ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

Technique for Secured Search on Encrypted Information based on Similarity Join Queries

Shital P. Patil [1], Prof. P. B. Mali [2]

P.G. Student, Department of Computer Engineering, SKN COE, Savitribai Phule University of Pune, Pune, India[1]

Asst. Professor, Department of Computer Engineering, SKN COE, Savitribai Phule University of Pune, Pune, India [2]

ABSTRACT: Today many users and organisations store their data on third party storage like public cloud to reduce storage as well as maintenance cost. Third party storage provide public access to outsourced information because of this data leakage may occur frequently. In order to protect outsourced data must be stored in secured format it reduces data breaches. To access outsourced data exiting system try to tackle this problem by performing search operation on encrypted data without disclosing it, but existing work has lack of security because data owner share secrete key with data user and token generation process requires more time. To overcome this problem this proposed technique is designed. It employs similarity search which discovers pair-wise similar data points from two datasets without sharing secret key.

KEYWORDS: Intermittently Cloud computing, Encrypted search, Security issues, Privacy preserving, Similarity search, Cloud storage, Data security.

I. INTRODUCTION

Many users and organisations stored data on cloud it allows users to outsource their data on cloud storage to reduce storage cost and cost of maintenance. Third party storage has public access users facing security issues of their data which is stored on cloud storage because any one can access data on public cloud. Cloud service providers try to provide primary security i.e. protection of firewall and virtualization but these mechanisms are not able to provide strong protection. In order to protect that data must be stored in secured format or in encrypted format but this mechanism has one problem how to perform searching on secured information in order to resolve this problem existing technique provide mechanism by using similarity join but existing work was not able to provide high level of security because it involves sharing of secret keys between data owner and data user. Because of this reason information leakage may happen.To solve this problem this proposed work is designed in this work interaction between data user and data owner is reduced here we are creating application software which will be responsible to perform encryption and decryption task here data owner and data user will not perform these task, because of this reason security level increase and time required to process data or sharing of secret key is reduced.

II. PROPOSED METHOD

To designed this work Locality-Sensitive-Hashing and Symmetric Searchable Encryption is employed. Locality-Sensitive-Hashing is used to generate hash index for file and Symmetric Searchable Encryption is used to employ secured search on encrypted information. Fig 1 shows proposed system architecture. It has three parts client, data owner who owns data, and public cloud server. Owner of dataset uploads his dataset to application software to generate secure dataset and to generate indexed file for data set. Resulted file of application software includes encrypted data set and hash index file to create index for each row. Then data owner upload both data set and index file on cloud server.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

