IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 12, Issue 1 (May. - Jun. 2013), PP 16-22
www.iosrjournals.org

Generate an Encryption Key by using Biometric Cryptosystems to secure transferring of Data over a Network

Sayani Chandra[1], Sayan Paul[2], Bidyutmala Saha[3], Sourish Mitra[4]
[1](Department of Computer Science & Engineering, Guru Nanak Institute of Technology, India)
[2](Department of Computer Science & Engineering, Guru Nanak Institute of Technology, India)
[3](Department of Computer Science & Engineering, Guru Nanak Institute of Technology, India)
[4](Department of Computer Science & Engineering, Guru Nanak Institute of Technology, India)

Abstract: *Our modern day era faces an inevitable problem of securing our most integrated data and messages. The chief problem is to protect our data in a unique way that could only be worked upon by the sender and the recipient. Most of the traditional techniques probably in-use today, emphasizes on keys that are generated by generic function, algorithms or in random key generators. But question remains whether this key is unique and authentic in nature. Moreover how can these keys be unique to one and one person only? The answer to this would be Biometric Cryptosystems. Biometric Cryptosystems are the newest members in the field of security. The very basis of this Biometric Cryptosystem lies on the very fact that some features of human body are significantly unique to each and every human in the world, such as fingerprint, DNA sequence, Iris, etc. Using those biometric we can generate an exclusive key that will be unique for each and every individual. Now having generated these keys we can use them for encrypting our message. And as because these keys are uniquely generated for individual persons there's no chance of there will be a matching keys. Moreover as we use RSA algorithm based encryption technique so the encryption lies on two basic sets of keys to decrypt the message. Hence, eavesdropper or third unwanted parties have to acquire two set of keys which adds up to security level of the encryption and hence protect the message from unwanted third parties from acquiring our secret message.*

Keywords - *Bio-metric cryptosystem, Encryption, Fingerprint, Minutiae point, Private Key, Public Key, RSA*

I. Introduction

As the Internet and other forms of electronic communication become more prevalent, electronic security is becoming increasingly important. Cryptography is used to protect e-mail messages, credit card information, and corporate data. The art of protecting information by transforming it (encrypting it) into an unreadable format, called cipher text. Only those who possess a secret *key* can decipher (or decrypt) the message into plain text. Encrypted messages can sometimes be broken by cryptanalysis, also called *codebreaking*, although modern cryptography techniques are virtually unbreakable. Cryptography systems can be broadly classified into symmetric-key systems that use a single key that both the sender and recipient have, and public-key systems that use two keys, a public key known to everyone and a private key that only the recipient of messages uses.

1.1 Symmetric Encryption

Symmetric encryption is the oldest and best-known technique. A secret key, which can be a number, a word, or just a string of random letters, is applied to the text of a message to change the content in a particular way. This might be as simple as shifting each letter by a number of places in the alphabet. As long as both sender and recipient know the secret key, they can encrypt and decrypt all messages that use this key.

1.2 Asymmetric Encryption

The problem with secret keys is exchanging them over the Internet or a large network while preventing them from falling into the wrong hands. Anyone who knows the secret key can decrypt the message. One answer is asymmetric encryption, in which there are two related keys--a key pair. A public key is made freely available to anyone who might want to send you a message. A second, private key is kept secret, so that only you know it. Any message (text, binary files, or documents) that are encrypted by using the public key can only be decrypted by applying the same algorithm, but by using the matching private key. Any message that is encrypted by using the private key can only be decrypted by using the matching public key. This means that you do not have to worry about passing public keys over the Internet (the keys are supposed to be public). A problem with asymmetric encryption, however, is that it is slower than symmetric encryption. It requires far more processing power to both encrypt and decrypt the content of the message.

