HYBRID DATA ENCRYPTION AND DECRYPTION USING RSA AND RC4

EZEKIEL Bala [1], AJIBOLA Aminat [2], and EBELOGU Christopher U [3]

ABSTRACT

This research study "Hybrid Data Encryption and Decryption using RSA and RC4" is one which combines the convenience of a public-key (asymmetric-key) cryptosystem with the efficiency of a symmetric-key (private-key) cryptosystem thereby developing a hybrid cryptosystem. Here, a two-way secured data encryption system, the concerns of user privacy, authentication and accuracy is addressed. The system has two different encryption algorithms which was used for both Encryption and Decryption sequence. One is public key cryptography based on linear block cipher, the second one is private key cryptography based on simple symmetric algorithm. This cryptographic algorithm provides a more robust and secures authentication system compared to other existing cryptographic algorithms such as the AES, MD5, RSA, RC4 etc. The Hybrid encryption will be considered a highly secure type of encryption as long as the public and private keys will be fully secured. The hybrid encryption is achieved through data transfer using unique session keys along with symmetrical encryption. Public key encryption was implemented for random symmetric key encryption. The recipient is able to use the public key encryption method to decrypt the symmetric key. Once the symmetric key is recovered, it is then used to decrypt the message. The result shows a tremendous improvement on security of data and it gives overall improvement of the system performance in terms of security and efficiency. The system is designed using C# programming language.

Keywords: *Encryption, Cipher, Ciphertext, Hybrid, Cryptography, Cryptosystem, Block Ciphers, Cryptanalysis, Cryptology, Symmetric cryptosystems, Asymmetric cryptosystems, Plaintext,*
Espionage

1. INTRODUCTION

For different reasons humans have been interested in protecting their messages. Right from the earliest days, people have been attempting to conceal certain information that they wanted to keep to their own possession by substituting parts of the information with symbols, numbers and pictures. The Assyrians for instance were interested in protecting their trade secret of manufacturing of the pottery. The Chinese were interested in protecting their trade secret of manufacturing silk. The Germans were interested in protecting their military secrets by using their famous Enigma machine.
Curiosity is one of the most common human traits, matched by the wish to conceal private information. People often resort to information hiding to pass messages securely, sometimes deliberately including misleading information.

Cryptography is the science of devising methods that allow information to be sent in a secure form in such a way that the only person able to retrieve this information is the intended recipient [1].

With the advancement of computers and technology, interconnectivity, businesses, individuals, and industries have become subject to attacks, attacks such as cyber attacks, intrusion and industrial espionage; there is an increasing need to secure data and transactions. Hence, a mechanism is required to guarantee the security and privacy of information that is transmitted, especially over electronic communications media and in the modern era of inexpensive internet connections, data computing and global communications. There is a high demand for energy efficiency, computational speedup and data security. Many efforts have been taken in response to the ever-increasing need for data security, to protect the data from unwanted action, unauthorized user and destructive forces. Thus, cryptography plays an important role [2][3].

2. LITERATURE REVIEW

A. Encryption/Decryption

Encryption is the process of trans-forming plaintext data into cipher text to conceal its meaning thereby preventing any unauthorized recipient from retrieving the original data. This is usually done with the use of an encryption key, which specifies how the message is to be encoded. Encryption works by running the data through a special encryption formula called a key.
Decryption on the other hand is the reverse process of converting encrypted data (cipher text) to its original un-encoded form [4].

Cryptography

Cryptographyor cryptology (from Greek *kryptós*, "hidden, secret"; and *graphein*, "writing", or *-logia*, "study", respectively) is the practice and study of techniques for secure communication in the presence of third parties called adversaries (wikipedia). Cryptography is about constructing and analyzing protocols that prevent third parties or the public from reading private messages; various aspects of information security such as data confidentiality, data integrity, authentication, and non-repudiation are central to modern cryptography. Modern cryptography exists at the intersection of the disciplines of mathematics, computer science, and electrical engineering. Applications of cryptography include military communications, electronic commerce, ATM cards, and computer passwords [5].

Cryptographic Techniques

Just as there are different types of household keys for the car, front door of the house, garage door, etc., keys also serve different functions in the world of digital communications. In general, cryptographic keys are categorized according to their properties and usage. There are several ways of classifying cryptographic techniques but for purposes of this work, they will be categorized based on the number of keys that are employed for encryption and decryption, and further defined by their application and use *[6]*. Below is a classification of cryptographic algorithms based on the number of keys used.

