March - April 2020
ISSN: 0193-4120 Page No. 10092 - 10099

A Novel Knowledge-Based Authentication Framework using Correlation Coefficient and LU Factorization approaches

Namratha Pullalarevu 1, Shoba Bindu Chigarapalle 2 and Srilakshmi Sirivaram 3
1 Research Scholar, Dept of Computer Science & Engineering, JNTUA, Anantaphuramu-515001, Andhra Pradesh, India
2 Dept of Computer Science & Engineering, JNTUA, Anantaphuramu-515001, Andhra Pradesh, India
3 Dept of Mathematics, JNTUA, Anantaphuramu-515001, Andhra Pradesh, India

Article Info
Volume 83
Page Number: 10092 - 10099
Publication Issue:
March - April 2020

Article History
Article Received*: 24 July 2019*
Revised*: 12 September 2019*
Accepted*: 15 February 2020*
Publication*: 12 April 2020*

Abstract:
Secure Cloud storage is achieved by applying encryption. The prevailing method for encrypting a huge amount of text is symmetric encryption. The proposed cryptographic key management service offers Envelope encryption. Envelope encryption is a process of encrypting one key with another key. User can protect secret key against compromise attack by using threshold cryptography. Data Encryption Key is divided into a number of shares. A new Key exchange technique based on LU-factorization is proposed. Now the shares are distributed securely i.e. encrypted by the key exchanged. Users are verified by trusted third-party before acquiring a minimum number of shares for key reconstruction. A Knowledge-based authentication method using the Correlation coefficient is designed to verify the authenticated users. Hybrid cloud model comprising a private cloud and a public cloud is considered in the proposed method. Private cloud is used to store sensitive data like user details, shares of key and master keys. Encrypted data files are stored in public cloud. Comparative analysis on the proposed techniques and existing techniques is done.

Keywords: *Key Exchange, Threshold Cryptography, LU-Factorization, Knowledge-based authentication, Correlation Coefficient.*

I. INTRODUCTION

An enormous amount of data is being uploaded to the cloud by the public and private sectors as well as individuals. Uncertainties arise in the cloud due to remote operations and shared management. Data security is one of the contentious issues in cloud computing. Cryptography is a necessary technique to protect cloud data. Practice of cryptography entails the use of cryptographic keys. These keys need to be fortified. Key management comprises of key generation, key storage, key exchange, and key destruction. Instead of storing the full key it would be better to split the key and store the shares. Suppose if a user needs key he has to collect minimum right number of shares and compute the key. This technique was proposed by Adi Shamir[7]. Key exchange is a technique of exchanging the secret key between two entities. Diffie- Hellman algorithm was the most popular algorithm for key exchange. But it suffers from Man in the middle attack. Different variations of Diffie Hellman key exchange algorithms have been proposed. Prior to key exchange, the two participating entities are to be verified. Verification is done by an authentication mechanism. Authentication is done in one or combination of four methods: Something user KNOWS, something user IS, something user POSSESS and something user DOES. Different authentication methods are in practice like password and PIN-based authentication, Symmetric key and Public key authentication, Hash functions, Multi-factor authentication, Biometric, and Digital signature. Cloud service providers have access to the user's account details stored in the cloud. For every cloud service, the user has to exchange his authentication information and it may lead to an exploit of authentication technique. Knowledge-based authentication mechanisms include the use of an alphanumeric password, a personal identification number (PIN) or a graphical secret. Users are using very simple passwords, writing them down or reusing the same passwords for different service thus breaking basic security rules. Users should be verified based on both static and dynamic information.

March - April 2020
ISSN: 0193-4120 Page No. 10092 - 10099

II. PRELIMINARIES

2.1 LU-Factorization

LU-Factorization of a matrix is the process of factoring.

A given square matrix into two triangular matrices, one upper triangular matrix, and one lower triangular matrix.

Original matrix is calculated as the product of these two matrices. By applying Gauss Elimination method L and U are formed.

