ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

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

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

Privacy-Preserving Over Encrypted Data using Similarity Joins

Tambe Bhagyashri Haribhau, Prof.Sachin V.Todkari

Department of Computer Engineering, JSPM'S JSCOE, Hadapsar, Pune, India

ABSTRACT: To search the similarity join over encrypted data,we proposed a new system. privacy-preserving similarity join queries, i.e., a pivotal primitive of similarity search that finds pair-wise similar data points across two datasets. We start from locality-sensitive hashing (LSH) and searchable symmetric encryption (SSE), i.e., the most practical techniques for similarity search and encrypted search respectively. The data owner will build an encrypted LSH-based index I, encrypt the dataset S, and upload them to the cloud server.User search any keyword then Hash value match with index keyword then user got this encrypted file if user want to download this file then user can request to the data owner after receiving the request data owner send the key which is enter by the user by using secret key. After receiving the key user can download that file.If user can enter that wrong key for three times then ,user become attackers ,then user cannot download that file. Data user can this file at a particular time as well as particular location .We can also view mostly how many user can download a particular file.

KEYWORDS: Searchable Symmetric Encryption, Similarity joins, Datasets.

I. INTRODUCTION

In this system contain mainly three modules Data Owners,Data User And Cloud Server modules. Data owner can login the system,after login The data owner will build an encrypted LSH-based index I, encrypt the dataset S, and upload them to the cloud server.Data owners can generate the secure token.Data owner can response the different file download which is request by the different user.Data owner can view the how many user can download a particular files and also view the attackers. Data owner can registration first with proper authentication. With proper checking authentication data users can authentication, it can be login. After login data user can search the query. After searching a query, data users get the result which is user want to search a file. We can check here a 2 condition, Any data user can search any file from a particular location and particular time span only. If all validation are done then user can request for file download. Data owner can send the secret key for a download a file. If data user can enter this secret key correctly then user can download that file. But if user can enter this secret key more than 3 times wrong then that user become a attackers, then that user cannot download this file for particular month. Cloud server is module in our system. This cloud server system can be used for storing any information. When data owners can upload different dataset,it can store to the cloud server.On a cloud server we can store data users information also.So cloud server can view the how many user can download a particular files and also view the attackers.

II. OBJECTIVE AND SCOPE

1. To provide a high security on a similarity join data information. We proposed privacy-preserving similarity join queries
2.We proposed locality-sensitive hashing (LSH) and searchable symmetric encryption (SSE), i.e., the most practical techniques for similarity search and encrypted search respectively.
3. We show Top-k ranking keywords which is search by the user.
4. We can search the information in a particular location.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

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

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

III. PROBLEM STATEMENT

Similarity search is used for studied for data processing and analytics.For searching any information time required is more .While issue of data security and privacy preserving occure.this data security and privacy concerns along the trend of data outsourcing have not been fully addressed.Security issue is also occure.

IV. REVIEW OF LITERATURE

A. Boldyreva and N. Chenette [1] have introduced, used to solve problem of efficient (sub-linear) fuzzy search on encrypted outsourced data. Using fuzzy-searchable encryption (EFSE). Problem of efficient (sub-linear) fuzzy search on encrypted outsourced data, in the symmetric-key setting. Show how to construct schemes that are more efficient and satisfy a weaker security notion. time required more for large data for searching data.

M. Kuzu, M. S. Islam, and M. Kantarcioglu [2] have introduced, to solve an efficient scheme for similarity search over encrypted data**.** A state-of-the art algorithm for fast near neighbor search in high dimensional spaces called locality sensitive hashing.Provide a real world application of the proposed scheme and verify the theoretical results with empirical observations on a real dataset. Cryptographic techniques are inefficient due to their reliance on costly cryptographic operations Cryptographic techniques do not scale well for real world data sources.

R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky [3] have introduced, used for Searchable symmetric encryption: improved definitions and efficient constructions. We begin by reviewing existing notions of security and propose new and stronger security definitions .Interestingly, in addition to satisfying stronger security guarantees, our constructions are more efficient than all previous constructions .All PIR m schemes is that the data is always unencrypted, unlike the previous two settings on searching on *encrypted* data.

K. Ren, C. Wang, and Q. Wang.[4] have introduced, solve Security challenges for the public cloud. Security and privacy are perceived as primary obstacles to its wide adoption. Cloud computing provides literally unlimited computation powers while reducing costs, how to prevent malicious cloud users from abusing cloud resources is still an issue.Adopting stricter monitoring of cloud resource usage could be one way to mitigate this concern, but it's inevitably in conflict with legal users' privacy rights.

Y. Zheng, H. Cui, C. Wang, and J. Zhou.[5] have introduced to study Privacy-preserving image denoising from external cloud databases using external techniques. we initiate the first endeavor toward privacy-preserving image denoising from external cloud databases.Our design can achieve the denoising quality close to the optimal performance in plaintext. Time consuming and space consuming process.

