# A Protected & Lively Multi-Keyword Hierarchical Search Arrangement over Encoded Cloud Data

**D. Agalya[1] N. Vasunthara Devi[2]**
[1,2]Department of Computer Science & Engineering
[1,2]A.V.C Arts and Science College, Bharathidasan University, Trichy, India

*Abstract*— Research focus on a secure multi-keyword ranked search scheme over encrypted cloud data, which simultaneously supports dynamic update operations like deletion and insertion of documents. The research propose a “Greedy Depth- first Search” algorithm to provide efficient multi-keyword ranked search. A secure tree-based search scheme over the encrypted cloud data, which supports multi-keyword ranked search and dynamic operation on the document collection. Specially 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, construct a tree-based index structure and propose a “Greedy Depth-first Search” 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. The secure 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. Moreover, parallel search can be flexibly performed to further reduce the time cost of search process. Extensive experiments are conducted to demonstrate the efficiency of the proposed scheme.

***Key words:*** TF, IDF

## I. Introduction

The research propose 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, construct a tree-based index structure and propose a “Greedy Depth-first Search” 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. 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.

Cloud Computing is a new but increasingly mature model of enterprise IT infrastructure that provides on-demand high quality applications and services from a shared pool of configuration computing resources. The cloud customers, individuals or enterprises, can outsource their local complex data system into the cloud to avoid the costs of building and maintaining a private storage infrastructure, as the cloud server possesses powerful functionality and flexibility. However, some problems may be caused in this circumstance since the Cloud Service Provider (CSP) possesses full control of the outsourced data. Unauthorized operation on the outsourced data may exist on account of curiosity or profit. To protect the privacy of sensitive information, sensitive data (e.g., emails, photo albums, personal health records, financial records, etc.,) should be encrypted by the data owner before outsourcing, which makes the traditional and efficient plaintext keyword search technique useless. The simple and awkward method of downloading all the data and decrypting locally is obviously impractical. So, three aspects should be concentrated on to explore privacy- preserving effective search service.

Firstly, ranked search, which can enable data users to find the most relevant information quickly, is a very important issue. The number of documents out sourced to the cloud is so large that the cloud should have the ability to perform search result ranking to meet the demand for effective data retrieval. Secondly, multi-keyword search is also very important to improve search result accuracy as single keyword search often return coarse search results. The last but not least, dynamic update is a useful functionality for a good data management system which should provide as more as possible convenience for the data owner. It is common that sometimes the data owner wants to add a document to the dataset or delete a document from the dataset. So, a data management system that supports insertion and deletion update is a more integrated one. In recent years, many researchers have engaged in the field of searchable encryption over encrypted cloud data and put forward a series of outstanding achievements. Nevertheless, there are still many challenging and important problems need to be solved. There does not exist scheme that supports both multi-keyword ranked search and dynamic update.

How to design a scheme supporting both multi-keyword ranked search and dynamic update is still a challenging open problem. Propose a practically efficient and flexible searchable encrypted scheme supporting dynamic update. To address multi-keyword search and result ranking, use Vector Space Model (VSM) to build document index. To improve search efficiency, use a tree-based index structure which is a balanced binary tree. Construct the searchable index tree based on the document index vectors. And our tree-based index tree also can support dynamic update, including insertion update and deletion update, without any privacy leakage. Our encryption scheme can meet the privacy requirements in the threat model.

Our contributions are summarized as follows: (1) For the first time, study the problem of multi-keyword ranked search supporting dynamic update over encrypted cloud data while supporting strict privacy requirements. (2) With the index tree designed in the research, our scheme can support dynamic update. And the time complexity is the worst case for updating a document. (n is the size of the keyword dictionary and m is the whole number of documents in the dataset). (3) Our scheme can meet the privacy requirements in the threat model. Make security analysis for our scheme which proves privacy guarantees. And experiments on the real-world dataset show that proposed scheme is indeed efficient.