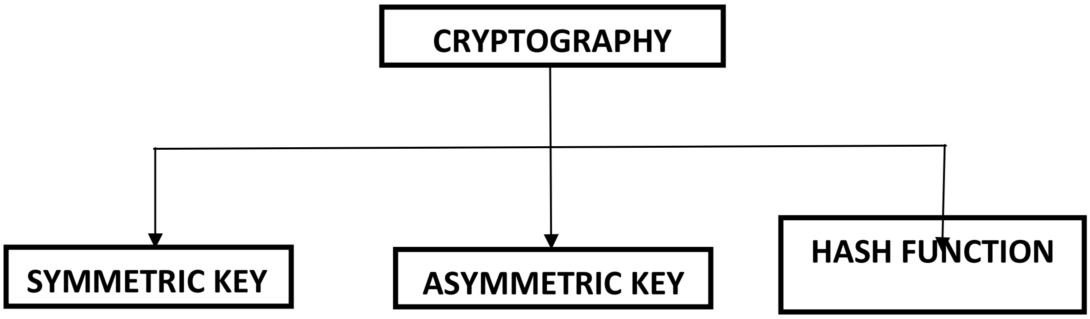

Figure 1: Types of Cryptography

Figure 1 shows that there are three basic types of cryptographic keys namely; Symmetric keys, asymmetric keys and hash function. Each one is discussed in details in subsequent sections of this article.

Symmetrical (Private) Keys Encryption

Symmetrical keys are also known as Private keys or secret keys. In secret key or symmetric cryptography there is only one key. It is used for both encryption and decryption. A key refers to any code that yields plain text when applied to cipher text. This key is shared by both sender and receiver. In private key encryption technology, both the sender and receiver have the same key and use it to encrypt and decrypt all messages. This makes it difficult to initiate communication for the first time. How does one securely transmit the single key to each user? If the key is disclosed the secrecy of the information is compromised. Lengthy keys are used to increase the security and to decrease the chances of identifying the key through brute force. It is relatively fast as it uses the same key for encryption and decryption [7].

However, more damage can occur if the key is compromised, such can decrypt everything that was encrypted with that key. Since symmetric encryption is used for two-way communication, both sender and receiver end data get compromised. Data Encryption Standard (DES), Rivest Cipher 4 (RC4), Advance Encryption Standard (AES), Carlisle Adams and Stafford Tavares (CAST) Algorithm, Blowfish, Twofish, International Data Encryption Algorithm (IDEA) are some examples of symmetric technique. Fig 2 shows symmetric encryption.

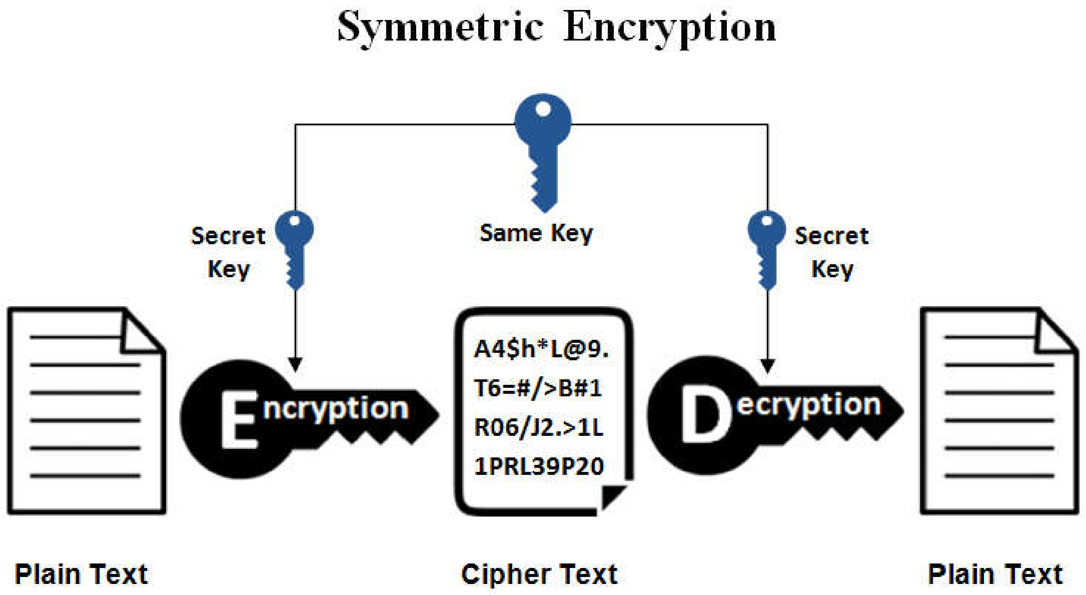

Figure 2: Symmetric Encryption (Source: William, 2012)

Asymmetric (Public) Key Encryption

Public-key /two-key/ asymmetric cryptography, uses two keys to encrypt and decrypt data: it involves the use of two keys: a public-key, which may be known to everyone, used to encrypt messages and verify signatures and a private-key, known only to the recipient, used to decrypt messages and sign (create signatures). It is called asymmetric cryptography because the key used to encrypt messages or verify signatures cannot be used to decrypt messages or create signatures.

Asymmetric key ciphers increase the security and convenience as private keys never have to be transmitted or revealed to anyone. Damage due to loss of private keys are mostly irreparable.

Rivest Shamir Adleman algorithm (RSA), Diffie-Hellman, Digital Signature Algorithm (DSA), Elgama and Elliptic Curve cryptography (ECC) are some public asymmetric algorithms or technique [7].

