IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 1, January 2015.
www.ijiset.com
ISSN 2348 – 7968

A Modified RSA Algorithm for Security Enhancement and Redundant Messages Elimination Using K-Nearest Neighbor Algorithm

Dr. Abdulameer K. Hussain

Computer Science Department, Jerash University
Jerash, 00962-02, Jordan

Abstract

This paper presents an effective solution to enhance the security of RSA scheme. The objective of this proposed method is to eliminate the redundant messages which occurred in some values of n, the product of two prime numbers, and this problem is considered as a weak point in the RSA method. The solution of this problem depends on replacement of this value of n using a secure agreement distance in a set of all available prime numbers. The next step is selecting either one of the primes responsible for generating an alternative n or both primes from that set.
This method also eliminates one parameter of the public key which is the value of n, and therefore we get a more secure method to prevent most common attacks against RSA method.

Keywords:*Cryptography, RSA Algorithm, K-Nearest Neighbour Algorithm.*

1. Introduction

Because of the rapid growth of telecommunication and internet, information security becomes more and more significant. Cryptography is the best way for protecting secret information. Cryptosystems can be divided into two types , secret-key cryptosystem and public-key cryptosystem.

The first type (secret-key cryptosystem), uses the same encryption key to encipher the plaintext and decipher the ciphertext .For this reason, this types is also called as symmetric cryptosystem. Though secret-key cryptosystem is easily to implement due to less computation, , it has several drawbacks, too many keys, key distribution problem, authentication and nonrepudiation problem.

The important type which is the public-key cryptosystem is developed to solve the problems of symmetric cryptosystem, and RSA cryptosystem is the most popular approach. The RSA cryptosystem was developed in 1977 by Ronald L. Rivest, Adi Shamir, and Leonard Adleman at MIT and first published in 1978 [1].

Despite the RSA algorithm is considered a very secure, t is rarely used in smart card, because it take long computation time. The main usage of RSA is in the digital signature. In addition RSA cryptosystem is relatively slow and therefore it is unsuitable for encryption of large messages [2], [3],[4], [5].

R.L. Rivest, A. Shamir, and L. Adleman developed the RSA public-key cryptosystem in 1978. The RSA cryptosystem is simply a modular exponentiation. The modulus n is the product of two large prime's p and q,
Public key and private key are obtained by:

$$e = d^{-1} \bmod Ø(n) \quad (1)$$

The encryption operation is performed using the public key n and e as follows:

$$C = M^e \pmod n \quad (2)$$

Where M is the plaintext such that 0<M<n and C is the ciphertext which can be decrypted using the private key n and d as follows:

$$M = C^d \pmod n \quad [1]$$

In recent days, data security has become an important issue for public, private and defense organizations because of the large losses of illegal data access. To protect confidential data or information from unauthorized access, illegal modifications and reproduction, various types of cryptographic techniques are used. One of these important techniques is cryptography which is the science of writing in secret form and it is divided into two types: symmetric and asymmetric cryptography [6], [7].

Symmetric algorithms are typically considered fast and they are suitable for processing large stream of data. Some of the famous and efficient symmetric algorithms include Twofish, Serpent, AES, Blowfish and IDEA [1]. In addition, there are generic algorithms which offer an alternative technique for encryption [3]. In general, genetic algorithms contain three basic operators: reproduction, crossover and mutation [8], [9].

On the other hand, there are different popular and efficient asymmetric algorithms including RSA, NTRU, and

ISSN 2348 – 7968

Elliptic curve cryptography. But RSA algorithm is a cryptosystem, which is known as one of the first practicable public -key cryptosystems and is widely used for secure data transmission [6]. The important property of RSA algorithm is that the encryption key is public and differs from the decryption key which is kept secret. This property gives the RSA algorithm asymmetry property which is based on the practical difficulty of factoring the product of two large prime numbers which is known as the factoring problem. The main procedure of the RSA algorithm is that allows the user to creates and then publishes the product of two large prime numbers, along with an auxiliary value, as their public key. RSA stands for Ron Rivest, Adi Shamir and Leonard Adleman, who first publicly described the algorithm in 1977 [10]. For more security, the prime factors must be kept secret. Any user can use the public key to encrypt a message (M), but with currently published methods, if the public key is large enough, only someone with knowledge of the prime factors can feasibly decode the message [1]. Breaking RSA encryption is known as the RSA problem. It is an open question whether it is as hard as the factoring problem [11], [1].

The main operation of many public-key cryptosystem (PKC) such as RSA is the using of the large integers modular exponentiation which is implemented by repeating modular multiplication [12],[13],[14]. For this reason, the efficiency of many PKCs is determined by the efficiency of the modular multiplication algorithm [15][16][17].

One of the efficient algorithms for modular multiplication is Montgomery modular multiplication (M3) algorithm because it avoids division by the modulus [18], [19].

