2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) | 978-1-6654-0521-8/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICACCS51430.2021.9441962

Securing Distributed Database Using Elongated RSA Algorithm

Sangeetha Radhakrishnan
Department of Computer Science,
School of Computing Sciences,
Vels institute of Science, Technology and Advanced Studies
Chennai, India
sangeetha0988@gmail.com

A.Akila
Department of Computer Science
School of Computing Science
Vels Institute of Science, Technology & Advanced Studies (VISTAS),
Chennai, India
akila.scs@velsuniv.ac.in

***Abstract*- Securing data, management of the authorised access of the user and maintaining the privacy of the data are some of the problems relating with the stored data in the database. The security of the data stored is considered as the major concern which is to be managed in a very serious manner as the users are sensitive about their shared data. The user's data can be protected by the process of cryptography which is considered as the conventional method. Advanced Encryption Standard (AES),Data Encryption Standard(DES), Two Fish, Rivest Shamir Adleman Algorithm (RSA),Attribute Based Encryption (ABE), Blowfish algorithms are considered as some of the cryptographic algorithms. These algorithms are classified into symmetric and asymmetric algorithms. Same key is used for the encryption and decoding technique in symmetric key cryptographic algorithm whereas two keys are used for the asymmetric ones. In this paper, the implementation of one of the asymmetric algorithm RSA with the educational dataset is done. To secure the distributed database, the extended version of the RSA algorithm is implemented as the proposed work.**

Keywords:-Database security, query processing, Encryption, Decryption, RSA Algorithm, Elongated RSA

I. INTRODUCTION

Cryptography is considered as the major way of securing the sensitive data. Encryption provides the confidentiality of the user information and work with the digital signature, user authentication etc [1]. So by adopting the techniques of decryption and encryption, the confidentiality ,integrity of the data user will be provided and it secure the data from leaking and being forged. While considering the asymmetric encryption algorithm, RSA is known as the most used and best algorithm till now. RSA was proposed by Rivest Shamir Adleman and hence the name RSA. [2]. As it is an asymmetric algorithm, it uses two keys for encryption and decryption of data. The public key is used to encrypt the plain text and for decrypting it to cipher text uses the private key [3]. The security of the RSA algorithm technique is purely lying on the prime number factorization which is a mathematical concept having no solution yet. RSA algorithm is the most widely used cryptosystem algorithm in digital signature standards and in encryption.

No Eaves Dropping

Fig 1:Encryption Process

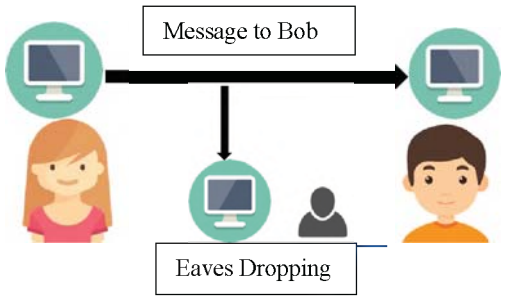
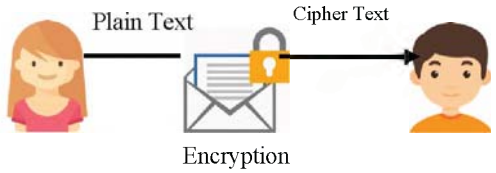


Fig 2: Without Encryption Process



II. CRYPTOGRAPHIC ALGORITHMS

A. Advanced Encryption Standard (AES)

AES is a symmetric key cryptographic algorithm which uses single key for the process of encryption and decryption. The AES is an encryption algorithm trusted by the US government and other organizations. AES is in the 128-bit form and is efficient. The key size of 192 bits and 256 bits are used for heavy duty in AES[4].AES is impermeable to the intruders attacks and the security experts considered it as the standard for the data

978-1-6654-0521-8/21/$31.00 ©2021 IEEE

encryption. Even though, the exemption of brute force is there as it decipher the text using all possible combinations.

B. Data Encryption Standard (DES)

DES is a block cipher algorithm by Davis R. which uses key which is private for both the encryption and decryption techniques. Block size of DES is 64 bits and it takes 56 bits key for encryption technique. The decryption process can be done by someone who is having the key which is used for encryption.

DES takes fixed length of plain text and transmutes through a sequence of some difficult operations into string cipher data bit of same length. Key apparently have 64 bits but 56 bits are only used for the process of encryption[5]. For parity checking, 8 bits are used and disposed then. So the effective key length used is 56 bits.

