Secure Encryption With Digital Signature Approach For SMS And Spam Detection Using SVM

[1]Shaheda Bano ,[2]Abu Rehan
[1]Research scholar, [2]Lecturer
[1]Department of Electronics and Communication Engineering,
Al-falahUniversity ,Dhauj ,Faridabad, Haryana, India

***Abstract* - Global System for Mobile (GSM) is a second generation cellular standard developed to cater voice services and data delivery using digital modulation. Short Message Service (SMS) text messages are indispensable in our lives today, but along with the convenience of using SMS messages in our daily lives, we also face a serious problem caused by SMS spamming. However their security is a critical issue cumbering their application and development. The current SMS hasn't achieved secure transmission of plaintext between different mobile phone devices. SMS doesn't have its own build-in mechanism to secure the transmitted data because security isn't considered as a priority application for mobile devices. This paper analyses the most popular digital signature algorithms such as RSA and ElGamal technique for encryption and decryption. SMS spam checking is also done by using Simple vector machine (SVM) technique and simulation is carried on MATLAB .SVM proves to be a great classifier and yields an overall accuracy of about 97-98%.**

***Key words* - GSM, SMS, spam, encryption, decryption, RSA, ElGamal, SVM.**

I. INTRODUCTION

Short text messages (SMS) are an important means of communication today between millions of people around the world. SMS services, which are a must-have service nowadays for telecom operators, transmit their messages using standardized communications protocols. At the same time, SMS messaging has become a perfect target for abuse via what are known as spamming - the misuse behavior of SMS to achieve some harmful purposes. Spamming is a serious problem for SMS today, as it is for emails and social networking services, because it disrupts people's daily life and harms well-being of telecom operators. In order to secure messages cryptography techniques are applied. The security framework solution allow us to provide strong customer authentication and non-repudiation by employing digital signature and spam detection.

II. SMS ARCHITECTURE

Short messages are delivered in GSM signaling channels between the Mobile Station (MS) and the Base Transceiver Station (BTS). The messages flow as normal calls, but they are routed from the MSC to a Short Message Service Center (SMSC). The SMSC stores the message until it can be delivered to the recipient(s) or until the message's validity time has elapsed. The recipient can be a normal MS user or a SMS gateway. The gateways are servers that are connected to one or more SMSCs to provide SMS applications for the MS users. These applications include ring tone and icon delivery, entertainment, bank services, and many other beneficial services.

Figure 1: SMS Architecture

III. SMS SECURITY

Data security has at least four important requirements to meet, as listed below:

(1) Secrecy: Also known as confidentiality. It is the effect of keeping unauthorized parties from accessing private information. Interception is the typical direct attack on secrecy.

(2) Integrity: It is preventing anybody other that authorized parties from modifying the computer system assets like writing, changing status and deleting and creating files. Among the methods of attacking integrity we found modification, replay and reordering of messages.
(3) Availability: It is the fact of being able to access information when needed and the prevention of unauthorized parties from withholding access to information. Inception and denial of service are the attacks over availability.
(4) Authenticity: Prevents that unauthorized parties can change the content of message or place random messages in the network.

IV. CRYPTOGRAPHY

Cryptography is an art of information security, whose fundamental objective is the security of the message which is being transferred through an insecure medium between two parties i.e. sender and the receiver. There are two types of cryptography algorithms; namely private or symmetric key algorithm and public or asymmetric key algorithm. In private or symmetric key algorithm the key of encryption and decryption is same. While in public or asymmetric key algorithm the key of encryption is the public key (which is mutually discussed between sender and the receiver) and the decryption key.

Public Key and Private Keys: The Public and Private Key pair comprise of two uniquely related cryptographic keys basically long random numbers. The Public Key is what its name suggests – Public, It is made available to everyone via a publicly accessible repository or directory. On the other hand, the Private Key must remain confidential to its respective owner. Because the key pair is mathematically related, whatever is encrypted with a Public Key may only be decrypted by its corresponding Private Key and vice versa . Public key cryptography is used where each user has a pair of keys, one called the public key and the other private key. Each user's public key is published while the private key is kept secret and thereby the need for the sender and the receiver to share secret information (key) is eliminated. The only requirement is that public keys are associated with the users in a trusted (authenticated) manner using a public key infrastructure (PKI).

