International Journal of Modern Computer Science (IJMCS)
Volume 6, Issue 3, June, 2018
ISSN: 2320-7868 (Online)

Secure Search on Encrypted Information through Secure Similarity Joins

Shital P Patil
PG Student,
Department of Computer Engineering
SKN COE SPPU
Pune, India
E-mail: shtlptl7593@gmail.com

Prof P. B. Mali
Asst. Professor,
Department of Computer Engineering
SKN COE SPPU
Pune, India
E-mail: pbmali@sinhgad.edu

Abstract: Cloud storage system is a storage system which includes collection of storage servers. When user store their data on cloud storage because of public access to data it may cause data security issues. To improve data security must be stored in secured format but the main problem will occurred in search operation on data. Existing work try to tackle this issue but it was not totally secured because data owner share secrete keys with data user it causes data leakage by unauthorised user. To resolve this problem proposed work is designed in which we can perform search operation on encrypted data without disclosing of information by employing Symmetric Searchable Encryption and Locality Sensitive Hashing.

Keywords: Cloud Computing, Encrypted Search, Security issues, Privacy Preserving, Similarity Search, Symmetric Encryption

I. INTRODUCTION

On cloud system users are allowed to store their data on third party storage to reduce storage cost and cost of maintenance. Data which is stored remotely requires security because of public access user must store date in secured format but main issue is searching on secured data. Existing work try to tackle this problem but it was not totally secured to improve security level proposed work is designed. Vendors of cloud service provide primary security by providing firewall protection and virtualization but which are not able to protect privacy of outsourced data. Data breaches may occur on outsourced data. To solve this problem Secure Similarity Joins approach is used. Secure Similarity Joins allows users to perform search operation on information which is in secured format.

The main purpose of Locality-Sensitive-Hashing and Symmetric Searchable Encryption is to provide strong protection to outsourced data. LSH hash values are used to create index for each file and searchable symmetric encryption (SSE) to provide secure similarity join on encrypted data. Fig 1 shows proposed system architecture. It consist of three parties i.e. the client of the authorized user, the data owner who owns source dataset, and cloud server in the public cloud. In this approach data owner upload file to application result of application will be encrypted file which generates hash value to create hash index for each row. After completion of dataset encryption and index generation data owner export resulted file and upload it on cloud server.

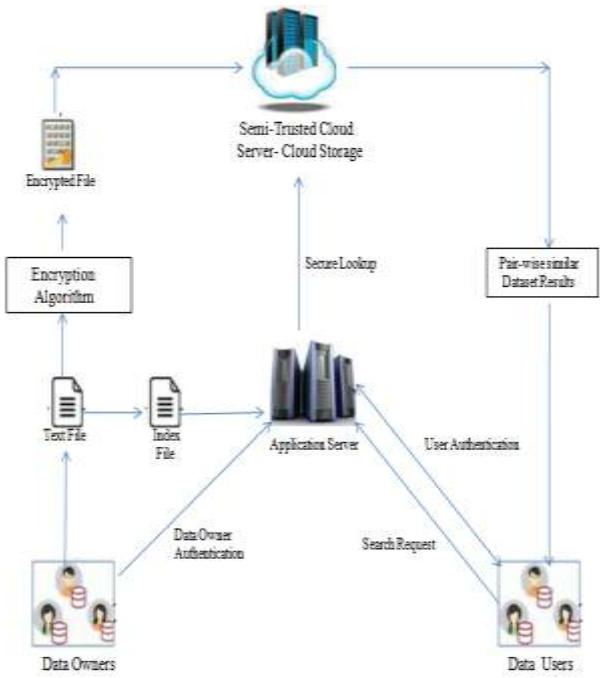

Figure 1: Proposed System Architecture

To access information authorised user generate query request and upload it to application software it return resulted file which contain only one column i.e. hash value column. Then user upload this file on cloud server for query search request cloud server will perform similarity search over encrypted datasets and return only matched records and put

 ISSN: 2320-7868 (Online)

those matched records into resulted file which will then send to user. After receiving resulted file user will upload to application software resulted file of application software will be in original format. In existing work this process was performed by sharing of secret key between user and the data owner and then between data user and cloud server because of this key sharing process provides low security level this proposed work reduces this security issue without sharing of any secret key.