C. TRIPLE DES

Triple DES was developed to restore the symmetric algorithm DES Data Encryption Algorithm which is prone to intruders attacks by time. Triple DES is considered as the standard at some point of time and is used to a wide extent by the professionals. It takes three independent key with each having 56 bits. The length of key is up to 168 bits even though the experts would say that the strength of key is 112 bits. Triple DES makes a solution for hardware encryption for the financial services. Triple DES is slower than other block cipher[6].

D. TWO FISH

Two fish is the successor of the blowfish algorithm which is developed by an expert Bruce Schneier. It is a symmetric algorithm having the block size of 128 bit and can have the varying length up to 256 bits. This symmetric algorithm is considered as the fastest among other algorithms. Two fish is ideal to be used in any environment such as software environment and hardware environment. Two fish symmetric algorithm is also available free to all like the another symmetric algorithm Blowfish. The Feistel networks are having 16 rounds and the function of 4 prior computed dependent key S-boxes. Half of n bit key is taken as the key for encryption and the remaining half for the alteration of the process of encryption.

E. Blowfish

Blowfish algorithm is the predecessor of Two fish algorithm which is introduced by the security expert Bruce Schneier. This is a private key encryption which takes single key for the technique of encryption and decryption. It is a block cipher in which the length of the key differ from 32 to 448 bits. The size of the block is 64 bits. Blowfish algorithm deals with the encryption and the expansion of key. The expansion of key converts one key up to 448 bits in to different sub keys up to 4168 bytes. The data encryption happens by a Feistel network having 16 rounds[7]. The application which is having the key unchanged like automatic file encryption is appropriate. The blow fish algorithm is considered as the fastest of any kind when implemented with large caches on a 32 bit processor. Following are the phases of blowfish encryption algorithm:-

- Input data is 64 bits. i.e, X
- X parted into 2 similar halves.let it be x1 &x2.
- Iteration from 0 to 15
- Y=x1 xor Ki Z=f(x1) xor x2.
- Interchange x1 & x2.
- Interchange x1 & x2 undo prev.
- Y =x2 xor k18.
- Z=x2 xor k17.
- Merge x1 & x2.

III. Existing Algorithm- RSA

RSA algorithm with the educational dataset is implemented. Ron Rivest, Leonard Adleman and Adi Shamir propose a cryptographic algorithm in 1978 to replace not so secure NBS algorithm. RSA algorithm RSA algorithm is inspired from the Diffie Hellman published works from some years before, in which they explained similar cryptographic algorithm but never developed. It is defined when the era of email was about to come. RSA algorithm implemented two main aspects:

A. *Public key encryption:* It defines the need of a dispatcher for the delivery of keys to the receiver upon other secure communication channel before transmitting the original data. In RSA, encrypted key is public where the decrypted key is private [8]. So, someone having the exact key of decryption can decode the un readable(encrypted) data. Every party has their own sets of encrypt and decrypt keys. These keys have to be created in a way in which decrypt key may not be deduce from the encrypt key which is public.

B. *Digital signatures:* The recipient has to verify that the transmitted information is developed actually by the signature of the sender of information. This is done by applying decode key of the sender, and anyone can later verify the signature by using the corresponding encryption key which is public. So, these signatures can't be forged and the person who signs cannot deny later about the content of the data signed [9]. By implementing the public key algorithm, to impose the integrity, information stored is encrypted before storage. The RSA algorithm is public key algorithm having a public and private key for encrypt and decrypt.

Each user creates a unique public and non-public pair of keys by:

- Choose 2 large numbers which is prime p and q
- Calculate the modulus=p*q.
- E(n) = (p-1) (q-1).

- Select arbitrary encrypt key e where 1<e<E(n),gcd(e,E(n))=1.
- e*d=1 mod E(n),0<=d<=n.
- Public encrypt key KU=[e,n].
- *Private decrypt key KR=[d,p,q].*

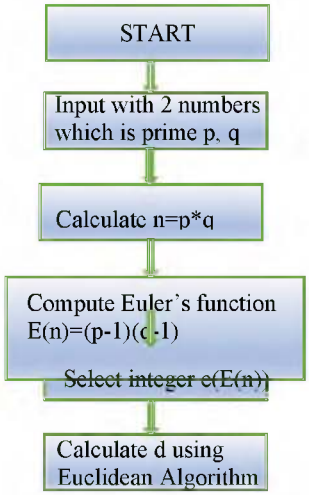
Key usage: For encoding the message M, the sender must :

- Get recipient's public key KU=[e,n].
- Calculate C=Me mod N, where 0<=M<n.

