decryption . Ours' is a window based application using Java swing as frontend. And all the details on RSA and AES algorithms have already been discussed in the above sections.

2.1 SHORTCOMINGS OF THE LITERATURE REVIEWS

The shortcomings of the literature reviews are discussed in this section. Firstly, it often fails to provide details of the overall strategy. Secondly, it often lacks details on how the analysis was conducted.Thirdly, it is often very time consuming. Fourthly, literature reviews require good supervision. Lastly, Google search doesn't show relevant information/results.

3. BACKGROUND OF THE PROPOSED SCHEME

The background of the proposed scheme consists of two sections:
i) System Module
ii) Design Goals

3.1 SYSTEM MODULE

The system module consists of three entities:

- **USERS**: Users, who have data to be stored in the cloud and rely on the cloud for data computation, consist of both individual consumers and organizations.
- **CLOUD SERVICE PROVIDER**: A CSP, who has significant resources and expertise in building and managing distributed cloud storage servers, owns and operates live Cloud Computing systems.
- **DATA OWNERS**: They store user's confidential data as well as big data on the database of the cloud environment and depend on the CSP to manage the data.

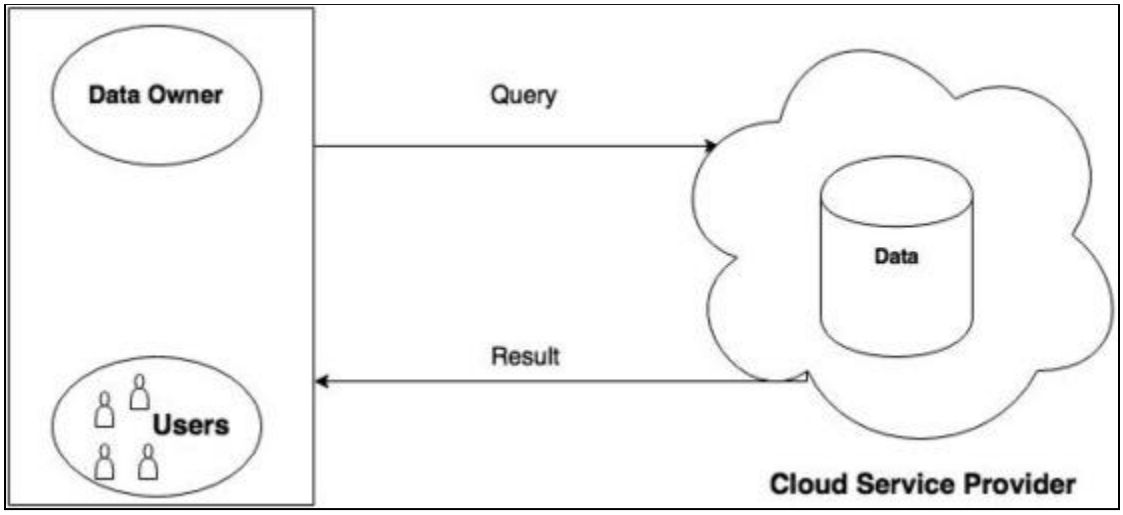

FIGURE 3: SYSTEM MODULE OF DATA SECURITY

3.2 DESIGN GOALS

The main design goals of this project are as follows:

1. To achieve an efficient and scalable data storage scheme cloud environment that provides a strong security to the users' confidential big data. That is why we have used a local cloud storage system for our project.

2. To provide the data access by generating the key using the RSA algorithm.

3. To encrypt the data file using the AES algorithm, of course with the help of a generated key.

4. PROPOSED SCHEME

In this report, we propose an effective and flexible distributed scheme with dynamic data support to ensure the correctness of users' data in the cloud. In other words, the main purpose of the proposed scheme is to provide a strong security of big data for data sharing by using cryptography. The three entities in the proposed system are as follows:

1) **Key-Pair Generator**: This entity provides the public and private keys to the sender and receiver during the registration process.

2) **Sender**: This entity stores confidential or normal data.

3) **Receiver**: This entity sends requests to the sender for accessing any required data.