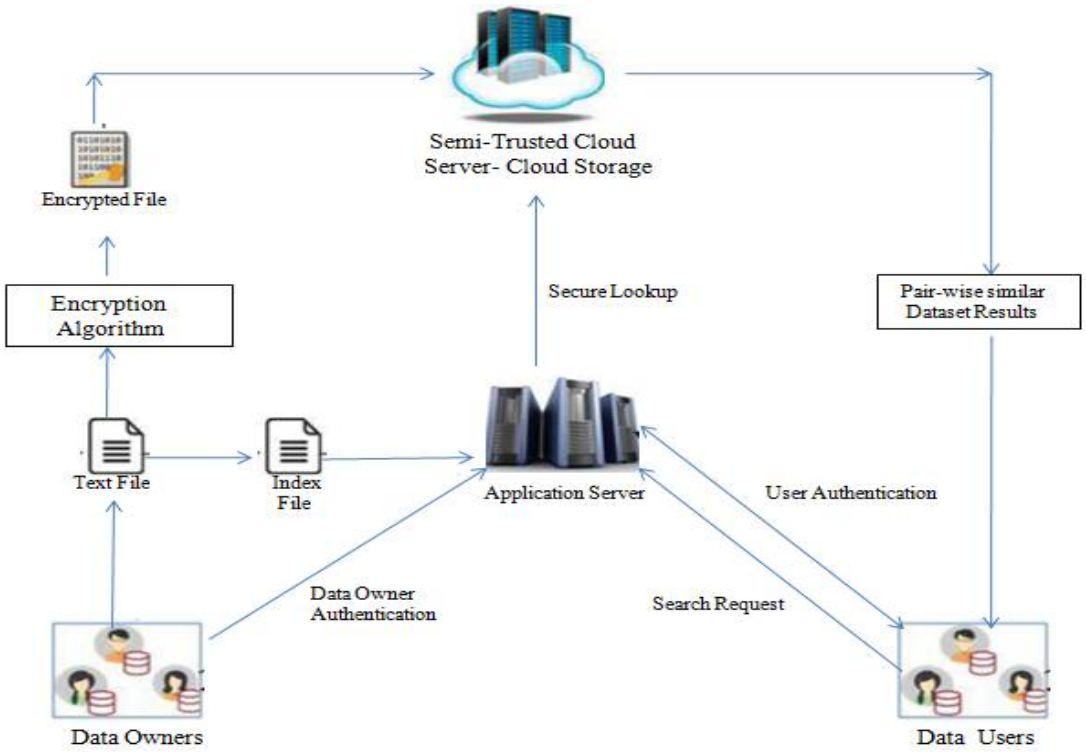

Fig 1: Proposed System Architecture

When data user wants to access data he generates a query set and upload it to application software it will generate index for each row resulted file will upload to cloud server by data user. Here in this proposed work data user and data owner don't have knowledge about which encryption, decryption and index creation mechanism is used only application software has knowledge about it. This technique is secured because here data user and data owner are not sharing any kind of information or not sharing any secrete keys. Here we are not sending any kind of information to loss to it provides high security level and reduces query processing time.

III. RELATED WORK

In [1] paper characterize a fitting primitive, for a general closeness work on the message space called as efficiently fuzzy-searchable encryption (EFSE) and recognize an ideal security thought for EFSE it enables a fuzzy search on information this technique returns large set of results it requires more time to process. In [2] paper provides order preserving technique to perform search on information in which order is preserved either alphabetically or numerically based on order performs search operation by employing Order Preserving Encryption. In Order Preserving technique order of data set is preserve so it has information leakage limitations. In [3] paper to address Leakage-profile attack against accessible information introduce a description of the spill profiles of the wild accessible encryption items and SE plans and present violation models in view of an disposed servers earlier information.
In [4] paper explore a wealthier setting in which the information owner D outsources its information to a server E yet D is captivated to permit customers (outsiders) to look through the database to such an extent that customers take in the data D approves them to learn yet nothing else while E still does not send out information or questioned admire as in the fundamental SSE. In [5] paper empower a protected and proficient cloud assisted image sharing architecture for mobile devices, by utilizing outsourced encoded picture datasets with security. This work is introduced to perform secured images sharing on mobile devices. In [9] paper propose a productive plan for similarity search over encrypted information. To do as such, use a state-of-the-art algorithm for quick near neighbor look in high dimensional spaces

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

called locality sensitive hashing which generates hash indexes for large data set. In 10] paper layout a few basic security challenges and facilitate examination of security solutions for a reliable open cloud condition. In public cloud reasons of data breaches and some techniques to protect this outsourced information is provided. In [11] paper, depicts cryptographic plans for the issue of looking on encoded information and give verifications of security to the subsequent crypto frameworks. In [14] paper presents methods that change the information preceding providing it to the specialist for comparability questions on the changed data. In [15] paper start the principal try towards protection saving picture denoising from outer cloud databases. This configuration empowers the cloud facilitating scrambled databases to give secure question based picture denoising administrations.