For cipher text C decryption, the receiver must:

- Use the private key KR = [d, p, q]. Calculate M C*d mod N.

It imposed the privacy of client data and enforces privacy of the client information and produces all the users not using the original information or data as it is encrypted [10]. Figure 3 illustrates the RSA algorithm.



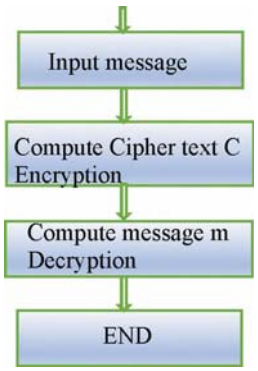


Fig 3: RSA Algorithm

IV. PROBLEM DEFINITION

The existing RSA encryption algorithm can be breached over the data network channel. To secure the data we have to make the algorithm more efficient so that the network through which the data is travelling will be safe and hence the data will be confidential. For making the algorithm more efficient, we have to enhance the algorithm by introducing a randomly generated prime number and compute the Euler's function.

V. DATASET

Educational dataset is collected for the performance evaluation of RSA algorithm in distributed database. 1441 records of dataset are used for the performance evaluation to show the efficiency of the proposed algorithm.

The fields are gender, Nationality, Place of birth, Stage ID, Grade ID, Section ID, Topic, Semester and Relation etc.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	gender	NationalIT	PlaceofBir	StageID	GradeID	SectionID	Topic	Semester	Relation	raisedhanc	VisITedRes	Announcei	Discussion	ParentAns	Parentsch	StudentAb	Class
2	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	15	16	2	20	Yes	Good	Under-7	M
3	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	20	20	3	25	Yes	Good	Under-7	M
4	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	10	7	0	30	No	Bad	Above-7	L
5	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	30	25	5	35	No	Bad	Above-7	L
6	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	40	50	12	50	No	Bad	Above-7	M
7	F	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	42	30	13	70	Yes	Bad	Above-7	M
8	M	KW	KuwaIT	MiddleSch	G-07	A	Math	F	Father	35	12	0	17	No	Bad	Above-7	L
9	M	KW	KuwaIT	MiddleSch	G-07	A	Math	F	Father	50	10	15	22	Yes	Good	Under-7	M
10	F	KW	KuwaIT	MiddleSch	G-07	A	Math	F	Father	12	21	16	50	Yes	Good	Under-7	M
11	F	KW	KuwaIT	MiddleSch	G-07	B	IT	F	Father	70	80	25	70	Yes	Good	Under-7	M
12	M	KW	KuwaIT	MiddleSch	G-07	A	Math	F	Father	50	88	30	80	Yes	Good	Under-7	H
13	M	KW	KuwaIT	MiddleSch	G-07	B	Math	F	Father	19	6	19	12	Yes	Good	Under-7	M
14	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	5	1	0	11	No	Bad	Above-7	L
15	M	lebanon	lebanon	MiddleSch	G-08	A	Math	F	Father	20	14	12	19	No	Bad	Above-7	L
16	F	KW	KuwaIT	MiddleSch	G-08	A	Math	F	Mum	62	70	44	60	No	Bad	Above-7	H
17	F	KW	KuwaIT	MiddleSch	G-06	A	IT	F	Father	30	40	22	66	Yes	Good	Under-7	M
18	M	KW	KuwaIT	MiddleSch	G-07	B	IT	F	Father	36	30	20	80	No	Bad	Above-7	M
19	M	KW	KuwaIT	MiddleSch	G-07	A	Math	F	Father	55	13	35	90	No	Bad	Above-7	M
20	F	KW	KuwaIT	MiddleSch	G-07	A	IT	F	Mum	69	15	36	96	Yes	Good	Under-7	M
21	M	KW	KuwaIT	MiddleSch	G-07	B	IT	F	Mum	70	50	40	99	Yes	Good	Under-7	H
22	F	KW	KuwaIT	MiddleSch	G-07	A	IT	F	Father	60	60	33	90	No	Bad	Above-7	M
23	F	KW	KuwaIT	MiddleSch	G-07	B	IT	F	Father	10	12	4	80	No	Bad	Under-7	M
24	M	KW	KuwaIT	MiddleSch	G-07	A	IT	F	Father	15	21	2	90	No	Bad	Under-7	M
25	M	KW	KuwaIT	MiddleSch	G-07	A	IT	F	Father	2	0	2	50	No	Bad	Above-7	L
26	M	KW	KuwaIT	MiddleSch	G-07	B	IT	F	Father	0	2	3	70	Yes	Good	Above-7	L
27	M	KW	KuwaIT	MiddleSch	G-07	A	IT	F	Father	8	7	30	40	Yes	Good	Above-7	L
28	M	KW	KuwaIT	MiddleSch	G-07	B	IT	F	Father	19	19	25	40	Yes	Bad	Under-7	M