For A $=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

Where L $=\begin{pmatrix} 1 & 0 & 0 \\ L_{21} & 1 & 0 \\ L_{31} & L_{32} & 1 \end{pmatrix}$ U$=\begin{pmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{pmatrix}$

Such that A = L*U

2.2 Correlation Coefficient

A correlation coefficient measures a statistical relationship between two variables. Karl Pearson Correlation coefficient 'r' is a measure of correlation between two variables x and y.

$$r=\frac{n(\sum xy)-(\sum x\sum y)}{\sqrt{(n\left(\sum x^2-\left(\sum x^2\right)\right)*(n\left(\sum y^2\right)-(\sum y)^2\ )}}$$

2.3 The inverse of a matrix

The multiplicative inverse of a matrix is known as its inverse matrix. An Inverse exists for a matrix if it is Non-Singular i.e determinant of the matrix should be non-zero. The inverse of a matrix A is denoted as A^{-1}. A^{-1} satisfies the following property:

$A * A^{-1} = I$, where I is the Identity matrix.

Envelope Encryption

Envelope Encryption is a technique of encrypting data with a Data Encryption Key(DEK) and then encrypting DEK with a Key Encryption Key(KEK). AWS, Google cloud and IBM cloud are using envelope encryption in the key management service.

2.5 Shamir's Secret sharing

Adi Shamir's threshold sharing scheme is based on the idea that in order to determine a polynomial of degree k-1, k points are sufficient.

To share our secret S in a (k, n) threshold scheme, Choose at random k-1 positive integers $a_1, \ldots, a_{k-1}$ with ai, and let a_0=S. Build the polynomial

$f(x)=a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots + a_{k-1}x^{k-1}$.

Let us construct any n points out of it, for instance, set j = 1. . . n to retrieve (j, f(j)). Every user is given a point (an integer input to the polynomial and the corresponding integer output). Given any subset of k of these pairs, the user can find the coefficients of the polynomial using Lagrange interpolation. The secret is the constant term a0.

III. LITERATURE REVIEW

Gadhavi et al., in [5] proposed a variation of the Diffie-hellman key. Exchange model based on arbitrary numbers and logarithms. This key exchange protocol prevents man in the middle attack and provides three basic security services - confidentiality, authentication, and integrity.

Jaikar et al., in [4] implemented a secure data distribution scheme using secret splitting. Secret splitting is accomplished by performing XOR operation between the original file and a dummy file, such that dummy file size is equal to original file size.

Karen Renaud et al., in [3] introduced a Snippet knowledge-based Authentication where users were asked to give 5 snippets of code in java. Users were authenticated by asking them to select their code among all the snippets of code. This method tested the memorability, observability and guessability of an authentication mechanism that relied on an expert programmer identifying his/her own code snippets and is resistant to shoulder surfing attack and guessing attack.

Sonia Chiasson et al., in [1] presented an evaluation of persuasive cued click- points. Users have to select a part of an image as a password and remember it. For initial logins i.e. within 2 or 4 days user-memorized password. But after two weeks success rates were very low.

IV. PROPOSED WORK

The proposed framework is designed into four phases:

1. Registration phase
2. Key Splitting Phase

March - April 2020
ISSN: 0193-4120 Page No. 10092 - 10099

3. Key exchange Phase
4. Knowledge-based Authentication Phase

Registration Phase

Users get registered with the cloud service provider (CSP) for data access by providing their own details. The user can specify the owner details whose data he wants to access. After successful registration, CSP provides credentials for the user. In the future, if any user wants to access data then he can log in with credentials provided by CSP.

Key Splitting Phase

It is recommended that full key should not be stored as well as transformed because of compromise attack. The solution for this problem is to split the key into a number of shares using Shamir's secret sharing technique and let the user store only his share.

In the proposed model owner initially splits the key into n shares (s_1, s_2. . . s_n) and the owner stores one share among the shares and securely distributes remaining shares to Trusted Third Party (TTP) and registered users. Secure distribution of shares to registered users is achieved by encrypting share with the key exchanged using LU factorization and to TTP is achieved by encrypting share with K_s (Master key between Owner and TTP). Here all users belong to one group identified by group id.