The receiver’s public key is used to encrypt a message; this message is then sent to the receiver who can decrypt it using its own private key. This is a one-way communication. If the receiver wants to send a return message, the same principle is used. The message is encrypted with the original sender’s public key (the original sender is now going to be the receiver of this new message) and can only be decrypted with his or her private key. If the original sender does not have a public key, a message can still be sent with a digital certificate (also sometimes referred to as a digital ID). The digital ID verifies the sender of the message. Fig. 3 below shows public key– encrypted communication between two units, User X and User Y.

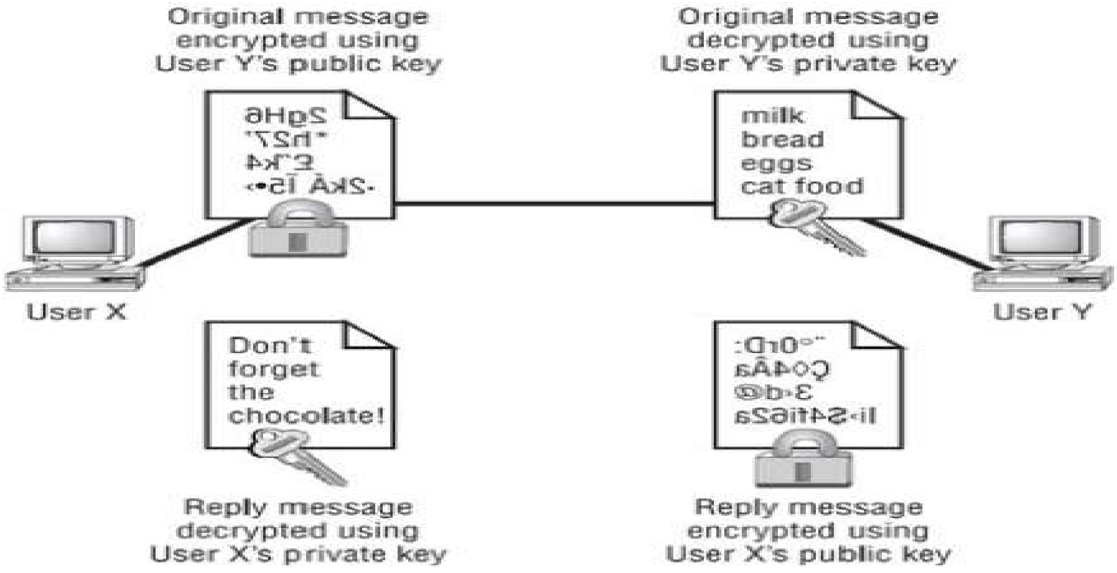
Original message encrypted using User Y's public key
Original message decrypted using User Y's private key
milk
bread
eggs
cat food
User X
User Y
Don't forget the chocolate!
Reply message decrypted using User X's private key
Reply message encrypted using User X's public key

Figure 3: Public key encryption (Source: Mousa et al, 2010)

Hash Function

The Hash Function uses a mathematical transformation to irreversibly "encrypt" information. This algorithm does not use keys for encryption and decryption of data. It rather uses a fixed-length hash value which is computed based on some plaintext that makes it impossible for either the contents or the length of the plaintext to be recovered [7]. These algorithms are typically used to provide a digital fingerprint of a file's contents, often used to ensure that the file has not been altered by an intruder or virus. Hash functions are also commonly employed by many operating systems to encrypt passwords to provide some amount of integrity to a file, examples are; (MD5), SHA-1, SHA-2.

B. THE RC4 ALGORITHM

RC4 is a cryptographic algorithm that was created by Ronald Rivest of RSA Security. RC4 is known to be one of the simplest and widely adopted cipher. However, the simplicity of RC4 makes it vulnerable to different security attacks. RC4 has two basic constituents; Key scheduling algorithm (KSA) and Pseudo random number generator (PRGA). It is observed that PRGA generates a pseudorandom output sequence (bytes) from the permuted internal state which itself is a random sequence. Statistical weaknesses are the biases in the random keystream that can be exploited with a very high probability of success, to differentiate the generated RC4 keystream from a truly random sequence of bytes [8].

3.1 The steps for RC4 algorithm are as follows:

- Get the data to be encrypted and selected key
- Create two strings
- Initiate one array with numbers from 0 to 255
- Fill the other array with the selected key
- Randomize the first array depending on the array of the key
- Randomize the first array within itself to generate the final key stream
- XOR the final key stream with the data to be encrypted to give ciphertex

Fig 4 depicts the flow chat algorithm of RC4.

