|| Volume 3 || Issue 6 || June 2018 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

Impregnable and Active Multi-keyword Ranked Search System over Encrypted Cloud Data

Santosh Rahane[1], Yogesh Chikane[2]
Student, M.E., Information Technology, AVCOE, Sangamner,India[1]
Assistant Professor, Information Technology, AVCOE, Sangamner, India [2]
santoshrahane4@gmail.com[1], yogeshchikane2006@yahoo.com[2]

ABSTRACT: ***The increasing importance of distributed computing, an ever increasing number of information proprietors are persuaded to outsource their information to cloud servers for incredible accommodation and diminished cost in information administration. Notwithstanding, touchy information ought to be scrambled before outsourcing for protection prerequisites, which obsoletes information usage like watchword based record recovery. In this paper, we show a safe multi-watchword positioned seek plot over scrambled cloud information, which at the same time underpins dynamic refresh tasks like erasure and addition of records. In particular, the vector space show and the generally utilized TF_IDF display are consolidated in the list development and question age. We build an exceptional tree-based record structure and propose a "Covetous Depth-first Search" calculation to give productive multi-catchphrase positioned look. The protected kNN calculation is used to encode the list and question vectors, and in the interim guarantee precise pertinence score figuring between scrambled list and inquiry vectors. Protection in mind the end goal to oppose quantifiable assaults, apparition terms are added to the record vector for blinding list items. Because of the utilization of our extraordinary tree-based list structure, the proposed plan can accomplish sub-direct hunt time and manage the cancellation and inclusion of archives flexibly. Extensive investigations are led to show the productivity of the proposed conspires.***

Keywords: *- Searchable encryption, multi-keyword ranked search, dynamic update*

I INTRODUCTION

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 [1]. Concerned 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 [2]. However, this will cause a huge cost in terms of data usability. For example, the existing techniques on keyword-based information retrieval, which are widely used on the plaintext data, cannot be directly applied on the encrypted data. Downloading all the data from the cloud and decrypt locally is obviously impractical. In order to address the above problem, researchers have designed some general-purpose solutions with fully homomorphic encryption [3]. However, these methods are not practical due to their high computational overhead for both the cloud sever and user. On the contrary, more practical special-purpose solutions, such as searchable encryption (SE) schemes have made specific contributions in terms of efficiency, functionality and security.

Searchable encryption schemes enable the client to store the encrypted data to the cloud and execute keyword search over cipher text domain. So far, abundant works have been proposed under different threat models to achieve various search functionality, such as single keyword search, similarity search, multi-keyword boolean search, ranked search, multi-keyword ranked search, etc. Among them, multi-keyword ranked search achieves more and more attention for its practical applicability. Recently, some dynamic schemes have been proposed to support inserting and deleting operations on document collection. These are significant works as it is highly possible that the data owners need to update their data on the cloud server. But few of the dynamic schemes support efficient multi-keyword ranked search. 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

(GDFS)" algorithm based on this index tree. 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.

