Enhancing Confidentiality and Integrity in Cloud Computing using RSA Encryption Standard and MD5 Hashing Algorithm

Katende Nicholas
University of Kigali (UOK)
PhD candidate in Information Technology

Cheruiyot Wilson
Professor
Jomo Kenyatta University of Agriculture of Agriculture and Technology, Kenya(JKUAT),

Ann Muthoni Kibe, PhD
Jomo Kenyatta University of Agriculture and Technology, Kenya(JKUAT)

ABSTRACT
Cloud computing has revolutionized how services are rendered and used by some many people in the world like providing hardware, software and infrastructural storage to many users at any time. This is in terms of software as a service, platform as a service and infrastructure as a service, hence providing room for convinience to the cloud consumers to choose what they want presently and catering for their future needs since its elastic. With the company's or individual's data held by a third party that is the cloud provider, it brings out the security issues in respose to confidentiality, availability and integrity of the data at the cloud provider's side. In this reseach paper solution is provided to maintain confidentiality of data and integrity of data at the cloud provider's side. This framework contains RSA encryption standard and MD5 hashing alogorithm. In this solution data is encrypted using RSA which generates both public and private keys used in the encryption and the decryption then using MD5 to generate the hash value which is stored before the data is sent to the cloud provider. The hash value is checked upon retriving of data from the cloud and if its still the same then the data was not modified or tempered with if else then the cloud provider has bleeched the contract. All these approaches undergo through the following steps Encryption, Hashing, Data uploading on a cloud, Verification and Decryption.

Keywords
Cloud Computing,; RSA, MD5

1. INTRODUCTION
Cloud computing has merged out as a popular field in networking. It is gaining popularity in all areas. The National Institute of Standards and Technology (NIST)[1] defines cloud computing as a model for enabling ubiquitous ,convenient, on demand network access to a shared pool of configurable computing resources. In other words cloud computing is a computing model in which resources are provided to the cloud consumers on the basis of their current and future needs as they so wish. In cloud computing resources are provided by the cloud service provider known as CSP. Many software companies like Google, Microsoft, Yahoo, Amazon, Salesforce, etc. is providing cloud services on different parameter.

By using the services of cloud service provider user's transfer their burden of installing the software, Data maintenance, Infrastructure, Storage space etc. on the cloud service provider. Theses providers offer to their clients the possibility to store, retrieve and share data with other users in a transparent way [2].As data is shared among various users in the cloud there may be possibilities that data may be lost or misuse by other users. It is the biggest matter of concern in cloud computing. Many people afraid to share their data on cloud as they are not aware of the people who will be accessing their data and for what purpose.

In this paper a new approach is created in which first, the data is encrypted by RSA algorithm and then the hash value is generated using MD5 before the data is sent to the cloud servers. In this approach each client can generate their public and private key. In which public key is known to all and private key is only known to the client and authorized users.

2. LITERATURE REVIEW
Priyanka Ora and Dr.P.R.Pal [4], proposed a solution to maintain data security and data integrity. The scheme contained a combination of RSA Partial homomorphic and MD5 hashing algorithm. The data was encrypted by RSA Partial before uploading it on cloud server. After uploading its hash value was calculated by MD5 hashing scheme. All these approaches under went through the following steps Encryption/Decryption, Data uploading on a cloud, Hashing and Verification

Sudhansu Ranjan Lenka and Biswaranjan Nayak [5], proposed a solution where the architecture provides a mechanism through which we can get secure communication as well as hiding the information from unauthorized users. The model was implemented using a combination of RSA encryption and digital signature technique which can easily with all types of cloud computing features like: PaaS, SaaS and IaaS. That combination mechanism provided three way security i.e. data security, authentication and verification. They proposed RSA encryption algorithm for confidentiality of data and for authentication MD 5 algorithm have been implemented.

NesrineKaaniche et al[2], has proposed ID based cryptography in which the data is firstly encrypted and stored on the public cloud server. This concept also offers access control so that only authorized users can use the data. With the help of this approach unauthorized user even not get the data without client permission.