II. Methodology

2.1 Biometric Crypto Systems

Cryptography provides the secure manner of information transmission over the insecure channel. It authenticates messages based on the key but not on the user. It requires a lengthy key to encrypt and decrypt the sending and receiving the messages, respectively. But these keys can be guessed or cracked. Moreover, Maintaining and sharing lengthy, random keys in enciphering and deciphering process is the critical problem in the cryptography system. The above mentioned problem is solved by a Biometric cryptosystems. Biometric cryptosystems combine cryptography and biometrics to benefit from the strengths of both fields. In such systems, while cryptography provides high and adjustable security levels, biometrics brings in non-repudiation and eliminates the need to remember passwords or to carry tokens etc. In biometric cryptosystems, a cryptographic key is generated from the biometric template of a user stored in the database in such a way that the key cannot be revealed without a successful biometric authentication.

2.2 RSA Algorithm

RSA is an algorithm for public-key cryptography that is based on the presumed difficulty of factoring large integers, the factoring problem. RSA stands for Ron Rivest, Adi Shamir and Leonard Adleman, who first publicly described the algorithm in 1977. Clifford Cocks, an English mathematician, had developed an equivalent system in 1973, but it was classified until 1997. A user of RSA creates and then publishes the product of two large prime numbers, along with an auxiliary value, as their public key. The prime factors must be kept secret. Anyone can use the public key to encrypt a message, 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. Whether breaking RSA encryption is as hard as factoring is an open question known as the RSA problem.

2.3 Operation

The RSA algorithm involves three steps: key generation, encryption and decryption.

2.4 Key generation

RSA 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 in a reasonable amount of time using the private key. The keys for the RSA algorithm are generated the following way:

1. Choose two distinct prime numbers p and q.
   - For security purposes, the integers p and q should be chosen at random, and should be of similar bit-length. Prime integers can be efficiently found using a primality test.
2. Compute n = pq.
   - n is used as the modulus for both the public and private keys. Its length, usually expressed in bits, is the key length.
3. Compute $\varphi(n) = \varphi(p)\varphi(q) = (p - 1)(q - 1)$, where φ is Euler's totient function.
4. Choose an integer e such that $1 < e < \varphi(n)$ and $\gcd(e, \varphi(n)) = 1$; i.e., e and φ(n) are coprime.
   - e is released as the public key exponent.
   - e having a short bit-length and small Hamming weight results in more efficient encryption – most commonly $216 + 1 = 65{,}537$. However, much smaller values of e (such as 3) have been shown to be less secure in some settings.[4]
5. Determine d as $d^{-1} \equiv e \pmod{\varphi(n)}$, i.e., d is the multiplicative inverse of e (modulo φ(n)).
   - This is more clearly stated as solve for d given $de \equiv 1 \pmod{\varphi(n)}$
   - This is often computed using the extended Euclidean algorithm.
   - d is kept as the private key exponent.

By construction, $d \cdot e \equiv 1 \pmod{\varphi(n)}$. The public key consists of the modulus n and the public (or encryption) exponent e. The private key consists of the modulus n and the private (or decryption) exponent d, which must be kept secret. p, q, and φ(n) must also be kept secret because they can be used to calculate d.

2.5 Encryption

Alice transmits her public key (n, e) to Bob and keeps the private key secret. Bob then wishes to send message M to Alice.

He first turns M into an integer m, such that $0 \leq m < n$ by using an agreed-upon reversible protocol known as a padding scheme. He then computes the cipher text c corresponding to:

$$c \equiv m^e \pmod{n}. \qquad (1)$$

This can be done quickly using the method of exponentiation by squaring. Bob then transmits c to Alice.

2.6 Decryption

Alice can recover m from c by using her private key exponent d via computing

$$m \equiv c^d \pmod{n}. \qquad (2)$$

Given m, she can recover the original message M by reversing the padding scheme.

(In practice, there are more efficient methods of calculating cd using the precomputed values below.)