Suppose any user (u_1) wishes to decrypt a file then he requests a minimum (threshold) a number of shares from the arbitrary users where owner and TTP share is mandatory. Other users belonging to the group send their shares using the asymmetric encryption algorithm (RSA) i.e. User encrypts his share using the public key of u_1 and sends an encrypted share to u_1. User u_1 will decrypt the share using his private key.

After getting, a minimum number of shares user can reconstruct full key and can decrypt the file.

Suitable threshold value will be 31 % of all shares 'n'. Janaratchkool et al., in [2] have given this value by collecting key distribution time and key reconstruction time to achieve optimal threshold value.

Key Exchange

Key exchange phase is executed twice. Once when an owner wants to distribute the shares of users. During the registration phase, the user can mention owner details whose data he wants to access. As per the details those owners send shares to users. Twice when the registered user wants remaining shares to reconstruct the full key. For the second time, the user needs to be verified. So, once authentication is successful, TTP starts key exchange algorithm. Authentication is explained in the further section.TTP and individual users share master keys. TTP distributes the secret key among the parties.

LU factorization factors a matrix as a product of Lower triangular matrix (L) and Upper triangular matrix (U). If user R wants to access file owned by owner S, S sends a request to R. File owner S sends a request to TTP for generating a secret key. TTP generates a secret key matrix randomly and decomposes that secret matrix into L and U matrix. TTP arbitrarily sends L, U to both users A and B by encrypting them with the corresponding master keys (K_S, K_R).

Algorithm for LU factorization is given below.

Algorithm 1: LU-Factorization(A)

Input: Keymatrix A
Output: L,U

```
1  n ← rows[A]
2  for k ← 1 to n
3        do U_kk ← A_kk
4              for i ← k + 1 to n
5              do L_ik ← A_ik/U_kk
6                 U_ki ← A_ki
7              for i← k + 1 to n
8                 do for j← k + 1 to n
9                 do A_ij ← A_ij − L_ik U_kj
10 return L and U
```

Steps involved in Key exchange:

1. TTP generates a random key matrix K_n with size nXn.
2. TTP applies LU factorization on K_n matrix and results in two matrices K_{Ln} and K_{Un}.
3. TTP encrypts $K_{Ln,\ Lifetime}$ of the key and Nonce N_2 with K_S and sends cipher text to the data owner.
4. TTP encrypts K_{Un} , Lifetime of the key, Nonce N_2 and N_1 with K_R and send cipher text to the user.
5. Owner sends $(K_{Ln}*A_{n)} + N_3$ to the user.
6. User sends $(K_{Un}*A_{n)} + N_4$ to the owner.
7. Owner calculates $K_n = K_{Ln}*(K_{Un}*A_n)*A_n^{-1}$.
8. User calculates $K_n = (K_{Ln}* A_n)*A_n^{-1} *K_{Un}$.

Owner sends one of the shares encrypted with K_n to the user. AES encryption algorithm is implemented for encryption.

March - April 2020
ISSN: 0193-4120 Page No. 10092 - 10099

Following diagram depicts the key exchange phase:

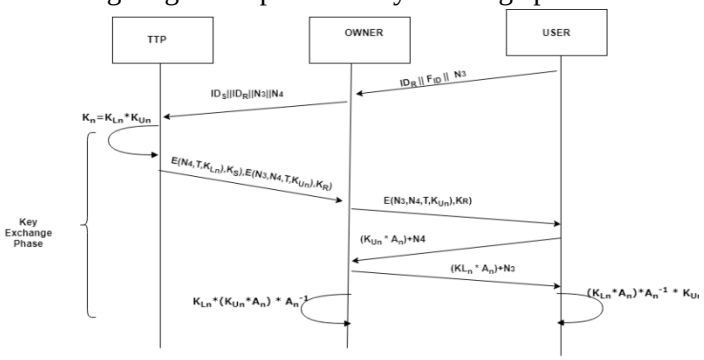

Fig 1: LU-Key Exchange

Knowledge-based Authentication Phase

Knowledge-based Authentication (KBA) protocol using correlation coefficient and a novel key exchange algorithm is proposed. Trusted Third party (TTP) service is provided by the cloud and authentication is performed by TTP. The proposed KBA method utilizes both static KBA and dynamic KBA. Static KBA consists of queries on user data during registration. Dynamic KBA consists of queries based on the previous session of the user. Queries of static KBA and dynamic KBA are further classified in to hard and medium. Now there are four types of queries static hard, static medium, dynamic hard and dynamic medium. After verifying the user, key exchange algorithm is initiated by the TTP.