Cloud computing has been considered as another model of enterprise IT infrastructure, which can compose gigantic resource of computing, storage and applications, and empower users to appreciate pervasive, helpful and on-demand network access to a mutual pool of configurable computing resources with incredible efficiency and insignificant economic overhead. Pulled in by these engaging features, both individuals and enterprises are roused to outsource their data to the cloud, rather than buying software and hardware to deal with the data themselves. In spite of the different points of interest of cloud services, outsourcing delicate information, (for example, e-mail, individual health records, organization account information, government archives, and so forth.) 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 way to deal with secure the data privacy is to encrypt the data before outsourcing. On the other hand, this will bring about a gigantic expense in terms of data ease of use.

For example, the current techniques on keyword-based information retrieval, which are broadly utilized on the plaintext data, can't be straightforwardly connected on the encrypted data. Downloading all the data from the cloud and decrypt locally is clearly unrealistic. With a particular final objective to address the above issue, analysts have illustrated some all-around helpful arrangements with totally homomorphic encryption or missing RAMs. In any case, these schedules are not down to earth in light of their high computational overhead for both the cloud server and user. In spite of what may be normal, more useful unique reason arrangements, for instance, searchable encryption (SE) plan have made specific responsibilities to the extent productivity, value and security. Searchable encryption scheme engage the user to store the encrypted data to the cloud and execute unequivocal word look for over cipher text domain. As being what is indicated, abundant works have been proposed under assorted risk models to finish distinctive interest value, for instance, single keyword search, closeness look, multi-keyword Boolean search, ranked search, multi-keyword ranked search, etc. Among them, multi-keyword positioned quest finishes more thought for its pragmatic propriety. Starting late, some component arrangements have been proposed to reinforce embedding and erasing operations on archive gathering. These are colossal goes about as it is exceptionally possible that the data owner need to overhaul their data on the cloud server. Yet, few of the dynamic plan support successful multi-keyword situated look. Inverse document recurrence (IDF)" model are joined in the list development and inquiry era to give multi-keyword positioned seek. Keeping in mind the end goal to get high search Effectiveness, develop a tree based list structure and based on this tree list propose a “Greedy Depth first Search” calculation. Because of the uncommon structure of our tree-based list, the proposed search scheme can flexibly accomplish sub-straight search time and manage the deletion and insertion of reports. The protected kNN algorithm is used to encrypt the index and query vectors, and in the interim guarantee relevance score calculation between encrypted index and query vectors. To oppose distinctive attacks in different threat models, build 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. Our commitments are condensed as takes after:

Design a searchable encryption scheme that underpins both the precise multi-keyword ranked search and flexible dynamic operation on document collection.

The proposed scheme is in a general sense kept to logarithmic for search complexity in the uncommon structure of tree-based index. What's more, practically speaking, the proposed scheme can accomplish higher search proficiency by executing our "Greedy Depth-first Search” algorithm. Additionally to reduce the time cost of search process parallel search can performed.

## *A. Clouds Classification*

Cloud computing is a new computing prototype that is built on virtualization, parallel and distributed computing, utility computing and service oriented architecture. For the past many years, cloud computing has emerged as one of the most important and used technology in the IT industry and academic world. The benefits of cloud computing include reduction in costs and in capital expenditures, increased operational efficiencies, scalability, flexibility, immediate time to market and many more. The various cloud computing models are Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). It is a subscription-based service, where in one can obtain networked storage space and computer resources. According to the different needs of the user, one can subscribe different types of cloud. Clouds are classified.

### *1) Public Cloud*

A public cloud can be used by any user with an internet connection and access to the cloud space.

### *2) Private Cloud*

A private cloud is established for a specific group or organization and access to this cloud is limited just to group or organization alone.

### *3) Community Cloud*

A community cloud is shared between two or more organizations which have common cloud requirements.

### *4) Hybrid Cloud*

A hybrid cloud is fundamentally a combination of at least two clouds, where the clouds included are a blend of public, private, or community.

Cloud computing is a term used to describe a set of IT services that are provided to a customer over a network on a leased basis and with the ability to scale up or down their service requirements. Clouds are large pools of easily usable and accessible virtualized resources. These resources can be dynamically reconfigured to adjust to a variable load (scale), allowing optimum resource utilization. It’s a pay-per-use model in which the Infrastructure Provider by means of customized Service Level Agreements (SLAs) offers guarantees typically exploiting a pool of resources. Organizations and individuals can benefit from mass computing and storage centers, provided by large companies with stable and strong cloud architectures. To protect data privacy and combat uninvited accesses in the cloud and further than, thin skinned data, e.g., emails, personal health records, photo albums, tax documents, financial transactions, etc., may have to be encrypted by data proprietor before outsourcing to the commercial public cloud. This however, obsoletes the traditional data utilization service based on plaintext keyword search. The trivial solution of downloading all the data and decrypting locally is clearly impractical, due to the huge amount of bandwidth cost in cloud scale systems.