NehaTirthani et al[3]explain about cloud security issues and then proposed a security model for cloud in which Diffie Hellman Key Exchange and Elliptical Curve Cryptography algorithms are used. The whole model is described in four steps in which first step establish connection, the second is account creation, third is authentication and last step contain data exchange.

2.1 RSA algorithm

The RSA algorithm named after Ron Rivest, Adi Shamir, and Leonard Adleman. It is based on a property of positive integers. RSA uses modular exponential for encryption and decryption. RSA is an algorithm for public-key cryptography, involves a public key and a private key. The public key can be known to everyone and is used for encrypting messages. Messages encrypted with the public key can only be decrypted using the private key. The process is shown in figure below.

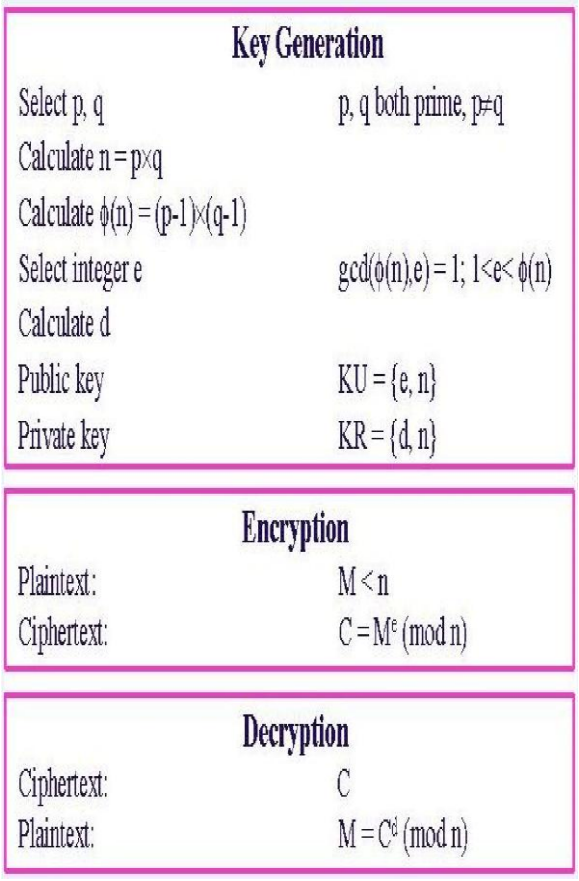

Figure. 1. RSA Algorithm

RSA uses two exponents, e and d, where e is public and d is private. Let the plaintext is M and C is cipher text, then at encryption

$C = M^e \bmod n$

And at decryption side

$M = C^d \bmod n$.

Where n is a very large number, created during key generation process.

2.2 MD5 HASHING ALGORITHM

MD5 is a message digest algorithm developed by Ron Rivest. MD5 is quite fast and produces 128-bit message digests. After some initial processing, the input text is processed in 512-bit blocks (which are further divided into 16 32-bit blocks). The output of the algorithm is a set of four 32-bit blocks which make up the 128-bit message digest (Priyanka and Vivek 2014). It contains various steps which includes Padding, append length, Divide the input into 512-bit blocks, initialize chaining variables, and Process blocks.

One MD5 operation

- A process P is first performed on b, c and d. This process is different in all the four rounds.
- The variable a is added to the output of the process P (i.e. to the register abcd).
- The message sub-block M[i] is added to the output of step2(i.e. to the register abcd).
- The constant t[k] is added to the output of step 3 (i.e. to the register abcd).
- The output of step 4 (i.e. the contents of register abcd) is circular-left shifted by s bits.(The value of s keep changing)
- The variable b is added to the output of step 5 (i.e. to the register abcd).
- The output of step 6 becomes the new abcd for the next step (Priyanka and Vivek 2014).