IV. PROPOSED ALGORITHMS

A. AES Encryption Algorithm:

State=M
AddRoundKey(state,&w[0])
Initialize state array and add the initial round key to the starting state array.
For i=1 step 1 to 9
Perform round = 1 to 9 : Execute Usual Round.
Usual Round: Execute following operation:-

1. Sub Bytes
2. Shift Rows
3. Mix Columns

Add Round key, using K(round)
Final Round : Execute following operations:-

1. Sub Bytes
2. Shift Rows
3. AddRoundKey,using k(10)

B. SHA 256:- In SHA-256 massage to be hashed first.

(1) Padded with its length in such a way that the result is a multiple of 512 bits long, and then
(2) Parsed into 512-bit message blocks $M^{(1)},M^{(2)},\ldots,M^{(N)}$.
The message blocks are processed one at a time: Beginning with a fixed initial hash value $H^{(0)}$, sequentially compute
$H^{(i)}=H^{(i-1)}+C_M^{(i)}(H^{(i-1)})$,
Where C is the SHA-256 compression function and + means word-wise mod 2^{32} addition $H^{(M)}$ is the hash of M.

PSEUDO CODE

Encrypted dataset and Encrypted Index Creation:
Step1. Dataset Input to Application
Step 2. Compute of hash value of each hash function
Step 3. Export Resulted File
Step 4. Generate counter for hash value in hash table
Step 5. Encryption of key-value pair of matched data points.
Step 6. Upload File on Server

Function of Secure Similarity join
Step 1. Dataset Input to application
Step 2. Compute hash value of each hash function
Step 3 Export resulted file
Step 4. Upload Query set
Step 5. Perform Similarity Search
Step 5. Output : pair-wise similar dataset.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

V. PERFORMANCE RESULT AND ANALYSIS

This Proposed approach is designed using c#. To evaluate its performance we used detailed bank transaction set. Performance and results of proposed technique are shown in following graphs In fig. 2 shows level of security improve in proposed approach than existing one. In existing system security level was so less because data owner and data user share secret key to access data user will use that key to access information. Security level is improved in proposed approach by using application software which performs

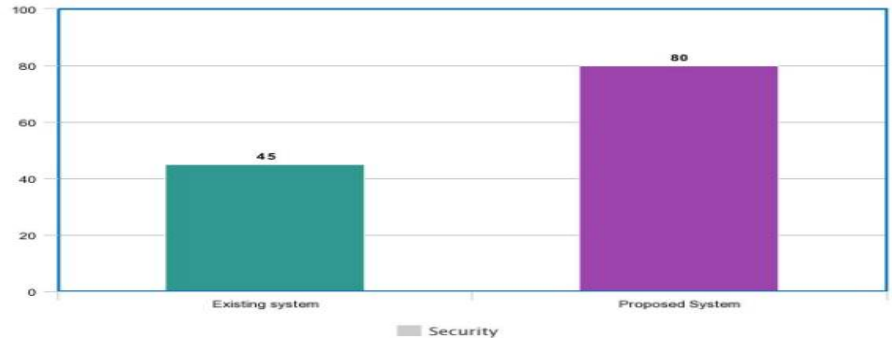

Fig. 2. File Security Evaluation

encryption and decryption of dataset because of this there is no need to share secret keys by data owner and date user it preserves the security. Fig 3. Shows the performance of the system in existing system performance was so low because token generation process in proposed work we eliminate the process of token generation between data user and data. In fig. 4 shows scalability of the system. Existing system was not that much scalable on large dataset proposed system overcome this by transferring only required query set in encrypted format.