Ha and Moon in [20] proposed a method that the common part of modular multiplication in modular exponentiation can be computed once rather than twice and called it common-multiplicand multiplication (CMM) method. Wu et al. in [21] proposed a method by using canonical recoding technique in order to recode the exponent. So the probability of the nonzero digit is reduced. Therefore the computational complexity of the modular exponentiation is decreased. In [21] CMM method proposed in [20] can be used in multiplication phase. Wu in [22] proposed a method to divide the signed-digit exponent into three equal lengths and use of CMM technique in order to compute common part of multiplications, once rather than several times.

In this paper we use the k-nearest neighbour algorithm. The definition of this algorithm is that: Suppose each sample in our data set has n attributes which we combine to form an n-dimensional vector:

x= (x1, x2, . . . ,xn)

These n attributes are considered to be the independent variables. Each sample also has another attribute, denoted by y (the dependent variable), whose value depends on the other n attributes x . We assume that Y is a categorical variable, and there is a scalar function, f, which assigns a class,
y=f(x) to every such vectors.

We do not know anything about f (otherwise there is no need for data mining) except that we assume that it is smooth in some sense.

We suppose that a set of T such vectors are given together with their corresponding classes:
$X^{(i)}, y^{(i)}$ for i= 1 , 2, . . . , T.
This set is referred to as the training set.

The problem we want to solve is the following. Suppose we are given a new sample where x = u. We want to find the class that this sample belongs. If we knew the function f, we would simply compute v = f(u) to know how to classify this new sample, but of course we do not know anything about f except that it is sufficiently smooth.

The idea in k-Nearest Neighbor methods is to identify k samples in the training set whose independent variables x are similar to u , and to use these k samples to classify this new sample into a class, v. If all we are prepared to assume is that f is a smooth function, a reasonable idea is to look for samples in our training data that are near it (in terms of the independent variables) and then to compute v from the values of y for these samples.

When we talk about neighbors we are implying that there is a distance or dissimilarity measure that we can compute between samples based on the independent variables. For the moment we will concern ourselves to the most popular measure of distance: Euclidean distance [23].

The Euclidean distance between the points x and u is:

$$d(u,x) = \sqrt{\sum_{i=1}^{n} (xi - ui)2}$$

2. Related Works

In [24], a proposal for introducing an approach which is more secure than original RSA algorithm, which is used for digital signatures and encryption in public key cryptography. This method eliminates the need to transfer n, the product of two random. The authors used

ISSN 2348 – 7968

large prime numbers which it becomes difficult for the intruder to guess the factors of n and hence the encrypted message remains safe from the hackers. For this reason, this approach provides a more secure path for transmission and reception of messages through public key cryptography.

A research was proposed to develop a factorization method which is used to obtain the factor of positive integer N. The present work focuses on factorization of all trivial and nontrivial integer numbers as per Fermat method and requires fewer steps for factorization process of RSA modulus N. By experimental results it has been shown that factorization speed becomes increasing as compare to traditional Trial Division method, Fermat Factorization method, Brent's Factorization method and Pollard Rho Factorization method [25].

To improve the security, a scheme presented a new cryptography algorithm based on additive homomorphic properties called Modified RSA Encryption Algorithm (MREA). MREA is secure as compared to RSA as it is based on the factoring problem as well as decisional composite residuosity assumptions which is the intractability hypothesis. The scheme is an additive homomorphic cryptosystem; this means that, given only the public-key and the encryption of m1 and m2, one can compute the encryption of m1 + m2. This scheme also presents comparison between RSA and MREA cryptosystems in terms of security and performance [26].

A new technique was proposed to provide maximum security for data over the network by using a modified RSA cryptosystem based on 'n' prime. This method involves encryption, decryption, and key generation.This technique used 'n' prime number which is not easily breakable. 'n' prime numbers are not easily decompose. So this modified RSA algorithm handled 'n' prime numbers and provides security [27].

3. Proposed System

When RSA is implemented, there is a situation in which the ciphertext is the same as the plaintext in some values of n which is the product of two prime numbers p and q. So it is very important to find an effective solution for such a problem because it is vulnerable for most common attacks.

In order to solve such a problem, the proposed method constructs a new method to change the value of n by constructing a large set of prime numbers from which the users can select alternative values of either one of the prime numbers or both used for changing the value of n and to ensure more security.

In this method, the parties in the communication session agree upon a secure set of alternative prime numbers (PR). From this set they can choose alternative values of prime number for p or q or both. In addition, this set of prime numbers is divided into different classes. Each class contains a specified numbers of primes .This choice is dependent on a secure distance (d_1) to select a certain neighbor of one of the classes in that set. This distance will be used to choose one prime number or both. This procedure is performed by assigning another secure distance (d_2) inside the selected class. Then when we get an equality of both message and its corresponding ciphertext due to a specific value of n, we can generate a new secure value of n to overcome these redundant values of messages.