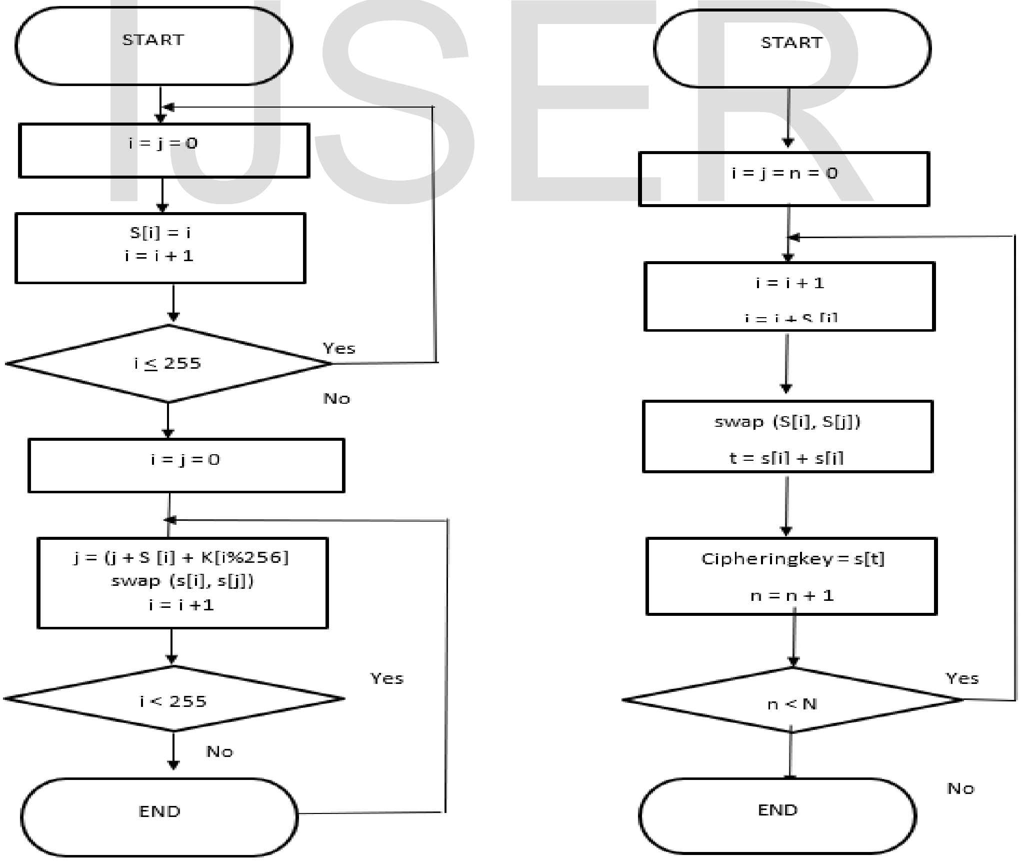

Figure 4: Encryption Algorithm flow chart for RC4 (Source: Okedola et al, 2015)

The algorithm can be broken into two stages: initialization, and operation.

In the initialization stage the 256-bit state table, **S** is populated, using the key, **K** as a seed. Once the state table is setup, it continues to be modified in a regular pattern as data is encrypted.

The initialization process can be summarized by the pseudo-code;

```
j = 0
for i = 0 to 255
S[i] = i
for i = 0 to 255
j = (j + S[i] + K[i]) mod 256
swap S[i] and S[j]
end for
end for
```

It is important to notice here the swapping of the locations of the numbers 0 to 255 (each of which occurs only once) in the state table. The values of the state table are provided.

Once the initialization process is completed, the operation process may be summarized as shown by the pseudo code below;

```
i = j = 0
(k = 0 to N-1)
i = (i + 1) mod 256
j = (j + S[i]) mod 256
swap S[i] and S[j];
pr = S[ (S[i] + S[j]) mod 256]
output M[k] XOR pr
```

Where M [0..N-1] is the input message consisting of N bits.

This algorithm produces a stream of pseudo-random values. The input stream is XORed with these values, bit by bit. The encryption and decryption process is the same as the data stream is simply XORed with the generated key sequence. If it is fed in an encrypted message, it will produce the decrypted message output, and if it is fed in plaintext message, it will produce the encrypted version.

Advantages

Simple to Implement, very quick in software (speed).

Disadvantages

Produce biased outputs towards certain sequences

C. THE RSA ALGORITHM

RSA stands for Ron Rivest, Adi Shamir and Leonard Adleman. It is a form of asymmetric cryptography. In RSA algorithm, the plain text/original data is encrypted in blocks, with each block having a binary value less than some number n. The public key in this cryptosystem consists of the value n, which is called the modulus, and the value e, which is called the public exponent. The private key consists of the modulus n and the value d, which is called the private exponent [9].

An RSA public-key / private-key pair can be generated by the following steps:

1. Generate a pair of large, random prime's p and q
2. Compute the modulus n as $n = p*q$.
3. Select an odd public exponent e between 3 and $n-1$ that is relatively prime to $p-1$ and $q-1$.
4. Compute the private exponent d from e, p and q. (See below.)
5. Output (n, e) as the public key and (n, d) as the private key.

The encryption operation in the RSA cryptosystem is exponentiation to the eth power modulo n:

$$c = \text{ENCRYPT}(m) = m^e \bmod n \text{ ----------------------------- } (1)$$

The input m is the message; the output c is the resulting cipher text/decrypted file. In practice, the message m is typically some kind of appropriately formatted key to be shared. The actual message is encrypted with the shared key using a traditional encryption algorithm. This construction makes it possible to encrypt a message of any length with only one exponentiation.

