Journal of Computer Science and Applications.
ISSN 2231-1270 Volume 6, Number 1 (2014), pp. 25-33
© International Research Publication House
http://www.irphouse.com

A comparative analysis of RSA and MD5 algorithms

Rashmi P. Sarode*, Piyush Gupta and Neeraj Manglani

Department of Computer Science, Jagan Nath University
Jaipur – 302 022, India
**E-mail for correspondence: rashmipsarode@gmail.com*

Abstract

This paper presents a comparative study between message digest algorithm, its versions and RSA algorithm. Also a comparison between it and the available algorithms in literature has been made. The main goal is to provide information about message digest in the prescriptive of hash as well as RSA in the prescriptive of asymmetric key concept. As an Information security measure this is important to understand about digital security. Since both algorithms are based on cryptography, only few topics which are related with cryptography are discussed in this paper. This paper is expected to provide good grounds for those who pursue want their research in the field of information security as well as cryptanalysis.

Keywords: Security, MD5, RSA, Cryptography, Cryptanalysis, Digital Security.

1. Introduction

The term Cryptography is a subset of Cryptology and the Cryptanalysis is an also subset of Cryptology.

In today's World Wide Web, cryptography can be defined as the technology of writing secret messages. It involves the way in which data can be communicated in an encoded manner to hide contents through eavesdropping methods like codes, ciphers, etc. This helps only certain intended people to view the actual message.

MD5 is used to measure data integrity by the help of 128-bit message which the users created by giving some input (variable length), this generates a master fingerprint or a hash code which is irreversible. The developer of this algorithm is Professor Ronald L. Rivest of MIT [1]. This algorithm runs best on 32 bit and 16 bit machines. It is the advancement of MD4 (Message Digest 4) which is quite faster than MD5. However, MD5 deals with security offers.

MD5 is a one way hash function, taking random data and transmitting set length hash values which are irreversible. Message Digest is also known as checksum. As known the current aspect on online file sharing like peer to peer (p2p) these files can contain some duplicity as well as it gives warning to a user for unauthorized downloads. Besides this, there are other message digest algorithms like SHA and CRC32 which are also based on hash.

Encryption is the typical method used to make all communications private. Any user who wants to send a private message to other user encrypts the message before transmitting it. Only the intended recipient knows how to decrypt the encoded message correctly. Any user eaves dropping on the communication would get only the encrypted message and would not be able to decrypt it successfully thus making no sense, hence, privacy needs to be safeguarded in electronic communication.

Security often means to keep data safe from unauthorized access with the best line of defense, as the physical security. However, this physical security is not always an option due to cost factor and/or efficiency considerations. Instead in today's era, most of the devices are interconnected openly with each other, thus giving exposure to the communication channels.

In symmetric cryptography, only one key (private key) is used for both encryption and decryption. The key is shared with the other party who is decrypting the data. It has to be shared in such a way that it is safeguarded to be a secret. A new key has to be created for every new communication with a new party which creates a problem in organizing the keys and safeguarding all these keys. Also since both the sending party and receiving party use the same key, we cannot confirm which message is coming from which particular user.

The problem of distributing keys for encryption has been solved by asymmetric encryption by a simple way. It is also known as public key encryption. Here two keys, public key and private key are used for encryption and decryption. Any user can use the public key to encrypt the message but the message can be decrypted only with the private key which only the receiver has access to. Public keys are published but private keys are kept confidential.

One example of asymmetric algorithm is RSA algorithm. This algorithm was developed by Ron Rivest, Adi Shamir and Leonard Adelman in 1977 as the first major asymmetric key cryptography algorithm [2,3]. The name RSA comes from the surnames of these three research scientists. Ron Rivest was a professor working in Massachusetts Institute of Technology, USA (MIT). He hired Shamir and Adelman to work on the notion of asymmetric key cryptography.

RSA requires keys of at least 1024 bits for security but keys of size 2048 are best for security purpose. It is widely used as secure communication channel and for authentication of service provider. [4] RSA is too slow for encrypting large amounts of data and is widely used for key distribution.

The approach used here is based on asymmetric algorithm which involves a key pair, a public key and a private key.

To communicate securely over any network, one needs to publish the public key. All these public keys are stored in a database which any one can refer to. But the private keys remain with the respective individuals only. It is a very challenging task to create private key from the public key so RSA is a very prevalent choice in data encryption.

