A SECURE AND DYNAMIC MULTI-KEYWORD RANKED SEARCH SCHEMA OVER ENCRYPTED CLOUD DATA

[1]P.Veeramuthu, [2]A.Sunitha
[1]Assistant Professor, [2]Student
[1]Computer science Department,
[1]B.T College, Madanapalle, India.

Abstract : ***Due to the increasing popularity of cloud computing, more and more data owners are motivated to outsource their data to cloud servers for great convenience and reduced cost in data management. However, sensitive data should be encrypted before outsourcing for privacy requirements, which obsoletes data utilization like keyword-based document retrieval. In this paper, we present a secure multi-keyword ranked search scheme over encrypted cloud data, which simultaneously supports dynamic update operations like deletion and insertion of documents. Specifically, the vector space model and the widely-used TF_IDF model are combined in the index construction and query generation. We construct a special tree-based index structure and propose a "Greedy Depth-first Search" algorithm to provide efficient multi-keyword ranked search. The secure Knn algorithm is utilized to encrypt the index and query vectors, and meanwhile ensure accurate relevance score calculation between encrypted index and query vectors. In order to resist statistical attacks, phantom terms are added to the index vector for blinding search results. Due to the use of our special tree-based index structure, the proposed scheme can achieve sub-linear search time and deal with the deletion and insertion of documents flexibly. Extensive experiments are conducted to demonstrate the efficiency of the proposed scheme.***

Index Terms - Cloud Computing, KNN algorithm, Greedy algorithm and RSA algorithm

---

I. INTRODUCTION

Cloud computing is a conversational phrase used to express a variety of dissimilar types of computing ideas that occupy large number of computers that are connected through a real-time communication network i.e Internet. In science, cloud computing is the capability to run a program on many linked computers at the same time. The fame of the term can be recognized to its use in advertising to sell hosted services in the sense of application service provisioning that run client server software on a remote location. Cloud computing relies on sharing of resources to attain consistency and financial system alike to a utility (like the electricity grid) over a network. The cloud also center on maximize the effectiveness of the shared resources. Cloud resources are typically not only shared by multiple users but as well as dynamically re-allocated as per demand. This can perform for assigning resources to users in dissimilar time zones. For example, a cloud computing service which serves American users during American business timings with a specific application (e.g. email) while the same resources are getting reallocated and serve Indian users during Indian business timings with another application (e.g. web server). The owner to register those details and also include login details. This module helps the owner to upload his file with encryption using RSA algorithm. This ensures the files to be protected from unauthorized user. Data owner has a collection of documents *F ={f1; f2; :::; fn}* that he wants to outsource to the cloud server in encrypted form while still keeping the capability to search on them for effective utilization. In our scheme, the data owner firstly builds a secure searchable tree index *I* from document collection *F*, and then generates an encrypted document collection *C* for *F*. Afterwards, the data owner outsources the encrypted collection *C* and the secure index *I* to the cloud server, and securely distributes the key information of trapdoor generation and document decryption to the authorized data users. Besides, the data owner is responsible for the update operation of his documents stored in the cloud server. While updating, the data owner generates the update information locally and sends it to the server.

This module includes the user registration login details. This module is used to help the client to search the file using the multiple key words concept and get the accurate result list based on the user query. The user is going to select the required file and register the user details and get activation code in mail email before enter the activation code. After user can download the Zip file and extract that file. Data users are authorized ones to access the documents of data owner. With *t* query keywords, the authorized user can generate a trapdoor *TD* according to search control mechanisms to fetch *k* encrypted documents from cloud server. Then, the data user can decrypt the documents with the shared secret key.

- Database
  - Tools needed/ setup (java script and MySQL).
  - Project details structure (Hierarchical structure of projects details).
  - Database design – Storing, updating and deleting the information in the database.
- Web application pages
  - HTML pages.
  - Project details structure (Hierarchical structure of projects details).
  - Database design (Storing, updating and deleting the information in the database).
  - Database implementation – SQL script with attributes.

II. Module Description

1. Registration form

The registration form helps to register the owner and users. The owner registration process is enter the username, password, email, date of birth.