The decryption operation is exponentiation to the dth power modulo n:

$$m = \text{DECRYPT}(c) = c^d \bmod n \text{ ------------------------ } (2)$$

The relationship between the exponent's e and d ensures that encryption and decryption are inverses, so that the decryption operation recovers the original message m. Without the private key (n, d) (or equivalently the prime factors p and q), it's difficult to recover m from c. Consequently, n and e can be made public without compromising security, which is the basic requirement for a public-key cryptosystem.

The fact that the encryption and decryption operations are inverses and operate on the same set of inputs also means that the operations can be employed in reverse order to obtain a digital signature scheme following RSA and RC4's model. A message can be digitally signed by applying the decryption operation to it, i.e., by exponentiation it to the dth power:

s = SIGN (m) = md mod n ---------------------------- (3)

The digital signature can then be verified by applying the encryption operation to it and comparing the result with and/or recovering the message:

m = VERIFY (s) = s^e mod$n$-------------------------------- (4)

In practice, the plaintext or decrypted file m is generally some function of the message, for instance a formatted one-way hash of the message. This makes it possible to sign a message of any length with only one exponentiation.
This algorithm is represented graphically in fig 3.3a and fig 3.3b below.

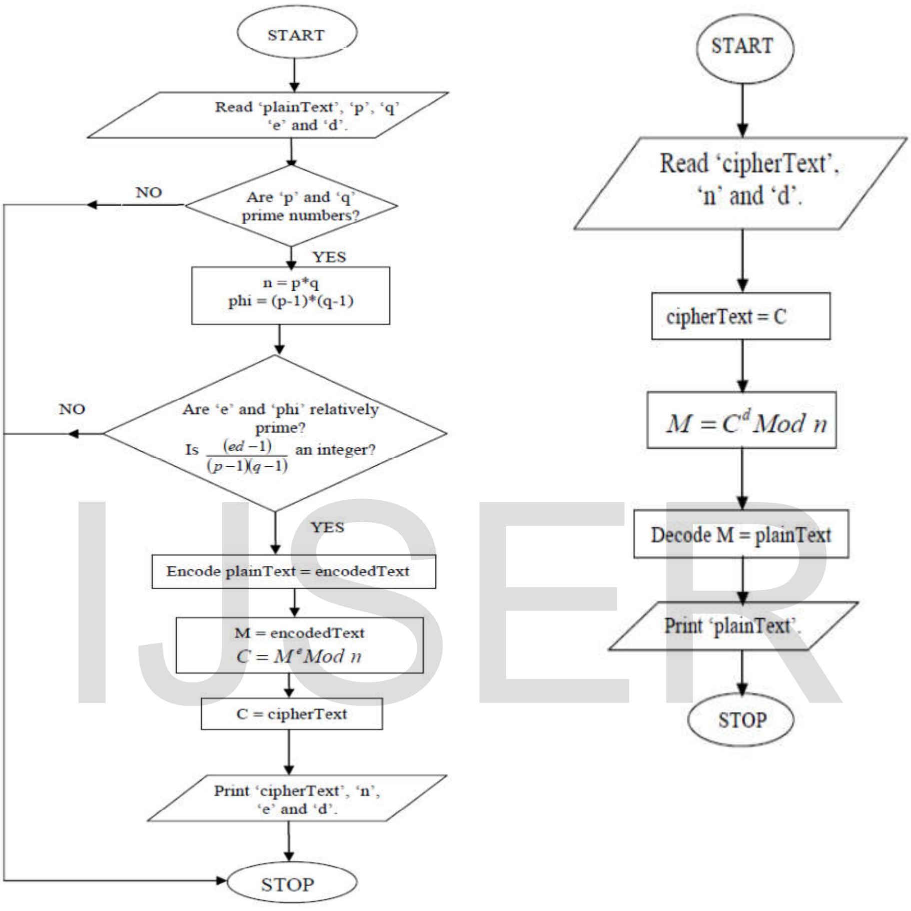
START
Read 'plainText', 'p', 'q' 'e' and 'd'.
NO
Are 'p' and 'q' prime numbers?
YES
n = p*q
phi = (p-1)*(q-1)
NO
Are 'e' and 'phi' relatively prime?
Is $\frac{(ed-1)}{(p-1)(q-1)}$ an integer?
YES
Encode plainText = encodedText
M = encodedText
$C = M^e Mod\ n$
C = cipherText
Print 'cipherText', 'n', 'e' and 'd'.
STOP
START
Read 'cipherText', 'n' and 'd'.
cipherText = C
$M = C^d Mod\ n$
Decode M = plainText
Print 'plainText'.
STOP

Figure 5: Encryption and Decryption Algorithm flow chart for RSA (Source: Okedola et al, 2015)

Advantages

There is no need for exchanging keys, thus eliminating the key distribution problem, the private keys do not ever need to be transmitted or revealed to anyone, can provide digital signatures that can be repudiated

Disadvantages

It is slow, two people using the same, receiving the same message, using small primes, Using primes that are very close.