Data user sends a request to the data owner for particular file accessibility. Data Owner in turn requests trusted third party (TTP) for authentication and key exchange. TTP send challenges to both the data owner and the data user. These challenges consist of static queries and dynamic queries which are again two types hard and medium. For these four types of queries, a fixed probability is assigned such that static has less probability when compared to dynamic and medium has less probability when compared to hard. Each query will be associated with a bounded time. Probability represents response function (y) and Time represents influential function (x). Karls Correlation coefficient (r) is calculated on probability and time values. ‘r’ value and is placed as a threshold.

For example, probabilities are assigned as shown in the following diagram:

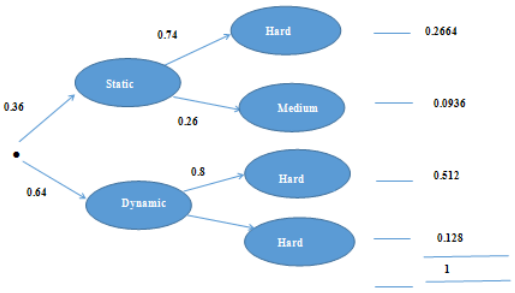

Fig 2: Probabilities of queries

Trusted third party randomly picks some queries from two sets and allows the user to answer them.

Suppose user response time exceeds the bounded time, correlation coefficient value increases otherwise value may be equal to or less than the threshold. The correlation coefficient for the responses is calculated and is denoted as ‘r`’, it is compared with the threshold value. If r` < = r and r` > 0, then accept user as authenticated otherwise reject the user as not authenticated. The Authentication process is explained in the following flowchart.

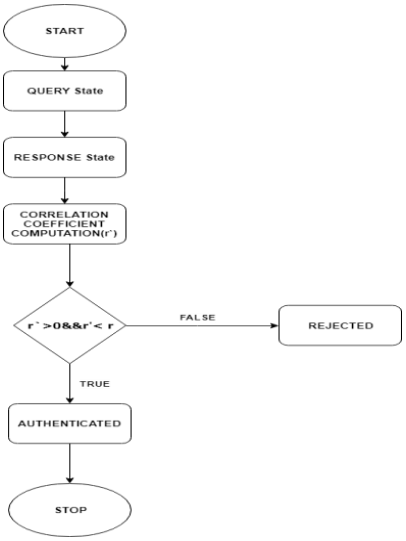

Fig 3: Flowchart of the Authentication process

The following table gives notations for authentication process and key exchange algorithm.

March - April 2020
ISSN: 0193-4120 Page No. 10092 - 10099

Table 1: Notations and their descriptions

Notation	Description
ID_S	*Identity of Owner*
ID_R	*Identity of User*
E	*Encryption*
A_n	*n X n Pre-agreed Matrix between Owner and user*
K_S	*Master Key between TTP and Data Owner*
K_R	*Master Key between TTP and Data User*
K	*n X n Key matrix*
K_{Ln}	*n X n Lower Triangular matrix*
K_{Un}	*n X n Upper Triangular matrix*
A_n^{-1}	*Inverse of n X n Pre-agreed matrix*
r_R`	*Correlation coefficient of User*
r_S`	*Correlation coefficient of Owner*
T	*Lifetime of Key*

After successful verification of the owner and the user, key exchange process is initiated.

Entire knowledge-based authentication process and key exchange process is explained in the following diagram.