For this purpose, the proposed system uses the k-nearest neighbor of p or q in a one of the neighbor class in the set of alternative prime numbers. The purpose of the distance (d_1) is to use an agreement secure parameter to choose one of the classes inside the set of all classes and this distance must be changed periodically to remove the redundant messages and to enhance more security for the RSA algorithm. There is another secure parameter which is the distance (d_2). The propose of this distance is to select the k-nearest neighbour of primes inside the selected class by distance (d_1).

Figure 1 illustrates the system model.

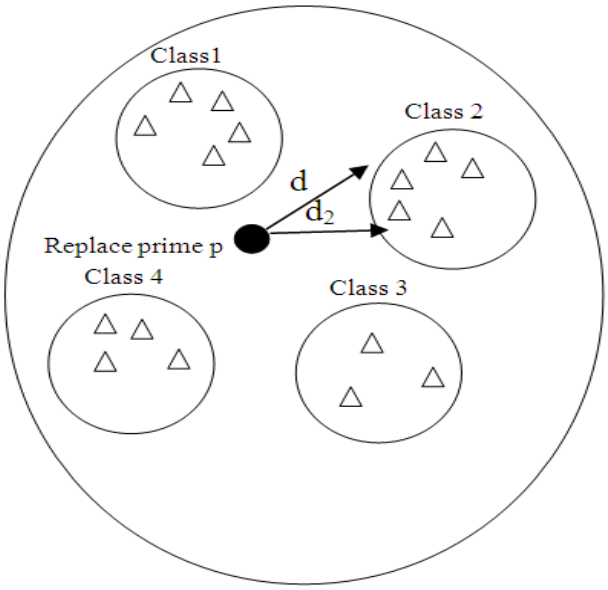

Fig.1 : The Nearest Prime Number System .

ISSN 2348 – 7968

Upon selecting the alternative prime number p , let's say p' , then we easily compute the alternative value of n denoted by n'. In such a case, this procedure produces agreement secure parameters. In order to acknowledge the receiver by changing n, the sent ciphertex must be appended by a secure agreement parameter, denoted by f, inside the ciphetext . This suggested parameter is used to prevent sending the value of alternative value as a public key, so we get a more secure procedure for RSA algorithm by reducing the public key into one parameter that is the public key of the user only because in the traditional method of RSA, the public key consists of two parameter; the value of n and the public of the user (e).

ALGORITHM

Let PR is the set of prime numbers p and q such that PR={p1,q1,p2,q2,…..p_n,q_n}
Divide PR into subsets S= {s1,s2,….Sn} such that each Si contains a limited numbers of primes .
Each S_i={p_{si1},p_{si2}, …..p_{sin}}
Choose two prime numbers from PR
*n= p*q*
Ø (n)=(p-1)(q-1)*
Let e be the public key
Let d be the private key
c= m^e mod n.
if c=m then
Sender operation:
1: Choose d_1 of the one of subsetssi in S for the secure class
2: Choose d2 inside si to pick one alternative prime p'
*3: Compute n'=p'*q*
4: Compute Ø (n')=(p'-1)(q-1)*
5: Choose alternative public key , lets e'
6: Generate the corresponding private key d'
7: Compute the ciphertext C'=$m^{e'}$ mod n'
8: Combine the agreement factor f with the new ciphetext and send C"" as :
C''=[C',f]

4. Conclusion

In this method we reduce the redundant messages occurred in RSA method. We see that for some values of n, there is a major problem in which the message and its corresponding ciphetext are the same. At the presence of recent active attacks, this problem can be exploited by many attackers. For this reason, this method presents an active solution by changing the value of n. In this case we need to apply the k-nearest neighbour values of either p or q or both. By applying this proposed system, the security of RSA algorithm is enhanced and we get a more secure algorithm by changing the value of n which can not be known except by the authorized users who only know the agreement factor. Another important property of this system is that we change the component of public key structure of the traditional RSA by sending only the public key (e) of the user without sending the value of n. To get a more secure system, it is important to change the distances applied to select the nearest neighbor periodically.