4.1 SYSTEM SETUP

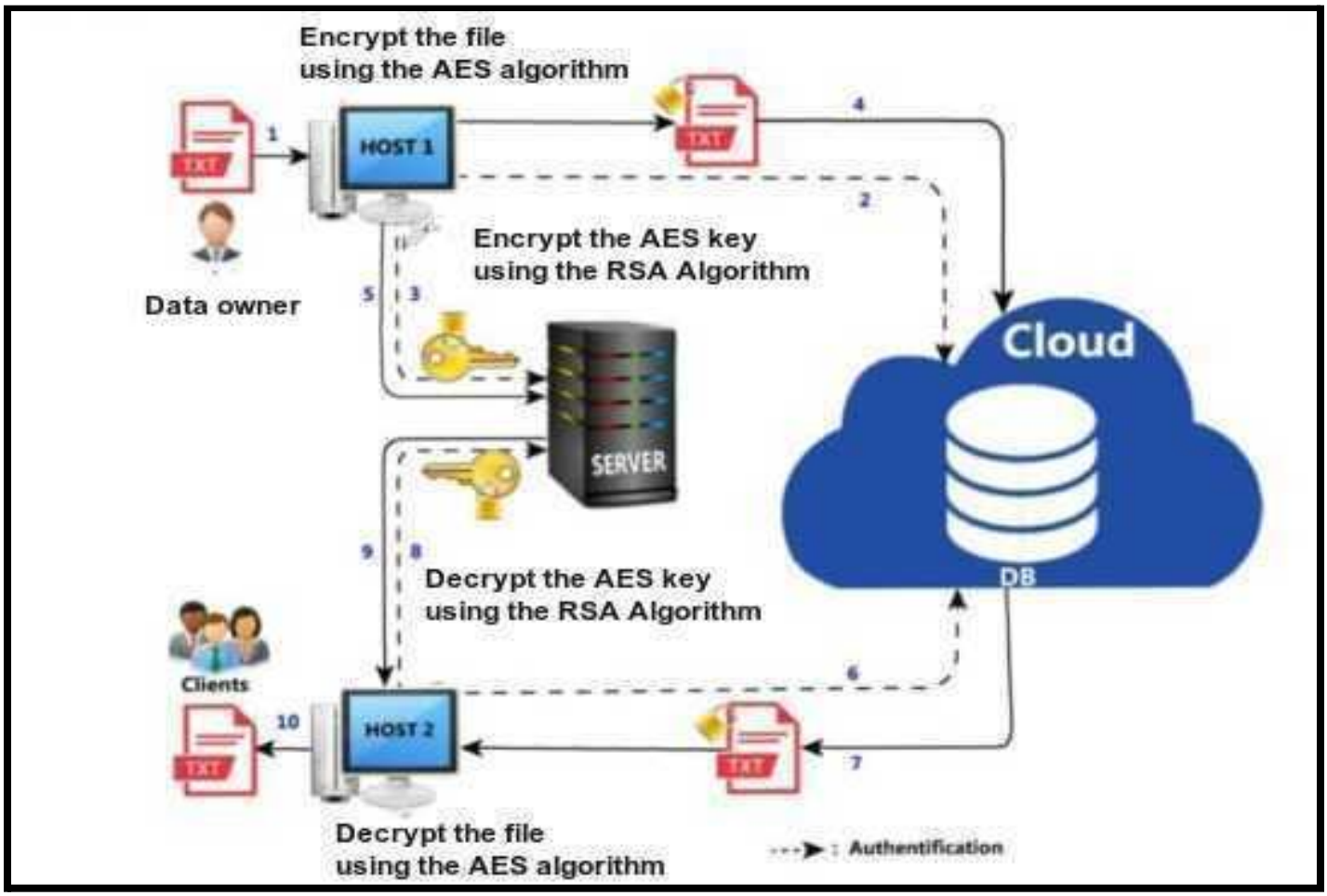

FIGURE 4: WORKFLOW OF THE PROPOSED SCHEME

4.2 DETAILS OF OUR APPLICATION

Our java windows application consists of four windows namely :

1. Authentication window
2. Home Window
3. Cloud window
4. Security window
5. About window

4.2.1 AUTHENTICATION WINDOW

Authentication window is the first window the user comes through and has to pass it.
Authentication window consists of two simple input zones.
One reserved for the user name, the other for the password on which the characters entered are hidden.
Also there is a button that allows access to the main interface if the password and the login are correct.
If the password is incorrect then an error message will be displayed on a dialog box.
Authentication window is shown in **figure 5**.