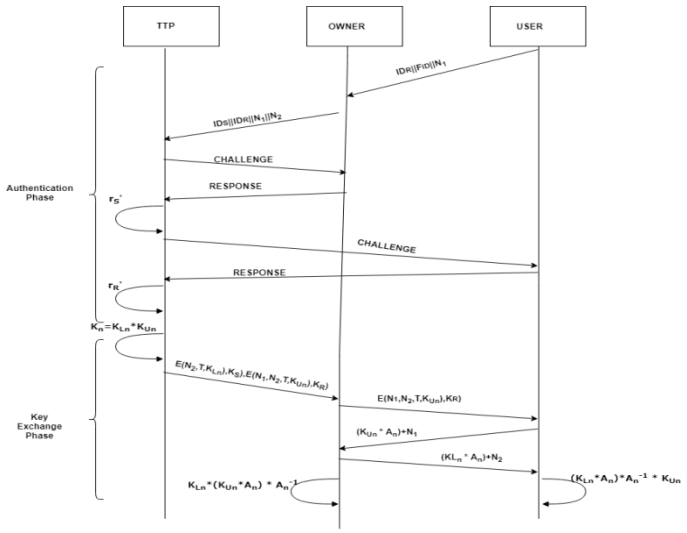

Fig 4: Authentication and Key exchange

In this way, authenticated key exchange is carried among owner and user. Owner sends his own share encrypted by the key exchanged with the user.

The reason for executing the key exchange phase twice is because of the far gap between registration phase and the authentication phase. Owner and user cannot use initial key exchanged for a long time as the key lifetime expires.

V. RESULTS AND DISCUSSION

A new LU-Factorization based Key exchange and KBA using

correlation coefficient algorithm are implemented in Python. KBA poses a mixture of time bounded static and dynamic challenges for the user to respond. Comparing user's response coefficient with threshold can detect and avoid unauthorized u access to the stored data.

Cryptanalysis

Cryptanalysis is an investigation of the system to obtain the meaning of encrypted information without knowing the secret information. Cryptanalysis of the two novel methods proposed in the framework is illustrated below.

Man in the middle attack

Man in the middle (MITM) attack can be launched at two sites. One site can be between TTP and Owner and another site is between Owner and user. MITM cannot be executed between TTP and owner because the messages are encrypted. MITM attack cannot be executed between Owner and user because the attacker has to guess nonce as well as An (pre-agreed matrix) to send his own product to owner and user.

March - April 2020
ISSN: 0193-4120 Page No. 10092 - 10099

Replay attack

A replay attack is a network attack in which a valid message is maliciously repeated. In the proposed system replay attack is not possible, since nonce n1 and n2 are used. The nonce can be used to identify whether a message is old or new.

In the case of authentication replay attack fails because response(dynamic) changes with time.

Factorization attack

It is infeasible to factor lower triangular matrix or upper triangular matrix from the resultant product i.e. KLn*An and KUn*An. Extracting one factor of a matrix from the product of two matrices is very difficult.

Guessing attack

It is a difficult task for the adversary to guess key matrix as it is decomposed into two matrices. An attacker cannot guess either the factored matrices or original matrix.

Brute force attack

Considered Key matrix size is 3X3,the key space is very large i.e. 26P9 (Alphabets are considered as key characters). An adversary has to guess all the nine elements of matrix correctly, which is infeasible.

In the case of authentication, the brute force attack cannot be executed. Since the challenge-response model applies both static and dynamic queries- which change with time and also, the correlation coefficient is computed for each response so that the attacker can be identified as unauthorized.

Performance analysis

A comparison of Diffie –Hellman key exchange [9] and LU-Factorization key exchange is done in terms of mathematical operations, computation complexity and types of attacks.

Table 2 gives the comparison details of two key exchange algorithms. Diffie-Hellman is a public key exchange algorithm. The proposed key exchange doesn't require a public key. Mathematical operation used in LU-factorization is matrix multiplication.As L and U matrix conatins zero's, L and U occupies less space and matrix multiplication is done very fast. None of the two key exchange algorithms can run in polynomial time.

Standard size of key exchanged for Diffie-Hellman is 1024 bits, RSA is 2048 bits, ECDH(Elliptic curve Diffie Hellman) is 256 bits and the proposed LU Factorization is 128 bits. Time required for key exchange among two users by different key exchange algorithms is shown in the following graph (Fig 5).

KBA using correlation is compared with SNIPPET[3] and Persuasive cued click- points KBA[1].Table 3 illustrates the comparison parameters of three methods. Two existing KBA methods[1] [3] apply static challenge. So, there is a chance of password guessing attack. The proposed KBA method applies both static and dynamic challenges and can overcome password guessing attack, Bucket Brigade attack, and Replay attack.