2. **Data Owner Module**

This module helps the owner to register those details and also include login details. This module helps the owner to upload his file with encryption using RSA algorithm. This ensures the files to be protected from unauthorized user. Data owner has a collection of documents *F ={f1; f2; :::; fn}* that he wants to outsource to the cloud server in encrypted form while still keeping the capability to search on them for effective utilization. In our scheme, the data owner firstly builds a secure searchable tree index *I* from document collection *F*, and then generates an encrypted document collection *C* for *F*. Afterwards, the data owner outsources the encrypted collection *C* and the secure index *I* to the cloud server, and securely distributes the key information of trapdoor generation and document decryption to the authorized data users. Besides, the data owner is responsible for the update operation of his documents stored in the cloud server. While updating, the data owner generates the update information locally and sends it to the server

3. Data User Module

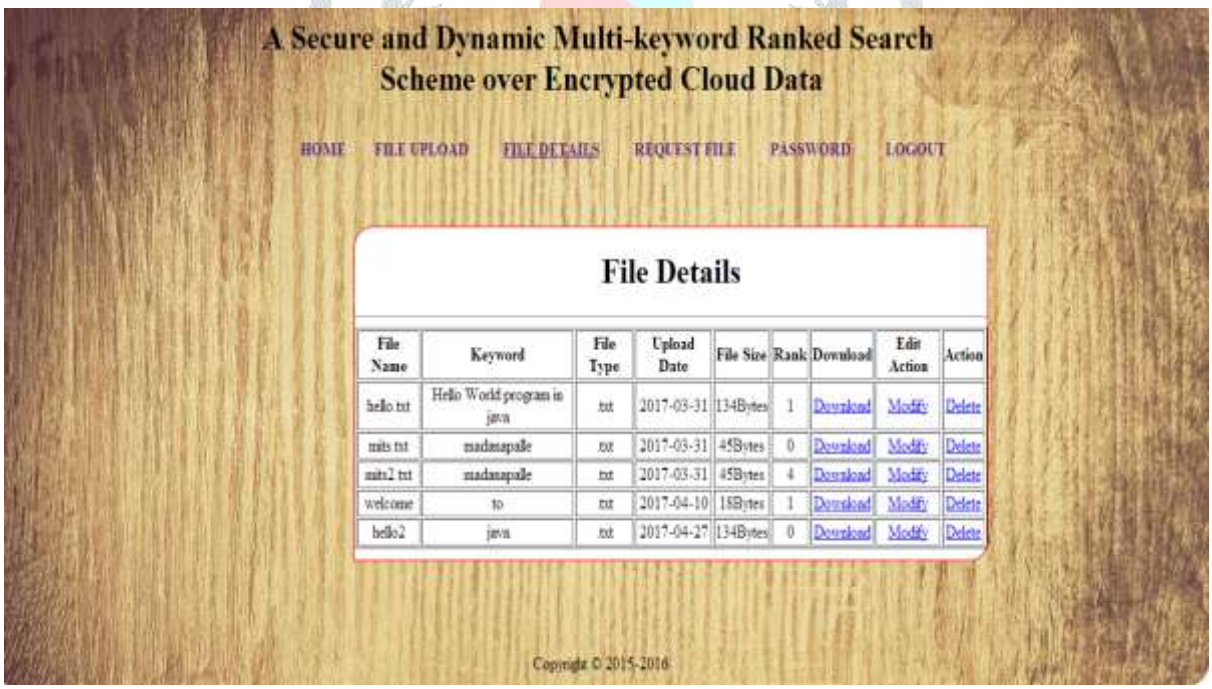
This module includes the user registration login details. This module is used to help the client to search the file using the multiple key words concept and get the accurate result list based on the user query. The user is going to select the required file and register the user details and get activation code in mail email before enter the activation code. After user can download the Zip file and extract that file. Data users are authorized ones to access the documents of data owner. With *t* query keywords, the authorized user can generate a trapdoor *TD* according to search control mechanisms to fetch *k* encrypted documents from cloud server. Then, the data user can decrypt the documents with the shared secret key.