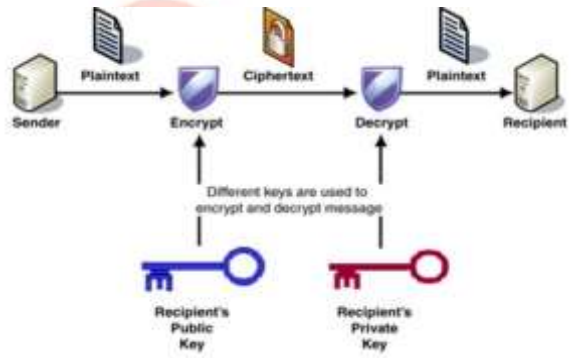

Figure 2: Public Key cryptography

V. APPROACH FOR SMS SECURITY

Our main approach is to provide security by encrypting and decrypting the desired message using RSA and ElGamal technique. For this we have to first used the Hashing function for sent and received messages. Then uses the RSA cryptosystem with MD5 for authentication and then ElGamal cryptosystem for encryption and decryption of the message. By using these algorithms the message becomes more secure and very tough to decrypt by any intruder. At the same time we are using the SVM classifier for spam detection that whether the message is altered by an intruder or not. Confidentiality and integrity of a message can be checked by using SHA-1,SHA-2 and MD5.

VI. RESEARCH METHODOLOGY

Public key cryptography is employed on the sent/received message.RSA and ElGamal encryption/decryption scheme is discussed below:

- **RSA Cryptosystem**
  This cryptosystem is one the initial system. It remains most employed cryptosystem even today. The system was invented by three scholars Ron Rivest, Adi Shamir, and Len Adleman and hence, it is termed as RSA cryptosystem. We will see two aspects of the RSA cryptosystem, firstly generation of key pair and secondly encryption-decryption algorithms.

1) Generation of RSA Key Pair

- Generate the RSA modulus (n)
  a. Select two large primes, p and q.
  b. Calculate n=p*q. For strong unbreakable encryption, let n be a large number, typically a minimum of 512 bits.
- Find Derived Number (e)
  a. Number e must be greater than 1 and less than $(p-1)(q-1)$.

b. There must be no common factor for e and (p − 1)(q − 1) except for In other words two numbers e and (p – 1)(q – 1) are coprime.

- Form the public key
  a. The pair of numbers (n, e) form the RSA public key and is made public.
  b. Interestingly, though n is part of the public key, difficulty in factorizing a large prime number ensures that attacker cannot find in finite time the two primes (p & q) used to obtain n. This is strength of RSA.
- Generate the private key
  a. Private Key d is calculated from p, q, and e. For given n and e, there is unique number d.
  b. Number d is the inverse of e modulo (p − 1)(q – 1). This means that d is the number less than (p − 1)(q − 1) such that when multiplied by e, it is equal to 1 modulo (p − 1)(q − 1).
  c. This relationship is written mathematically as follows:

$$ed = 1 \bmod (p-1)(q-1)$$

2) Encryption and Decryption

A. RSA Encryption

- Suppose the sender wish to send some text message to someone whose public key is (n, e).
- The sender then represents the plaintext as a series of numbers less than.
- To encrypt the first plaintext P, which is a number modulo n. The encryption process is simple mathematical step as:

$$C = Pe \bmod n$$

- In other words, the cipher text C is equal to the plaintext P multiplied by itself e times and then reduced modulo n. This means that C is also a number less than n.
- Returning to our Key Generation example with plaintext P = 10, we get ciphertext C:

$$C = 105 \bmod 91$$

B. RSA Decryption

- The decryption process for RSA is also very straightforward. Suppose that the receiver of public-key pair (n, e) has received a cipher text C.
- Receiver raises C to the power of his private key d. The result modulo n will be the plaintext P

$$Plaintext = Cd \bmod n$$

- Returning again to our numerical example, the cipher text C = 82 would get decrypted to number 10 using private key 29:

$$Plaintext = 8229 \bmod 91 = 10$$

➢ **ElGamal Cryptosystem**

ElGamal cryptosystem, called Elliptic Curve Variant, is based on the Discrete Logarithm Problem. It derives the strength from the assumption that the discrete logarithms cannot be found in practical time frame for a given number, while the inverse operation of the power can be computed efficiently. Let us go through a simple version of ElGamal that works with numbers modulo p. In the case of elliptic curve variants, it is based on quite different number systems.