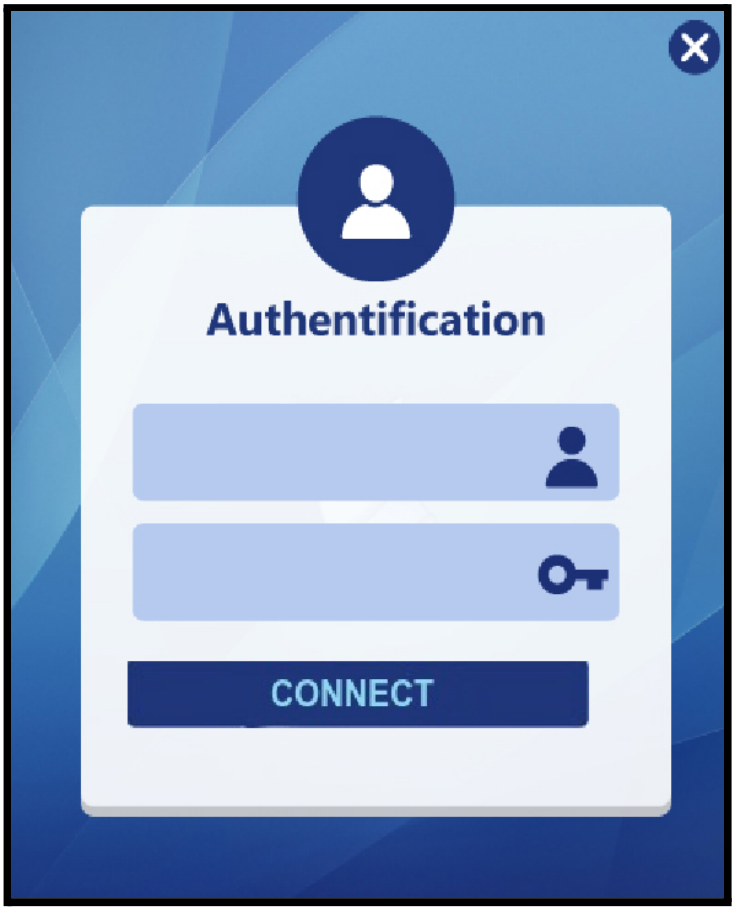

FIGURE 5: AUTHENTICATION WINDOW

4.2.2 HOME WINDOW

The home window here, is the window that the user would interact with as soon as he logins to the application. It provides the convenience to the user to navigate through various other features of the application that is Cloud Window, Security Window and the About Window and also come back to its own window that is Home Window.

FIGURE 4.2: HOME WINDOW

4.2.3 CLOUD WINDOW

In this window the customer can import a file (Word, Text) from his own computer. In this case where the customer plays the sender role, or from the cloud in the case where the client plays the recipient role. In both cases, the customer must import the key or enter it in the key area, to encrypt in the case of an issue. After the encryption process, the file should be exported to the cloud.
The cloud window is shown in **figure 6**.

FIGURE 6: CLOUD WINDOW

4.2.4 SECURITY WINDOW

The Security window is reserved for establishing a connection between the users of the application, in this case the server listens to a connection request.
After the clients are successfully connected, the transmitter encrypts the key and sends it to the server. The AES encrypted key is encrypted using RSA algorithm and sent to the receiver by the sender and then as soon as the receiver gets the message that is the encrypted key, the user can decrypt the key by clicking the button decrypt and save it in a file by exporting it.
The security window is shown down below in **figure 7**.