4. Cloud Server and Encryption Module:

This module is used to help the server to encrypt the document using RSA Algorithm and to convert the encrypted document to the Zip file with activation code and then activation code send to the user for download. Cloud server stores the encrypted document collection *C* and the encrypted searchable tree index *I* for data owner. Upon receiving the trapdoor *TD* from the data user, the cloud server executes search over the index tree *I*, and finally returns the corresponding collection of top- *k* ranked encrypted documents. Besides, upon receiving the update information from the data owner, the server needs to update the index *I* and document collection *C* according to the received information. The cloud server in the proposed scheme is considered as "honest-but-curious", which is employed by lots of works on secure cloud data search

5. Rank Search Module

These modules ensure the user to search the files that are searched frequently using rank search. This module allows the user to download the file using his secret key to decrypt the downloaded data. This module allows the Owner to view the uploaded files and downloaded files. The proposed scheme is designed to provide not only multi-keyword query and accurate result ranking, but also dynamic update on document collections. The scheme is designed to prevent the cloud server from learning additional information about the document collection, the index tree, and the query.



III. Original work for the Research

This paper proposes a secure tree-based search scheme over the encrypted cloud data, which supports multi-keyword ranked search and dynamic operation on the document collection. Specifically, the vector space model and the widely-used "term frequency (TF) × inverse document frequency (IDF)" model are combined in the index construction and query generation to provide multi-keyword ranked search. In order to obtain high search efficiency, we construct a tree-based index structure and propose a "Greedy Depth-first Search" algorithm based on this index tree.

The secure kNN algorithm is utilized to encrypt the index and query vectors, and meanwhile ensure accurate relevance score calculation between encrypted index and query vectors. To resist different attacks in different threat models, we construct two secure search schemes: the basic dynamic multi-keyword ranked search (BDMRS) scheme in the known cipher text model, and the enhanced dynamic multi-keyword ranked search (EDMRS) scheme in the known background model.

IV. ADVANTAGES OF PROPOSED SYSTEM:

- Due to the special structure of our tree-based index, the proposed search scheme can flexibly achieve sub-linear search time and deal with the deletion and insertion of documents.
- We design a searchable encryption scheme that supports both the accurate multi-keyword ranked search and flexible dynamic operation on document collection.
- Due to the special structure of our tree-based index, the search complexity of the proposed scheme is fundamentally kept to logarithmic. And in practice, the proposed scheme can achieve higher search efficiency by executing our "Greedy Depth-first Search" algorithm. Moreover, parallel search can be flexibly performed to further reduce the time cost of search process.

V. FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

- ECONOMIC FEASIBILITY
- TECHNICAL FEASIBILITY
- SOCIAL FEASIBILITY

ECONOMIC FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

SOCIAL FEASIBILITY

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

TECHNOLOGY AND SYSTEM FEASIBILITY

The assessment is based on an outline design of system requirements in terms of Input, Processes, Output, Fields, Programs, and Procedures. This can be quantified in terms of volumes of data, trends, frequency of updating, etc. in order to estimate whether the new system will perform adequately or not. Technological feasibility is carried out to determine whether the company has the capability, in terms of software, hardware, personnel and expertise, to handle the completion of the project. When writing a feasibility report the following should be taken to consideration:

- A brief description of the business to assess more possible factor/s which could affect the study
- The part of the business being examined
- The human and economic factor
- The possible solutions to the problems
- At this level, the concern is whether the proposal is both *technically* and legally feasible (assuming moderate cost).

VI. CONCLUSION

In this paper, a secure, efficient and dynamic search scheme is proposed, which supports not only the accurate multi-keyword ranked search but also the dynamic deletion and insertion of documents. We construct a special keyword balanced binary tree as the index, and propose 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 kNN algorithm. Experimental results demonstrate the efficiency of our proposed scheme.

There are still many challenge problems in symmetric SE schemes. In the proposed scheme, the data owner is responsible for generating updating information and sending them to the cloud server. Thus, the data owner needs to store the unencrypted index tree and the information that are necessary to recalculate the IDF values. Such an active data owner may not be very suitable for the cloud computing model. It could be a meaningful but difficult future work to design a dynamic searchable encryption 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. Actually, there are many secure challenges in a multi-user scheme.

VII. References

[1] A Platform Computing Whitepaper. ―Enterprise Cloud Computing: Transforming IT.‖ Platform Computing, pp6, 2010.

[2] B.P. Rimal, Choi Eunmi, I. Lumb, ―A Taxonomy and Survey of Cloud Computing Systems‖, Intl. Joint Conference on INC, IMS and IDC, 2009, pp. 44-51,Seoul, Aug, 2009. DOI: 10.1109/NCM.2009.218

[3] B. R. Kandukuri, R. Paturi V, A. Rakshit, ―Cloud Security Issues‖, In Proceedings of IEEE International Conference on Services Computing, pp. 517-520, 2009.