RSA also uses the technique of digital signatures, so we know the message is coming from which particular user. The use of digital signature also prevents the message to be altered in the transit. In RSA algorithm we need the keys same as number of participants so this algorithm scales up quite well. There is no problem of key agreement of key exchange here.

2. Versions of Message Digest Algorithms

There are three versions of Message Digest i.e. MD2, MD4, MD5 .There is a key difference between previous

Message digest algorithms which are shown in the tables below:

Table 1:-Difference Between MD2 and MD5

Message Digest 2	***Message Digest 5***
The algorithm is optimized for 8-bit PC	The algorithm is optimized for 8,16,32 bit PC
It was developed in 1989	It was developed in 1992
This uses less amount of CPU	This uses High Amount of CPU
This is Fast & Simple	This is also Fast but hard too
Contains small amount of data	Contains large amount of data

Table 2: Difference between MD4 and MD5

Message Digest 4	***Message Digest 5***
Precursor to MD5	This is the advance version of MD4
Also produces a 128-bit hash of message	Produces a variety of hash 64,128 bit hash code
Has 3 rounds of 16 steps	Has 4 rounds of 16 steps
Collision resistant	This May Produces collision on Digital CER.

Table 3: Difference between MD2 and MD4

Message Digest 2	***Message Digest 4***
It was developed in 1989	It was developed in 1990
The message is first padded so that its length in bytes is divisible by 16	The message is padded to ensure that its length in bits plus 448 is divisible 512
A 16-byte checksum is then appended to the message.	A 64-bit binary representation of the original length of message.

MD5 requires variable length message of at least 8 bits of message. It is widely used as secure communication channel and for authentication of identity of secure certificate over the wide network. MD5 is too fast but also takes large amount of data.

2.1 MD5

This algorithm is based on message length

```
// M= (Y0, Y1,………., Yn-1), Message to hash , after padding
// Each Yi is a 32-bit word and N is a multiple of 16
MD5 (M)
//initialize (A,B,C,D) = IV
(A,B,C,D) = (0x67452301 , 0xefab89 , 0x98badcfe , Ox10325476 )
For i=0 to N/16 -1
  // Copy block I to X
  Xj = Y16i+j for j = 0 to 15
// Copy X to W
  Wj = Xσ(j) , for j = 0 to 63
// initialize Q
  (Q-4 , Q-3 , Q-2 , Q-1) = (A , D , C , B)
  // Rounds  0 , 1 , 2 and 3
             Round0(Q , W)
             Round1(Q , W)
Round2(Q , W)
Round3(Q , W)
// Each addition is modulo $2^{32}$
(A , B , C , D)=(Q60 + Q-4 , Q63 + Q-1 , Q62 + Q-1 , Q61 + Q-3)
  next i
return A , B , C , D
end MD5
Round0(Q , W)
//steps 0 through 15
  for i = 0 to 15
Qi = Qi-1 + (( Qi-4 + F(Qi-1 , Qi-2 , Qi-3 ) + Wi +Ki ) <<< si )
next i
end Round()
```

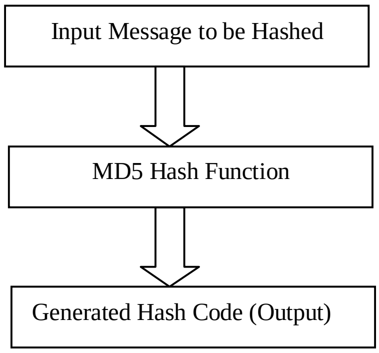

Fig 1: Basic Concept of Message digest

Step 1:- Padding bits and Append Length
The bits must be padded with the ‘1’ bit, and ‘0’ bits first and last respectively until the resulting is not equal to bit length which is equals 448 mod 512, and the end of bit length of the original message as a 64-bit integer. The final bit length of the message which is already padded is 512N for a true integer N.

Step 2:-Divide the input into 512-bit blocks
This padded message is partitioned into N successive 512-bit blocks M1, M2....... MN.

Step 3:- Initialize Channing variables
Initialization of 32-bit number in the form of chaining variables (A,B,C,D) these values are represented in hash only.

A	01 17 2d 43
B	89 AB CD EF
C	FE DC BA 98
D	76 54 32 10

Step 4 :- Process blocks
The four buffers (A, B, C and D) messages (content) are merged now with the input words, using the four auxiliary functions (W, X, Y and Z).

