ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 4, April 2017

Enhanced Multi-Keyword Ranked Search Manipulations via Cloud Virtualization

[1] S.Uma Maheswari, [2] p.priya ponnusamy, [3] dr.r.vidhyapriya
[1][2] sri shakthi institute of engineering and technology,Coimbatore
[3] psg college of technology,coimbatore

Abstract— **The enhancement and the fame of cloud computing are increasing nowadays which is driving the data owners to keep their personal and professional data on public cloud servers With the help of data outsourcing the organizations can provide reliable and efficient services to their users. The other advantage of outsourcing the data over cloud servers is for high benefit and lesser cost in managing the data and the data can be accessed from anywhere and at any time. However, for privacy concerns, the data that are highly sensitive should be encrypted before outsourcing. Taking into consideration the huge amount of data users and files that are present in the cloud, it is important that multiple keywords should be allowed in the searching request and retrieve the files relevant to those keywords. There are some methods and solutions offered to provide privacy and security for the data over the cloud server. However, the data that are sensitive should be encrypted and stored before outsourcing. In this paper, a secure scheme called the enhanced multi-keyword ranked search is presented. In the proposed scheme the search takes place as content based search along with ranking. Based on the content, data is searched and retrieved and then the rank is applied. During searching, attacks might be done by SQL injection to dump the database contents by the unauthorized users or attackers. In this project SQL injection is not allowed. The SQL parameterization technique is used to prove that SQL injection is not possible.**

Index Terms — **multi-keywords ranked searching; efficient retrieval; cloud server; data outsourcing**

.

I. INTRODUCTION

Cloud computing is described as a service model whichallows appropriate, on-demand network access to a huge shared pool of computing resources which can be quickly purveyed and liberated with minimum management exertion or service supplier interaction. This advanced information system architecture, which is basically altering the way in which the computing, storage and networking resources are assigned and controlled, delivers a large number of advantages to users, including but not confined to reduction in capital costs, easily accessing the information, better flexibility, automatic service desegregation, and fast deployment. In spite of these advantages, as a rising technology, cloud computing also faces a large number of security and privacy challenges which impede its quick adoption. Security has been accepted as the leading barrier for users to move to cloud computing.

Multi-Keywords Ranked Search

A data owner can outsource their data such as personal or Professional data to the cloud servers and he can either query on the outsourced data or attest client to execute query. The multi-keyword based search allows the data users for searching over the encrypted data with the aid of many keywords. There are three entities which form the layout for multi-keyword searching over the encrypted cloud data which Includes the data owner, the cloud server and the data users. The data owners outsource their sensitive data in the cloud servers after the data is encrypted. The data users can retrieve those data using multi-keyword ranked search techniques. A data owner can outsource their data such as personal or professional data to the cloud servers and he can either query on the data which is outsourced or attest client to perform query.

II. LITERATURE SURVEY AND RELATED WORK

M. Kuzu in [2] proposed that the secure index that is LSH based and a searching strategy for activating quick similarity searching over encrypted data. A strict security definition is provided and the security of the strategy to check confidentiality of the sensitive data is proved.

David Cash in [6] has proposed that the searchable encryption strategies that are dynamically symmetric are effectively searched server-held ciphered databases with billions of pairs of keywords are designed and implemented. Moreover, the implementation experience shows that even complicated queries, that go well apart from the single-keyword search capacity of SSE strategies, can be backed up in practical ways for these huge databases. The same is true for the complicate functional settings that support deputation and empowerment of queries to multiple clients as well as provisioning query privacy from the data owner. Regarding the security, it is necessary to spotlight that while the outflow of the constructions equates

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 4, April 2017

well with related work, there is informal information being exposed to the server about result sets. Moreover, the effectuation that uses the dynamic SSE strategies created here as the foundation for encouraging new SSE advances, including complicate search queries (e.g., Boolean queries) and richer functional settings (e.g., query deputation), in the terabyte-scale databases is reported.

Ning Cao in [5] has proposed the solution of the difficulty of privacy-preserving multi-keyword ranked searching over ciphered data in the cloud storage. A set of security demands for such a safe cloud data employment system is constituted. The similarity mensuration of unified matching to seizure the relevant data records to the search query is chosen. Then, the similarity of the inner product for assessing such similarity mensuration is used. An initial idea for the MRSE based on the calculation of the inner product, and giving two enhanced MRSE strategies for attaining different security demands in two threat models is proposed in this paper.

Philippe Golle has proposed in [8] that two protocols for connective search for which it is demonstrably difficult for the server to differentiate between the encrypted keywords of files of choosing by its own is presented. These protocols permits secure connective search with little capacity. The work only partly resolves the difficulty of secure Boolean search on ciphered data. Especially, an accomplished solution requires the skill to do disjunctive keyword search safely, in both across and within the keyword fields. A major problem that isn't found by the privacy games is the data seeped out by the capabilities. In both of the protocols given, the server studies the keyword fields that the capability activates the server to search. This alone may be sufficient to let the server to deduce unplanned information about the files.