III. Proposed Work

3.1 Cryptographic Key Generation from Biometrics

In our approach we have selected fingerprint as the biometrics feature for generating cryptographic key. We have put into the thought of representing the scanned fingerprint in form a matrix which contains information about the number of ridges and furrows in a small region and we have considered of marking those ridges and furrows into set of data for individual matrix elements which have been subdivided into smaller regions. So each element in the scanned fingerprint matrix is actually a set of number of ridges and furrows extracted by minutiae points recognition technique.

3.1.1 Extracting Minutiae Points From Fingerprint

Extracting minutiae points requires three-stage approach. These are

a) Preprocessing - At this stage the fingerprint of an individual is scanned using a Scanning machine using ink-pigment representation of the fingerprint and the information is fed into computer.
b) Minutiae Extraction - The stage includes recognizing the different minutiae points of the scanned fingerprint. There is different kind of minutiae of a fingerprint such as-Ridge ending, Ridge bifurcation, Crossover, Island, etc. We will be using Ridge and Furrows which are distinctive to each other.
c) Post processing - The stage includes ridge thinning which is required for minutiae recognition. [3] [13]

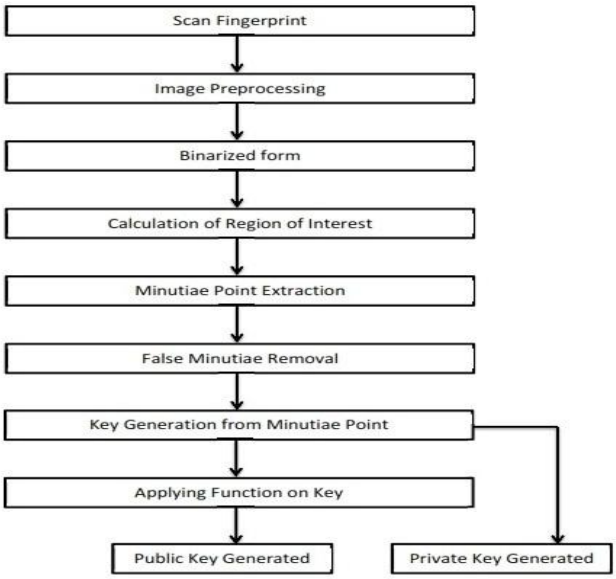

Fig 1 : Generation of Private and Public key from Fingerprint

3.1.2 Binarization

Binarization of the scanned fingerprint requires transforming the 8-bit Gray fingerprint image to a 1-bit image with 0-value for ridges and 1-value for furrows. The operation highlights ridges in the fingerprint with black colour while furrows are white. A locally adaptive binarization method is performed to binarize the fingerprint image.

Such a named method comes from the mechanism of transforming a pixel value to 1 if the value is larger than the mean intensity value of the current block (16x16) to which the pixel belongs.

3.1.3 ROI Extraction By Morphological Operations

The 'OPEN' operation can expand images and remove peaks introduced by background noise. The 'CLOSE' operation can shrink images and eliminate small cavities. The bound is the subtraction of the closed area from the opened area. Then the algorithm throws away those leftmost, rightmost, uppermost and bottommost blocks out of the bound so as to get the tightly bounded region just containing the bound and inner area.

3.1.4 Minutiae Points Extraction

Ridge Thinning is to eliminate the redundant pixels of ridges till the ridges are just one pixel wide. The method uses an iterative, parallel thinning algorithm. In each scan of the full fingerprint image, the algorithm marks down redundant pixels in each small image window (3x3). And finally removes all those marked pixels after several scans. After the fingerprint ridge thinning, marking minutia points is relatively easy. For each 3x3 window, if the central pixel is 1 and has exactly 3 one-value neighbours, then the central

pixel is a ridge branch. If the central pixel is 1 and has only 1 one-value neighbour, then the central pixel is a ridge ending. Suppose both the uppermost pixel with value 1 and the rightmost pixel with value 1 have another neighbour outside the 3x3 window, so the two pixels will be marked as branches too. But actually only one branch is located in the small region, so a check routine requiring that none of the neighbours of a branch are branches is added. [3] [13]

3.1.5 False Minutiae Removal

False Minutiae removal is very necessary for the key generation; otherwise false key would be generated. Such condition occurs due insufficient amount of ink impression and cross-overs of ridges. Actually all the earlier stages themselves occasionally introduce some artifacts which later lead to spurious minutia. This false minutia will significantly affect the accuracy of matching if they are simply regarded as genuine minutia. So some mechanisms of removing false minutia are essential to keep the fingerprint verification system effective.