There are 4 rounds, each involves 16 basic operations. The Processing block P is applied to the four buffers (A, B, C and D), using message word M[i] and constant K[i]. The item "<<<s" denotes a binary left shift by s bits.

The four type info related functions that each take as input three 32-bit words and produce same bits of output i.e. 32-bit word.

They apply the logical operators and, or, not and xor to the input bits.

W (D, E, F) = DE v not (D) F
X (D, E, F) = DE v E not (F)
Y (D, E, F) = D xor E xor F
Z (D, E, F) = E xor (D v not (F))

The bits of D, E, and F are autarchic and stabilized the each bit of W (D, E, F) will be autarchic and stabilized. The functions (D, E and F) are equal to the processing P, in that they do job in "bitwise parallel" to produce the reliable output from the bits of D, E and F.

In such a way that if the be similar bits of D, E and F are autarchic and balanced, then each bit of X (D, E, F), Y(D, E, F) and Z(D, E, F) will be autarchic and stabilized.

Step 5:- Hashed Output

There are 4 rounds performed in message digest 5 (MD5) which is of 128 bits. Fig 2 shows One MD5 Operation. [2,3]

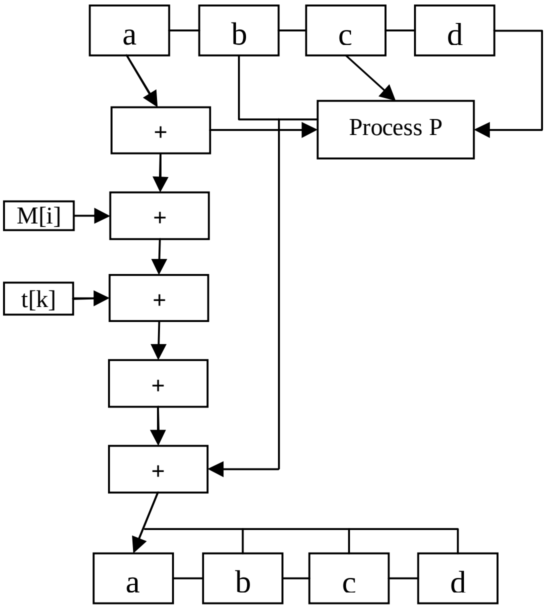

Fig 2: One MD5 Operation

2.2 RSA

This algorithm is based on the difficulty of factorizing large numbers that have 2 and only 2 factors (Prime numbers). [1]

1) Choose two large prime numbers P and Q

2) Calculate N = P x Q

3) Select the public key (i.e. the encryption key) E such that it is not a factor of (P-1) and (Q-1).

4) Select the private key (i.e. the decryption key) D such that the following equation is true:

$(D \times E) \bmod (P-1) \times (Q-1) = 1$

5) For encryption, calculate the cipher text CT from plain text PT as follows:

$CT=PT^{E} \bmod N$

6) For decryption, calculate the plain text CT as follows:

$PT = CT^{D} \bmod N$

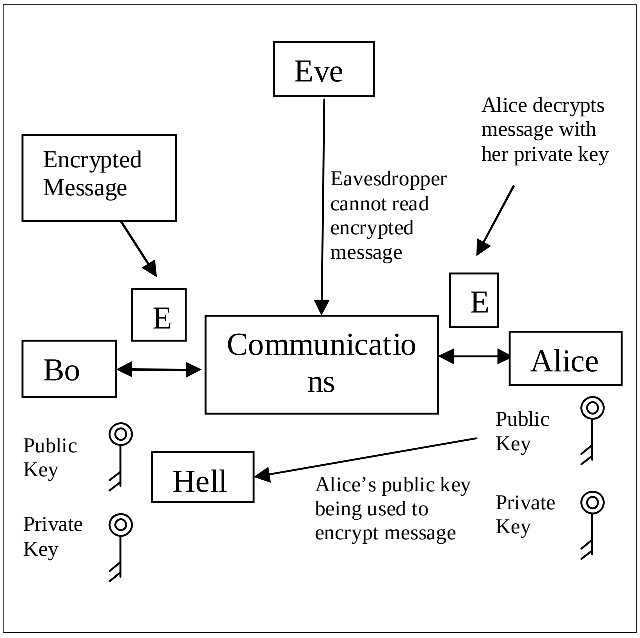
Eve
Alice decrypts message with her private key
Encrypted Message
Eavesdropper cannot read encrypted message
E
E
Communicatio ns
Bo
Alice
Public Key
Public Key
Hell
Alice's public key being used to encrypt message
Private Key
Private Key

