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. 5, Issue 12, December 2017

Privacy-Preserving over Encrypted Data using Similarity Joins

Tambe Bhagyashri, Prof. Todkari S.V

Dept. of Computer Engineering, Jayawantrao Sawant College of Engineering, 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. LITERATURE REVIEW

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 neighbour 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 endeavour 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 advert 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.

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.

III. EXISTING SYSTEM APPROACH

In old techniques we can Secure search over encrypted data has recently attracted the interest of many users. The problem of secure search over encrypted data. They propose the conception of searchable encryption, which is a cryptographic primitive that enables users to perform a keyword-based search on an encrypted dataset, just as on a plaintext dataset. Searchable encryption is further developed. This user can decrease the storage cost for secure keyword search over encrypted cloud data, and also enrich the category of search function, including secure ranked multi-keyword search, fuzzy keyword search, and similarity search However, all these schemes are limited to the single-owner model. As well as In existing system architecture, When different data owner can upload this any file encrypted index is generated this encrypted index goes to administrator system. Different data owners can upload files on a cloud so for every file generate encrypted indexes. Data Administrator can re-encrypted index then store on a cloud server. When user can search any file then after checking authentication user get file. If user want to download that file then data user request to data owner. After getting the request user can send the key for download the file. In this system we can show rank search result.

DOI: 10.15680/IJIRCCE.2017. 0512029

IV. PROPOSED SYSTEM APPROACH

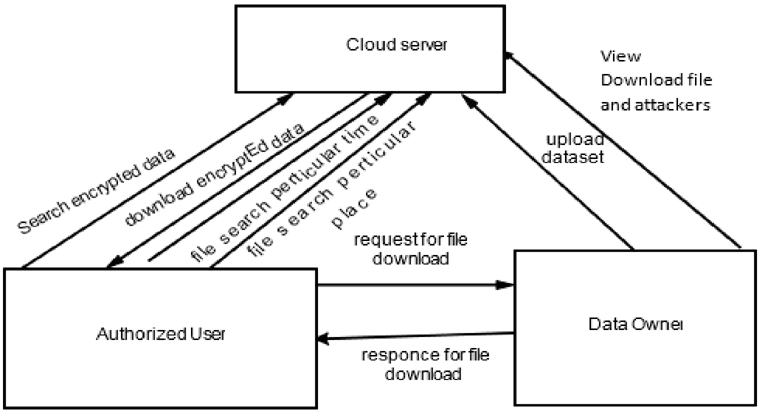

Fig.1 Block Diagram of Proposed System

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.

V. CONCLUSION

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.

ACKNOWLEDGMENT

This work is supported in a traffic management system of any state in india. Authors are thankful to Faculty of Engineering and
Technology (FET), SavitribaiPhule Pune University,Pune for providing the facility to carry out the research work.

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 cloudcomputing,"** in Proc. 44th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw, Jun. 2014, pp. 276–286

BIOGRAPHY

First Author – Tambe Bhagyashri, BE Computer, Jayawantrao Sawant College of Engineering, Hadapsar, Pune

Second Author – Prof.S.V.Todkari Jayawantrao Sawant College of Engineering, Hadapsar, Pune