Fig 4: Educational Dataset

VI. PROPOSED METHODOLOGY-ELONGATED RSA ALGORITHM

In the proposed methodology, a new prime number is generated. We compute the common modulus n by multiplying the two prime numbers like the traditional RSA algorithm and then randomly generated a prime

number named z. We compute the Euler's function with the 2 existing prime numbers and the randomly selected new prime number. E(n)=(p-1)*(q-1)*(z-1). Then calculated d using the Euclidean algorithm and computed the cipher text from the plain text.

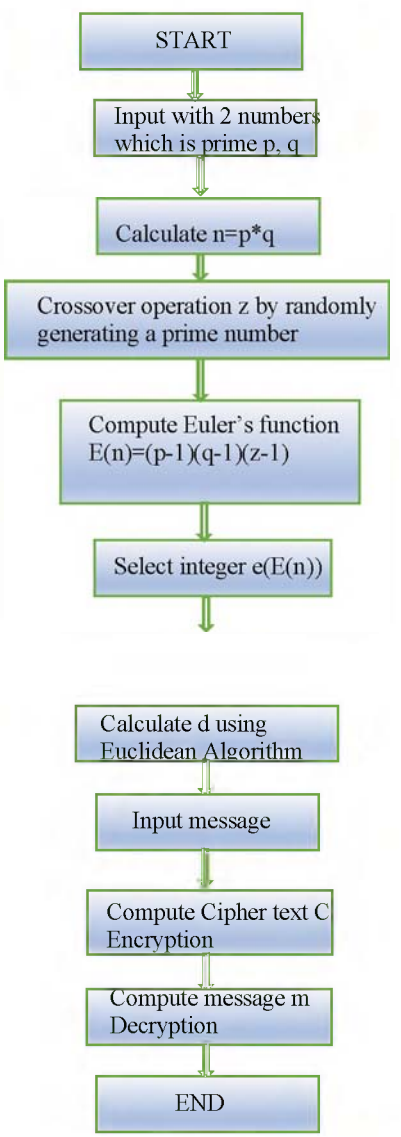


Fig 5: Elongated RSA Algorithm

A. *Psuedo Code*

STEP 1: Input with 2 numbers p and q which is prime.
STEP 2: Compute n=p*q
STEP 3: Add new prime via cross over by combining 2 primes (x, y) =z
STEP 4: Compute Euler's function
e= (p-1)(q-1)(z-1)
STEP 5: Select integer E(n, e)=1.
STEP 6: Calculate d using Euclidean algorithm
STEP 7: Input message
STEP 8: Compute cipher text c (encryption)
STEP 9: Compute message m (decryption)

B. *Advantages of Proposed Methodology*

In the proposed methodology, the robustness of the big number which is prime depends on the 3 variables q, p and z unless the traditional one. Thus making it hard for the intruder to decompose the large number into three. Generating the new random prime number makes the algorithm more secure. As the prime number is randomly generated, even the administrator doesn't have any idea of which key is used. So the key will be kept private which implies the security of the data travelling in the unsecured communication channel. The randomly generated key increases the security of the proposed algorithm from the traditional system hence making it more efficient than the traditional RSA algorithm.

VII. RESULTS AND DISCUSSIONS

A. *Performance Metrics*

1)*Accuracy*: Accuracy is taken for the performance evaluation in distributed environment for the RSA and Blowfish algorithm. Higher the accuracy, efficient the database is. Equation (1) illustrate the accuracy of RSA and elongated RSA.

$$\text{accuracy}=\frac{\text{length(encrypted data-original data)}}{\text{length(original data)}} *100; \quad (1)$$

2)*Encryption Time*: Another performance metrics taken into account for the performance evaluation of RSA and Blowfish is time. Time is considered as the important aspect in distributed environment. Lesser the time, efficient the database is. Equation (2) illustrate the formulae for time required for the encryption.