All the above steps are shown in the following figure 2 i.e. it shows the process of one MD5 operation. As it contains the process P, variables a, b, c, d, message sub-block M[i], constant t[k] for the operation of MD5 algorithm.

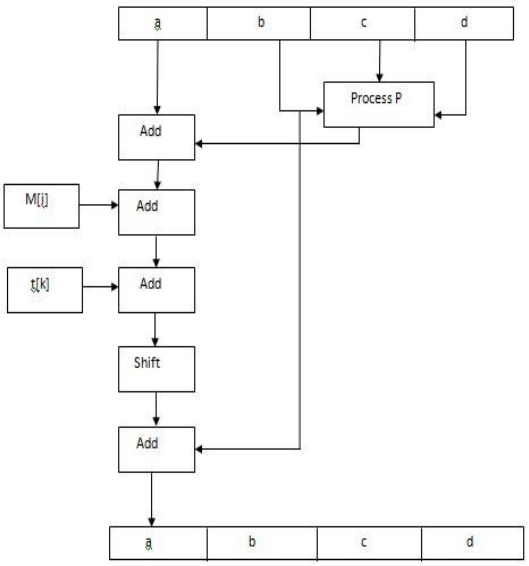

Figure 2: One MD5 operation Source: Shreya and Neeraj, 2015

We can mathematically express a single MD5 operation as follows:

a = b + ((a + Process P (b, c, d) + M[i] + t[k]) <<< s) where <<< s = Circular left shift by s bits Understanding the process P

As we can see, the most crucial aspect here is to understand the process P, as it is different in the four rounds. In simple terms process P is nothing but some Boolean operations on b, c and d as shown in the following table (Priyanka and Vivek 2014).

Table 1.0: Process P in each round

Round	Process P
1	(b AND c) OR ((NOT b) AND (d))
2	(b AND d) OR (c AND (NOT d))
3	b XOR c XOR d
4	c XOR (b OR (NOT d))

3. PROPOSED MODEL

In this model two major entities Data owner (cloud consumer) and cloud service provider (CSP) are the main entity. Cloud consumers are the person who uploads their data and CSP are cloud service provider who provide a cloud environment for the whole model .The steps of the proposed model are as follows:

A. Proposed Work

In the proposed work for keeping the data secure first, the data is encrypted with the RSA algorithm after that public and private key related with file is generated. As encryptions are performed, the encrypted file is subjected to the MD5 hash algorithm to generate the hash value which is saved in form of a number. After uploading Data owner gets its general detail like uploading time, date, hash generation of file and verification. On requesting the same file from the cloud, the hash value is checked to confirm with the already stored value and if it was accessed once with on modifications done the value will be the same. If the file submitted to the cloud has been tempered with then a prompt message will be generated automatically with an error message.

B. Prerequisites

Before uploading the file on cloud some basic steps have to be done which is encrypted of the specified file and generation of public and private key. In this model for encryption , RSA algorithm is used. For encryption user has to only upload their file and then generate its public key. After a generation of public key its related private key is also generated which is used for future decryption. Now the encrypted file is further subjected to hashing before uploading it to the cloud. The encrypted and hashed file is the one uploaded to the cloud

C. Secure Data Storage

As Data owner upload their file on cloud they get general detail of files like uploading date, time , etc. Some other steps are also performed which are as follows:

D. Data *Backup*

As data or file is uploaded on cloud it is necessary to keep this data secure for future usage so that the data cannot be damaged or lost. For keeping the data backup strong the following steps should be performed:

E. Hashing

For keeping the data secure on the cloud hash value of the uploaded file is calculated before uploading the file. This hashing is performed by Cloud Consumer with the help of the MD5 hashing algorithm. A copy of calculating hash value is backup by the data owner which is further used for verification purpose. Hence on Cloud Service Provider only encrypted and hashed data is present, which is beneficial for security purpose because the cloud service provider does not use it for malicious purpose.