N. Cao, in [4] has proposed in this paper, an effective strategy for search that is with similarity over encrypted data. For this, the algorithm called state-oSf-the art for quick closer neighbor searching in more dimensional spaces known as the locality sensitive hashing is utilized. For checking the confidentiality of the sensitive data, a strict privacy description and proving the security of the given strategy under the rendered definition is rendered.

B.Wang in [10] has proposed in this paper, the strategy to carry on with safe ranked multi-keyword searching in a multi-owner framework. For allowing cloud servers to execute secure search without cognizing the real data of both keywords and trapdoors, a new secure search protocol is relatively constructed. For ranking the results of the searching and upholding the privacy of relevant scores between keywords and files, a new Additive Order and Privacy Preserving Function family is proposed.

Zhihua Xia in [1] proposed that, a secure multi-keywords ranked searching strategy over encrypted cloud data, that encourages active update functioning like removing and adding of files is presented. The vector space model and the TF_IDF model are joined in the construction of the index and generation of the query. A peculiar index which is tree based tree-based construct and suggest an algorithm called "Greedy Depth-first Search" for providing an effective multi-keywords ranked searching is constructed. The algorithm for secure kNN is used for the encryption of the index and query vectors, and check accurate relevant score computation between encrypted indices and vectors of the query. For withstanding attacks that are statistical, phantom terms are included to the indices vectors to blind the results of the search. Because of the usage of the structure of the specialized tree-based indices, the given strategy can attain time of sub-linear search and carry on with the removal and addition of files flexibly.

Wenhai Sun et al in [3] proposed that a privacy-preserving multi-keyword text searching strategy with the ranking which is based on the similarity is presented. For encouraging multi-keyword searching and searched result ranking, for constructing the searching index based upon frequency of term and the model of vector space with similarity of cosine mensuration to attain higher search result accuracy is proposed. For advancing the efficiency of the search, a structure that is tree-based index and different adaptation method for algorithm of multi-dimensional for the efficiency of the search to be better than that of linear search is proposed. For the improvement of the searching privacy, two secure indices strategies for meeting the security demands under powerful models of threat, that is the cognize cipher text model and known background model are proposed.

Bing Wang in [9] proposed a new multi-keywords fuzzy searching strategy by victimizing the locality-sensitive hashing. It gains fuzzy matching with the help of algorithmic design than spreading out the index file. It extinguishes the want of pre- established dictionary and encourages multiple keywords fuzzy searching without enlarging the index or the searching complexness.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 4, April 2017

Cong Wang et al in [7], proposed a solution for the difficulty in secured ranked keywords searching over encrypted data. The ranked searches increases system's serviceability by activating result of the search relevant ranking instead of directing undifferentiating results, and then checks the accuracy of the retrieval of file. The statistical measure ideas, that is the relevant score, from retrieval of data to construct a secured searching indices, and building the technique of one-to-many order-preserving mapping for saving those delicate scored data. The out coming design will capable to provide effective server-sided ranking without the loss of privacy of the keyword.

III. PROPOSED SYSYEM

The earlier mechanisms support multi- keyword searches, for retrieving the data from the cloud storage. In the proposed scheme an efficient data retrieval algorithm is proposed to search the data encrypted and stored in the cloud server. The searching takes place as content based search along with ranking. Based on the content, data is searched and retrieved and then the rank is applied.
When the data searching takes place, attacks might take place by SQL injection which will dump the database contents by the unauthorized users or attackers. In this system SQL injection is not allowed. The SQL parameterization technique is used to prove that SQL injection is not possible.

User Authorization and Authentication

Proxy defines the security by means of user authorization and authentication. Proxy signature is a signature strategy, where the signer who is original can depute his/her signing capacity to another signer of proxy, and after that on the behalf of the original signer, the proxy signer produces a signature. From the proxy signature, a verifier can be persuaded by the original signer's agreement on the message which is signed.

The user authorization and authentication is performed. The data which are outsourced should not have any unauthorized access hence the authorization and authentication helps to have an authorized access of the data.

Trusted Data maintenance

This trusted data maintenance module describes the unencrypted dynamic multi-keyword ranked search (UDMRS) strategy that is built on the initial idea of model of vector space and KBB tree. On the basis of the UDMRS strategy, two secure searching strategies are built against two threat models.

Data security model

In this module, the cloud provider is considered as semi trusted, it intends that cloud servers would come after the proposed scheme, but try to come out as much unrevealed data as possible which is based upon inputs of each group member. Generally, it is assumed that cloud servers are concerned with the contents of the data and private information of the group member instead of other sensitive data. Cloud servers may conspire with few hostile members for the cause of getting subject of the data and group members' private information. The strategy must meet the privacy demands of secrecy that are forward and backward. The first checks that the reneged user is not able to decrypt new cipher texts. The second checks that the recently combined user can also make use of and decrypt the antecedently issued data.