II. RELATED WORK

In [1] paper provides a general similarity on message space implements an efficiently fuzzy searchable encryption (EFSE) also it distinguishes an ideal security for this technique. It enables a fuzzy search on encrypted data. In [2] paper provides mechanism to provide dynamic search on encrypted information using a Symmetric Searchable Encryption technique. In [3] paper represents an attack models with presence of servers prior history and utilizes. leakage profiles against searchable encryption. In [4] describes practical techniques like Pseudo random function, Sequential scan, Cryptographic scheme to perform search operation on encrypted information. In [5] paper provides techniques to improve security and provides different mechanism than existing mechanism for sharing images on mobile devices by creating encrypted sets of images by using secure similarity search, Yao's garbled circuits, and image denoising operations. In [6] paper provides guidelines to improve cloud security based on the users requirements and high lighten the security issues present in public cloud and mention different ways to remove these issues. In [7] paper In this paper provides an scheme to perform similarity search on encrypted data by utilizing a state-of-t heart algorithm for fast near neighbour search in high dimensional spaces called locality sensitive hashing it is an highly efficient scheme. In [8] paper provides an similarity search on data assets by using Order Preserving Encryption(OPE) in which order is preserve either alphabetically or numerically. This implementation also provides an interesting trade-offs between query cost and accuracy they are then extended to provide a guarantee about privacy. In [12] provides techniques to improve the security of processing images in image database by applying queries through external cloud data sources to perform secure denoising of images.

III. PROPOSED ALGORITHM

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.

IV. PSEUDO CODE

Build Encrypted dataset and encrypted index :

Input Dataset
Compute hash value of each hash function
Generate tokens
Maintain counter of hash value in hash table
Encrypt key-value pair of matched data points.

ISSN: 2320-7868 (Online)

Perform Secure Similarity join :
Input Dataset
Compute hash value of each hash function
Generate tokens
send token to server
matched set will decrypt
Output : pair-wise similar dataset.

V. PERFORMANCE RESULT AND ANALYSIS

To evaluate proposed technique here we have used detailed bank transaction set as a dataset. The proposed similarity join approach is designed in c#. Performance and results are shown in following graphs In fig. 2 shows security levels in existing system and in proposed system in existing system security level was so less because data owner send secret key with data user and data user will use that key to access information this security Final Stage When you submit your final version, after your paper has been accepted, prepare it in two-column format, including figures and tables. Send their final manuscript with filling form of copyright form and proof for manuscript charges submission. level is improved by creating system software which is responsible for encryption and decryption of dataset because of this process there is no need to share encryption and decryption keys by data owner and date user it preserves the security.

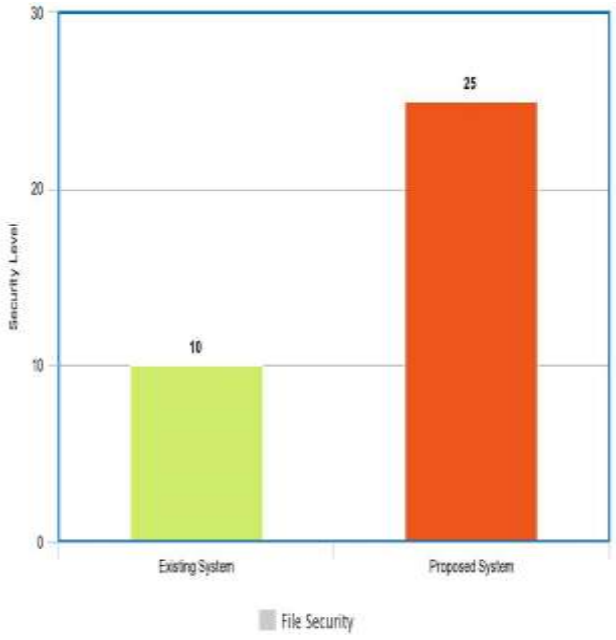

Figure 2. File Security Evaluation

Fig 3. Shows the query processing time how it is reduced than existing system because in proposed work we eliminate the process of token generation between data user and data owner existing system was increasing execution time because of token generation process.

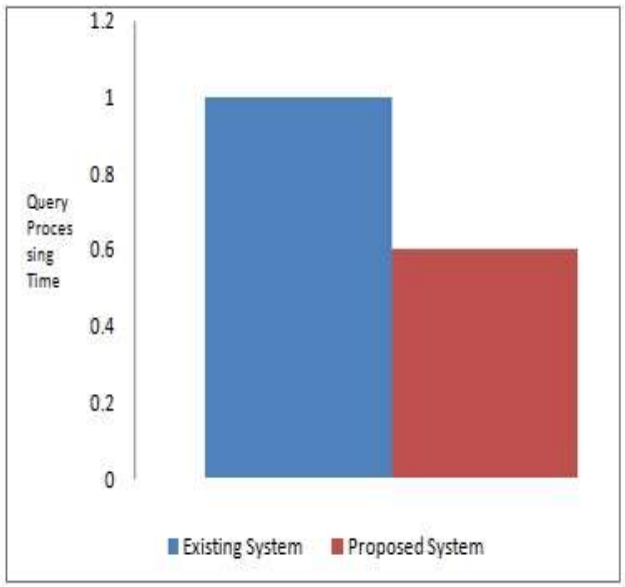

Figure 3. Query Processing Time

In fig.4 shows how proposed system is scalable than existing system existing system was not that much scalable on large dataset in proposed work we are transferring only required query set in encrypted format.