Moreover, aside from eliminating the local storage management, storing data into the cloud serves no purpose unless they can be easily searched and utilized. Thus, exploring privacy-preserving and effective search service over encrypted cloud data is of paramount importance. Considering the potentially large number of on-demand data users and huge amount of outsourced data documents in the cloud, this problem is particularly challenging as it is extremely difficult to meet also the requirements of performance, system usability and scalability. Need for data retrieval is the most frequently occurring task in cloud by the user to the server. The retrieval of the data should be fast enough. But the large amount of data space is used by the user, which in turn increases the time of search. Generally cloud server assigns ranks to document in order to make the search as faster. Such ranked search system enables data users to find the most relevant information quickly, rather than burdensomely sorting through every match in the content collection. Ranked Search can also elegantly eliminate unnecessary network traffic by sending back only the most relevant data, which is highly desirable in the “pay-as-you-use” cloud paradigm. For privacy protection, such ranking operation, however, should not leak any keyword related information.

On the other hand, to improve the search result accuracy as well as to enhance the user searching experience, it is also necessary for such ranking system to support multiple keywords search, as single keyword search often yields far too coarse results. As a common practice indicated by today’s web search engines (e.g., Google search), data users may tend to provide a set of keywords instead of only one as the indicator of their search interest to retrieve the most relevant data. And each keyword in the search request is able to help narrow down the search result further. “Coordinate matching”, i.e., as many matches as possible, is an efficient similarity measure among such multi-keyword semantics to refine the result relevance, and has been widely used in the plaintext information retrieval (IR) community. However, how to apply it in the encrypted cloud search system remains a very challenging task because of inherent security and privacy obstacles, including various strict requirements like the data privacy, the index privacy, the keyword privacy, and many others.

The research focuses on to the solution of multi-keyword ranked search over encrypted cloud data (MRSE while preserving strict system-wise privacy in the cloud computing paradigm. A variety of multi- keyword semantics are available, an efficient similarity measure of “coordinate matching”, i.e., as many matches as possible, to capture the relevance of data documents to the search query is used. Particularly “inner product similarity”, i.e., the number of query keywords appearing in a document, to quantitatively evaluate such similarity measure of that document to the search query is used in MRSE algorithm.

Cloud computing is being intensively referred to as one of the most prominent innovations in information technology in recent epoch. By using resource virtualization cloud delivers us computing resources and services in a pay-as-you-go mode. Today world is moving on digitization and cloud computing is best concept to handle big datasets. Various cloud computing services are categorized into Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and last one is Software-as-a-Service (SaaS). Cloud computing is the elongated dreamed hallucination of computing as efficacy, where cloud customers tenuously stock up their data into the cloud so as to take pleasure in the on-order far above the ground-eminence application and services from a public pool of configurable computing resources. Its great plasticity and financial savings are rousing both folks and enterprise to outsource subcontract their local multifaceted data management system into the cloud. To guard privacy of data and be in opposition to unwelcome accesses in the cloud and further than it, susceptible data, for illustration, e-mails, personal health records, photo albums, tax documents, and so on, may have to be encrypted by data owners before Outsourcing to the commercial public cloud; this, however, obsoletes the conventional data employment service based on plaintext keyword investigate.

The irrelevant way out of downloading all the data and decrypting locally is obviously unreasonable, due to the big quantity of bandwidth cost in cloud scale systems. Images also be full of practical and vital information, so anticipated system also provides image cataloging in MRSE scheme. Moreover, aside from eliminating the local storage management, storing data into the cloud doesn't hand out any point except they can be with no trouble searched plus utilized. Hence, explore privacy preserving and effective search service over encrypted cloud data is of immense consequence. Ranked search can also stylishly get rid of needless network traffic by sending back only the majority germane data, which is exceedingly enviable in the “pay-as-you-use” cloud paradigm. For privacy protection, such ranking process, however, should not leak any keyword related information. Besides, to improve search result accuracy as well as to enhance the user searching experience, it is also necessary for such ranking system to support multiple keyword search, as single keyword rummage around often yields far too coarse results. As a common practice indicate by today’s web search engines (ex. Google search), data users may tend to provide a set of keywords instead of only one as the indicator of their search interest to retrieve the most applicable data. Along with the privacy of data and efficient penetrating schemes, real privacy is obtained only if the user's identity remains hidden from the Cloud Service Provider (CSP) as well as the third party user on the cloud server.