D. Cash, P. Grubbs, J. Perry, and T. Ristenpart[6] have introduced study of leakage-abuse attacks and demonstrate their effectiveness for varying leakage profiles and levels of server knowledge. characterization of the leakage profiles of in-the-wild searchable encryption products and SE schemes in the literature, and present attack models based on an adver serial server's prior knowledge. The usage of SE with L1 level leakage over weaker variants.seek leakage levels even lower than L1.

D. Song, D. Wagner, and A. Perrig[7] ,have introduced to show how to support searching functionality without any loss of data confidentiality. This describe our cryptographic schemes for the problem of searching on encrypted data and provide proofs of security for the resulting crypto systems. Searching time is more in a encrypted data.

Man Lung Yiu, Ira Assent, Christian S. Jensen [8] ,have introduced to used for similarity search techniques for sensitive metric data. The data prior to supplying it to the service provider for similarity queries on the transformed data. Techniques provide interesting trade-offs between query cost and accuracy. Flexible Distance-based Hashing methods finishes in just a single round of communication, but does not guarantee retrieval of the exact result.

H. Cui, X. Yuan, and C. Wang[9] ,have introduced to study for the correlated encrypted image datasets to enable a secure and efficient cloud-assisted data sharing service for mobile devices with privacy assurance. To enable a secure and efficient cloud-assisted image sharing architecture for mobile devices, by leveraging outsourced encrypted image datasets with privacy assurance. design two specialized .Encryption mechanisms that support the secure image reproduction inside the cloud directly from the encrypted candidate selection. Preserving more the image content privacy.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

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

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

D. Cash, J. Jaeger, S. Jarecki, C. Jutla, H. Krawczyk, M.-C. Rosu, and M. Steiner.[10] have introduced to study dynamic searchable encryption in very large databases while using different structures and implementations. This is to extend the OXT protocol of Cash et al. to support arbitrary boolean queries. Implementation is difficult as compare to existing techniques.

V.PROPOSED SYSTEM APPROACH

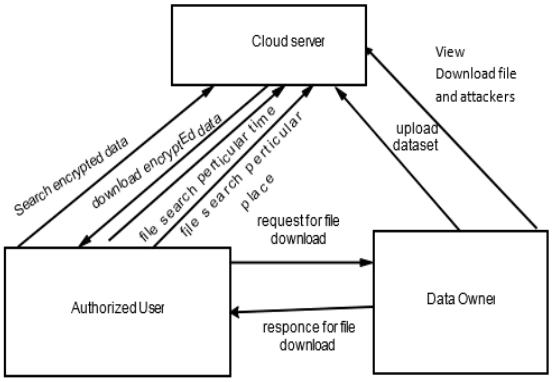


Fig.1 Block Diagram of Proposed System

In this system contain mainly three modules Data Owners,Data User And Cloud Server modules.

Data Owners:-

Data owner can login the system,after login The data owner will build an encrypted LSH-based index I, encrypt the dataset S, and upload them to the cloud server.Data owners can generate the secure token.Data owner can response the different file download which is request by the different user.Data owner can view the how many user can download a particular files and also view the attackers.

Data Users:-

Data owner can registration first with proper authentication. With proper checking authentication data users can authentication, it can be login. After login data user can search the query. After searching a query, data users get the result which is user want to search a file. We can check here a 2 condition, Any data user can search any file from a particular location and particular time span only. If all validation are done then user can request for file download. Data owner can send the secret key for a download a file. If data user can enter this secret key correctly then user can download that file. But if user can enter this secret key more than 3 times wrong then that user become a attackers, then that user cannot download this file for particular month.

Cloud Server:-

Cloud server is module in our system. This cloud server system can be used for storing any information. When data owners can upload different dataset, it can store to the cloud server. On a cloud server we can store data users information also. So cloud server can view the how many user can download a particular files and also view the attackers.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

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

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

VI.MATHAMATICAL MODEL

Notation:-
NU-Normal User
AU-Attackers User
TU-Total User
ESK-Enter Secret Key
EISK-Enter Incorrect Secret Key
Normal User
NU= total no. of user –user enter correct secret key
i.e. **NU**=TU-ESK
Attackers User
AU= total no. of user –user enter incorrect secret key.
i.e. **AU**=TU-EISK

VIII.COMPARATIVE RESULTS

In our experimental setup, In table 1,find out number of file upload and file download. In our experimental setup, in our system number file upload and download of files.

Sr.No	Number of File Upload	Number of File Download
1	35	15

Table1: No. Upload and download files

In our experimental setup, In table 2, find out number of file upload and file download. In our experimental setup, in our system number file upload and download of files.

Sr.No	No of Search Keyword 1	No of search Keyword 2	No of Search Name
1	15	16	18

Table1: No. file Search by keywords

IX.RESULTS

From above data, In graph 1,we can see the no. of file upload and no of file download in the graph; we see 35 files upload by different data owners and 15 different users are download in the graph.