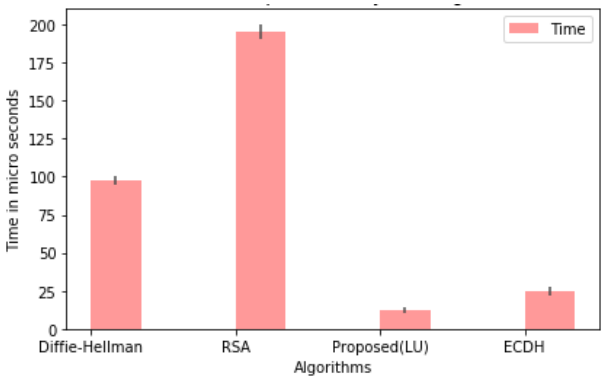

Fig:5 Time for key exchange

Table 2: Comparison of Diffie –Hellman and LU- Factorization

Key Exchange	Mathematical operation	No of Computations	Strength	MITM	Public key Exchange	Time Complexity
Diffie-Hellman	Exponentiation	4	Discrete Logarithm problem	Yes	Yes	O(M(n) k)
LU-Factorization	Matrix Multiplication	6	Factorization Problem	No	No	$O(n^{2.373})$

March - April 2020
ISSN: 0193-4120 Page No. 10092 - 10099

Table 3: Comparison of KBA mechanisms

Authentication mechanism	Method	Password Guessing attack	Replay attack	Bucket Brigade attack	Challenge-Response type
SNIPPET	Code Snippets	Yes	Yes	Yes	Static
Persuasive –cue-clicked points KBA	Click based graphical passwords	Yes	Yes	Yes	Static
KBA-Correlation coefficient	Probability based Challenge –Response method	No	No	No	Static & Dynamic

VI. CONCLUSION

Cloud offers storage as a service to users. Data owners can store their information in cloud storage data centers which are located far away from users. Data security has received much attention over the last decade. Encryption is obviously a necessary technique applied to secure data. Symmetric encryption algorithm needs a secret key which is to be securely distributed to users. To defend key compromise attack Shamir's secret splitting is used to divide secret key into shares. A new key exchange algorithm using LU factorization is implemented. A new knowledge-based authentication method using correlation coefficient is proposed to verify the authenticity of the user. Cryptanalysis shows that LU-Key exchange can defend the man in the middle attack, factorization attack, and brute force attack. Replay attack can be avoided by both key exchange and authentication techniques. Performance analysis concludes that the proposed key exchange algorithm's strength is factorization problem and can defend man in the middle attack, where as KBA-correlation coefficient is a challenge response type of authentication which can counter attack replay and bucket brigade attacks. Time required for key exchange using new algorithm is less when compared to existing algorithms. Future research might integrate these techniques into a cloud environment.

VII. REFERENCES

1. Chiasson, Sonia & Stobert, Elizabeth & Forget, Alain & Biddle, Robert & C. van Oorschot, Paul. (2012). Persuasive Cued Click-Points: Design, Implementation, and Evaluation of a Knowledge-Based Authentication Mechanism. IEEE Trans. Dependable Sec. Comput.. 9. 222-235. 10.1109/TDSC.2011.55.
2. Janratchakool, Weena & Boonkrong, Sirapat & Smanchat, Sucha. (2016). Finding the Optimal Value for Threshold Cryptography on Cloud Computing. International Journal of Electrical and Computer Engineering (IJECE). 6. 2979-2988. 10.11591/ijece.v6i6.11573.
3. Gupta, Alisha, and Vivek Sharma. "Implementation of LEACH protocol using Homomorphic Encryption." International Journal of Electrical and Electronics Engineering (IJEEE) 2.4 (2013).
4. Renaud, Karen & Kennes, Demetris & van Niekerk, Johan & Maguire, Joseph. (2013). SNIPPET: Genuine knowledge-based authentication. 2013 Information Security for South Africa - Proceedings of the ISSA 2013 Conference.1-8. 0.1109/ISSA.2013.6641059.
5. Jaikar, Sagar & C Maheshwar, Ramkrushna & P Mamadapure, Sameer & Bhosle, Anand. (2017). Secure Data Distribution using Secret Splitting Over Cloud. Global Journal of Computer Science and Technology : B Cloud and Distributed. 17.
6. Bangar, Ashwini, and Swapnil Shinde. "Study and comparison of cryptographic methods for cloud security." Int J Comput Sci Eng Inf Technol Res 4.2 (2014): 205-213.
7. Gadhavi, Lata & Bhavsar, Madhuri & Bhatnagar, Monica & Vasoya, Shivani. (2016). Design of efficient algorithm for secured key exchange over Cloud Computing.180-187. 10.1109/CONFLUENCE.2016.7508110.
8. Abbas Othman, Abdelahaliem. (2013). Binary LU encryption. 192-195. 10.1109/ICCEEE.2013.6633931.
9. https://en.wikipedia.org/wiki/Shamir%27s_Secret_Sharing
10. Praptodiyono, Supriyanto & Furqon, Moh & Maulana, Alief & Hasbullah,Iznan & Rehman, Shafiq. (2018).
11. Performance Analysis of Internet Key Exchange Algorithms on IPSec Security Association Initiation.