## II. Related work

### A. Multiuser Multi-Keyword Ranked Search over Encrypted Cloud Using MHR & KP-ABE

The benefit of storage as a service several enterprises area unit moving their valuable information to the cloud, since it prices less, simply scalable and can be accessed from anyplace any time. However, delicate information like email, personal health record, Gov. Record ought to be encoded before outsourcing for security prerequisites, which obsoletes information utilization like keyword based document retrieval. Present Multiuser Multi-keyword Ranked Search scheme over Encrypted Cloud using MHR Tree and KP-ABE. The MHR tree (Markel hash tree) algorithm can achieve logarithmic search time and deal with the deletion and insertion of documents flexibly. Keyword Policy-Attribute based encryption (KP-ABE) is secure encryption technique, it provide the fine grained access control and high security on document collection.

The cloud server can then support queries as follows, a data consumer submits a Multi-keyword query and attributes to the cloud server. The cloud server conduct search using Markel tree search and returns only list of files name in which keyword is present. Once receiving encrypted files, the user decrypts the files if the set of attributes of the user matches the policy of data owner. This approach guarantee the information security and preserve the data privacy. Throughout the complete process, no plain text data or keywords are visible to the cloud servers. Our contribution are summarized as follows:

- Design a secure encryption scheme that supports accurate multi-keyword search using Markel hash tree and provide the security by using KP-ABE encryption algorithm.
- Due to special structure of our tree based the search complexity of implemented scheme is logarithmic.
- Implementation scheme achieves the scalability, if number of user increases then also system is not affected.

The research implemented an improved technique for multi-keyword searching over secured cloud server. Make contribution mainly in two aspects: Markel hash tree search algorithm for fast searching and key policy attribute based encryption for providing more security on document over cloud. In terms of efficiency, implement the hash tree, in that all data present in at leaf node in the hash code format, so searching over the hash tree is easy, it gives conjunctive as well as disjunctive search result. This tree search only in leaf nodes hence its time complexity is O(log n). KP-ABE provide one extra level of security. It achieve fine grained access control, scalability and data confidentiality simultaneously.

### B. DRSMS: Domain and Range Specific Multi-Keyword Search over Encrypted Cloud Data

One of the most fundamental services of cloud computing is Cloud storage service. Huge amount of sensitive data is stored in the cloud for easy remote access and to reduce the cost of storage. It is necessary to encrypt the sensitive data before uploading to the cloud server in order to maintain privacy and security. All traditional searchable symmetric encryption (SSE) schemes enable the users to search on the entire index file. Propose the Domain and Range Specific Multi- keyword Search (DRSMS) scheme that minimizes the search time and Index storage space. This scheme adopts collection sort technique to split the index file into D Domains and R Ranges. The Domain is based on the length of the keyword; the Range splits within the domain based on the first letter of the keyword. A mathematical model is used to search over the encrypted indexed keyword that eliminates the information leakage. Binary search is used to select the range within the domain with time complexity O(RlogD) and linear search is used to find the keyword within the range with O(R). The space complexity of the index storage space is $O(N_T \text{ X } 3)$ and search time complexity is O(1)+O(RlogD)+O(R), while the complexity of index generation is $O(N_T \text{ X } 3)$. Extensive experiments on real-world dataset validate our analysis and shows that the proposed DRSMS scheme is more efficient and secure than RSSE Scheme.

