International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 09 | September - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

EFFICIENT AND EXPRESSIVE SEARCH OVER ENCRYPTED DATA WITH MULTI FACTOR AUTHENTICATION SCHEME

Wasiba Begam M[1]

[1]*Assistant Professor, Dept. of Computer Science, Thassim Beevi Abdul Kader College forWomen, Kilakarai, Tamil Nadu, India.*

ABSTRACT

More data owners are tempted to outsource their data to cloud services as a result of the development of cloud storage. Sensitive data should be encrypted before being outsourced due to privacy concerns. To assure data availability, a variety of searchable encryption techniques are used. However, the effectiveness of data consumers' queries is not given much consideration by the current search algorithms, particularly in the case of several owners. In this study, linguistic coefficients are used to create an encrypted search index. In particular, the encrypted search index utilizing Lagrange co-efficient makes use of secure inner-product calculation for both search and relevance assessment by taking into account a sizable amount of data in the cloud. A system architecture of preferred search over encrypted data in the cloud scene is built, and the requirements in terms of efficiency and privacy are specified. This enables the clo ud servers to do a secure search without knowing any sensitive data (e.g., keywords and trapdoors).For each data owner, the Lagrange coefficient is used to build indices that can support searches over numerous keyword fields and enable accu rate relevance computation in order to produce an efficient search.To protect user data privacy, a multifactor authenticati on technique might be utilised.After logging in to the system, the user will receive three different sorts of authentications, including an OTP utilizing a number, a captcha, and an OTP image. This picture will serve as a key. In order to accomplish secure search and relevance score calculation, the user's preference and the search query are expressed in vector form. Additionally, Preferred Search over Encrypted Data (PSED) is consistent with cloud computing's scalability and adaptability. Finally, a thorough security study verifies the security of our scheme, and a performance analysis highlights its effectiveness and efficiency.

Keywords: Encrypted Data, Inner product calculation, Multifactor authentication, Data privacy

I. INTRODUCTION

In cloud storage the data is stored in logical pools as digital data. In multi-owner scenario, the same data will contain several owners. A main server will be there to handle the entire data. The cloud may contain multiple servers may be reside in multiple locations. The main server or the cloud storage providers will be responsible for protection and handling of the stored data. The cloud users will buy or lease the

storage capacity from these cloud storage providers. Cloud storage enables distributed and scalable network access to the digital data. A problem that has to be faced in cloud storage is the secured search over the encrypted data.

The most challenging task in cloud storage is secured search on encrypted cloud data. There are various search schemes exist. But they results either in system overhead or sometimes those methods will be really hard to implement over large data sets. To prevent the unauthenticated access the data will be stored in cloud as in the encrypted form.

To provide an efficient search, a scheme for preferred search over encrypted data (PSED) is proposed that can take users' search preferences into the search over encrypted data. In the search process, it ensures the confidentiality of not only keywords but also quantified preferences associated with them. PSED constructs its encrypted search index using Lagrange coefficients and employs secure inner-product calculation for both search and relevance measurement. The dynamic and scalable property of cloud computing is also considered in PSED. A series of experiments have been conducted to demonstrate the efficiency of the proposed scheme when deploying it in real world scenarios.

To characterize both the frequency and uniqueness of each keyword in an index, PSED first assigns each keyword with Eight evaluated by the term frequency (TF) ×inverse document frequency (IDF) model. PSED then expresses the search query and the user's preference in vector form and employs secure inner-product calculation to perform search and relevance score calculation without leaking the index information (including keywords and keyword Eights) or the query information (including keywords and their preferences). Our contributions can be summarized as follows.

1) A system framework of PS over encrypted data in the cloud scene and specify the requirements in terms of efficiency and privacy.

2) Lagrange coefficients to construct indices is used that can support search over multiple keyword fields and enable correct relevance calculation. It expresses the query and user's preference in vector form and adopt secure inner-product calculation to securely perform search and relevance score calculation. Moreover, PSED is also compatible with the scalable and dynamic property of cloud computing.

3) A thorough analysis is conducted on efficiency and privacy protection provided by PSED, and carried out extensive experiments with a real-world dataset to demonstrate the applicability of PSED in real-world scenarios.

