Secure Email – Ensuring CIA (Confidentiality, integrity and Authentication) and Privacy for Organizations

R.Dhanalakshmi ,
[1]Department of computer science and Engg.
Anna University , Chennai- 60025,
TamilNadu ,India
dhanalakshmisai@gmail.com

Dr.C.Chellappan
[2]Department of computer science and Engg.
Anna University , Chennai- 60025,
TamilNadu ,India
drcc@annauniv.edu

Abstract:- *Email has created tremendous opportunity and a powerful medium of data transfer but has also resulted in the loss of revenue and productivity due to its universal openness and accessibility. Ensuring Security against confidential data and privacy has become the immediate needs of Email Security Solutions. Encryption is an essential technology for protecting data both inside and outside the network perimeter, but implementing encryption effectively in a large enterprise is no easy task. One essential aspect for secure communication is public key cryptography which involves the use of asymmetric key algorithms, where the key used to encrypt a message is not the same as the key used to decrypt it. The key selection mechanism expresses the efficiency of the cipher text generated. A secure communication uses a method of encryption to transfer the message from the sender to the recipient. The message to be encrypted, known as the plaintext, are transformed by a function that is parameterized by an encryption key followed by a random mapping. The output of the encryption process, known of the cipher text has to be transmitted. Only the intended recipient has to decipher the message. Any intruder in the middle should not be allowed to read the contents of the message. The sender of the message has to sign the message using the publicly revealed encryption key. The intended recipient has to verify the signature of the sender in order to ensure authenticity.*

Keywords:-Encryption, Decryption, Cipher text, Message digest, Digital signatures,RSA,Random Array

1. Introduction

E-Mail Security battle has two sides of the coins such as protecting your organization against inbound online threats like spam and viruses and prevention of data leakage as outbound threat. Content filtering prevents inappropriate, malicious, or confidential content from leaving the corporate email system, allowing organizations to monitor and enforce the appropriateness of outbound corporate messages. The two major techniques for protecting private data are Data Loss Prevention solutions and Encryption.Data Loss Prevention (DLP) solutions scan traffic on enterprise networks looking for the transit of private data through applications such as E-Mail. Upon discovering private data (by matching keywords, phrases, hashes, or some other pattern), a DLP solution takes action. It can either block the data from leaving the network entirely or encrypt the data so that it leaves in secure format that only the intended recipients can access. It reduces the risk of confidential data to be accessed by hackers and malicious users. Encryption is an essential component of DLP failing makes the private

data to be blocked entirely is an unwise solution. Encryption ensures that even if a sensitive message is intercepted by a malicious party, the message's contents will remain unreadable and hence untampered with.
E-Mail Encryption proposed in this paper plays a very important role of ensuring the essential security features such as

- **Authentication:** The process of proving one's identity.
- **Confidentiality Privacy:** Ensuring that no one can read the message except the intended receiver. The user's identity is not revealed to others except the intended ones (i.e.) hiding or preserving identity.
- **Integrity:** Assuring the receiver that the received message has not been altered in any way from the original.

In this paper, we have used the following two levels of encryption/decryption:

- The first level encryption and decryption is implemented using RSA algorithm.
- The second level encryption and decryption is implemented using random array method in which each of the numbers from 0 to 9 is mapped to a number from 0 to 9 randomly using the random number generation algorithm.

Here, the main focus is on Public Key Cryptography. **Public-key cryptography** refers to a widely used set of methods for transforming a written message into a form that can be read only by an intended recipient. This cryptographic approach uses **asymmetric key algorithms which has** the public and private keys. The public key is used to transform a message into to an unreadable form, decryptable only by using the private key. The users in such a system create a key pair i.e., a public and a private key. By publishing the public key, the key producer empowers anyone who gets a copy of the public key (required for encryption) to produce messages only he can read -- because only the key producer has a copy of the private key (required for decryption). When someone wants to send a secure message to the creator of those keys, the sender encrypts it using the intended recipient's public key; to decrypt the message, the recipient uses the private key. Unlike symmetric key algorithms, a public key algorithm does not require a secure initial exchange of one, or more, secret keys between the sender and receiver. These algorithms work in such a way that, while it is easy for the intended recipient to generate the public and private keys and to decrypt the message using the private key, and while it is easy for the sender to encrypt the message using the public key, it is extremely difficult for anyone to figure out the private key based on their knowledge of the public key. The use of this algorithm also allows authenticity of a message to be checked by creating a digital signature of a message using the private key, which can be verified using the public key.