Encryption Time= Average time for encrypting data from plain text to cipher. (2)

3)*Bandwidth*: Bandwidth is a challenge in running distributed computations. Higher Bandwidth, efficient the database is. Bandwidth is the maximum capacity of the data encryption.

4)*Throughput*: Throughput should be higher, the database will be efficient. Equation (3) represents the formulae for throughput.

$$\text{Throughput}=\frac{\text{Average(Total Plain text)}}{\text{Average(Encryption time)}} \quad (3)$$

5)*Response Time*: The performance metrics Response time is a key metrics to evaluate the performance of RSA and Blowfish. Lesser the response time, efficient the

database is. Equation (4) represents the formulae for calculating response time.

$$\text{Response time} = \text{End time of execution} - \text{Start time of execution} \quad (4)$$

B. Performance Evaluation

Performance of the traditional RSA and elongated RSA algorithm is evaluated using the above mentioned performance metrics such as throughput, accuracy, bandwidth, response time and encryption time.
Accuracy of RSA and ERSA is computed and the elongated RSA is showing higher accuracy than the traditional RSA. Fig 6 illustrate the comparison of the accuracy of RSA and Elongated RSA.

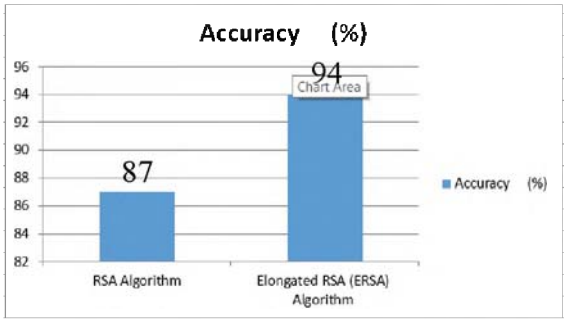


Fig 6: Evaluation of accuracy with RSA and ERSA

Bandwidth of RSA and Elongated RSA is calculated and is clearly showing that the Elongated RSA is having a higher bandwidth than of traditional RSA.

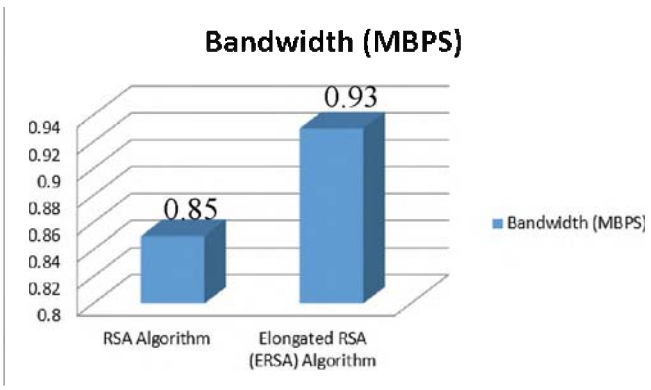


Fig 7:Evaluation of Bandwidth with RSA and ERSA

While evaluating the time required for encrypting the data using RSA and Elongated RSA, the traditional RSA takes higher time for encrypting the data. Fig 8 illustrate the comparison of the encryption time of RSA and Elongated RSA.

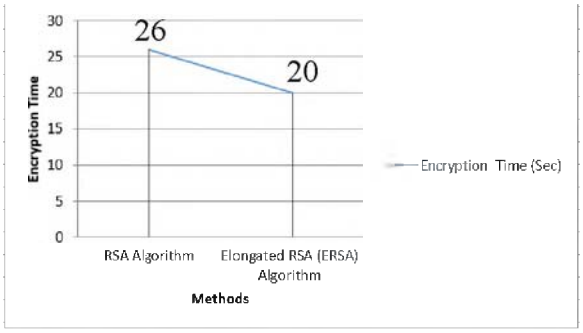


Fig 8:Evaluation of Encryption Time with RSA and ERSA

Evaluation of the throughput of RSA and Elongated RSA is done and the Elongated RSA is having a higher throughput. Fig 9 illustrate the throughput of RSA and Elongated RSA.