3.1.6 Key Generation From Minutiae Points

In this Section we explain the Key Generation
Algorithm Assumptions
Mp → Minutiae point set
Kl → Key length
Np → Size of Minutiae point set
S → Seed value
Sl → seed limit.
m → (x,y) – co-ordinate of a minutiae point
Kv → Key Vector

Step 1 : The Extracted minutiae points are represented as:
$Mp = \{ m_i \} i=1,\ldots, Np$ (3)

Step 2 : The initial key vector is defined as follows:
$Kv = \{ x_i : p(x_i) \}\ i=1, \ldots.., Kl$ (4)
Where, $p(x) = Mp[I \% Np] + Mp[(i+1) \% Np] + S$ (5)
$i=1,\ldots. , Kl$

Step 3 : Initial value of S is equal to total Number of Minutiae points. The value of S will be dynamically changed as follows:
$S = Kv (i) \% Sl , -1 < i< Kl$ (6)

Step 4 : Initial key vector (Kv) is converted in to a matrix Km of size Kl / 2 * Kl / 2 as follows:
$Km = (aij)\ Kl / 2 * Kl / 2$ (7)

Step 5 : An intermediate key vector is generated as follows:
$KIV = \{ Ki : (m(k_i) \}\ i=1,\ldots. Kl$ (8)
Where
$m(k) = | A_{ij} |$,
$A_{ij} = Km_{i,j} : i+size,j+size , -1 < i< Kl/2$
A_{ij} is a submatrix formed from the key matrix.

Step 6 : Final key vector (Private key) formed is:
Kv = 1, if KlV [i] > mean(KlV) (9)
0, otherwise

3.1.7 Mapping each binary data precisely to each region

Although we know, that no two fingerprints are similar to one another but there's a whole lot of chances where the number of ridge may be equal to the number of furrows. So we use a nxn matrix that precisely stores each furrows and ridges markings in the byte pattern accordingly to the scanned image of the fingerprint. This helps to recognise individual fingerprint distinctively as the data in the matrix form will have different patterns of data set (i.e. matrix containing key vector elements from the above algorithm).

3.1.8 Calculation of public key

Let,
d→ be the total number of 1s in the data set due to furrows in the matrix $A_{i,j}$.
e → be the total number of 0s in the data set due to ridges in the matrix $A_{i,j}$.

i,j → 1,2,3,…Np.
s = (d – e)
Pb → Public key vector
Pb = Kv * (mod (s)) * e (10)
e → public key vector

3.2 Speech (Voice Signal) to Digital Data Conversion

In our paper we would like to propose the new technology of speech to text conversion as a method of extracting message and data rather than normal traditional digital data techniques of text, image sharing. We accessorized the technique of speech to text conversion used in modern day technologies such as Voice Recognition, Google speech search, etc. which uses a whole bunch of algorithms and probability theorems such as acoustic model and language model that determines the set of words and probabilistically determines other words for the text conversion.