1) Generation of ElGamal Key Pair

- Choosing a large prime p.
  Generally a prime number of 1024 to 2048 bits length is chosen
- Choosing a generator element g.
  a) This number must be between 1 and p − 1, but cannot be any number.
  b) It is a generator of the multiplicative group of integers modulo p. This means for every integer m co-prime to p, there is an integer k such that gk=a mod n.
- Choosing the private key.
  The private key x is any number bigger than 1 and smaller than p−1.
- Computing part of the public key.
  The value y is computed from the parameters p, g and the private key x follows:

$$y = gx \bmod p$$

- Obtaining Public key. The ElGamal public key consists of the three parameters (p, g, y)
  For example, suppose that p = 17 and that g = 6 (It can be confirmed that 6 is a generator of group Z17). The private key x can be any number bigger than 1 and smaller than 71, so we choose x = 5. The value y is then computed as follows:

$$y = 65 \bmod 17 = 7$$

- Thus the private key is 62 and the public key is (17, 6, 7).

2) Encryption and Decryption

The generation of an ElGamal key pair is comparatively simpler than the equivalent process for RSA. But the encryption and decryption are slightly more complex than RSA.

A. ElGamal Encryption

Suppose sender wishes to send a plaintext to someone whose ElGamal public key is (p, g, y), then:

- Sender represents the plaintext as a series of numbers modulo p.
- To encrypt the first plaintext P, which is represented as a number modulo p. The encryption process to obtain the ciphertext C is as follows:
  a) Randomly generate a number k;
  b) Compute two values C1 and C2, where:

$$C1 = g^k \bmod p$$

$$C2 = (P * y^k) \bmod p$$

  c) Send the ciphertext C, consisting of the two separate values (C1, C2), sent together
  d) Referring to our ElGamal key generation example given above, the plaintext P = 13 is encrypted as follows:
- Randomly generate a number, say k = 10
- Compute the two values C1 and C2, where:

$$C1 = 610 \bmod 17$$

$$C2 = (13*710) \bmod 17 = 9$$

- Send the ciphertext C = (C1, C2) = (15, 9).
- Send the ciphertext C = (C1, C2) = (15, 9).

B. ElGamal Decryption

- To decrypt the cipher text (C1, C2) using private key x, the following two steps are taken:
  a) Compute the modular inverse of $(C1)^x$ modulo p, which is $(C1)^{-x}$, generally referred to as decryption factor.
  b) Obtain the plaintext by using the following formula:

$$C2 \times (C1)^{-x} \bmod p = \text{Plaintext}$$

  c) In our example, to decrypt the ciphertext C = (C1, C2) = (15, 9) using private key x = 5, the decryption factor is

$$15^{-5} \bmod 17 = 9$$

  d) Extract plaintext $P = (9 \times 9) \bmod 17 = 13$.

➢ SECURITY MECHANISM

Security mechanism such as Hash functions are used to tackle the active modification threats. Hash functions are extremely useful and appear in almost all information security applications. A hash function is a mathematical function that converts a numerical input value into another compressed numerical value. The input to the hash function is of arbitrary length but output is always of fixed length. Values returned by a hash function are called message digest or simply hash values.

