Implementation and Comparative Analysis of RSA and MD5 Algorithm

Saikat Das
VIT University
Vellore, India

Saugata De
VIT University
Vellore, India

Rahul Kumar
VIT University
Vellore, India

ABSTRACT

In this paper, we have made a comparative analysis of RSA Algorithm and a hash algorithm i.e., MD5 (Message Digest 5). Although the commonly used hash algorithms in the field of information security are MD5 and SHA-1 which evolved from MD4. We have implemented the logic of these algorithms in Java, and try to find something new and ended with the conclusion.

Keywords

Cryptography, Hash Algorithm, MD5, RSA.

1. INTRODUCTION

Cryptography is the practice of hiding the message in a secret code. In this modern world, to provide integrity, authenticity of the messages send over the communication channel needs to be provided with security. Features of using cryptography are privacy of the data only the intended receiver can read the data send from the sender, authenticity can be defined as a process of not denying one’s identity and most important is the privacy of the data i.e., the data should not be altered in any way. The three types of cryptographic algorithms used are public key cryptography, private key cryptography and the hash algorithms. Encryption can be defined as converting a message or a data to a secret code known as cipher text and when the cipher text is converted back to the original message is called decryption.

A. Symmetric key cryptography also called as a private key cryptography. It uses only one key for the encryption and decryption process. It takes the plain text then transforms it into cipher text using the secret key.

B. Asymmetric Key Cryptography is also known as Public key Cryptography. Here, the message is taken and then converted into cipher text using secret key and is deciphered at the receiver’s end using another key. Most commonly used Asymmetric Key is the RSA algorithm.

C. Hash Function is a one way cryptography where no key is used. It is also known as message digest. It has wide range of applications like password encryption.

D. As stated earlier RSA algorithm uses two keys, one for encryption and another for decryption. It takes two prime numbers initially and computes n and e. e is kept as private key whereas d is the public key.

E. MD5 algorithm uses checksum for 128 bit value of a file. It may be possible that two different files can generate the same hash value. It evolved from the older version i.e., MD4. MD5 requires message of at least 8 bits. MD5 is fast but it can take a large amount of the data. It is widely applied in the secure communication channel. MD5 function gives a 32 digit hexadecimal number. While comparing MD5 with MD4 the former provides more assurance of data security.

2. LITERARY SURVEY

A. The paper [1] focuses on proposing the attack on RSA algorithm. If we take a small value of e it works on simple RSA algorithm but as soon as an arbitrary value is taken it does not work so factoring modulus [1] is used to get an efficient solution and is implemented in C++.

B. In the paper [2], the paper demonstrates and compares between the two hashing algorithms SHA1 and MD5. Message digest leads to higher collision rates of the above two algorithms as they consist of at least 128 bits. Simulating technology is used to check the collision rates and establishes a statistical evaluation model [paper no.] to compare if there is any change with the hash value. However, the conclusion drawn is SHA1 is more secure than MD5.

C. The paper [3] focuses on a new image retrieval method [3] which was introduced based on the combination of Content Based Image Retrieval and MD5 [3]. It is basically applied by extracting the message digest of the saved image and is being stored in the database [3] at the same time. Here the unique MD5 values of different images helps in the betterment of the image accuracy and message-digest of fixed length reduces search complexity [3].

D. The paper [4] is entirely based on the research of the MD5 algorithm, thereby analyzing the different comparative measures of the algorithm from the program codes. The analysis thus result in the sum up of a good number of recent approaches.

E. In the paper [5], the paper demonstrates how we can recover the password of a PDF file which is in encrypted form thereby decreasing the instruction quantity which is to be executed in the recovery process of the hashing operation.

The paper [15] is entirely based on the comparative study of MD5 and SHA-1 algorithm. The paper compares the two hashing algorithms in different aspects like security point of view, length of hash value and execution speed and finally concludes that SHA-1 algorithm is much more secure than MD5 algorithm since MD5’s algorithm’s message-digest is 32 bit longer than SHA-1 algorithm. RSA is a public-key cryptosystem based on mathematical problems. It is based on asymmetric cryptography. It is used in encryption and decryption of the message through network. The security is provided in the network to send the data in securely manner to the designation point without being hacked. In this algorithm keys are needed which should be of at least 1024 bit for security purpose and for maximum is your choice if you want more security then you take more no bits like 2048,4096, etc. But 2048 bit size is used for best security. It is used worldwide for best security purpose. It is too slow because it takes large amount of bit and data, so calculation takes more time for encryption.