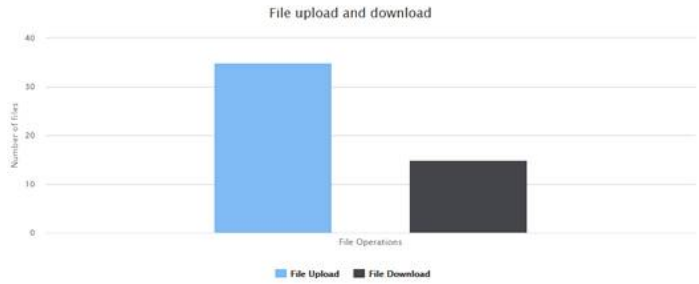
ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

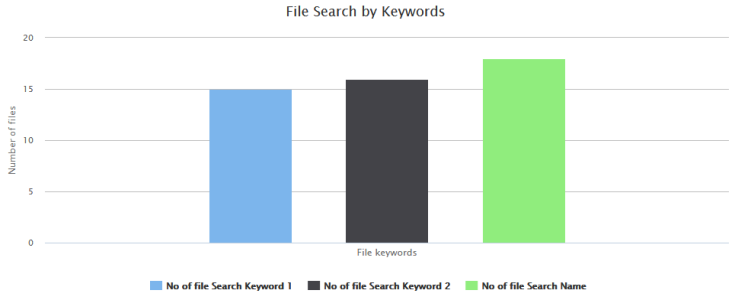
(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018



From above data, In graph 2,we can see the no. of file search keyword and file name also and no of file keyword 2 in the graph; we see 15 files search by keyword and 16 files search by keyword2 by different users and 18 files search by file name are shown in the graph..



X. CONCLUTION

We study the problem of privacy-preserving similarity join queries over encrypted high-dimensional data.We start from the state-of-the-art approaches that combine LSH and SSE to realize secure similarity search. In particular, three different secure query schemes are proposed to solve the problem regarding different practical requirements on security, efficiency, accuracy and deployability.In Our proposed system searching of any file is fast compare than old technique. In this system we can detect the attackers which is attempt wrong security key for file download time and user can download a file from particular time period as well as particular location.

FETURE SCOPE

In this system if user enter wrong secret key for more than 3 times then user become attacker. If this user are attack again and again then we can block that user for particular months or In feature we can block that user permanently.

ACKNOWLEDGMENT

This work is supported in a privacy-preserving similarity join queries over encrypted data of any state in india. Authors are thankful to Faculty of Engineering and Technology (FET), Savitribai Phule Pune University,Pune for providing the facility to carry out the research work.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

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

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

REFERENCES

1. M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski,G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, **"A view of cloud computing,"** Commun. ACM, vol. 53, no. 4, pp. 50–58, 2010.
2. D. Song, D. Wagner, and A. Perrig**, "Practical techniques for searches on encrypted data,"** in Proc. IEEE Int. Symp. Security Privacy, Nagoya, Japan, Jan. 2000, pp. 44–55.
3. R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, **"Searchable symmetric encryption: Improved definitions and efficient constructions,"** in Proc. 13th ACM Conf. Comput. Commun. Security, Oct. 2006, pp. 79–88.
4. C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, **"Secure ranked keyword search over encrypted cloud data,"** in Proc. IEEE Distrib. Comput. Syst., Genoa, Italy, Jun. 2010, pp. 253–262.
5. . Xu, W. Kang, R. Li, K. Yow, and C. Xu**, "Efficient multikeyword ranked query on encrypted data in the cloud,"** in Proc. IEEE 19th Int. Conf. Parallel Distrib. Syst., Singapore, Dec. 2012, pp. 244–251.
6. J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, **"Fuzzy keyword search over encrypted data in cloud computing,"** in Proc. IEEE INFOCOM, San Diego, CA, USA, Mar. 2010, pp. 1–5.
7. W. Zhang, Y. Lin, S. Xiao, Q. Liu, and T. Zhou, **"Secure distributed keyword search in multiple clouds,"** in Proc. IEEE/ACM 22nd Int. Conf. Quality Service, Hong Kong, May 2014, pp. 370–379.
8. Q. Liu, C. C. Tan, J. Wu, and G. Wang, **"Efficient information retrieval for ranked queries in cost-effective cloud environments,"** in Proc. IEEE INFOCOM, 2012, pp. 2581–2585.
9. W. Zhang, S. Xiao, Y. Lin, T. Zhou, and S. Zhou, **"Secure ranked multi-keyword search for multiple data owners in cloud computing,"** in Proc. 44th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw, Jun. 2014, pp. 276–286
10. Cash, J. Jaeger, S. Jarecki, C. Jutla, H. Krawczyk, M C. Rosu, and M. Steiner. **"Dynamic searchable encryption in very large databases: Data structures**
11. **implementation."** In Proc. of NDSS, 2014