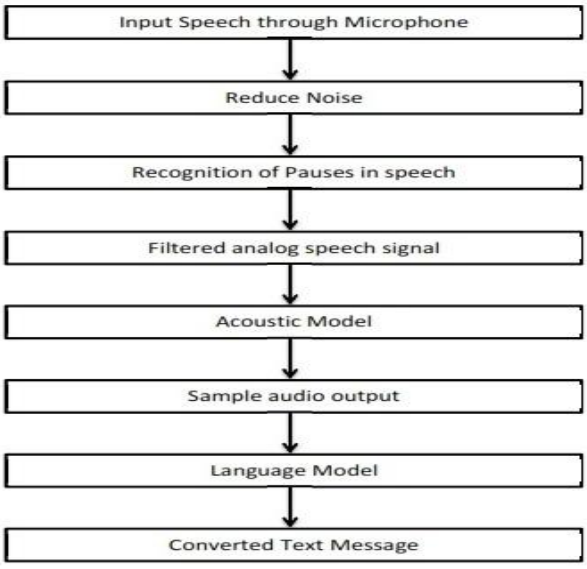

Fig 2 : Generation of Text message from Speech

3.2.1 Input Speech through Microphone

First of all we have to input the speech i.e. the analog signal using a microphone. This signal is very necessary for the digitalization of our secret message that needs to be saved temporarily for further conversion.

3.2.2 Reduce Noise

Next we have to filter the analog signal so as to remove the noise and disturbances from our message, otherwise this would produce inefficient message sequence.

3.2.3 Recognition of Pauses in Speech

Pauses are integral part of a speech. This provides us with the sense of correct sequence of the sentence, so we have to correctly detect the pause sequence in our speech. Such recognition can be assumed properly by many algorithms or sampling methods.

3.2.4 Acoustic Model

An acoustic model is created by taking audio recordings of speech, and their text transcriptions, and using software to create statistical representations of the sounds that make up each word. It is used by a speech recognition engine to recognize speech. Speech recognition engines require two types of files to recognize speech. They require an acoustic model, which is created by taking audio recordings of speech and their transcriptions (taken from a speech corpus), and 'compiling' them into a statistical representations of the sounds that make up each word (through a process called 'training').

3.2.5 Language Model

A statistical language model assigns a probability to a sequence of m words by means of a probability distribution.

Language modeling is used in many natural language processing applications such as speech recognition, machine translation, part-of-speech tagging, parsing and information retrieval.

In speech recognition and in data compression, such a model tries to capture the properties of a language, and to predict the next word in a speech sequence. When used in information retrieval, a language model is associated with a document in a collection.

3.2.6 Text Message

As the analog Signal i.e. the speech passes all through this model a text is generated. We shall use this text message as the message to be sent in the Encryption model that uses RSA algorithm based technique for encryption.

3.3 Encrypt Digital Data by using Proposed Cryptographic Key

In our paper we have slightly revised the technique of using p, q two co-primes as the in case of RSA algorithm. We have generated two keys using Biometric Fingerprint System:

1. Private Key : This is generated directly from the biometric fingerprint using the algorithm mentioned earlier
2. Public Key : This is generated by superimposing the function as described in Step 3.1.8.

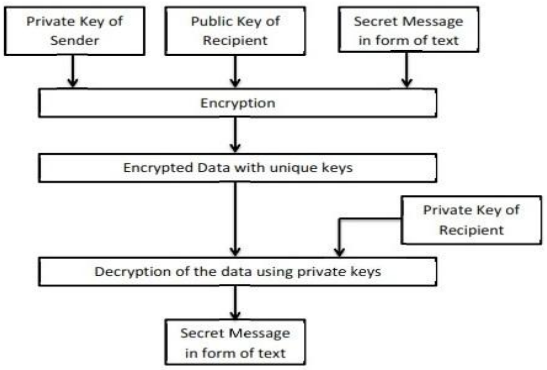

Fig 3 : Encryption and Decryption of the message

3.3.1 Private key of the Sender

This key is generated by the Biometric Fingerprint algorithm as described before. This is an unique key which is exclusive for any individuals. This key must be kept secret and should be only generated when required. This way the encryption becomes unique, authenticated and surprisingly difficult for hacking.

3.3.2 Public key of the recipient

This key is generated by superimposing the defined function (Step 3.1.8) which is applied on the private key of the recipient. The recipient of the secret message have to privately send this key to the sender using face-to-face or any other technique. This key adds upto the difficulty level of the encryption.