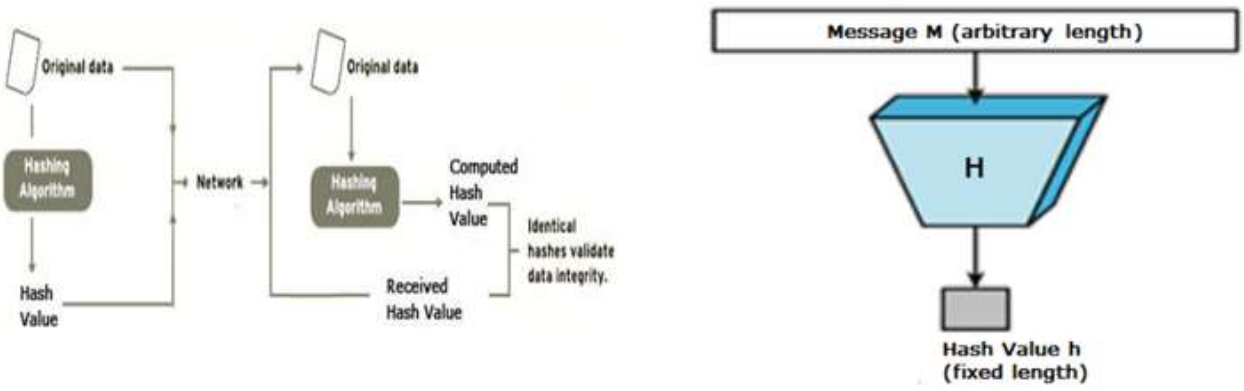

Figure 3. Data integrity Check & Message Digest

➢ SMS Spam checking using SVM technique

In this paper we design and implement a technique to check spam in SMS by using Simple Vector Machine (SVM) technique and then the result will shown through MATLAB programming. SVM is a supervised learning method in which the data used for training of the algorithm is labeled as to which class it belongs. Using the labeled data, the

algorithm learns the relationship between the feature sets and the output, and hence it then classifies the unlabeled data from the learned relationship.

VII. RESULT AND DISCUSSION

This project gives security to the transmitted message by using RSA and ElGamal encryption and decryption techniques. We observed a good improvement in spam classification precision using the SVM classifer.SVM yields an overall accuracy of 97-98%.

Figure 5: Authentication checking by Hash Function

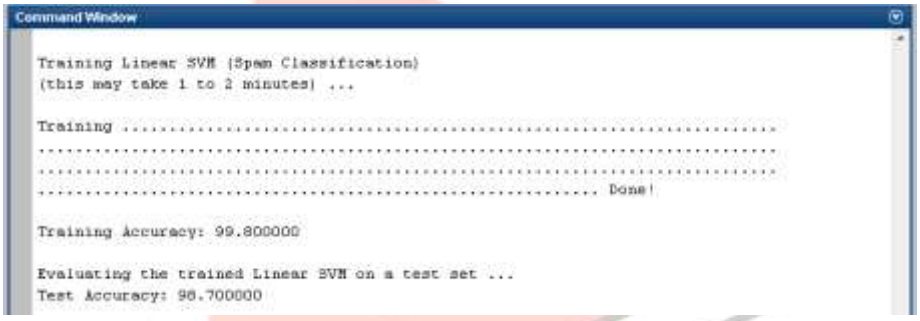

Figure 6: Spam detection by SVM classifier

VIII. CONCLUSION

The approach for securing of SMS has been designed and implemented.RSA is used for authentication and ElGamal encryption/decryption technique has been applied to make it more secure. For making the project more impressive ,a qualitative approach has adopted. Use of SVM technique for spam detection makes the project more useful for future use also. It yields the accuracy of about 97-98%.

REFERENCES

[1] Zhang Wei and Wang Zi-Xuan. GSM spam SMS filtering solution. Telecom Express: Networking and Communications (In Chinese), 3, 2009.
[2] Mary Agoyi, Devrim Seral, "SMS Security: An Asymmetric Encryption Approach", Sixth International Conference on Wireless and Mobile Communications, 2010@IEEE, pp. 448-452.
[3] A. K. Nanda and L. K. Awasthi, XTR Cryptosystem for SMS Security, International Journal of Engineering and Technology, IJET, 4(6), 2012, 1793-8244.
[4] Urbana Ivy B.Persis, Mandiwa Purshotam. Kumar Mukesh "A modified RSA cryptosystem based on 'n' prime numbers" Volume1 Issue 2 Nov 2012, IJECS.
[5] J. Gordon, "Strong RSA Keys" IEEE Electronics Letters, 20, (12), 1984, pp. 514-516.
[6] SMS Spam and Mobile Messaging Attacks Introduction, Trends and Examples. [Online], Jan. 2011.
[7] A. K. Nanda and L. K. Awasthi, XTR Cryptosystem for SMS Security, International Journal of Engineering and Technology, IJET, 4(6), 2012, 1793-8244.
[8] M. T. Mohammed, A. E. Rohiem, A. Elmoghazy, and A. Ghalwash, Chaotic Encryption Based PGP Protocol, 2013.