2. Literature survey and related work

Various researches have been carried out in the areas to data loss prevention and information leakage which focuses on the solution Email Encryption. In [6], it is shown that it is possible to protect a user's privacy from risks by exploiting mutually oblivious, competing communication channels. They create virtual channels over online services (e.g., Google's Gmail, Microsoft's Hotmail) through which messages and cryptographic keys are delivered. The message recipient uses a shared secret to identify the shares and ultimately recover the original plaintext. In so doing, they create a wired "spread-spectrum" mechanism for protecting the privacy of web-based communication. They also discuss the design and implementation of our open-source Java applet, Aquinas, and consider ways that the myriad of communication channels present on the Internet can be exploited to preserve privacy. Based on the behavioural perspective on the use of email they treat email policy as an embodiment of managerial beliefs and

values about the employer-employee relationship and the role of communication in the workplace [10]. They examine employee attitudes towards email, their perceptions and expectations regarding the privacy and ownership of email, and a variety of work environment characteristics. Their results have interesting implications for organizations desirous of constructing an email policy. Theoretical implications as well as guidelines for practicing managers are offered. The purpose of [8] is to propose a privacy compliance engine that takes email messages as input and filters those that violate the privacy rules of the organization in which it is deployed. Their system includes two main parts: an information extraction module that extracts the names of the sender and recipients as well as sensitive information contained in the message; and an inference engine that matches the email information against a knowledge base owned by the organization. This engine then applies compliance rules to the information obtained from the extraction and database matching steps of the process. In a university setting, it was shown to obtain a precision score of 77%.In [2]; it is shown that many widely deployed email encryption systems reveal the identities of Blind-Carbon-Copy (BCC) recipients. Additionally, several implementations of PGP expose the full name and email address of BCC recipients. In this paper, they present a number of methods for providing BCC privacy while preserving the existing semantics of email. Their constructions use standard public key systems such as RSA and ElGamal and suggest that BCC privacy can be implemented efficiently without changing the underlying broadcast semantics of the email system. The Simple Encryption/Decryption [7] application states that it is able to work with any type of file; for example: image files, data files, documentation files…etc. The method of encryption is simple enough yet powerful enough to fit the needs of students and staff in a small institution. The application uses simple key generation method of random number generation and combination. The final encryption is a binary one performed through rotation of bits and XOR operation applied on each block of data in any file using a symmetric decimal key. The key generation and Encryption are all done by the system itself after clicking the encryption button with transparency to the user. The same encryption key is also used to decrypt the encrypted binary file.

3. Proposed Algorithm

In public key cryptography, keys and messages are expressed numerically and the operations are expressed mathematically. The private and public key of a device is related by the mathematical function called the one-way function. One-way functions are mathematical functions in which the forward operation can be done easily but the reverse operation is so difficult that it is practically impossible. In public key cryptography the public key is calculated using private key on the forward operation of the one-way function. Obtaining of private key from the public key is a reverse operation. If the reverse operation can be done easily, that is if the private key is obtained from the public key and other public data, then the public key algorithm for the particular key is cracked. The reverse operation gets difficult as the key size increases. The public key algorithms operate on sufficiently large numbers to make the reverse operation practically impossible and thus make the system secure.

This algorithm is basically divided into two levels as shown in Fig 1:

- Two levels of encryption and decryption
    - Public key encryption (First level encryption and decryption) using RSA algorithm.
- Random Array Method (Second level encryption and decryption) in which

ISBN: 978-1-61804-044-2

each of the numbers from 0 to 9 is mapped to a number from 0 to 9 randomly using the random number generation algorithm.