### C. Dynamic Multi-Keyword Rank Scheme using Top Key over Encrypted Cloud Data

Many data user are encouraged to outsource their data to cloud servers for great portable and reduced costs in data management which is increased popularity of cloud computing. However, sensitive data should be encrypted before outsourcing of data protection requirements, the use of data as keyword- based document obsoletes can be retrieved. It presents a secure multi-keyword search Place Scheme over encrypted cloud data that simultaneously supports dynamic update operations such as deleting and inserting documents. In particular, the vector space model and the widespread TF_IDF model are migrated in the index evaluation and query generation. It construct a special tree- based index structure and propose a "Greedy search" code which is efficient multi-keyword search on the place and formed tree index structure. The safe kNN algorithm used to encrypt the index and query vectors, and evaluated calculation between encrypted index and query vectors for efficiency. To withstand statistical attacks are added Search results for Phantom dazzle with respect to the index vector. By using our special tree-based index structure, the proposed rule flexibly sub-linear search time and deal with the deleting and inserting documents reach.

Cloud Computing, a critical pattern for advanced data services, has to outsource a necessary feasibility for Data Users Data. Controversies on privacy, but were presented as outsourcing of sensitive information, including e-mail, medical records and personal photos unceasingly expands explosively. Reports of data loss and data breaches in cloud computing systems from time to time appear. The biggest threat to privacy roots when users outsource their private data to the cloud in the cloud itself. The cloud service providers capable of the data and the communication between the users and the cloud will, lawful or unlawful to control and monitor.

### D. A Secure & Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data

The advent of cloud computing, data owners are motivated to outsource their complex data management systems from local sites to commercial public cloud for great flexibility and economic savings. But for protecting data privacy, sensitive data has to be encrypted before outsourcing, which obsoletes traditional data utilization based on plaintext keyword search. Thus, enabling an encrypted cloud data search service is of paramount importance. Considering the large number of data users and documents in cloud, it is crucial for the search service to allow multi-keyword query and provide result similarity ranking to meet the effective data retrieval need. Related works on searchable encryption focus on single keyword search or Boolean keyword search, and rarely differentiate the search results. The first time, define and solve the challenging problem of privacy-preserving multi-keyword ranked search over encrypted cloud data (MRSE), and establish a set of strict privacy requirements for such a secure cloud data utilization system to become a reality. Among various multi-keyword semantics, choose the efficient principle of "coordinate matching", i.e., as many matches as possible, to capture the similarity between search query and data documents, and further use "inner product similarity" to quantitatively formalize such principle for similarity measurement. First propose a basic MRSE scheme using secure inner product computation, and then significantly improve it to meet different privacy requirements in two levels of threat models. Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given, and experiments on the real-world dataset further show proposed schemes indeed introduce low overhead on computation and communication.

### *E. A Secure and Dynamic Multi-Keyword Ranking Search on Encrypted Cloud Data using GDFS*

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. A secure multi-keyword ranked search scheme over encrypted cloud data, which simultaneously supports dynamic update operations like deletion and insertion of documents is proposed. A special tree-based index structure is constructed and a Greedy Depth-first Search algorithm to provide efficient multi-keyword ranked search is proposed. The secure kNN algorithm is utilized to encrypt the index and query vectors, and ensures accurate relevance score calculation between encrypted index and query vectors. Due to the use of the special tree-based index structure, the proposed scheme can achieve sub-linear search time and deal with the deletion and insertion of documents flexibly.

### *F. Fastened Multi keyword Search over Encrypted Cloud Data*

Data owner incentivizes to outsource their data on the cloud to get more flexibility. Cloud computing provides data outsourcing and high quality accommodation. For data security, the data owner provides encryption on their data. The data owners outsource their data on the cloud through which they reduces a cost and computational overhead quandaries. Considering the sizably voluminous number of data users and documents in the cloud, it is obligatory to sanction multiple keywords in the search request and return documents in the order of their suggestion to these keywords. Cognate works on searchable encryption fixate on single keyword search or Boolean keyword search, and virtually not ever sort the look for results. In subsisting system, for the first instance, the conundrum of privacy-preserving multi-keyword ranked search over encrypted data in cloud computing (MRSE) is define and solve. A set of several privacy desiderata for such a bulwarked cloud data utilization system are defined by proposed work. The basic idea of MRSE mechanism is to perform a safe inner product calculation, and then gives two appreciably improved MRSE schemes (Coordinate matching and Inner product similarity) to achieve various privacy requirements in two different models. Advance tree based index structure and numerous adaptive approaches for multi-dimensional (MD) algorithm are propose to increase the search efficiency so that the technical search efficiency is more than that of linear search.