In RSA algorithm two key is used, private and public key. Public key is used to encrypt the data which is in cipher text or finger print, public keys are stored in database where everyone can see these, while private key is used to crack or decrypt the data. Private Key is with the person who will decrypt the message. It is very difficult to create secret key from public key therefore RSA is prevalent choice for data encryption. This algorithm is used in email, remote login, e-banking etc.

A. RSA Algorithm Logic

STEP 1: START

STEP 2: Select two prime numbers randomly. (Let A and B)

STEP 3: Multiply prime numbers (A*B), store in another variable (Let N)

STEP 4: Multiply (A-1) and (B-1), store in another variable (Let M)

STEP 5: Input any Integer (Let I) whose gcd (I, M) = 1 and 1 < I < M

STEP 6: Compute the secret exponent D, 1<D<M such that I.D = 1(mod M)

STEP 7: Set PUBLIC KEY (I, N) and PRIVATE KEY (D, N)

STEP 8: Input Plain text (Let X), such that X < N

STEP 9: Encryption Process: Calculate Cipher text (Let C), C = X^{I} (mod N)

STEP 10: Decryption Process: X = C^{D} (mod N)

STEP 11: STOP

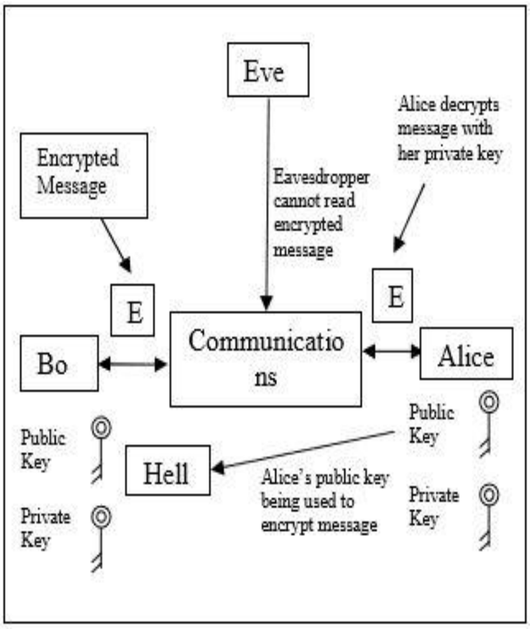
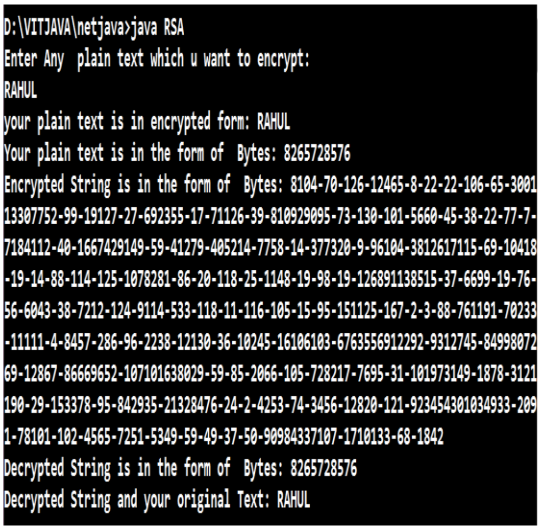

Eve
Encrypted Message
Eavesdropper cannot read encrypted message
Alice decrypts message with her private key
E
E
Bo
Communications
Alice
Public Key
Private Key
Hell
Alice's public key being used to encrypt message
Public Key
Private Key