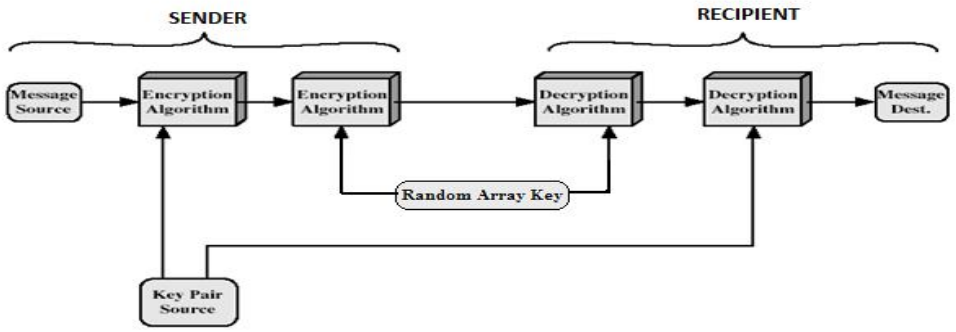


Fig 1: TWO LEVELS OF ENCRYPTION AND DECRYPTION

3.1. RSA based Encryption and Decryption

- **Key distribution**

A message encrypted with a recipient's public key cannot be decrypted by anyone except a possessor of the matching private key—presumably, this will be the owner of that key and the person associated with the public key used. This is used for confidentiality. In this method, different keys are used for encryption and decryption. Each user has a pair of cryptographic keys—

- public encryption key - which may be known by anybody, and can be used to **encrypt messages**, and **verify signatures**
- private decryption key - known only to the recipient, used to **decrypt messages**, and **sign** (create) **signatures**

An analogy to public-key encryption is that of a locked mailbox with a mail slot. The mail slot is exposed and accessible to the public; its location (the street address) is in essence the public key. Anyone knowing the street address can go to the door and drop a written message through the slot; however, only the person who possesses the key can open the mailbox and read the message. The most obvious application of a public key encryption system is confidentiality. Only the intended recipient can read the sent message because he alone will have the corresponding private key for decryption.

3.2. Random Array based Encryption and Decryption

In this method, the encrypted content by the first level encryption is again encrypted using the random array method. That is, each of the numbers from 0 to 9 is mapped or assigned a random number from 0 to 9 by using random number generation algorithm as shown in Fig 2. Only this second level encrypted content is sent to the recipient. Similarly, the decryption at the recipient's end is performed by once again retrieving the original number associated with every mapped number. This decrypted content is once again decrypted using the first level decryption in order to retrieve the original contents. This is a symmetric key algorithm superimposed on the first asymmetric algorithm to further strengthen the security.

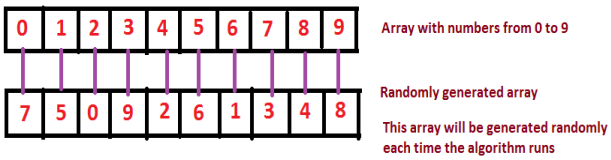


Fig 2: Random Array based Encryption and Decryption

3.3.Steps for two levels of encryption and decryption

- Key generation
- Encryption using RSA algorithm
- Encryption using Random Array Method
- Decryption using Random Array Method
- Decryption using RSA algorithm

3.3.1. Key Generation Algorithm

- Generate two large random primes, *p* and *q*, of approximately equal size such that their product n = pq is of the required bit length, e.g. 1024 bits.
- Compute n = pq and (φ) phi = (p-1)(q-1).
- Choose an integer *e*, 1 < e < phi, such that gcd(e, phi) = 1.
- Compute the secret exponent *d*, 1 < d < phi, such that ed ≡ 1 (mod phi).
- The public key is (n, e) and the private key is (n, d). Keep all the values d, p, q and phi secret.

Where

- n is known as the *modulus*.
- e is known as the *public exponent* or *encryption exponent* or just the *exponent*.
- d is known as the *secret exponent* or *decryption exponent*.

3.3.2. Encryption using RSA Algorithm

Sender A does the following:-

- Obtains the recipient B's public key (n, e).
- Represents the plaintext message as a positive integer *m* .
- Computes the cipher text c1 = m^e mod n.