METHODOLOGY OF THE RESEARCH

Cloud computing has been considered as a new model of enterprise IT infrastructure, which can organize huge resource of computing, storage and applications, and enable users to enjoy ubiquitous, convenient and on demand network access to a shared pool of configurable computing resources with

great efficiency and minimal economic overhead. Attracted by these appealing features, both individuals and enterprises are motivated to outsource their data to the cloud, instead of purchasing software and hardware to manage the data themselves.

Despite of the various advantages of cloud services, outsourcing sensitive information (such as e-mails, personal health records, company finance data, government documents, etc.) To remote servers brings privacy concerns. The cloud service providers (CSPs) that keep the data for users may access users' sensitive information without authorization. A general approach to protect the data confidentiality is to encrypt the data before outsourcing. However, this will cause a huge cost in terms of data usability.

Today data security plays an important role in the field of software and quality of service, Cloud computing focuses a new challenging security threats. Therefore, a data security model must solve the most challenges of cloud computing security. Cloud computing is a capable technology to make easy development of large-scale, on-demand, flexible computing infrastructures. The concept of cloud computing has changed the view of infrastructure architectures, software delivery and development models. Cloud computing incorporates elements from grid computing, utility computing and autonomic computing to an innovative deployment architecture. This paper presents an overview and the study of the cloud computing. Also include the several security and challenging issues, emerging application and the future trends of cloud computing. To overcome this problem, the concept of Email OTP is used for providing authentication, whose purpose is to ensure security in cloud environment. Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. The concept of Cloud Computing revolves around distributed architecture that centralizes server resources on a scalable platform so as to provide on demand computing resources and services. Cloud platforms are offered by the Cloud service providers (CSP's) for their customers to use and create their web services, much like internet service providers offer costumers high speed broadband to access the internet.

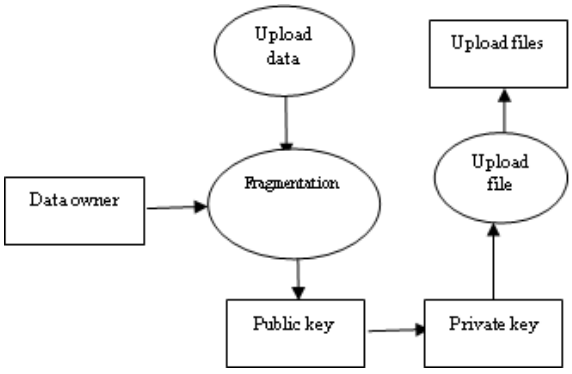

Fig. No. 1 – System Flow diagram

II. EXISTING SYSTEM

A tree-based ranked multi-keyword search scheme is existing technique, which supports not only the accurate multi-keyword ranked search but also the dynamic deletion and insertion of documents and construct a special keyword balanced binary tree as the index, and existing technique was a "Greedy Depth-first Search" algorithm to obtain better efficiency than linear search. In addition, the parallel search process can be carried out to further reduce the time cost. The security of the scheme is protected against two threat models by using the secure top k-retrieval algorithm. Experimental results demonstrate the efficiency of our existing technique. There are still many challenge problems in symmetric SE schemes. In the existing scheme, the data owner is responsible for generating updating information and sending them to the cloud server. Such an active data owner may not be very suitable for the cloud computing model. It could be a meaningful but difficult**.**

III. PROPOSED SYSTEM

Proposed scheme to design a the encrypted search index using Lagrange coefficients scheme whose updating operation can be completed by cloud server only, meanwhile reserving the ability to support multi-keyword ranked search. In addition, as the most of works about searchable encryption, our scheme mainly considers the challenge from the cloud server. I can revoke the user in this scheme. If it is needed to revoke a user in this scheme, need to rebuild the index and distribute the new secure keys to all the authorized users. Secondly, symmetric SE schemes usually assume that all the data users are trustworthy. It is not practical and a dishonest data user will lead to many secure problems. For example, a dishonest data user may search the documents and distribute the decrypted documents to the unauthorized ones. Even more, a dishonest data user may distribute his/her secure keys to the unauthorized ones. Proposed scheme to improve the SE scheme to handle these challenge problems.

IV. PROPOSED ALOGORITHM

Algorithm & Protocol Used

There are three types of algorithm used in this system.

Encryption:

This is used to encrypt the data files. This converts the plain text into the cipher text. This uses the AES (Advanced Encryption Standard) algorithm.

Decryption:

This is used to decrypt the data files. This converts the cipher text into the plain text. This uses the ADS (Advanced Decryption Standard) algorithm.