F. Verification

Verification is an important domain of this model. In this part data present on cloud is verified by data owner just to check data integrity. This task is performed at cloud consumer end. As described in the previous section after calculation of the hash value of the encrypted data its value is backed up by the data owner. In future data owner can verify their data by retriving the file from the cloud and verify the hash value. This calculated hash value is matches with the old hash value which is present at owner end. If this value match's then data present at the cloud is safe and no modification has been done if it does not match then there are some changes on cloud data.

4. IMPLEMENTATION

The whole model is implemented using MYSQL server to depict the cloud server. For coding JAVA version and the jdk-8u161-nb-8_2-windows-x64 version is used. Some important snapshots are as follows.

A. Encryption

In this form the first process of the model , i.e. generation of public key and private key through encryption and decryption of file is done. The file is encrypted using RSA and MD5 algorithm, after a security code is inserted which will be used in the verification process later.

Plain text file from the cloud consumer

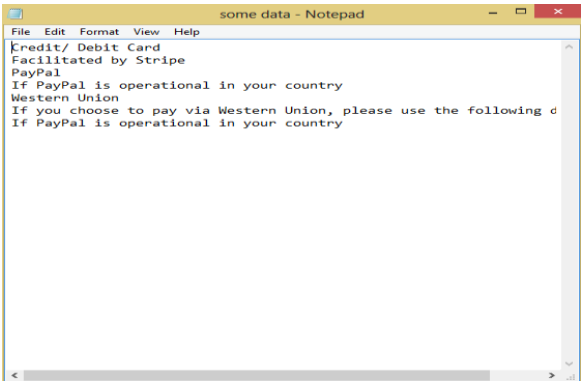

Figure 3. plain text

Encryption for both RSA and MD5 algorithm takes place

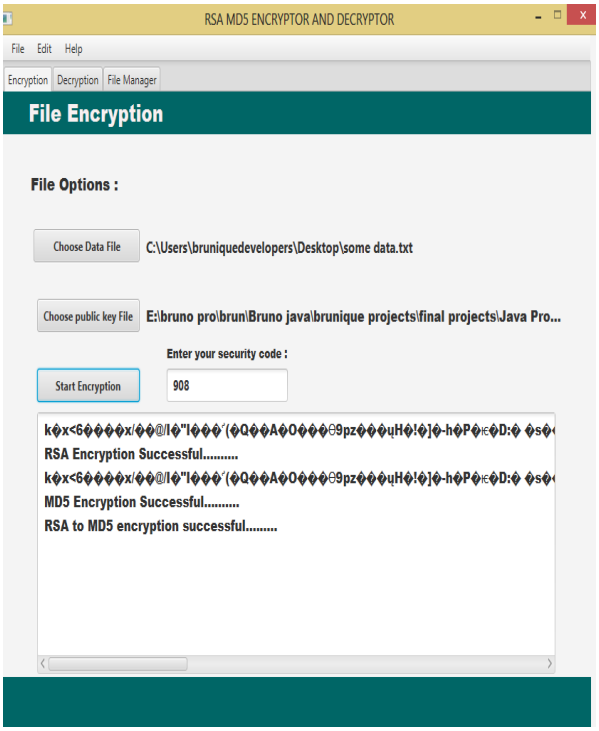

Figure 4: Encryption

B. Data Uploaded On Cloud Server

After encryption Owner simply upload encrypted files cloud. After uploading owner get general details like file ID, name, uploaded date, and the content of the encrypted.