### *G. An efficient Multi-Keyword Synonym Ranked Query over Encrypted Cloud Data using BMS Tree*

Cloud computing is embryonic technology as a new computing model in several business domains. Large numbers of large-scale organizations are starting to shift the information on to the cloud environment. However, shifting the data to the cloud relieves the organizations from the monotonous tasks of organization management, minimizes maintenance costs and also less hands on management. However, security and privacy become key security issues when data owners outsource their private data onto un trusted public cloud servers. However, traditional keyword plain text search is obsolete. Several techniques have been developed to preserve the privacy of sensitive data in the cloud environment. But existing searching techniques over encrypted cloud data considers only exact or fuzzy keyword or multi-keyword, but not synonym based ranking searching including multi-keyword. Propose an efficient multi-keyword synonym query over encrypted cloud data by retrieving top k scored documents. The vector space model and TFIDF model are used to construct index and query generation. The KNN algorithm used to encrypt index and query vectors. Construct a special tree called Balanced M-way Search (BMS) Tree for indexing and propose a Depth First Search Technique (DFST) algorithm to achieve efficient multi-keyword synonym ranked search. The efficiency and accuracy of DFST algorithm are illustrated with an example, BMS tree, it takes sub-linear time complexity.

### *H. A Secure & Dynamic Multi Keyword Ranked Search Scheme over encrypted*

The major aim of this research is to solve the problem of multi-keyword ranked search over encrypted cloud data (MRSE) at the time of protecting exact method wise privacy in the cloud computing concept. Data holders are encouraged to outsource their difficult data management systems from local sites to the business public cloud for large flexibility and financial savings. However for protecting data privacy, sensitive data have to be encrypted before outsourcing, which performs traditional data utilization based on plaintext keyword search. As a result, allowing an encrypted cloud data search service is of supreme significance. In view of the large number of data users and documents in the cloud, it is essential to permit several keywords in the search demand and return documents in the order of their appropriate to these keywords. Similar mechanism on searchable encryption makes center on single keyword search or Boolean keyword search, and rarely sort the search results. In the middle of various multi-keyword semantics, deciding the well-organized similarity measure of "coordinate matching," it means that as many matches as possible, to capture the appropriate data documents to the search query.

### *I. A Survey on Multi-Keyword Ranked Query Search over Encrypted Cloud Storage*

Advancement in cloud computing have revamped the view of modern information technology which is motivating the data owners to outsource their data to the public cloud server like Amazon, Microsoft Azure, Google Drive, etc. With the help of data outsourcing, the organizations can provide reliable data services to their users without any concerns for the data management overhead. One more advantage of outsourcing the data over cloud as SAAS (Storage as a Service) is its cost-effectiveness, scalable and it can be accessed from anywhere and anytime. Normally, CSPs (Cloud Service Providers) take care of the data and its privacy, but there are some of the factors because of which the data privacy and user identity may be violated like an apostate employee, etc. Therefore, data owners should encrypt their respective sensitive data before outsourcing it to the public cloud server. Because the data is getting encrypted before outsourcing which may affect the performance of some important data accessing operations like searching of a document, etc. As know CSPs plays a vital role for data privacy, but is it sufficient for the sensitive data like account figures, budgeting data, photos, health care files, etc. So, to answer the question, there are some methods/solutions offered to provide security and privacy to the data over cloud server. In the survey few of the searching techniques have been studied to find an effective method/solution for the retrieval of data/files over the encrypted cloud data.

With the remunerative option of pay-as-you-use, general and private data are outsourced by many individual users and organizations to third party CSPs. A data owner can outsource their data to the cloud and either can query on that outsourced data or can authenticate a client to perform query. Various domains where searching is performed on outsourced Cloud data are:

*1) Search Engine*

Where a document collection is outsourced to cloud storage and client can retrieve documents which contain the query keywords.

*2) Personalized Medication*

Where patient's medical record is outsourced to hospital's server and an authorized doctor can perform secure searching on patient's medical record for diagnosis.

*3) Email Server*

Where a collection of private emails is outsourced to email server and client can retrieve pertinent emails based on the content of the mail/sender names/receiver names or email IDs.