March - April 2020
ISSN: 0193-4120 Page No. 10092 - 10099

MATEC Web of Conferences. 218. 03001. 10.1051/matecconf/201821803001.

12. Karthik, P., P. S. Ranjith, and M. Jayagnesh. "SCCE: Secure Communication Based on A Chaotic System for Modern Wireless Communication." International Journal of Research in, Engineering & Technology (IMPACT: IJRET), ISSN (E) (2014): 2321-8843.
13. W.Diffie and M.E. Hellman. New directions in cryptography. IEEE Transactions on Information Theory,IT-22(6):644-654 1976.
14. P. Dileep Kumar Reddy, R. Praveen Sam, C. Shoba Bindu "Tripartite partite key assignment scheme for Security of cloud data classes" Journal of Theoretical and Applied Information Technology, 15th July 2017. Vol.95.No 13, ISSN: 1992-8645, E-ISSN: 1817-3195 Pg.No:3116-3126.
15. Wadhwani, Priyanka, Akanksha Gaur, and Vipin Jain. "Cryptanalytic JH and Blake Hash Function for Authentication and Proposed Work Over Blake-512 on C Language." International Journal of Computer Science Engineering and Information Technology Research 4.3 (2014): 187-198.
16. P. Dileep Kumar Reddy, R. Praveen Sam, C. Shoba Bindu "Optimal Blowfish Algorithm based Technique for Data Security in Cloud" Int. J. Business Intelligence and Data Mining, ISSN online 1743-8195, ISSN print 1743-8187, Vol. 11, No. 2, 2016.Pp.171–189.DOI: 10.1504/IJBIDM.2016.10001484. (Inder Science)(UGC Approved). Journal No: 16481
17. P. Dileep Kumar Reddy, C. Shoba Bindu,R. Praveen Sam "A Tripartite Partite Key generation and management for Security of Cloud Data Classes",2nd International Conference on Computational Intelligence And Informatics (ICCII 2017), JNTUHCEH, Department of Computer Science &Engineering, Hyderabad., September 25th to 27th2017. Springer AISC.
18. Bishoi, Tanmoy Kumar, Ramkrishna Ghosh, And Tanmoy Sinha Roy. "An algorithm on text based security in modern cryptography." J Comput Netw Wirel Mobile Commun 5.1 (2015).
19. Namratha, P and Bindu, C Shoba and SriLakshmi, S, Securing Cloud Data Using Laplace Transform Based Encryption (February 7, 2018). 2018 IADS International Conference on Computing, Communications & Data Engineering (CCODE).
20. Khami, Mohammed Jawar. "Unlimited Size of English Plain Text-in-Text Hiding Algorithm." International Journal of Computer Science and Engineering (IJCSE) 6 (2017).
21. P. Namratha, C. Shoba Bindu and S. Sri Lakshmi Key Generation and Encryption Using Laplace Transform for Data Security in Cloud JARDCS volume 11 issue 6 pp 174-180