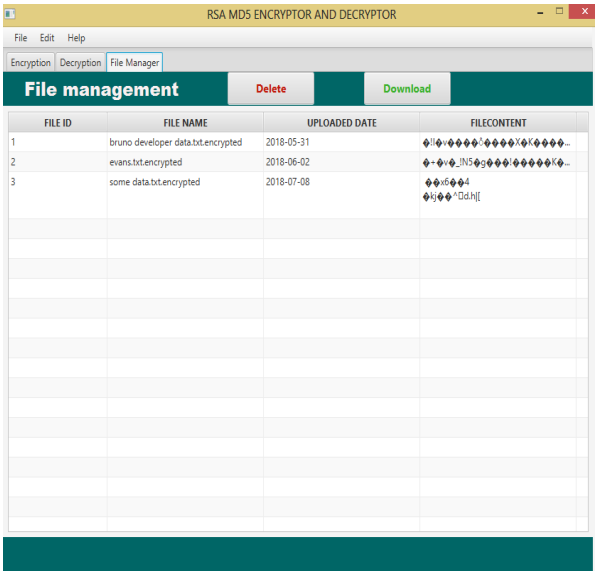

Figure 5: uploaded file

C. Verification

Encrypted files can be downloaded from the cloud and the verification process is done by confirming the security code that was used to before submitting the original files to the cloud, this is done before data is decrypted.

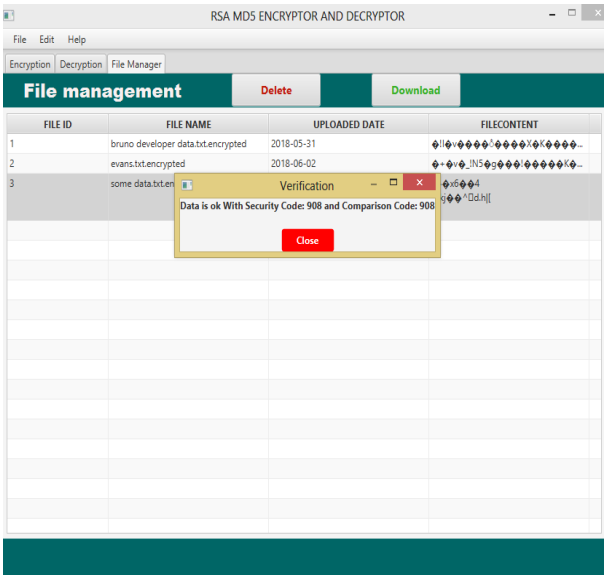

Figure 6: verification

D. Decryption

After the encrypted files have been downloaded and verified, then the decryption process starts with the reverse of the encryption.

Copy of the encrypted file

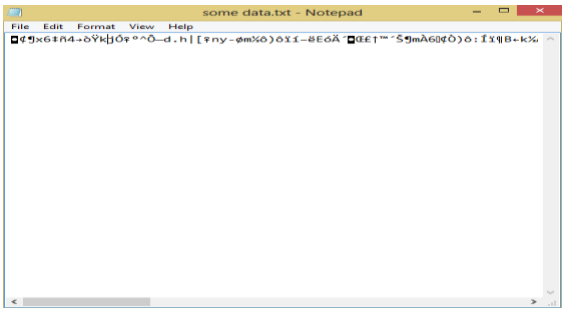

Figure 7 Decryption

Encryption process done successfully

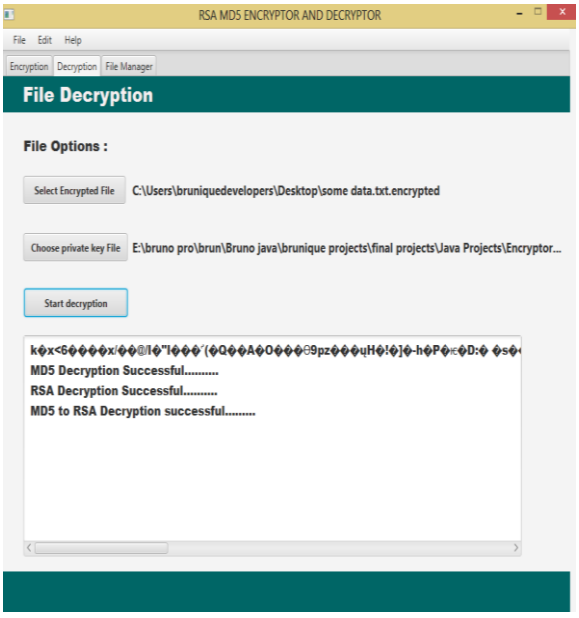

Figure 8: Confirmation

E. Modified file

Incase the encrypted file was modified by any one on the cloud service providers side then the message showing that the file was modified will be generated and the file can not be saved again