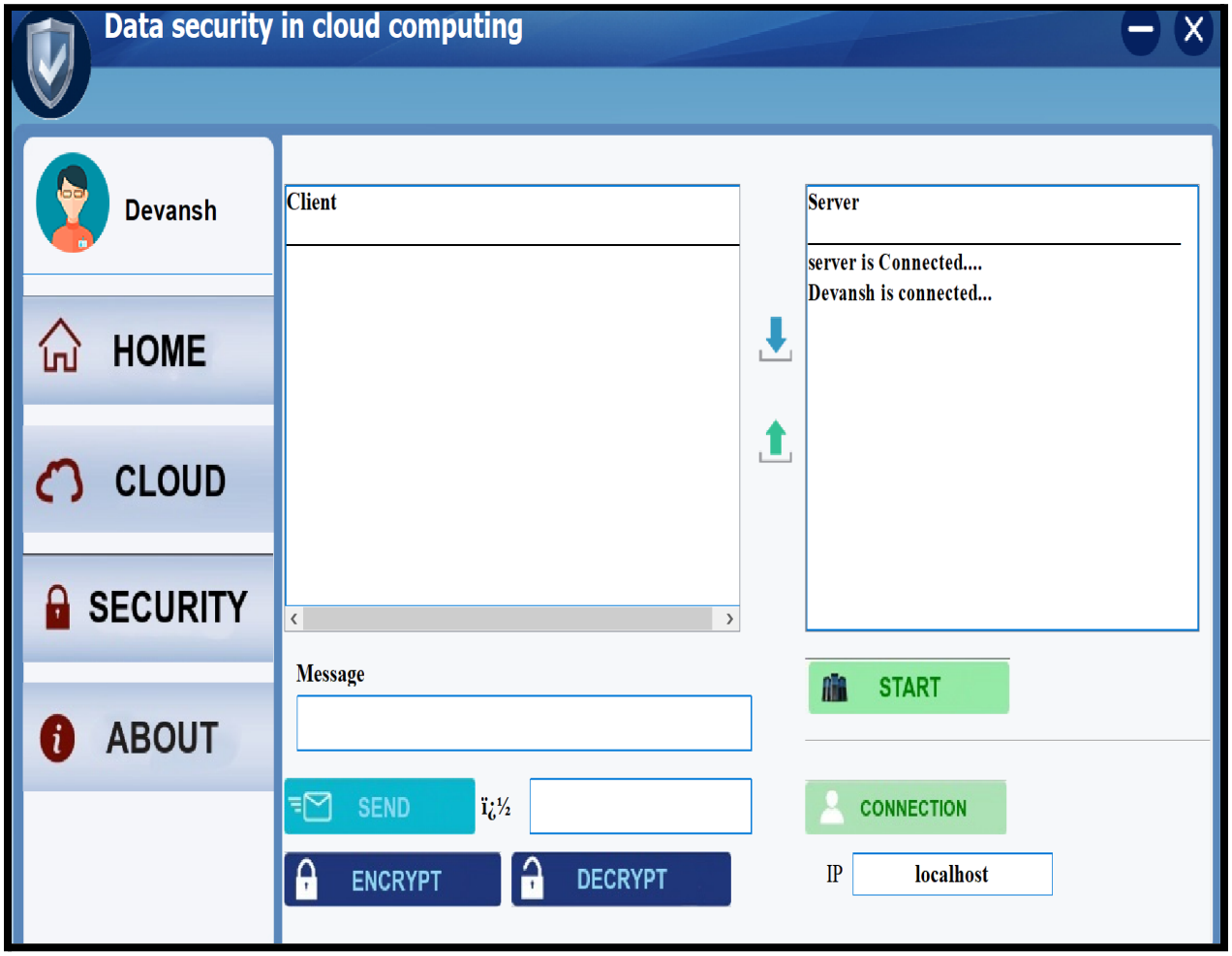

FIGURE 7: SECURITY WINDOW

4.2.5 ABOUT WINDOW

The About window is literally the about us page where we have given information such as the project name(Data security in cloud computing) and our mentor(Dr. Suyel Namasudra Sir) and the members working on the project i.e. Devansh Jain(1806174), Mrinal(1806149), Gaurav Antil(1806146).
The About Window is shown in **figure 8**.