File search

Encrypted Search Index Using Lagrange Coefficients is used to search files

SMTP:

Simple Mail Transfer Protocol is used to send mail to the users.

V. IMPLEMENTATION

Authentication and Authorization

In this module, the data owner and the data user register with their user id, Password, email-Id, mobile number and gender then user can access the database. After registration completed, data owner and the user access the database by giving the user-id and their password.

Cloud Storage

Admin is responsible for cloud storage. In this module, admin can view the details of all the data owner and registered users. Admin can see the status of the shared files among multiple users.

Data Owner

Data owner is responsible for uploading file on cloud as well as view files uploaded by him or others. Data owner has information about the placed fragment and its replicas with their node numbers in cloud. Owner can best split the file such that no meaningful data is available. Method proposed is not to store the entire file at a single node. Compromised data on attacks can be reduced by storing fragments on separate nodes. A successful intrusion only provides access to a portion of data. The probability of finding fragments on all of the nodes is very low. The selection of nodes is made by keeping an equal focus on both security and performance. Choose the nodes that are most central to provide better access time. Uses the concept of centrality to reduce access time. If intruder compromises a node, then location of other fragments cannot be determined. In cloud systems with more nodes, probability for an attacker to hack data reduces.

Data User

Data user is the one who is responsible for downloading files or view files uploaded by others. To download file from cloud he has to be authenticated user otherwise he will be considered as attacker. when the user shares a particular data file, the private key is sent to all the shared users. When a user wants to decrypt the data file then he/she request the decryption key. Then the server sends the decryption key to the authenticated user to their registered mail Id. By using this decryption key, the user decrypts the data files and then he/she can download the file. The user can decrypt the data by entering the decryption key. The fragmented files will be merged during this process

Uploading and Downloading

In this module, Files are uploaded to the server after file is encrypted by the encryption method. This encryption is done by AES (Advanced Encryption Standard) Algorithm and generate key. This Encrypted Data is in the form of Binary and stored in Cloud. User needs decryption key to download the data files.

File Sharing

In this module, the uploaded files are shared to the multiple users. In this system, the private key of the Data which is shared will be

send via a secure channel called Gmail. This decryption key is used by the user at the time of download the files.

Key Generation

In this module, when the user wants to access the data files then the server send the decryption key. Through this decryption key, the user who wants to access the data file, uses this decryption key to decrypt the files with the help of private key sent at the time of file sharing.

VI. CONCLUSION

The problem of preferred search over encrypted cloud data is investigated. First a set of designed goals are established and used the TF×IDF model for keyword weight measurement. The user's query and preference and keywords and their weights in vector form are expressed. The secure inner product computation was then employed to perform search and measure the relevance between files and the user's preference. Thorough analysis concerning privacy and efficiency was presented, and the intensive evaluation on a modern server demonstrated its suitability.

VII. FUTURE WORK:

Future work will consider the problem of secure fuzzy keyword search in a multi-owner paradigm. On the other hand, plan to implement extend our work on images, video in commercial clouds. This study addresses these issues by proposing Visual Cryptography Scheme (VCS) technique for securing the files. Then the files are encrypted using Advanced Encryption Standard (AES). Then the encrypted files are securely sent to the cloud. This research work can be extended to implement image storage and retrieval.

VIII. RESULTS AND DISCUSSION

A onetime password (OTP) is just what the names implies, a password that is only valid for one login. The benefit of OTPs is that it offers much higher security than static passwords, in expense of user friendliness and configuration issues. OTPs is immune against password sniffing attacks, if an attacker use software to collect your data traffic, video records you when you type on your keyboard, or use social engineering, it doesn't matter since the password that the attacker gets hold on will not be valid to use. On your keyboard, or use social engineering, it doesn't matter since the password that the attacker gets hold on will not be valid to use. A OTP can be generated using different methods.

OTP using numbers

A one-time password (OTP) is an automatically generated 4 digit number that authenticates the user for a single transaction or session. A onetime password as the word indicates is only valid for a specific time interval or one-time usage. If the user credentials are valid, a 4 digit one time password is sent to your registered email and you are required to enter it when prompted. It is the Second Factor Authentication. If the session of OTP number

expires, the user is able to receive a new OTP number automatically

Disadvantages of number OTP

The Random OTP password generation technique is easily predictable for hacker and man in middle attack.