Modified encrypted file

Figure 9: modified file

Figure showing the error message due to modification of the file

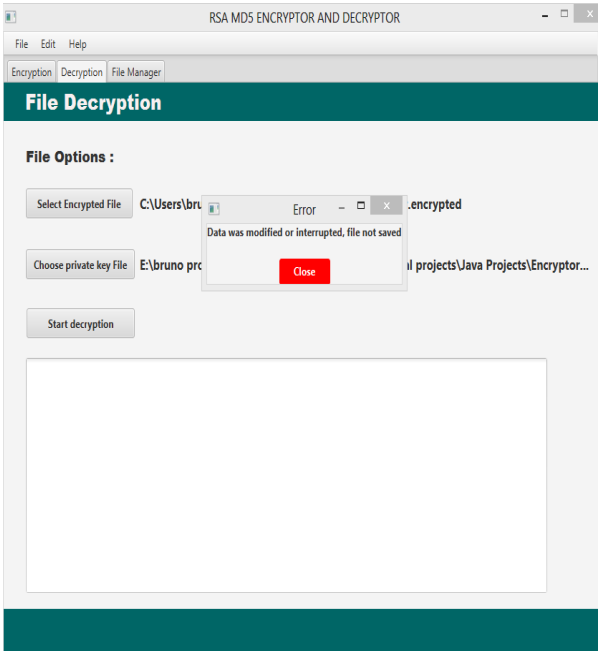

Figure 10: error message

5. SECURITY DISCUSSION

In this section an informal security discussion is proposed which are as follows.

A. Confidentiality

In this proposal encryption is performed for data confidentiality. Firstly, the data is encrypted with RSA which results in generation of public and private key and the MD5 algorithm. This encrypted file is then uploaded on cloud servers. In case if any, malicious get the data they are not able to get the original data as they don't have a private key.

B. Integrity

It is the most important approach uploaded data cannot be modified by other users. As Data owner can perform verification at any time. In this approach the security code created by the data owner is verified every time that file is accessed from the cloud. If these values match, then data is safe if it does not match, then modification has been done on the data. In this way the owner is having hold on their data, they can easily verify it. CSP also control the things very wisely as control of the data is provided by the owner. In this way the integrity of the data is maintained.

6. CONCLUSION

In this paper a solution is proposed to maintain data confidentiality and data integrity on cloud servers. For this approach RSA and MD5 hashing algorithm is used. Encryption and decryption is done by RSA algorithm, MD5 hashing algorithm and security code for verification.

With future emphasis is given to implement the proposed architecture with different comparison to show the effectiveness of our approach.

7. REFERENCES

[1] P.Mell and T.Grance, "The NIST Definition of Cloud Computing," National Institute of Standards and Technology,Vol.53,no.6,p.50,2009. [Online]. Available: http://csrc.nist.gov/groups/SNS/cloud computing/clouddefv15.doc.

[2] NesrineKaaniche,AymenBoudguiga, Maryline Laurent, "ID Based Cryptography for Secure Cloud Data Storage,"Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference .

[3] NehaTirthani, GanesanR,"Data Security in Cloud Architecture Based on diffie Hellman and Elliptical Curve Cryptography," International Association for Cryptologic Research, Nov 2013.

[4] Priyanka Ora and Dr.P.R.Pal, "Data Security and Integrity in Cloud Computing Based On RSA Partial Homomorphic and MD5 Cryptography", IEEE International Conference on Computer, Communication and Control (IC4-2015).

[5] Sudhansu Ranjan Lenka and Biswaranjan Nayak, "Enhancing Data Security in Cloud Computing Using RSA Encryption and MD5 Algorithm", International Journal of Computer Science Trends and Technology (IJCST) – Volume 2 Issue 3, June-2014