3.3.3. Encryption using Random Array Method

- Each of the numbers from 0 to 9 is mapped to a different number from 0 to 9 using random generation algorithm and stored in an array.
- The cipher text obtained from the first level encryption is once again encrypted by mapping the individual digits to a different number by using stored array.
- This new cipher text c2 is sent to the recipient B.

3.3.4. Decryption using Random Array Method

Recipient B does the following:-

- Uses the array containing the mapped number for each of the individual digit of the ciphetext c2 to convert into the original number.
- Sends this decrypted contents c1 to the next level decryption algorithm.

3.3.5. Decryption using RSA Algorithm

- Uses his private key (n, d) to compute m = $c1^d$ mod n.
- Extracts the plaintext from the message representative *m*.

4. Ensuring Authentication

Authentication plays a vital role in secure email systems ensuring the claimed identity is true and not being spoofed by a malicious user as a legitimate entity. Authentication is verified by different approaches such as Digital signatures, PGP, S/MIME, SPF, SID, Domain Keys and IIM. The paper discusses the following process of Digital signatures

ISBN: 978-1-61804-044-2

4.1. Digital Signature

A message signed with a sender's private key can be verified by anyone who has access to the sender's public key, thereby proving that the sender had access to the private key (and therefore is likely to be the person associated with the public key used), and the part of the message that has not been tampered with. An analogy for digital signatures is the sealing of an envelope with a personal. The message can be opened by anyone, but the presence of the seal authenticates the sender. Digital signature schemes are used for sender authentication and non-repudiation. Here, a user who wants to send a message computes a digital signature of this message and then sends this digital signature together with the message to the intended receiver. Digital signature schemes have the property that signatures can only be computed with the knowledge of a private key. To verify that a message has been signed by a user and has not been modified the receiver only needs to know the corresponding public key.

4.2. Steps for Digital Signatures

- Digital signing
- Signature verification

4.2.1. Digital signing algorithm

Sender A does the following:-

- Creates a *message digest* of the information to be sent.
- Represents this digest as an integer *m* between 0 and *n*-1.
- Uses her *private* key (n, d) to compute the signature $s = m^d \bmod n$.
- Sends this signature *s* to the recipient, B.

4.2.2. Signature Verification Algorithm

Recipient B does the following:-

- Uses sender A's public key (n, e) to compute integer $v = s^e \bmod n$.
- Extracts the message digest from this integer.
- Independently computes the message digest of the information that has been signed.
- If both message digests are identical, the signature is valid.
-

5. TEST CASES

Sample Emails have been generated and tested with Gmail. Screenshots of a test email being sent after encryption and received and decrypted as shown in Fig 3 and Fig 4.

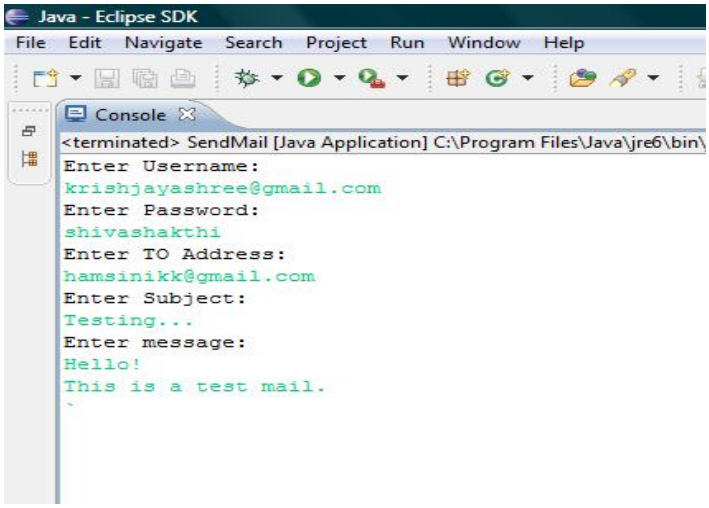


Fig 3 : Input Message before sending

This will be the encrypted contents of the mail sent by the sender which will be stored in the inbox of the recipient.