3.3.3 Message from the speech

The text message generated from the Speech-To-Text conversion is the secret message that requires to be encrypted. This text uses varied model of speech recognition but never in any case should modify the original message speech. Modification may result to inappropriate message transferring.+

3.3.4 Encryption

In this paper we propose a similar technique to that of the RSA algorithm that uses two set of keys i.e. a private key (of the sender) and public key (of the recipient) rather than traditional computation of random prime numbers as in case of RSA algorithm.

Pb_r →the public key of the recipient

e_r → the public key of the recipient that was chosen by him for calculation of the function

Pri_s → the private key of the sender

e_s → the public key of the sender that was chosen by him for calculation of the function

The recipient sends its public key along with its public key Pb_r and e_r to the sender. The sender uses a padding sequence i.e. a set of recognized protocols to first encrypt the text message (T) using his private key Pri_s to form an encrypted message (M_e). Then the sender uses the public key of the recipient, Pb_r and e_r to encrypt the M_e.

$C = M_e^{er} \pmod{Pb_r}$ (11)

This C is the cipher text that could be sent to the recipient for sending the text message.

3.3.5 Decryption

The recipient can recover the text using the private key of its own.

$Pri_r \rightarrow$ the private key of the recipient

$T = C^{Prir} \pmod{Pb_r}$ (12)

IV. Conclusion

In our proposed paper we have tried to scheme out an authenticated way of securing our data i.e. speech message using our biometric fingerprint cryptosystem that uses a two set of keys for encryption. The encryption technique we are using is very similar to the RSA encryption technique although we are using a unique set of keys that can only be generated by any one individual using its fingerprint. Moreover the two set of keys adds upto the security level and thus protect our invaluable data from third unwanted parties.

References

Journal Papers:

[1]. B. Goode, "Voice Over Internet Protocol (VOIP)". *Proceedings of three IEEE, VOL. 90, NO. 9,* Sept. 2002.

[2]. Announcing the "ADVANCED ENCRYPTION STANDARD (AES)" – *Federal Information, Processing Standards Publication 197*, November 26, 2001

[3]. D.Maio and D. Maltoni, "Direct gray-scale minutiae detection in fingerprints ". *IEEE Trans. Pattern Anal. And Machine Intell., 19(1)*:27-40, 1997

[4]. L.C. Jain, U.Halici, I. Hayashi, S.B. Lee and S. Tsutsui. "Intelligent biometric techniques in fingerprint and face recognition.", 1999, *the CRC Press*

Websites and PDFs:

[5]. "Breaking Through IP Telephony "http://www.nwfusion.com/reviews/2004/0524voipsecurity.html

[6]. "Voice over Internet Protocol" from http://en.wikipedia.org/wiki/Voice_over_IP

[7]. "Cisco IP Phones Compromise "http://www.syssecurity.com/archive/papers/The_Trivial_Cisco_IP_Phones_Compromise.pdf.

[8]. "Security Risk Factors in IP Telephony Based Networks "http://www.syssecurity.com/archive/papers/Security_Risk_Factors_with_IP_Telephony_based_Networks.pdf

[9]. "Security Testing of Protocol Implementations at the University of Finland "http://www.ee.oulu.fi/research/ouspg/protos/

[10]. "Advanced Encryption Standard "from http://en.wikipedia.org/wiki/Advanced_Encryption_ Standard

[11]. Umut Uludag, Sharath Pankanti, Salil Prabhakar, Anil K. Jain "Biometric Cryptosystems Issues and Challenges" *Proceedings of the IEEE 2004*.

[12]. "GaborFilter"from http://en.wikipedia.org/wiki/Gabor_filter

[13]. "Fingerprint Minutiae Extraction Based On FPGA and MatLab", Víctor López Lorenzo, Pablo Huerta Pellitero, José Ignacio Martínez Torre, Javier Castillo Villar, http://www.escet.urjc.es/~phuerta/pdf/dcis_2005.pdf