Figure: RSA Algorithm Architecture

B. Implementation of RSA algorithm in java

First six variable needs to be taken of type of Big Integer and after that bit length is to be taken and initialized the bit length of our own choice. Initializing the bit length by 1024 bit, a random value has been created and two prime no's has been taken and then a function called Prime() function has been called of Big Integer class by passing bit length and random value then it returns 309 digits prime no and we have stored in variable A and B. After multiplying A and B and storing the result invariable N, by using multiply function of Big. After that the calculation (A-1)*(B-1) is done and stored it in M and then random and by probable Prime function is created again, and prime no is generated and stored it in I (public). Then the calculation of D (private) is done by using the mod Inverse function and then it also returns big integer after calculating (I^-1 mod M). Finally input is taken from the user and converted them into bytes and call encrypt method. Under this method I and N has been used to encrypt data (P^I mod N) for this Mod POW has been used. This function calculates power and after calculating power it calculates mod and return big integer and return the data in byte array format (encrypted data) and (P^D mod N) has been used for decrypting the message after getting the original message. The execution result of the java program is shown in figure 3:



```
D:\VITJAVA\netjava>java RSA
Enter Any  plain text which u want to encrypt:
RAHUL
your plain text is in encrypted form: RAHUL
Your plain text is in the form of  Bytes: 8265728576
Encrypted String is in the form of  Bytes: 8104-70-126-12465-8-22-22-106-65-3001
13307752-99-19127-27-692355-17-71126-39-810929095-73-130-101-5660-45-38-22-77-7-
7184112-40-1667429149-59-41279-405214-7758-14-377320-9-96104-3812617115-69-10418
-19-14-88-114-125-1078281-86-20-118-25-1148-19-98-19-126891138515-37-6699-19-76-
56-6043-38-7212-124-9114-533-118-11-116-105-15-95-151125-167-2-3-88-761191-70233
-11111-4-8457-286-96-2238-12130-36-10245-16106103-6763556912292-9312745-84998072
69-12867-86669652-107101638029-59-85-2066-105-728217-7695-31-101973149-1878-3121
190-29-153378-95-842935-21328476-24-2-4253-74-3456-12820-121-923454301034933-209
1-78101-102-4565-7251-5349-59-49-37-50-90984337107-1710133-68-1842
Decrypted String is in the form of  Bytes: 8265728576
Decrypted String and your original Text: RAHUL
```

Figure : Output of RSA Algorithm

3. THE INTRODUCTION AND WAY OF IMPLEMENTATION OF MD5 ALGORITHM

Message-digest algorithm is also cryptography but it uses hash function to produce 128-bit hash value which is usually expressed in 32 digit Hexadecimal number. It is known as one way transmission that means if by the help of hash function which gives the result in the message digest form. It is also called as finger print or hash value of plain text. Then we cannot get original message by the help of hash function. In this algorithm it is very difficult to get original text but in RSA we can get original message then we can say that RSA is Reversible but MD5 is irreversible.MD5 can also take in to plain text format and even in binary digits to encrypt the data. Check sum is generated with the form that if the hash function gets the value between 0 and 256 it returns the same value but if it gets higher value it takes the mod by 256 and gives the quotient value in return. It runs faster on 32 bit or 16 bit machine. There are many algorithms which are based on Hash function like SHA, CRC32, etc.

A. MD5 Algorithm Logic

STEP 1: Append padding bits.

STEP 2: Join Length
STEP 3: Take input from user
STEP 4: Partitioned the input into 512 bit blocks.
STEP 5: Initialized MD Buffer
STEP 6: Process blocks means, we have to take 4 MD5 buffer and then we have performed 16 basic operations on it.
STEP 7: To get output in the form of hash value in this step, it performs 4 operations to get 128 bit hash value.

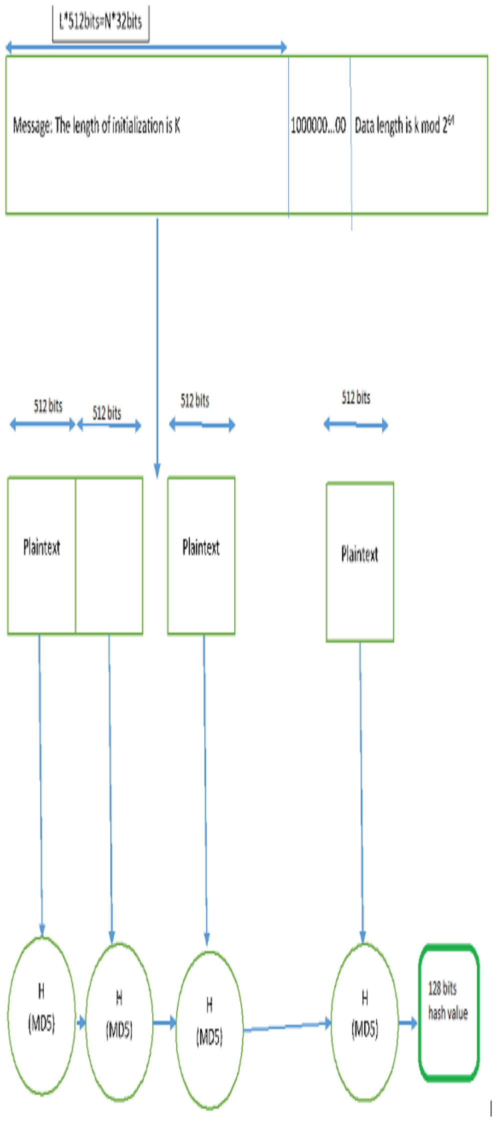
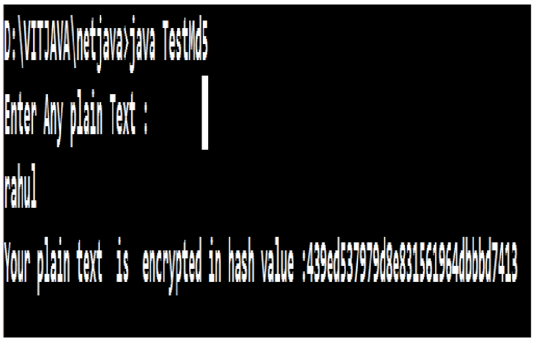