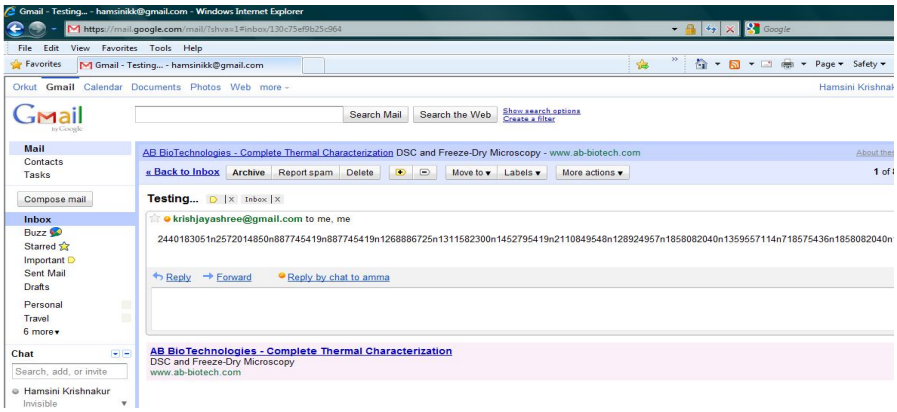


Fig 4: Decrypted Message viewed by the receiver.

6. Performance Evaluation

An "RSA operation," whether encrypting, decrypting, signing, or verifying is essentially a modular exponentiation. The computational complexity of the RSA algorithm completely depends upon the key length and the length of the modulus n. Hence the complexity of the encryption and decryption depends on the length of the key. The computational complexity of RSA encryption and decryption of a single n bit block is approximately O(n^3), with n is denoting both the block length and key length (exponent and modulus) and it is in polynomial time. The computational complexity of finding the encryption component e and a decryption component d. This is done using the Extended Euclidian algorithm as follows

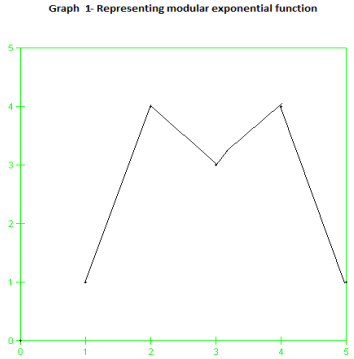
Find a number e, such that gcd(e, $\phi(n)$) = 1. All the powering and gcd calculations are clearly in polynomial time in the number of bits of n. To find a number e such that gcd(e, $\phi(n)$) = 1, the fraction of elements, which are relatively prime to N, is (1/ logN). So setting N = $\phi(n)$, after O(logN) random trials for e which is prime to $\phi(n)$ and is still all polynomial in the number of bits of n.Hence the complexity of the RSA algorithm is polynomial in time with respect to the length of the key and the modulus, n. The speed and efficiency of the many commercially available software and hardware implementations of the RSA algorithm are increasing rapidly.

GRAPHICAL REPRESENTATIONS:

The modular exponential function. The graph of the discrete modular exponential function

$$f(m) = m^e \bmod n,$$

which is used for the RSA algorithm, is plotted for $0 \; m < n$. Small values of e and n are considered for simplicity. The following graph is plotted for the values: e=4, n=6

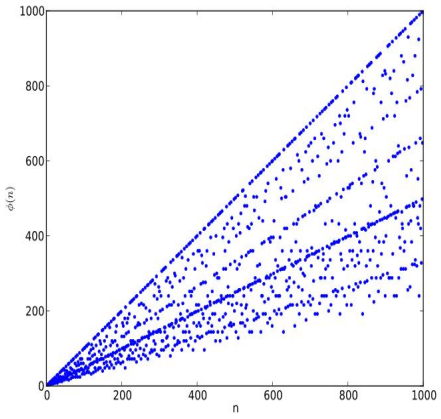


Representing Modular exponential function

Euler totient function : The totient φ(n) of a positive integer n is defined to be the number of positive integers less than or equal to n that are coprime to n. The first 1000 values of φ(n) is plotted below:

ISBN: 978-1-61804-044-2

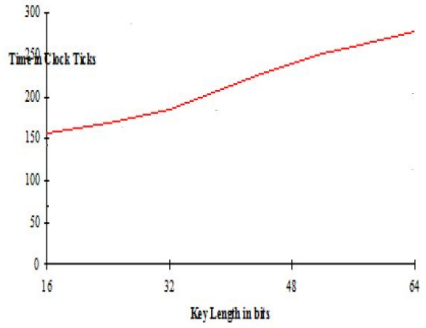
Graph 2 - Representation of euler totient function