References

[1] R. L. Rivest, A. Shamir and L. Adleman “A method for obtaining digital signatures and public – key cryptosystems” Communications of the ACM, vol. 21,pp. 120 - 126, 1978.
[2] W. Stallings “Network and internetwork security: principles and practice” Prentice - Hall, Inc., 1995.
[3] W. Stallings “Network security Essentials: Applications and Standards” Pearson Education India, 2000.
[4] J. Joshi, et al. “Network Security” Morgan Kaufmann, 2008.
[5] W. Stallings “Cryptography and network security vol. 2” prentice hall, 2003.
[6] R. Stinson, “Cryptography - Theory and Practice”, CRC Press, 1995.
[7] A. J. .Menzes.,C.Paul, , V. Dorschot, S. A. Anstone, “Handbook of Applied Cryptography”, CRS press 5th Printing; 2001.
[8] A. Tragha., F. Omary, and A. Mouloudi,”ICIGA: Improved Cryptography Inspired by Genetic Algorithms”, Proceedings of the International Conference on Hybrid Information Technology (ICHIT'06), pp. 335 -341, 2006.IEEE.
[9] B.Harsh,and A. Nakul, “Reliability Infocom Technology and Optimization ”, Conference Proceedings pages 226 – 230, 2010.
[10] G. Nils, P.Arun, W.Arvinderpal, E.Hans, and C.Sheueling, “Comparing Elliptic Curve Cryptography and RSA on 8- Bit CPUs ” M. Joye and J.- J. Quisquater (Eds.): CHES 2004, LNCS 3156, pp. 119–132, 2004.
[11] H. Jeffrey, P.Jill, and H .Joseph “An Introduction to Mathematical Cryptography ”, Springer Science +Business Media, LLC, 2008.
[12] N. Nedjah and L.M. Mouller, “Hihgh-performance hardware of the sliding-window method for parallel computation of modular exponentitions,” international journal of parallel programming, Springer Netherlands, vol.37, pp.537-555, 2009.
[13] S. R. Dusse, B. S. Kaliski , “A cryptographic library for the Motorola DSP 56000,” Advance in Cryptology Proceedings of UROCRYPT’90, LNCS,vol.73, pp. 230-244, 1990.
[14] P. Keshavarzi and C. Harrison, “A new modular multiplication algorithm for VLSI implementation of public-key cryptography,” Proceedings of First International Symposium on Communication Systems and Digital Signal Processing, pp.516-519, 1998.
[15] K. Sakiyama, L. Batina, B. Preneel and I. Verbauwhede, “ High-performance publik-key cryptoprocessor for wireless mobile applications,” Mobile networks and applications , vol. 99, pp. 245-258, 2007.
[16] A. Rezai and P. Keshavarzi, “Improvement of high-speed modular exponentiation algorithm by optimum using smart methods, ” Proceedings of 18th Iranian Conference on Electrical Engineering, Iran, pp.2104-2109, May 2010.
[17] A. Rezai and P. Keshavarzi, “Speed Improvement in elliptic curve cryptosystem scalar multiplication algorithm,” proceedings of 7th International ISC Conference on Information Security and Cryptology2010, Iran, pp.181-188, September 2010.
[18] A.F.Tenca and C.K.Koc, “A scalable architecture for modular multiplication based on Montgomery’s algorithm,”IEEE Trans. On computer, vol.52, no.9, pp. 1215-1221, 2003.
[19] N.Pinckney, P. Amberg and D. Harris, “Parallelized Booth-encoded radix-4 Montgomery multipliers,” proceeding of 16th IFIP/IEEE IntlernationalConferene on Very Large Scale Integration, Oct. 2008.
[20] J.C. Ha, S.J. Moon, “A common-multiplicand method to the Montgomery algorithm for speeding up exponentiation,” Information Processing Letters, vol.66, no.2, pp.105–107,1998.
[21] C.Wu, D.Lou and T.Chang, “An efficient Montgomery exponentiation algorithm for public-key cryptosystem,” Proceedings of

ISSN 2348 – 7968

IEEE international conference on intelligence and security information, pp.284-285, June 2008.
[22] C.Wu, “An efficient common-multiplicand-multiplication method to the Montgomery algorithm for speeding up exponentiation,” Information Sciences, vol.179, pp.410-421, 2009.
[23] K. Ming Leung , k-Nearest Neighbor Algorithm for Classification , POLYTECHNIC UNIVERSITY Department of Computer Science / Finance and Risk Engineering , 2007.
[24] C .Aayush ,and M . Srushti ," Modified RSA Algorithm: A Secure Approach ", CSDL HomeCCICN2011Computational Intelligence and Communication Networks, International Conference on 2011.
[25] V.Kuldeep, Kr.Rajesh,and C .Ritika , "Modified Prime Number Factorization Algorithm (MPFA) For RSA Public Key Encryption", International Journal of Soft Computing & Engineering 2012 .
[26] R.S. Dhakar, and P. Sharma, " Modified RSA Encryption Algorithm (MREA)" , Advanced Computing & Communication Technologies (ACCT), 2012 Second International Conference on , Page(s): 426 – 429, 7-8 Jan. 2012 .
[27] B.Persis Urbana Ivy, PurshotamMandiwa.MukeshKumar , A modified RSA cryptosystem based on ‘n’ prime numbers , International Journal Of Engineering And Computer Science ISSN:2319-7242 , Volume1 Issue 2 Page No.63-66, Nov 2012.