Hybrid Encryption/Decryption

Hybrid encryption is a system of encryption that merges two or more encryption systems. It incorporates both asymmetric and symmetric encryption to benefit from the strengths of each form of encryption. These strengths are respectively defined as speed and security.

By combining public and private key cryptosystems, it is possible to overcome some of the disadvantages of each. Public key pairs are used to set up a secure session, and then data is exchanged using a secret key system. This provides both the security and authentication processes associated with public key systems and the bulk data encryption capabilities of secret key systems. Pretty Good Privacy (PGP) is a well-known security system used by computer enthusiasts to encrypt their email; it is an example of a practical hybrid encryption system which uses both secret key and public key [10][11].

It provides the functionalities of both public and private cryptography. Data and applications can be secured in an efficient manner.

In this research work RC4 will be adopted as the private key and RSA the public key.

D. PROPOSED HYBRID SYSTEM ARCHITECTURE OF RSA AND RC4

The proposed hybrid system will be achieved by combining RSA and RC4 algorithms. The diagram of the proposed hybrid technique is shown in Figure. 3.1.

In this technique, at the transmitting side, RC4 encryption algorithm E will be used to encrypt the data to be transmitted (Original file, P) with the help of a randomly generated session key k, turning it into an encrypted file, C. The resulting encrypted file is given by:

$$C = E(k, P) \text{ -----------------------------------(1)}$$

The session key will be generated using pseudorandom number generator. Since RC4 is a secret key algorithm, this key has to be kept secret and this is achieved by encrypting the session key using RSA algorithm with the help of the recipient public key, e. This produces an encrypted session key u, given by:

$$u = k^{e}(\text{mod(n)}) \text{ ------------------------------------- (2)}$$

Where n is the product of two randomly generated large prime numbers p and q used in the RSA algorithm. The encrypted file C and the encrypted session key u would be sent to the receiver over a communication channel. The figure below shows the proposed hybrid architecture.

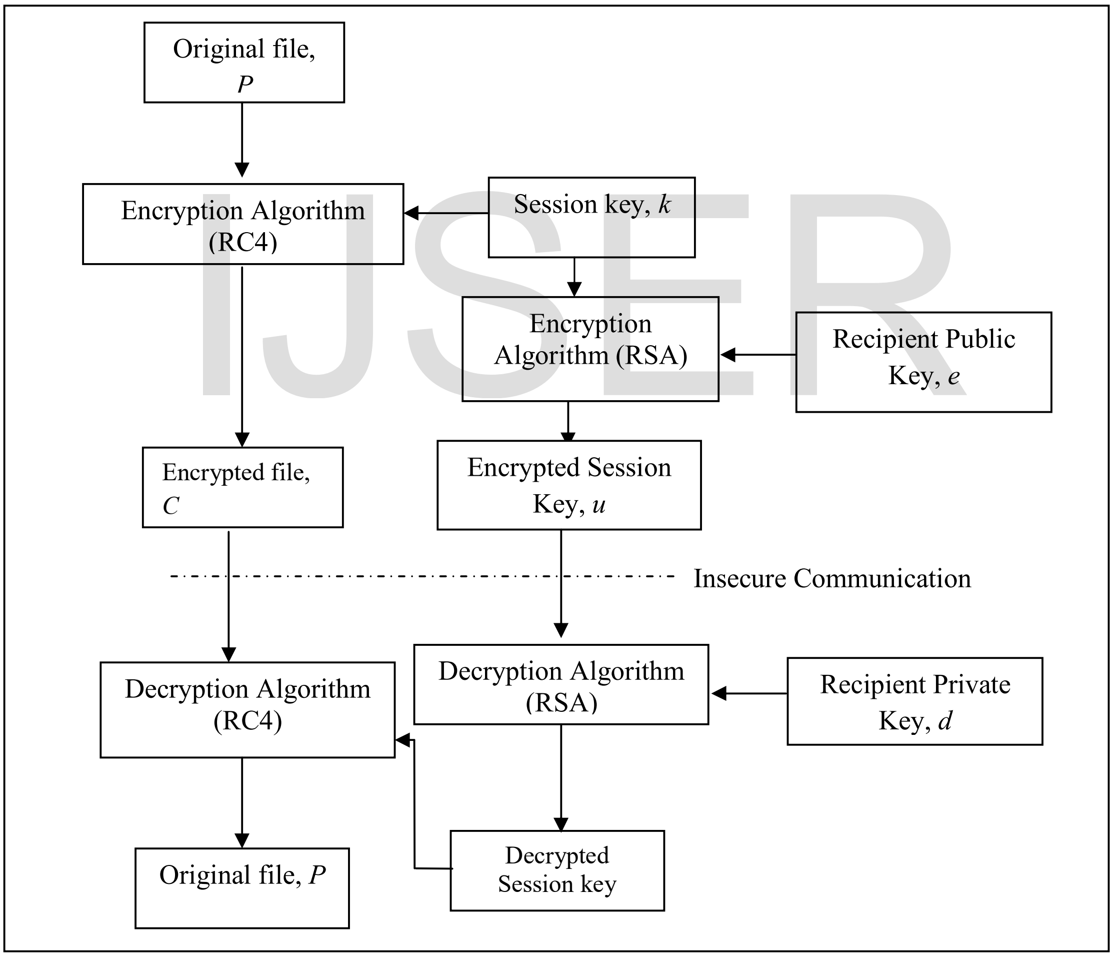

Figure 6 Architecture of the Hybrid RSA and RC4

At the receiving end, the recipient private key d, will be used with the RSA decryption algorithm to decrypt u, thereby producing the session key as given by