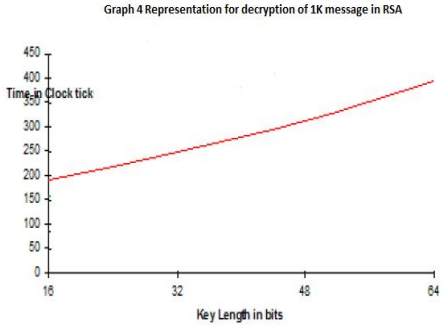
The following graph shows the time taken for the encryption of 1K message using RSA algorithm.

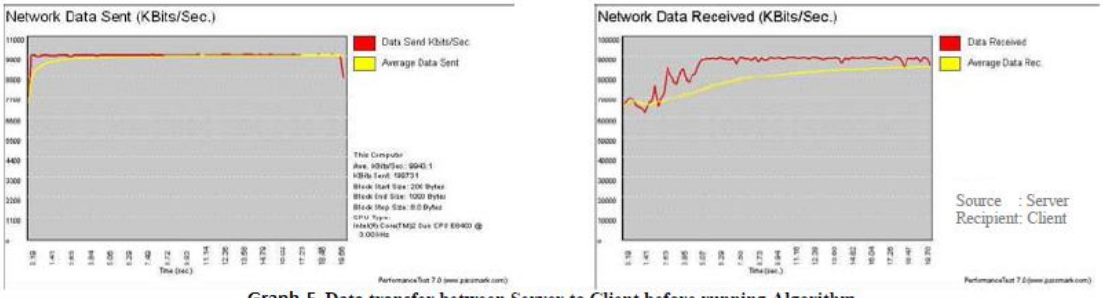
The following graph shows the time taken for the decryption of 1K message using RSA algorithm

Graph 3 Representation for encryption of 1K message in RSA

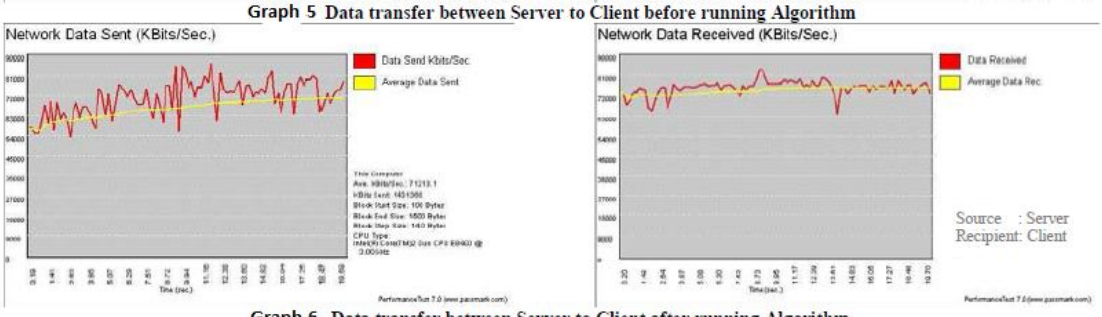


Graph 4 Representation for decryption of 1K message in RSA



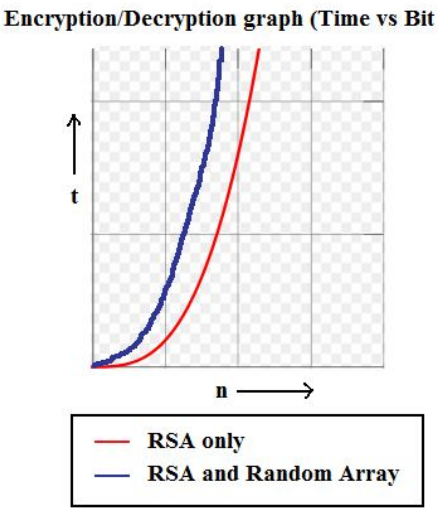


Graph 5 Data transfer between Server to Client before running Algorithm



Graph 6 . Data transfer between Server to Client after running Algorithm

ISBN: 978-1-61804-044-2



7. The Challenge of Implementing Encryption in the Enterprise

Encryption is an essential technology for protecting data mitigating inbound and outbound threats within the network perimeter, but has challenging issues in implementing encryption effectively in a large enterprise as follows