Figure: MD5 Chart of Processing

B. Implementation of MD5 algorithm in java

First input is taken from user in binary or text/String format then the MD5 function is called by passing this input. In MD5 function the instance of Message Digest class is created by passing the algorithm name and then it returns an object of class Message Digest that implements the algorithm which is passed to the parameter of that class. After that the Message Digest object is updated by calling the function update and passing the array of bytes of input, starting point and input length. Then the hash calculation is completed by performing the last operation like padding. After that the hash value is converted into Big Integer by calling Big Integer constructor and by passing the bit and array of bytes and finally it is converted into String and then it is returned to the function.

The execution result of the java program is shown in figure:



```
D:\VITJAVA\netjava>java TestMd5

Enter Any plain Text :

rahul

Your plain text  is  encrypted in hash value :439ed537979d8e831561964dbbbd7413
```

Figure : Output of MD5 Algorithm

4. COMPARISON OF RSA AND MD5 ALGORITHM

A. Compare of bit length

RSA uses large number of bits in I and D whose range falls between 1024 and 2048 which generates 309 digits code which is considered strong enough for virtually all applications while MD5 uses very little no bit comparison to RSA. It uses 128 bits which generates 32 digits.

B. Attack on RSA surpassed by MD5

There are mainly two kinds of attacks on RSA

1. Mathematical Attacks

2. Implementation Attacks

Mathematical Attack:

It attacks on the structure of the RSA functions. The intuitive attack efforts to factor the mod N. because knowing the factor one can easily decrypt the process and the D can be determined. There are three more categories under Mathematical Attack

a) Elementary Attack: This attack is done openly to misuse of RSA.

b) Small Private Key Attack: To increase the performance of decryption in RSA, value is taken smaller i.e. small private key and is decrypted easily by the attackers.

c) Small Public Key Attack: Similar to private key, to improve the performance of encryption under RSA, value is taken smaller i.e. small public key. The base attack is done by the Hastad's Broadcast attack and the same encryption code is distributed among the different users.

d) Implementation Attack: This is known as time attack. The attack which falls under this category implements hidden cryptosystem which is a trivial task.

While in MD5 algorithm all attack is surpassed and there is no issue of cipher text, private or public key attack.

C. Comparison of execution time

MD5 takes very less time in comparison to RSA because of large prime numbers. In RSA algorithm there is need of two large prime numbers which will be converted into 1024 bits and then processor takes more time for calculation and hence it becomes slow.

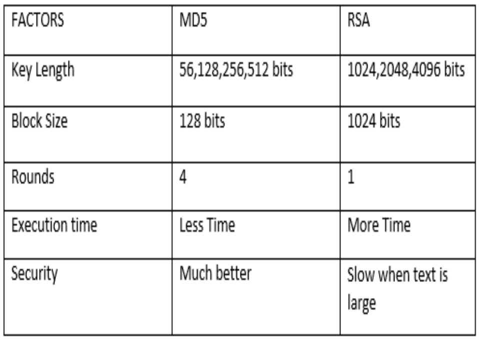
D. Comparison of Utilization of CPU

MD5 utilizes CPU better than RSA because of fast calculation algorithm while RSA utilizes CPU better than MD5 because RSA uses more computer resources and hence its speed goes down.

E. Comparison of Security

In security point of view, MD5 is better than RSA because RSA provides powerful security to small file but in the case of large file, it is not compressed well compared to MD5 and hence in the case of large fiMD5 is better.