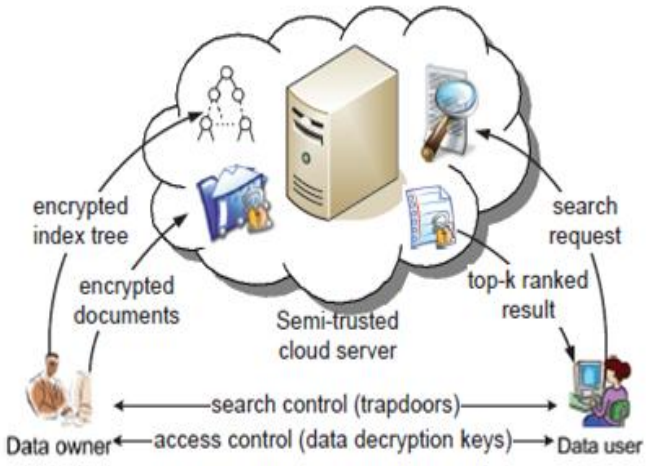

Figure. 1. Architecture of the search over encrypted cloud data

II MULTI-KEYWORD RANKED SEARCH OVER ENCRYPTED CLOUD DATA

Now a day's cloud computing has become essential for many utilities, where cloud customers can slightly store their data into the cloud so as to benefit from on-demand high quality request and services from a shared pool of configurable computing resources. Its huge suppleness and financial savings are attracting both persons and enterprise to outsource their local complex data management system into the cloud. To safe guard data privacy and struggle. unwanted accesses in the cloud and away from, sensitive data, for example, emails, personal health records, photo albums, videos, land documents, financial transactions, and so on, may have to be encrypted by data holder before outsourcing to the business public cloud; on the other hand, obsoletes the traditional data use service based on plaintext keyword search. The insignificant solution of downloading all the information and decrypting nearby is clearly impossible, due to the enormous amount of bandwidth cost in cloud scale systems. Furthermore, apart from eradicating the local storage management, storing data into the cloud supplies no purpose except they can be simply searched and operated. Thus, discovering privacy preserving and effective search service over encrypted cloud data is one of the supreme importance. In view of the potentially large number of on-demand data users and vast amount of outsourced data documents in the cloud, this difficulty is mostly demanding as it is really difficult to gather the requirements of performance, system usability, and scalability. On the one hand, to congregate the efficient data retrieval requirement, the huge amount of documents orders the cloud server to achieve result relevance ranking, as an alternative of returning undifferentiated results. Such ranked search system allows data users to discover the most appropriate information quickly, rather than burdensomely sorting during every match in the content group. Ranked search can also gracefully remove redundant network traffic by transferring the most relevant data, which is highly attractive in the "pay-as-you-use" cloud concept. For privacy protection, such ranking operation on the other hand, should not reveal any keyword to related information. To get better the search result exactness as well as to improve the user searching experience, it is also essential for such ranking system to support multiple keywords search, as single keyword search often give up far too common results. As a regular practice specifies by today's web search engines i,e Google search, data users may lean to offer a set of keywords as an alternative of only one as the indicator of their search interest to retrieve the most relevant data. And each keyword in the search demand is able to help narrow down the search result further. "Coordinate matching", as many matches as possible, is an efficient resemblance measure among such multi-keyword semantics to refine the result significance, and has been widely used in the plaintext information retrieval (IR) community. Though, the nature of applying encrypted cloud data search system remains a very demanding task in providing security and maintaining privacy, like the data privacy, the index privacy, the keyword privacy, and many others. Encryption is a helpful method that treats encrypted data as documents and allows a user to securely search through a single keyword and get back documents of interest. On the other hand, direct application of these approaches to the secure large scale cloud data utilization system would not be necessarily suitable, as they are developed as crypto primitives and cannot put up such high service-level needs like system usability, user searching experience, and easy information discovery.

Even though some modern plans have been proposed to carry Boolean keyword search as an effort to improve the search flexibility, they are still not sufficient to provide users with satisfactory result ranking functionality. The solution for this problem is to secure ranked search over encrypted data but only for queries consisting of a single keyword. The challenging issue here is how to propose an efficient encrypted data search method that supports multi-keyword semantics without privacy violation. In this paper, we describe and solve the problem of multi-keyword ranked search over encrypted cloud data (MRSE) while preserving exact system wise privacy in the cloud computing concept. Along with various multi keyword semantics, select the efficient resemblance measure of "coordinate matching," it means that as various matches as possible, to confine the significance of data documents to the

search query. Particularly, inner product similarity the numbers of query keywords show in a document, to quantitatively calculate such similarity assess of that document to the search query. For the period of the index construction, each document is associated with a binary vector as a sub-index where each bit signifies whether matching keyword is contained in the document. The search query is also illustrates as a binary vector where each bit means whether corresponding keyword appears in this search request, so the resemblance could be exactly calculated by the inner product of the query vector with the data vector. On the other hand, directly outsourcing the data vector or the query vector will go against the index privacy or the search privacy. To face the challenge of cooperating such multi keyword semantic without privacy breaches, we propose a basic idea for the MRSE using secure inner product computation, which is modified from a secure k nearest neighbour (kNN) method, and then give two considerably improved MRSE method in a step-by-step way to accomplish different severe privacy needs in two risk models with enlarged attack competence[4].

1. We suggest two MRSE schemes based on the Similarity calculation of "coordinate matching" at the time of assembling different privacy needs in two different threat models.
2. We examine some further improvements of our ranked search method to maintain more search semantics and dynamic data process.
3. we determine the problem of multi keyword ranked search over encrypted cloud data, and set up a set of privacy needs for such a secure cloud data operation system.
4. Detailed analysis investigating privacy and Efficiency assurance of the proposed schemes is known, and testing on the real-world data set further show the proposed schemes certainly bring in low overhead on calculation and communication. In this paper we propose two new methods to maintain more search semantics. These methods also study the support of data/index dynamics in the system design.