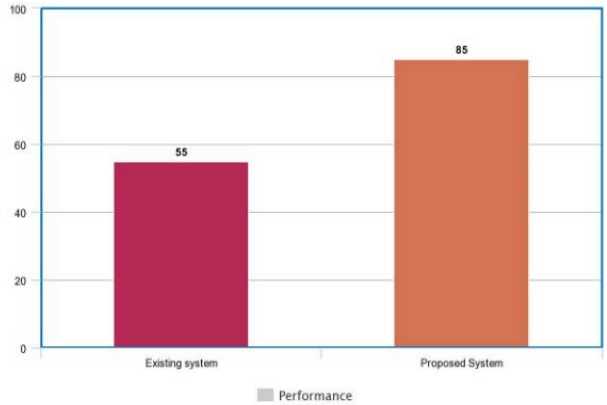

Fig 3. Performance Evaluation

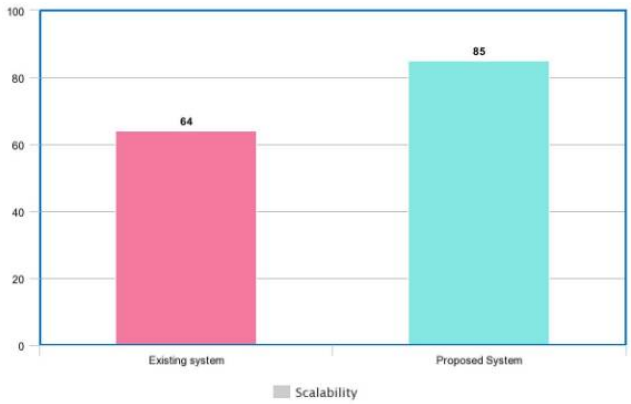

Fig 4. Scalability Evaluation

VI. CONCLUSION

The proposed approach employ secure search on encrypted information using similarity join queries which find out matched set across two data. This approach provides secured search without knowledge of encryption and decryption to data user and data owner and without sharing of secrete keys.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

REFERENCES

1) Boldyreva and N. Chenette. "*Efficient fuzzy search on encrypted data*" In Proc. of FSE, 2014.
2) A. Boldyreva, N. Chenette, and A. ONeill. "*Order-preserving encryption revisited: Improved security analysis and alternative solutions*" In Proc. of CRYPTO. 2011.
3) D. Cash, P. Grubbs, J. Perry, and T. Ristenpart. "*Leakage-abuse attacks against searchable encryption*" In Proc. of ACM CCS, 2015.
4) D. Cash, J. Jaeger, S. Jarecki, C. Jutla, H. Krawczyk, M. C. Rosu, and M. Steiner. "*Dynamic searchable encryption in very large databases: Data structures and implementation*" In Proc. of NDSS, 2014.
5) H. Cui, X. Yuan, and C. Wang. "*Harnessing encrypted data in cloud for secure and efficient mobile image sharing*" IEEE TMC, 16(5):13151329, 2017.
6) R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky. "*Searchable symmetric encryption improved definitions and efficient constructions*" In Proc. of ACM CCS, 2006.
7) Y. Elmehdwi, B. K. Samanthula, and W. Jiang. Secure k-nearest neighbor query over encrypted data in outsourced environments. In Proc. of IEEE ICDE, 2014.
8) R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky. Searchable symmetric encryption: improved definitions and efficient constructions. In Proc. of ACM CCS, 2006.
9) M. Kuzu, M. S. Islam, and M. Kantarcioglu. E_cient similarity search over encrypted data. In Proc. of IEEE ICDE, 2012.
10) K. Ren, C. Wang, and Q. Wang. Security challenges for the public cloud. IEEE Internet Computing, 16(1):6973, 2012. 54
11) D. Song, D. Wagner, and A. Perrig. Practical techniques for searches on encrypted data. In Proc. of IEEE SP, 2000.
12) W. Wang. Similarity join algorithms: An introduction In SEBD, volume 8, page 2, 2008.
13) B. Yao, F. Li, and X. Xiao. Secure nearest neighbor revisited. In Proc. of IEEE ICDE, 2013.
14) M. Yiu, I. Assent, C. Jensen, and P. Kalnis. Outsourced similarity search on metric data assets. IEEE TKDE, 24(2):338352, 2012.
15) Y. Zheng, H. Cui, C. Wang, and J. Zhou. Privacy-preserving image denoising from external cloud databases. IEEE TIFS, 12(6):12851298, 9 2017.