Fig 3: Working of RSA Algorithm

3. Parameters used

3.1 Parameters of MD5

Below equation shows a single MD5 operation.

a = b + ((a + Process P (b, c, d) + M[i] + t[k]) <<< s)

Here :-

a, b, c, d = are Chaining variables

Process P=A non linear operation

M[i] =For M[q x 16 + i], which is the ^{ith} 32-bit word in the ^{qth} 512-bit block of the message

t[k]=a constant

<<<s =circular-left shift by s bits [1]

3.2 Parameters of RSA

Consider the equation for encryption and decryption as follows:

$CT=PT^{E} \bmod N$ and $PT = CT^{D} \bmod N$ [1]

Here:
CT and PT are Cipher text and Plain text respectively
E is the encryption key
D is the Decryption key
N is the product of the two chosen prime numbers for factorization

4. Comparisons between MD5 and RSA

RSA and MD5 have few similarities like they both follow the concept of key generation, they are used in digital certificates and 3d transaction protocols [2-6].

MD5 solves the problem of password management by sending URL to the user instead of old password and RSA solves the problem of Key distribution by using two separate keys – one for encryption and other for decryption.

Interestingly, MD5 can be used for finger print generation and this technique can be used to encode images whereas RSA is used in technologies like satellite TV's and radios.

MD5 utilizes CPU in an efficient manner by using a fast computation algorithm whereas RSA uses most of the computer resources thus slowing down the speed of encryption.

MD5 is an irreversible algorithm as it provides one way has function whereas RSA is a reversible algorithm as we can encrypt as well as decrypt.

MD5 though very secure, some companies have stopped its usage due to security flaws whereas RSA has many secure versions.

Table 4: Comparisons between MD5 and RSA.

FACTORS	**MD5**	**RSA**
Key Length	64 bits, 128 bits , 256 bits , 512 bits	1024 bits , 2048 bits , 4096 bits
Block Size	128 bits	1024 bits
Developed	1992	1977
Cryptanalysis Resistance	Strong against Digital Certificate and very fast on 32 bit machines	Strong against Digital Certificate and data
Security	Secure	Secure but slow with large amounts of text
Rounds	4	1

As observed from the table no 4, the key length of MD5 is comparatively less to RSA. MD5 uses key length starting from 64 bits to 512 bits and RSA usually uses key length of 1024 bits and even higher for more security. Moreover, the block size of MD5 is 128 bits which is much more as compared to RSA which is 1024 bits. RSA was developed much before MD5; both had been developed by Ron Rivest. Both the algorithms have a strong cryptanalysis resistance; MD5 is even faster on 32 bit machines. The weak point of RSA is that it is much slow with large extent of texts.

MD5 needs four rounds to complete one full operation as compared to 1 round in RSA.

4. Conclusions

In this paper, a new comparative study between MD5 and RSA were presented into 6 factors viz. key length, block size, developed, cryptanalysis resistance, security and rounds. These proves that MD5 is quite secure with large extend of data; but has certain security flaws. On other hand, RSA is a very secure for smaller amount of data.

References

[1] Rivest R., 1992, “The MD5 Message-Digest Algorithm,” RFC 1321, MIT LCS and RSA Data Security, Inc.

[2] Rivest, R., Shamir A, .and Adleman L ., 1978,. "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems". Communications of the ACM **21** (2) pp 120–126.

[3] Kahate, Atul, 2003, "Cryptography and Network Security", Tata McGraw-Hill , India.

[4] Rayakar R, Upadhay S., and Pimpale P., 2012, “SMS Encryption using AES Algorithm on Android “International Journal of Computer Applications (0975 – 8887) Volume 50(19),pp 12-16

[5] Kasgar A. K., Agrawal Jitendra, Sahu Santosh, 2012, “New Modified 256-bit MD5 Algorithm with SHA Compression Function” , International Journal of Computer Applications (0975 – 8887) Volume 42(12), pp 47-51

[6] William Stallings, Cryptography and Network Security: Priciples and Practice, 5th EditionPrentice Hall; 5 edition (January 24, 2010)