[4] Cloud Computing. Wikipdia. Available at http://en.wikipedia.org/wiki/Cloud_computing

[5] Cong Wang, Qian Wang, KuiRen, and Wenjing Lou, ―Ensuring Data Storage Security in Cloud Computing,‖ 17th International workshop on Quality of Service, USA, pp.1-9, July 13-15, 2009, ISBN: 978-1-4244-3875-4

[6] C. Weinhardt, A. Anandasivam, B. Blau, and J. Stosser.―Business Models in the Service World.‖IT Professional, vol. 11, pp. 28-33, 2009.

[7] Daniel Oliveira and Eduardo Ogasawara. Article: Is Cloud Computing the Solution for Brazilian Researchers?. International Journal of Computer Applications 6(8):19–23, September 2010.

[8] D. Oliveira, F. Baião, and M. Mattoso, 2010, "Towards Taxonomy for Cloud Computing from an e-Science Perspective", Cloud Computing: Principles, Systems and Applications (to be published), Heidelberg: Springer-Verlag.

[9] Dr. Gurdev Singh, ShanuSood, Amit Sharma, ―CM- Measurement Facets for Cloud Performance‖, IJCA, , Lecturer, Computer science & Engineering, Eternal University, Baru Sahib (India), Volume 23 No.3, June 2011.

[10] Ertaul, L. and Singhal, S. 2009. Security Challenges in Cloud Computing. California State University, East Bay. Academic paper http://www.mcs.csueastbay.edu/~lertaul/Cloud%20Security%20CamREADY.pdf

[11] Gaoyun Chen, Jun Lu and Jian Huang, Zexu Wu, ―SaaAS - The Mobile Agent based Service for Cloud Computing in Internet Environment‖, Sixth International Conference on Natural Computation, ICNC 2010, pp. 2935-2939, IEEE, Yantai, Shandong,China, 2010. ISBN: 978-1-4244-5958-2.

[12] Global Netoptex Incorporated. ―Demystifying the cloud.Important opportunities, crucial choices.‖ pp4-14. Available: http://www.gni.com [Dec. 13, 2009].

[13] Hanqian Wu, Yi Ding, Winer, C., Li Yao, ―Network Security for Virtual Machines in Cloud Computing,‖ 5th Int'l Conference on Computer Sciences and Convergence Information Technology, pp. 18-21, Seoul, Nov. 30- Dec. 2, 2010. ISBN: 978-1-4244-8567-3.

[14] Harjit Singh Lamba and Gurdev Singh, ―Cloud Computing-Future Framework for e-management of NGO's‖, IJoAT, ISSN 0976-4860, Vol 2, No 3, Department Of Computer Science, Eternal University, Baru Sahib, HP, India, July 2011.

[15] Harold C. Lin, Shivnath Babu, Jeffrey S. Chase, Sujay S. Parekh, ―Automated Control in Cloud Computing: Opportunities and Challenges‖, Proc. of the 1st Workshop on Automated control for data centres and clouds, New York, NY, USA, pp. 13-18, 2009, ISBN: 978-1-60558-585-7.

[16] Hoang T. Dinh, Chonho Lee, DusitNiyato, Ping Wang, ―A Survey of Mobile Cloud Computing: Architecture, Applications and Approaches‖, Wireless Communications and Mobile Computing, Wiley Inc.

[17] I. Foster, Y. Zhao, I. Raicu, and S. Lu, 2008, Cloud Computing and Grid Computing 360-Degree Compared, In: Grid Computing Environments Workshop, 2008. GCE '08, p. 10, 1.

[18] Joachim Schaper, 2010, ―Cloud Services‖, 4th IEEE International Conference on DEST, Germany.

[19] K.Mukherjee and G.Sahoo. Article: Cloud Computing: Future Framework for e-Governance. International Journal of Computer Applications 7(7):31–34, October.

[20] Kuyoro S.O., Ibikunle F., Awodele O., ―Cloud Computing Security Issues & Challenges‖, IJCN, Vol. 3 Issue 5: 2011, pp. 247-255.