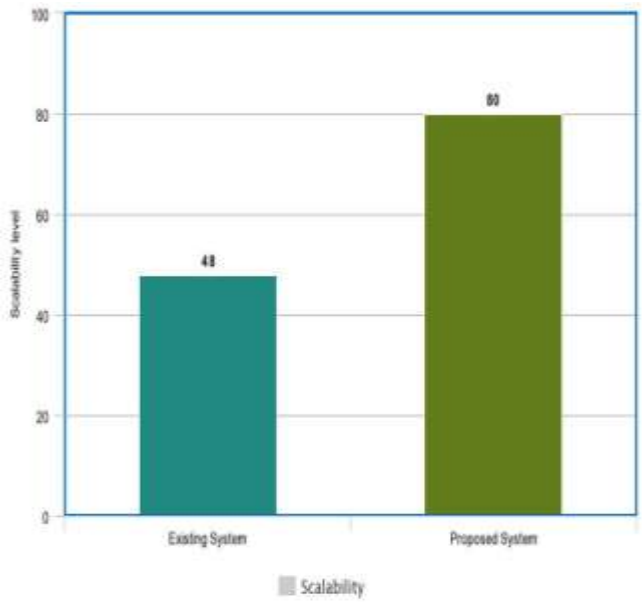

Figure 4. Scalability

 ISSN: 2320-7868 (Online)

VI. CONCLUSION AND FUTURE WORK

The Proposed system provides secure, scalable efficient searching technique on encrypted information. Secure Similarity Joins it is a pivotal primitive of similarity search on encrypted data it finds pair-wise similar data points across two datasets. In this work proposed a mechanism to perform search on encrypted information without token generation or sharing of secret key between data owner and data user which improves the security level.

ACKNOWLEDGMENT

The First and foremost, I would like to thank my guide, Prof. P.B.Mali, for his direction and support. I will perpetually remain thankful for the consistent help and guidance stretched out by my guide, in making this project. Through our nu- merous exchanges, he helped me to frame new thoughts. The important discourses I had with him, the entering questions he has put to me and the steady inspiration, has all prompted the advancement of the thoughts displayed in this task. Also, I would like to thank my parents for their continual encouragement and the positive support. I would also like to thank my friends for listening my ideas, asking questions and providing feedback and suggestions for improving my ideas.

REFERENCES

[1]. A. Boldyreva and N. Chenette "Efficient fuzzy search on encrypted data" In Proc. of FSE, 2014.

[2]. 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.

[3]. D. Cash, P. Grubbs, J. Perry, and T. Ristenpart. "Leakage-abuse attacksagainst searchable encryption" In Proc. of ACM CCS, 2015

[4]. D. Song, D. Wagner, and A. Perrig. "Practical techniques for searches on encrypted data".In Proc. of IEEE S&P, 2000.

[5]. H. Cui, X. Yuan, and C. Wang. "Harnessing encrypted data in cloud for secure and efficient image sharing from mobile devices". In Proc. Of IEEE INFOCOM, 2015.

[6]. K. Ren, C. Wang, and Q. Wang. "Security challenges for the public cloud. IEEE Internet Computing", 16(1):69–73, 2012.

[7]. M. Kuzu, M. S. Islam, and M. Kantarcioglu. "Efficient similarity search over encrypted data" In Proc. of IEEE ICDE, 2012.

[8]. M. Yiu, I. Assent, C. Jensen, and P. Kalnis."Outsourced similarity search on metric data assets". IEEE TKDE, 24(2):338–352, 2012.

R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky. "Searchable symmetric encryption improved definitions and efficient constructions". In Proc of ACM CSS, 2006.

[9]. S. Kamara, C. Papamanthou, and T. Roeder. "Dynamic searchable symmetric encryption". In Proc. of ACM CCS, 2012.

[10]. W. Wang. "Similarity join algorithms: An introduction" In SEBD,volume 8, page 2, 2008 Y. Zheng, H. Cui, C. Wang, and J. Zhou. "Privacy-preserving image denoising from external cloud databases". IEEE TIFS, 12(6):1285–1298, 92017.

[11]. Y. Zheng, H. Cui, C. Wang, and J. Zhou. Privacy reserving image denoising from external cloud databases. IEEE TIFS, 12(6):12851298, 9 2017.

AUTHOR'S BIOGRAPHIES

1. **Shital P Patil** is a ME student in Department of Computer Engineering SKN COE Pune. Her research interests include Cloud Computing.

2. **Prof. P. B. Mali** is working as an Assistant Professor in Depart of Computer Engineering SKN COE Pune. He has completed his M.Tech in Computer Engineering. He has published papers in more than 20 International Journals and Conference. His current research interest includes Data Mining. He has membership of LMIST.