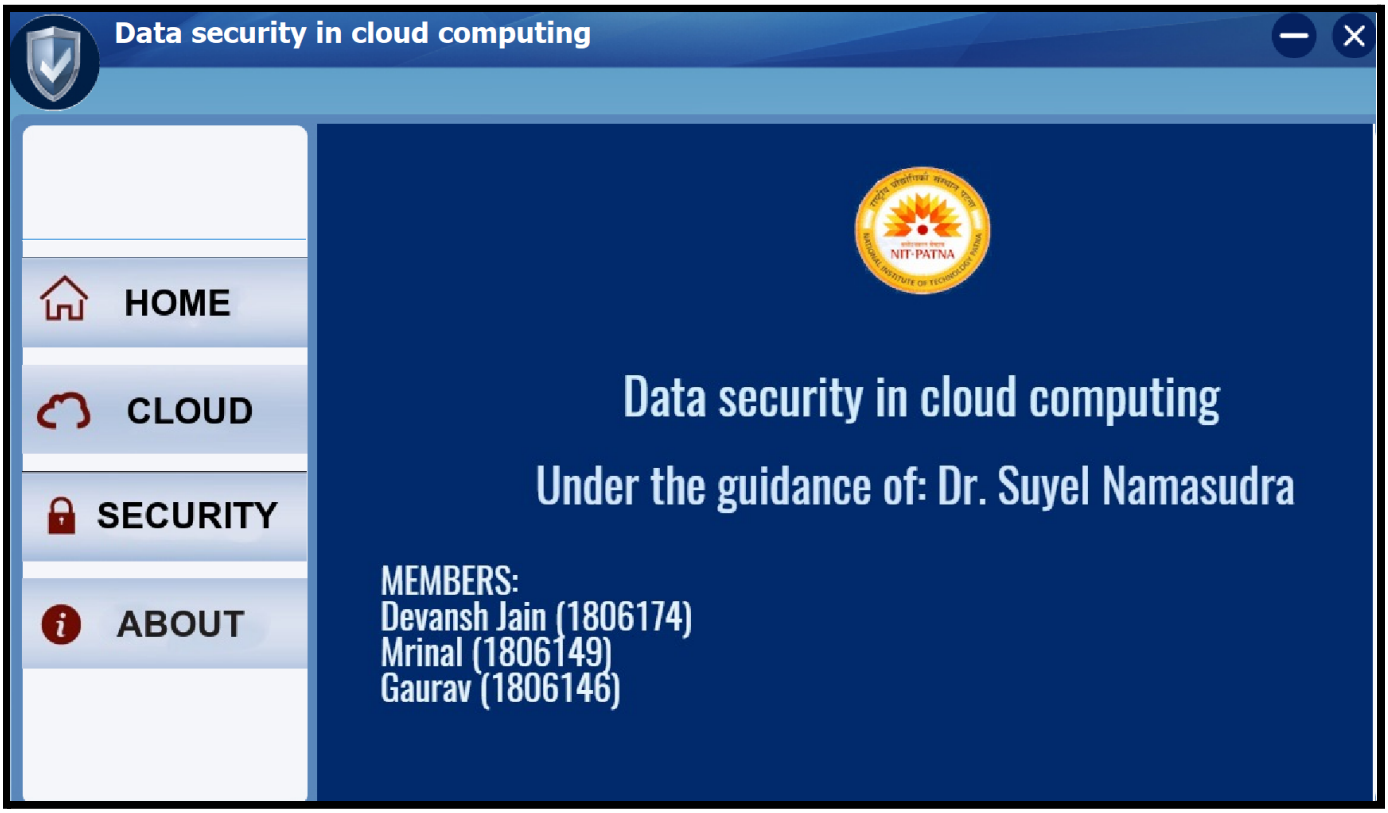

FIGURE 8: ABOUT WINDOW

5. SECURITY ANALYSIS

The security analysis of data security has been presented in this section.

1. **Plain text attacks**: It is classified into 3 subcategories:-
    - **Short message attack**: In this we assume that the attacker knows some blocks of plain text and tries to decode cipher text with the help of that. So, to prevent this pad the plain text before encrypting.
    - **Cycling attack**: In this attacker will think that plain text is converted into cipher text using permutation and he will apply right for conversion. But the attacker does not write plain text. Hence will keep doing it.
    - **Unconcealed Message attack**: Sometimes it happens that plain text is the same as cipher text after encryption . So it must be checked and it cannot be attacked.