- Involving a vast database and thousands of email transactions in a large enterprise carries both private data and non-private data. How to categorize for protecting the private data?
- Encrypting absolutely every outbound communication degrades network performance and consumes computing resources.
- Encrypted data becomes inscrutable to network monitoring tools; the more network traffic that's encrypted; the less network traffic can be optimized for performance using protocol optimization and other optimization techniques.

8. Conclusion and Future Work

In order to enhance the security of the Email, models have been developed to ensure authentication for the incoming mails, maintaining integrity and confidentiality for the outgoing emails. The proposed system has got various overheads such as

- A constant time overhead is required for generating the random array and sending it to the sender and receiver along with their keys.
- Encryption and Decryption overhead: An extra over head of complexity O (n) for an n-bit block is required to perform the mapping.

Future extensions shall involve encrypting the attachments also in the Email which are to be sent. To enhance privacy, new image and linguistic steganography techniques can be included allowing users to more fully obfuscate their communications.

9. Acknowledgment

This work is supported by the NTRO, Government of India. NTRO provides the fund for collaborative project "Smart and Secure Environment" and this paper is modeled for this project.Authors would like to thanks the project coordinators and the NTRO members

10. References

1. Anti-Phishing Working Group, Digital Signatures to Fight Phishing Attacks. http://www.antiphishing.org/smim-dig-sig.htm

2. Adam Barth and Dan Boneh,"Correcting Privacy Violations in Blind-Carbon-Copy (BCC) Encrypted Email "

3. B. Adida, S. Hohenberger and R. L. Rivest. Fighting Phishing Attacks: A Lightweight Trust Architecture for Detecting Spoofed Emails, http://people.csail.mit.edu/rivest/AdidaHo

henbergerRivest-FightingPhishingAttacks.pdf, 2005.

4. B. Adida, S. Hohenberger and R. L. Rivest. Seperable Identity-Based Ring Signatures: Theoretical Foundations For Fighting Phishing At tacks, presented at the DIMACS Workshop on Theft in E-Commerce, Piscataway, New Jersey February 2005.

5. Dhanalakshmi R,L.Kavisankar,C.Chellappan ,"Enhanced E-Mail Authentication Against Spoofing Attacks to Mitigate Phishing ",European Journal of Scientific Research Vol 54 Issue 1,165-175.

6. Kevin Butler, William Enck, Jennifer Plasterr, Patrick Traynor, and Patrick McDaniel, "Privacy Preserving Web-Based Email" , Systems and Internet Infrastructure Security Laboratory, World Scientific Review Volume , 2007.

7. Majdi Al-qdah and Lin Yi Hui Simple Encryption/Decryption Application , International Journal of Computer Science and Security, Volume (1) : Issue (1),2007

8. Quintin Armour, William Elazmeh, Nour El-Kadri, Nathalie Japkowicz, and Stan Matwin ,Privacy Compliance Enforcement in Email Proceedings of Canadian Conference on AI 2005,194-204

9. R. L. Rivest, A. Shamir, and Y. Tauman. How to leak a secret. In C. Boyd, editor, *Advances in Cryptology — ASIACRYPT '01*, volume 2248 of *LNCS*, pages 552–565. Springer Verlag, 2001.

10. Ritu Agarwal, Florence Rodhain ,"Mine or Ours: Email Privacy Expectations, Employee Attitudes, and Perceived Work Environment Characteristics", Proceedings of the 35th Hawaii International Conference on System Sciences – 2002

11. S. M. Bellovin. Spamming, phishing, authentication, and privacy. Inside Risks, Communications of the ACM,47:12, December 2004.

12. S. L. Garfinkel. Email-Based Identification and Authentication: An Alternative to PKI? IEEE Security &Privacy, 1(6):20–26, Nov. 2003.