$$k= u^{d}(\text{mod}(n)) \text{-------------------------------------(3)}$$

Having retrieved the session key, this session key will be used with the RC4 decryption algorithm *D* to obtain the original data (Original file, *P*) using the equation below:

$$P = D(C, k) = D(E(P, k),k)\text{------------------------------}(4)$$

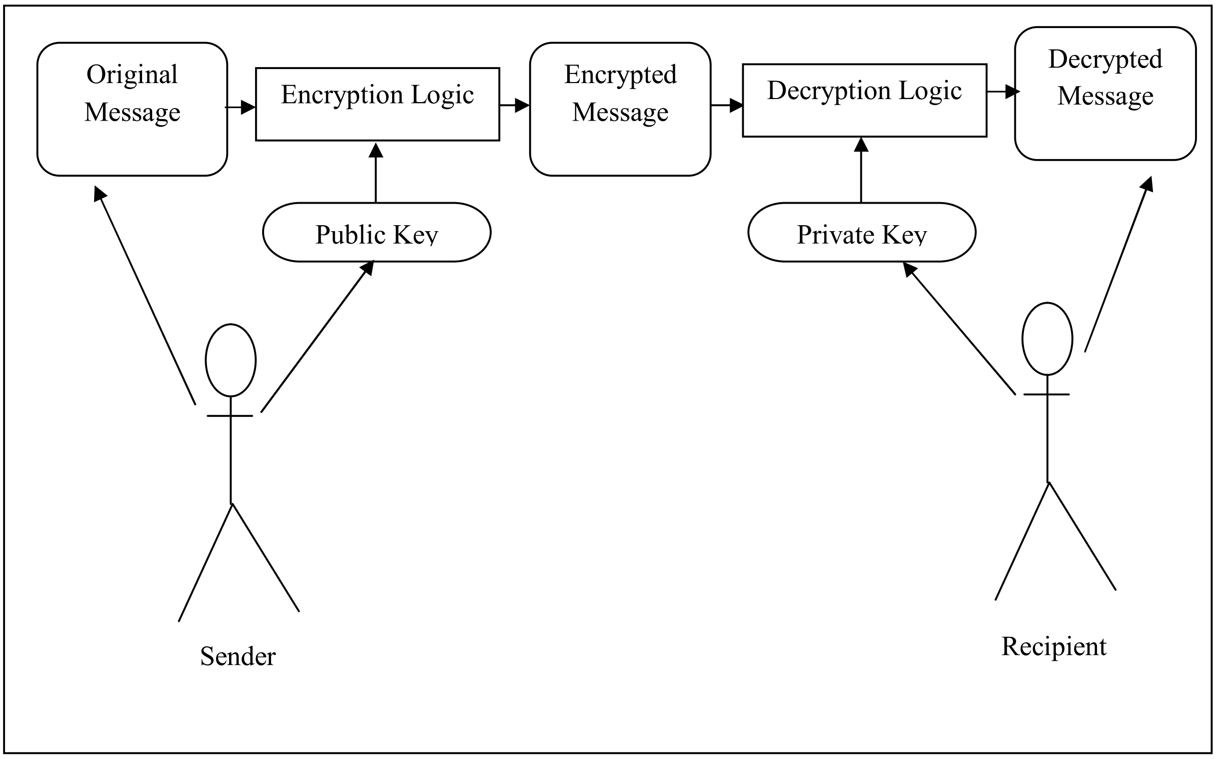

Figure 7: Model of the hybrid system

Figure 7 shows the model of the proposed system where the sender encrypts the message using the encryption key and send to the receiver, the receiver at the other end will receive the ciphertext and session key and decrypt it using the decryption key and manifest.

Sender's System Steps are as follows:

In the sender's system we are encrypting the data with the RSA encryption algorithm with the help of a randomly generated session key. Theis session key is then encrypted with the RC4 encryption algorithm and together with the encrypted cipher text is sent through a suitable channel.

1. Take a plain text or any file as input.
2. Generate private and public key by RSA algorithm.
3. Applying the RSA encryption algorithm on plain text by using key to generate cipher text.
4. The session key is encrypted using RC4
5. An encrypted file and encrypted session key are obtained.
6. Now, Send encrypted file along with encrypted session key (manifest) to destination.

Sender's system sends the two file (1) Encrypted data, (2) Encrypted session key.

Receiver's System Steps are as follows:

The receiver system decrypts the received encrypted data by applying the private key and session key. To decrypt the ciphertext, apply the receiver's session key on this ciphertext data. Follow up by applying the receivers private key. The resulting message we get is the plain text.