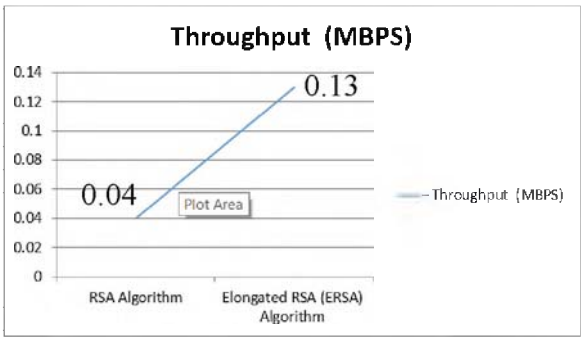


Fig 9:Evaluation of Throughput with RSA and ERSA

While calculating the response time of the RSA and Elongated RSA, Response time is higher for the traditional RSA when compared to the elongated RSA. Fig 10 illustrate the response time of both RSA and Elongated RSA.

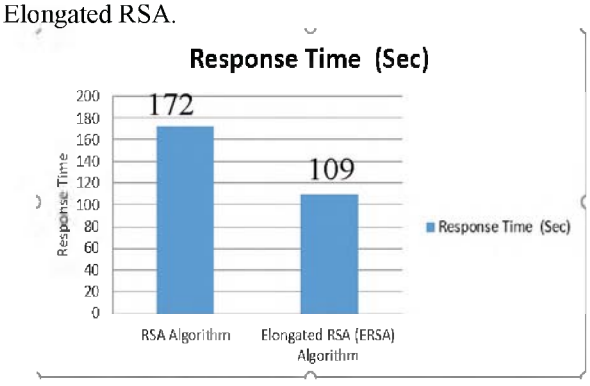


Fig 10: Evaluation of Response Time with RSA and ERSA

VIII. CONCLUSION

In data security system, cryptography plays a vital role. The cryptographic algorithms secure the confidential data which is sending through the unsecured communication channel by encryption and decryption technique. The encrypted data can be decrypted only by the authorised person thus by securing the data. This paper deals with the implementation of one such algorithm which is an

asymmetric algorithm RSA for the process of encoding and decoding of data with the educational dataset.

The proposed work is the elongated version of the traditional RSA algorithm to secure the data in the database. In the proposed work, a new variable z is introduced by generating a prime number randomly. As the generated key is random, even the administrator is unaware of the key thus by making the data more secure. The evaluation of the performance of the traditional and the elongated RSA algorithm with various performance metrics are done. The accuracy ,Bandwidth and throughput of the elongated RSA is higher compared to the traditional RSA. While coming to the time comparison of elongated RSA, encryption and response time are lower and hence concluding that the elongated RSA algorithm is better in performance compared to the traditional RSA algorithm for the dataset.

REFERENCES

[1] RIMAN, C., and Abi-Char, P. E.: Comparative Analysis of Block Cipher-Based Encryption Algorithms: A Survey. Information Security and Computer Fraud, Vol.3, No.1, 1-7, (2015).

[2] R. Rivest, A. Shamir, and L. Adleman,"A Method for Obtaining Digital Signatures and Public-Key Cryptosystems", Communications of ACM, Vol.-21, Issue-2, February 1978.

[3] Elminaam, D. S. A., Kader, H. M. A., & Hadhoud, M. M.: Performance Evaluation of Symmetric Encryption Algorithms. International Journal of Computer Science and Network Security, Vol.8, No.12, 280-286, (2008).

[4] Singh, M. G., Singla, M. A., & Sandha, M. K.: Cryptography Algorithm Comparison for Security Enhancement in Wireless Intrusion Detection System. International Journal of Multidisciplinary Research, Vol. 1, No. 4, 143-151, (2011).

[5] Sridevi, C.: A Survey on Network Security. Global Journal of Computer Science and Technology (2018).

[6] Koko, S. O. A. F. M., & Babiker, A.: Comparison of Various Encryption Algorithms and Techniques for improving secured data Communication. IOSR Journal of Computer Engineering (IOSR-JCE), Vol. 17, No. 1, 62-69, (2015).

[7] Manku, S., & Vasanth, K.: Blowfish Encryption Algorithm for Information Security. ARPN Journal of Engineering and Applied Sciences, Vol.10, No.10, 4717-4719, (2015).

[8] Evgeny Milanov, "The RSA Algorithm", 3 June 2009

[9] Ms.RituPatidar, Mrs.RupaliBhartiya, "Modified RSA Cryptosystem Based on Offline Storage and Prime Number", IJCAT International Journal of Computing and Technology,ISSN : 2348 - 6090 , Vol. 1, Issue 2, March 2014.

[10] M. Gobi and R. Sridevi, "An Approach for Secure Data Storage in Cloud Environment", International Journal of Computer and Communication Engineering, Vol. 2, No. 2, March 2013.