F. Comprehensive Comparison



FACTORS	MD5	RSA
Key Length	56,128,256,512 bits	1024,2048,4096 bits
Block Size	128 bits	1024 bits
Rounds	4	1
Execution time	Less Time	More Time
Security	Much better	Slow when text is large

5. ACKNOWLEDGEMENT

We are highly thankful to our parents for their constant support and we also want to thank Prof. Navaneethan C for his guidance of the paper.

6. CONCLUSION

In this paper the logic of two algorithms i.e., RSA and MD5 has been implemented in Java. The differences between these two algorithms has been concluded and how is the working principle. In future, our aim will be to implement SHA algorithm. However, MD5 is a hashing function, with one way cryptography and due to the randomness characteristics, it is used for the data integrity purpose.

7. REFERENCES

[1] Sattar J Aboud, “An Efficient method for Attack RSA Scheme,” Iraqi Council of Representatives, pp. 587-591, 2009

[2] Ming Hu, Yan Wang, “The Collision Rate Test of Two Known Message Digest Algorithms,” International Conference on Computational Intelligence and Security, pp. 319-323, 2009

[3] Hancheng LIAO, “Image Retrieval based on MD5,”International Conference on Advanced Computer Theory And Engineering, pp. 987-991, 2008

[4] Zhao Yong-Xia, Zhen Ge, “MD5 Research,” Second International Conference on Multimedia and Information Technology, pp. 271-273, 2010

[5] Keonwoo Kim, Un Sung Kyong, “Efficient Implementation of MD5 Algorithm in Password Recovery of a PDF File,” Cyber Convergence Security Division, ETRI, Daejon, Korea, pp. 1080-1083

[6] Anak Agung Putri Ratna, Ahmad Shaugi, Prima DewiPurnamasari, Muhammad Salman, “Analysis andComparison of MD5 and SHA-1 Algorithm Implementation in Simple –O Authentication Based Security System,” Universitas Indonesia, IEEE, pp. 99-104, 2013

[7] Xiaoling Wei, “MD5 Encryption Algorithm and Application,” Yan’an University Computing Center, 2010

[8] Quist-Aphetsi Kester, Lauret Nana, Anca Christine Pascu, Sophie Gire, “A New Encryption Cipher for Securing Digital Images of Video Surveillance Device using Diffle-Hellman-MD5 Algorithm and RGB pixel shuffling,”European Modelling Symposium, pp. 305-311, 2013

[9] Wang Xiayoun, “How to Break MD5 and other Hash Functions,” 2005

[10] Kasgar A.K., Agrawal Jitendra, Sahu Santosh, “New Modified 256-bit MD5 Algorithm with SHA Compression Function,” International Journal ofComputer Applications, pp. 47-51, 2012

[11] A. Sinha, K. Singh, “A technique for image encryption using digital signature,” Optics Communication”, pp. 229-234, 2003

[12] Kahate, Atul, 2003, “Cryptography and Network Security,” Tata McGraw-Hill, India

[13] William Stallings, “Cryptography and Netwok Security: Principles & Practice,” 5th EditionPrentice Hall; 5 edition (January 24, 2010)

[14] R. Rivest, “The MD5 Message-Digest Algorithm,” Network Working Group, 1992

[15] Zhengi Wang, Lisha Cao, “Implementation and Comparison of Two Hash Algorithms,” International Conference on Computational and Information Sciences, pp. 721-725, 2013

[16] X. Wang, D.Feng, X.Lai and H.Yu, “Collisions for Hash Functions,” in Crypto, 2004

[17] Bonteh S, “Twenty years of Attacks on the RSA Cryptosystem,” Notices of the American Mathematical Society, 46(2):203-213,1999

[18] Rashmi P.Sarode, Piyush Gupta, Neeraj Manglani, “A Comparative analysis of RSA and MD5 Algorithm,” Journal of Computer Science and Applications, pp. 25-33,2014

[19] Itoh, K., Kunihiro N.,KurosawaK., “Small secret key attack on a variant of RSA,”volume 4964 of Lecture Notes in Computer Science , pp. 387-406, 2008

[20] Shahzad Alam, Amir Jamil, Ankur Saldhi, Musheer Ahmad, “Digital Image Authentication and EncryptionUsing Digital Signature,” International Conference on Advances in Computer Engineering & Applications, pp. 332-336, 2015