OTP using graphical image

After entering the OTP number, the user is asked to load the OTP images which he/she received on the email. I used similarity measure algorithm for image matching. An important problem in image processing is the comparison of images. The Verification of user's Image or Photo process is shown in the following figure.

Fig. No. 2 Scalability chart for result analysis

IX. REFERENCES

[1] Bing Wang, Wei Song, Wenjing Lou, Y. Thomas Hou, "Inverted Index Based Multi-Keyword Public-key Searchable Encryption with Strong Privacy Guarantee", 2015 IEEE Conference on Computer Communications (INFOCOM).

[2] Hongwei Li, Yi Yang, Tom H. Luan, Xiaohui Liang, Liang Zhou, Xuemin (Sherman) Shen, "Enabling Fine Grained Multi-Keyword Search Supporting Classified Sub-Dictionaries over Encrypted Cloud Data", IEEE Transactions on Dependable and Secure Computing, Vol. 13, No. 3, May/June 2016.

[3]Muhammad Saqib Niaz we et al, "proposed a forward secure searchable symmetric encryption**"**, IJRT vol 5 2018

[4] Ning Cao, Cong Wang, Ming Li, Kui Ren, Wenjing Lou, "Privacy Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data", IEEE Transactions on Parallel and Distributed Systems, Vol. 25, No. 1, January 2014.

[5] Shulan Wang, Junwei Zhou, Joseph K. Liu, Jianping Yu, Jianyong Chen "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data", IEEE Transactions on Information Forensics and Security, Vol. 11, No. 6, June 2016.

[6] Shulan Wang et al,[6] "entitled an efficient file hierarchy attribute-based encryption scheme in cloud computing", IEEE transactions on information forensics and security, vol. 11, no. 6, June 2016

[7] W. K. Wong, D. W. Cheung, B. Kao, and N. Mamalis, "Secure ken computation on encrypted databases", in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2009, pp. 139-152

[8] Wenhai Sun, Bing Wang, Ning Cao, Ming Li, Wenjing Lou, Y. Thomas Hou, Hui Li, "Verifiable Privacy Preserving Multi-Keyword Text Search in the Cloud Supporting Similarity-Based Ranking", IEEE Transactions

on Parallel and Distributed Systems, Vol. 25, No. 11, November 2014

[9] Wei Zhang, Yaping Lin, Sheng Xiao, Jie Wu, Siwang Zhou, "Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing", IEEE Transactions on Computers, Vol. 65, No. 5, May 2016.

[10] Zhihua Xia, Xinhui Wang, Xingming Sun, Qian Wang, "A Secure and Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data", IEEE Transactions on Parallel and Distributed Systems, Vol. 27, No. 2, February 2016.

[11] Zhangjie Fu, Xingming Sun, Zhihua Xia, Lu Zhou, Jiangang Shu, "Multikeyword Ranked Search Supporting Synonym Query over Encrypted Data in Cloud Computing", 2013

[12] Wenhai Sun, Xuefeng Liu, Wenjing Lou, Y. Thomas Hou, Hui Li, Catch You If You Lie to Me: "Efficient Verifiable Conjunctive Keyword Search over Large Dynamic Encrypted Cloud Data", 2015 IEEE Conference on Computer Communications (INFOCOM).

[13] Wei Zhang, Yaping Lin, Catch You if You Misbehave "Ranked Keyword Search Results Verification in Cloud Computing", VOL. 6, NO. 1, JANUARY 2015.

[14] C. Wang, N. Cao, J. Li, K. Ren, W. Lou, "Secure ranked keyword search over encrypted cloud data", in Proc. IEEE Distributed Computer System, Genoa, Italy, Jun. 2010, pp. 253262.

[15] Zhangjie Fu, Xingming Sun, Nigel Linge, Lu Zhou, "Achieving Effective Cloud Search Services: Multi keyword Ranked Search over Encrypted Cloud Data Supporting Synonym Query", IEEE Transactions on Consumer Electronics, Vol. 60, No. 1, February 2014.

[16] Zhangjie Fu, Jiangang Shu, Xingming Sun, Nigel Linge, "Smart Cloud Search Services: Verifiable Keyword based Semantic Search over Encrypted Cloud Data", IEEE Transactions on Consumer Electronics, Vol. 60, No. 4, November 2014.