2. **Chosen cipher attack**: In this attacker is able to find out plain text based on cipher text using Extended Euclidean Algorithm.

3. **Factorisation attack**: If an attacker will be able to know P and Q using N, then he could find out the value of the private key.This can be failed when N contains at least 300 longer digits in decimal terms, which the attacker will not be able to find. Hence it fails.

4. **Attacks on Encryption key**: If we take smaller value of E in RSA this may occuR so to avoid this take value of E = 2^16+1 (atleast).

5. **Attacks on Decryption key**:
    - **Revealed decryption exponent attack**: If attacker somehow guesses decryption key D, not only the cipher text generated by encryption the plain text with corresponding encryption key is in danger, but even future messages are also in danger. So, it is advised to take fresh values of two prime numbers (i.e; P and Q), N and E.
    - **Low decryption exponent attack:** If we take a smaller value of D in RSA this may occur so to avoid this take value of D = 2^16+1(atleast).

6. EXPERIMENTAL SETUP

The required setup for the development of the projectare:

1. **TOOLS:** Apache NetBeans
2. **LANGUAGE**: JAVA
3. **LIBRARY**: Swing, POI
4. **HARDWARE: The project was run on system with following specifications**
    - PROCESSOR: Intel i5-8th gen @ 2.3GHz
    - RAM: 8GB
    - STORAGE SPACE:128GB SSD

7. RESULTS AND DISCUSSION

Successful encryption and decryption of both key and message is achieved by decrypting the message and key successfully with their respective algorithms.
Server is starting successfully without any bugs at least for the test cases we have used in our application.
Users were connecting successfully and sending encrypted keys to each other over the local cloud. Users are successfully receiving the key and successfully decrypting the encrypted file. More complex and efficient user defined algorithms can be developed and used in the project.

8. CONCLUSION

In this project, we investigated the problem of data security in a cloud computing environment using cryptography.
Cloud computing here is achieved on a local system, using different java libraries and APIs. This project is an example for people who still have trust issues on cloud computing for their privacy as this can be extended to a higher level. Cloud computing is the future as we don't have to do anything but buy the services, which is a huge advantage. Through detailed security analysis we prove that our proposed scheme is highly efficient and resilient to Plain Text Attacks, Chosen cipher attacks, Factorisation Attacks, etc.

We believe that data storage security in Cloud Computing, an area full of challenges and of paramount importance, is still in its infancy now, and many research problems are yet to be identified. We envision several possible directions for future research on this area.

9. FUTURE WORK

As we are well aware of the use of DNA computing and its power to store a large amount of data consuming a very small piece. Using biological molecules to perform calculations rather than silicon chips, it could revolutionize the way we store data. So, our future works include:

- Incorporating DNA based data security.
- Implementing a dedicated cloud space rather than a local cloud.
- Working and developing a better algorithm for data security, which would be more secure to implement.
- Automating the cloud operations using AWS cloud automation.

10. REFERENCE

The reference were taken from:

- Towards DNA based data security in the cloud computing environment:
  Suyel Namasudra , Debashree Devi , Seifedine Kadry , Revathi Sundarasekar , A. Shanthini
- Data Security and Privacy in Cloud Computing :
  Yunchuan Sun, Junsheng Zhang, Yongping Xiong
- Data Security in Cloud Computing:
  Ahmed Albugmi University of Southampton Madini O. Alassafi King Abdulaziz University, University of Southampton, Robert John Walters University of Southampton Gary Wills University of Southampton
- http://web.mit.edu/6.933/www/Fall2000/aes/AES_final_6933.pdf
- https://docs.oracle.com/javase/7/docs/api/javax/crypto/Cipher.html
- https://docs.oracle.com/javase/8/javafx/api/javafx/application/Application.html
- http://java.sun.com
- http://www.sourcefordgde.com
- http://www.networkcomputing.com/
- http://www.java2s.com/