Receiver system receives two files 1) Encrypted data, 2) Encrypted Session Key (Manifest).

1. Receive encrypted file along with the encrypted session key (manifest) and perform cryptanalysis on it.
2. Then we will be having a Cipher text, private key, public key and Manifest
3. Apply the session key (manifest) on cipher text.
5. Apply private key on cipher text block which will give plain text.

3. SYSTEM IMPLEMENTATION

The implementation of the proposed system was accomplished using C# and XAML programming language. The reason for choosing this is because of its high security level and flexibility.

The proposed system will require the following hardware and software at minimum to work. Processor 1.0MHZ or higher, Hard disk 150GB or higher, RAM 2GB or higher and Windows 7 or higher.

4. SUMMARY

All though there are many encryption systems out there which provide some levels of security, there is always an opportunity to enhance whatever is available or come up with something more efficient and sophisticated, this can be done by developing something right from the scratch or building on existing systems, protocols and techniques. One technique of producing a more secure and sophisticated encryption and decryption system is the development of a hybrid RSA and RC4 encryption and decryption system. This system combines two different systems to produce something more robust.

The system developed in this research which combined RSA and RC4 algorithms works effectively by ensuring that only an authorized user can have access to information by making use of the new system.

The secrete message can be safely saved or transmitted without disruption from an intruder who does not have access to the encryption key.

5. CONCLUSION

Cryptography provides solution for data integrity, authentication and non-reproduction. The hybrid data encryption and decryption using RSA and RC4 has been successfully implemented. From the experience of the new system developed it has shown that the hybrid RSA and RC4 is easy to implement, fast and difficult to crack. The proposed system works effectively by providing maximum security for data transmission. This will proffer solution to security needs of both individuals and corporate organization by protecting their data from potential attack. Hence, it will allow user to communicate effectively without the fear of third-party interference with its communication.

An ideal encryption tries to provide maximum system security but can still be prone to attacks beyond prediction. The proposed protocol has been designed to be secure as possible. An added layer is one of the biggest advantages of this algorithm. However, if the attacker manages to penetrate through this complex layer by any means, the algorithm is prone to vulnerabilities as any other security algorithm available.

6. RECOMMENDATION

Considering how valuable sensitive data can be and the cost implication if lost. It is highly recommended that the new system will provide the much needed security for individuals, private and government organization who desire a secure and reliable means of transferring or transmitting data without the fear of attack of such information by a third party.

It is indeed necessary to realize that technology is extremely dynamic; therefore, it is vital to keep current on the latest and newest concept of information security.

REFERENCE

[1]. P. Gutmann, (2014), “Cryptographic Security Architecture: Design and Verification International Journal of Computer Networks & Communications (IJCNC) Vol.3, No.4, pp 68-99.

[2]. Glover, P. and M. Grant, (2013),” Digital Communications, 2nd edition, Person Education. Pp 143-167.

[3]. Joset Pieprzyk, Fundamentals of Computer Security, Springer, 2003

[4] A. Naser, H. Fatemeh and K. Riza (2013), “Developing a new hybrid cipher using AES, RC4 and SERPENT for encryption and Decryption”, International Journal of Computer Applications, vol. 69, no. 8, pp.53-62.

[5]. *Ilya Mironov (1 June 2002),* "(Not So) Random Shuffles of RC4", Advances in Cryptology – CRYPTO 2002 *(PDF), Lecture Notes in Computer Science,* ***2442****, Springer-Verlag, pp. 304–319,* doi:10.1007/3-540-45708-9_20, ISBN 3-540-44050-X*, Cryptology*

[6]. William Stallings (2015), “Cryptography and network security: Principles and practice”, Prentice Hall, Upper Saddle River, New Jersey, pp 12-32.

[7] Afolabi, A.O and E.R. Adagunodo, (2012), “Implementation of an improved data encryption algorithm in a web-based learning system,” International Journal of research and reviews in Computer Science. Vol. 3, No. 1 pp 32-49.

[8]. Meenakshi Shankar, Akshaya. P, (2014), “Hybrid Cryptographic Technique Using RSA Algorithm and Scheduling Concepts”, International Journal of Network Security & Its Applications (IJNSA) Vol.6, No.6, pp 20-78.

[9]. Akinyele A. Okedola, Yekini N. Asafe (2015), “RSA and RC4 Cryptosystem Performance Evaluation Using Image and Text File”, International Journal of Scientific & Engineering Research, Volume 6, Issue 5, pp16-36

[10]. A. P Shaikh and V. kaul, (2014), “Enhanced security algorithm using hybrid encryption and ECC”, IOSR Journal of Computer Engineering (IOSRJCE),
Vol. 6, Issue 3, pp. 80-85.

[11]. T. Charomie, (2010), “Implementation of Hybrid Encryption Method using Caesar Cipher Algorithm”, Dissertation submitted to Faculty of Computer System & Software Engineering Universiti Malaysia Pahang (UMP), Malaysia pp 48-59.