THE SQL INJECTION: SQL Injection (SQLi) refers to an injection attack wherein an attacker can execute malicious SQL statements (also commonly referred to as a malicious payload) that control a web application's database server. SQL injection is a code injection technique, used to attack data-driven applications, in which nefarious SQL statements are inserted into an entry field for execution (e.g. to dump the database contents to the attacker).

```
S = "Select  * From Reg Where Mail='" + TxtUN.Text +
"' and Pwd='" + TxtPwd.Text + "' And UType='Data
Owner'";
```

This SQL query is inserted into the code of the data owner for
SQL
injection.

THE SQL PARAMETERIZATION: By SQL parameterization SQL injection will not be permitted. (i.e) the unauthorized access by an attacker to the database will be completely eliminated.

```
S = "Select * From Reg Where Mail=@Ml and
Pwd=@Pwd
```

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 4, April 2017

And UType='Data Owner'";

This SQL query is inserted into the code of the data owner for SQL parameterization.

Content basedsearch

The Content Based Data Search module eliminates the problem of unwanted confusions and problems over data mining scenario, which is achieved by means of structured data maintenance module. The structured data storage scheme fully describes about the flow of data structure maintenance and the concept of implicit data mining and document maintenance schema. Once the resource owner uploads the respective document it checks for the reference schema of the existing document for reference, if the document is presented in the database server then the following document is sequenced under the existing document otherwise it creates a new schema for the following document, so that the data into the database server is maintained in the structured manner. All the data into the server is based on the cluster format and provides the frequent access for the user search between the server and the data client (user).

Page Ranking Scenario

In the page ranking scenario module the ranking methodology is used by Search engines to rank the results in their search results. The pages will be ranked based on the results that are obtained during the search based on the demand of the data user from the cloud server.

IV. RESULTS

The data owner and data user login/signup: The data owner and the data user should signup/login to have an authorized access to the file that is uploaded.

The data owner authentication

The data owner authentication takes place here. The data owner can login as an authorized user.

The uploading of the file:

The file which has to be saved in the cloud server can be uploaded here

The file has been secured sucessfully

The file that has been uploaded is encrypted and stored in the cloud server.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 4, April 2017

The search for data:

The file which has to be retrieved is searched by using the keyword.
The content based search:

The dataor the file has been retrieved based on the content based search.

TheSqlInjectionhastakenplac
The unauthorized users can access the by using the following code

S = "Select * From Reg Where Mail='" + TxtUN.Text + "' and Pwd='" + TxtPwd.Text + "' And UType='Data Owner'";

This SQL query is inserted into the code of the data owner and the SQL injection takes place.

The Sql parameterization has taken place:

S = "Select * From Reg Where Mail=@Ml and Pwd=@Pwd And UType='Data Owner'";

This SQL query is inserted into the code of the data owner for
SQL
parameterization.
Now there is a complete security due the sql parameterization.The unauthorized users cannot access the data/file that is stored in the cloud server.It becomes an invalid authentication.

V. CONCLUSION

In the proposed scheme, accessing the data that is stored in the cloud server is made authorized. The data is uploaded by the authenticated data owner. The multi-keyword search takes place after uploading the data to the cloud server. The status of the data is viewed whether it is approved. The uploaded data is retrieved by the data user. The search takes place as content based search along with ranking. Based on the content, data is searched and retrieved and then the rank is applied. During searching, attacks might be done by SQL injection to dump the database contents by the unauthorized users or attackers. In this project SQL injection is not allowed. The SQL parameterization technique is used to prove that SQL injection is not possible.

REFERENCES

[1] Zhihua Xia, "A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL: PP NO: 99 YEAR 2015

[2] M. Kuzu, M. S. Islam, and M. Kantarcioglu, "Efficient similarity search over encrypted data, " in Data Engineering (ICDE), 2012 IEEE 28th International Conference on. IEEE, 2012, pp. 1156–1167.

[3] Wenhai Sun et al., "Privacy-Preserving Multikeyword Text Search in the Cloud Supporting Similarity-based Ranking", the 8th ACM Symposium on Information, Computer and Communications Security, Hangzhou, China, May 2013.

[4] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preservingmulti-keyword ranked search over encrypted cloud data," in IEEE INFOCOM, April 2011, pp. 829–837.

[5] C.Wang, N. Cao, K. Ren, and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced cloud data,"Parallel and Distributed Systems, IEEE Transactions on, vol. 23, no. 8,pp. 1467–1479, 2012.

[6] 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,vol. 14, 2014.

[7] Cong Wang et al.,"Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data", IEEE Transactions on parallel and distributed systems, vol. 23, no.8, August 2012

[8] P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in Applied Cryptography and Network Security. Springer, 2004, pp.
31–45