III OBJECTIVES

The Proposed cloud storage systems that offer privacy, reliability and authentication of client data against a un-trusted cloud provider. This OTP used to see data in cloud and it can be used once only in a time, when you search a file and want to see the file, the OTP will send to the email or to the phone and getting the OTP use the OTP to utilize the file . Presently in the existing system the cloud server hosts third-party data storage and get back services. As information may have sensitive information, the cloud servers cannot be fully hand over in protecting data. For this cause, outsourced files must be encrypted. Any type of data leakage that would involve data privacy is considered as undesirable. To meet the effective data retrieval requirement, the huge amount of documents command the cloud server to achieve result relevance ranking, as an alternative of returning undifferentiated results. Such ranked search system allows data users to find the most appropriate information quickly, rather than burdensomely sorting through every match in the content collection. Ranked search can also gracefully eradicate avoidable network traffic by transferring back only the most appropriate data, which is highly attractive in the "pay-as-you-use" cloud concept. For privacy protection, such ranking process, yet, should not leak any keyword related information. On the other hand, to progress the search result correctness as well as to improve the user searching experience, it is also essential for such ranking system to maintain multiple keywords search, as single keyword search regularly yields far too common results.

IV PROPOSED SYSTEM

In the Proposed work, we will discover checking the integrity of the rank order in the search result analyzing the cloud server is un-trusted. To advise OTP (one Time Password) as our upcoming work. This OTP used to see information in cloud and it can be used once only in a time, when you search a file and be likely to see the file, the OTP will transmit to email and we receive the OTP and apply to see the file[5].

In this technique the following are the different things which we have to implement.

1) Setup
2) Cloud Storage
3) Vector Model

Cloud Setup Firstly, we have to setup data owner and cloud server. So the data owner will then push the data into the cloud servers. When users outsource their confidential data onto the cloud, the cloud service providers are capable to control and check the data and the communication between users and the cloud will be secured. Cryptography cloud Storage Secondly, while the data is uploaded into the Storage and retrieve services. Since data may have confidential information, the cloud servers cannot be fully hand over in protecting data. For this cause, outsourced files must be encrypted. Any kind of information leakage that would change data privacy are regarded as Unacceptable. Vector Model We used a series of searchable symmetric encryption systems that have been allowing search on cipher text. In the earlier, files are ranked only by the number of get back keywords, which damage search correctness.

V CONCLUSION AND FUTURE WORK

In this paper we describe and solve the problem of multi keyword ranked search over encrypted cloud data, and set up a range of privacy requirements. Among various multi-keyword semantics, we select the efficient similarity measure of "coordinate matching," i.e., as many equivalent as possible, to

effectively capture the relevance of outsourced documents to the query Keywords, and utilize "inner product similarity" to quantitatively calculate such comparison measure. In order to acquire the test of supporting multi-keyword semantic without privacy violation, we offer a basic idea of MRSE using secure inner product calculation. Then, we give two improved MRSE schemes to attain various severe privacy needs in two different threat models. The further enhancements of our ranked search method, including supporting more search semantics, i.e., TF _ IDF, and dynamic data process. Detailed analyses in investigating privacy and efficiency assurance of proposed schemes are mentioned, and testing on the real-world data set demonstrate our proposed schemes which introduces low transparency on both calculation and communication.

ACKNOWLEDGMENT

This research was supported by Amrutvahini college of engineering sangamner. We are thankful to our colleagues Dr. B. L. Gunjal, Prof. B. R. Borkar who provided expertise that greatly assisted the research.

REFERENCES

[1] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," Proc. IEEE INFOCOM, pp. 829- 837, Apr, 2011.

[2] E.-J. Goh, ―Secure Indexes, ‖ Cryptology ePrint Archive, http:// eprint.iacr.org/2003/216. 2003 [3] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, ―Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data, ‖ Proc. IEEE INFOCOM, Jan, 2014

[3] Towards Secure Multi-Keyword Ranked Search over Encrypted Cloud Data (MRSE) Authors: B. Jeeva, Dr.S.Rajalakshmi.

[4] Achieving Efficiency of Encrypted Cloud Data with Synonym Based Search and Multi-Keyword Ranked Search Dipika Chavan1, Dinesh Yadav2