*4) Crime Investigation*

Where the Interpol's criminal database acts as the server and clients are the authenticated crime investigation agencies like police departments.

The data that is being outsourced may or may not be sensitive. Some of the sensitive data might be like patient's medical records, financial data, etc. So, outsourcing plain data will raises some privacy issues. The data owner cannot afford to leak the private data to the CSPs or any unauthorized party. So, for such data owners, their data needs to be encrypted before outsourcing to the third party CSPs.

### *J. Privacy-Preserving Of Encrypted Cloud Data through Dynamic Multi-Keyword Search*

Recently, more and more people are interested to outsource their local data to public cloud servers for great convenience and reduced costs in data management and security. But in fact of consideration privacy issues, sensitive data should be encrypted here before outsourcing, which obsoletes traditional data utilization like keyword-based document retrieval policy. Here, present a secure and efficient multi-keyword ranked search scheme over encrypted data, which supports dynamic update operations like deletion and insertion of documents and security supports. Specifically, construct an index tree based approach on vector space model to provide multi-keyword search, which meanwhile supports flexible update operations. Here cosine similarity measure is utilized to support accurate ranking for search result. To improve effectiveness of search efficiency, propose a search algorithm based on "Greedy Depth-first Traverse Strategy". Moreover, to protect the search privacy, propose a secure scheme of various privacy requirements in the known cipher text threat model.

## III. METHODOLOGY

### *A. Efficient Secure Ranked keyword search Algorithms over outsource cloud data (ESRK) Method*

The research retrieve the documents based on broader conceptual entities, which will improve the efficiency of ranked keyword search. Traditional searchable encryption schemes allow a user to securely search over encrypted data through keywords without first decrypting it, these techniques support only conventional Boolean keyword search, without capturing any relevance of the files in the search result. When directly applied in large collaborative data outsourcing cloud environment, they may suffer from the following two main drawbacks. On the one hand, for each search request, users without pre-knowledge of the encrypted cloud data have to go through every retrieved file in order to find ones most matching their interest, which demands possibly large amount of post processing overhead. On the other hand, invariably sending back all files solely based on presence/absence of the keyword further incurs large unnecessary network traffic, which is absolutely undesirable in today's pay-as-you-use cloud paradigm. Consumers will be able to access applications and data from a "Cloud" anywhere in the world on demand. The consumers are assured that the Cloud infrastructure is very robust and will always be available at any time. Computing services need to be highly reliable, scalable, and autonomic to support ubiquitous access, dynamic discovery.

To make possible ranked searchable symmetric encryption for successful employment of outsourced cloud data under the mentioned model. Our system design should achieve the following security and performance assurance. Specifically have to reduce the size of index. A list of standard IR techniques can be adopted, including case folding, stemming, and stop words etc. Omit this process of keyword extraction and refinement and refer readers to for more details Ranked search. In order to rank the documents, a ranking function is required, which assigns relevancy scores to each document matching to a given search query. One of the most widely used metrics in information retrieval is the term frequency. Term frequency is denied as the number of times a keyword appears in a document. Instead of using term frequency itself, assign relevancy levels based on the term frequencies of keywords. To enable ranked keyword search for effective utilization of outsourced cloud data under the model, our system design should achieve the following security and performance guarantee.

*1) Ranked key word search*

For efficient searching process the process use the mechanism of Topic detection and tracking. The search time includes fetching the posting list in the index, decrypting, and rank ordering each entry.

*2) Security guarantee*

For providing the security in the cloud server, this process uses the privilege method.

Algorithm:

1) Step 1. Read the document F
2) Step 2. Segment the document term wise and encrypt with key
3) Step 3. Calculate term frequency (TF) and inverse document frequency (IDF) and publishing time (PT)
4) Step 4. Generate index table (table) and files upload to server.

*3) Algorithm for index table generation:*

```
For   all documents Ri do
Compare(level1 index of Ri , query index)
j = 1
while match do
increment j
Compare (levelj indices of Ri, query index)
end while
rank of Ri = highest level that match with
query index
end
```

*4) Algorithm for Ranked Search:*

A searching method to improve the efficiency of ranked keyword search Algorithms. Gave introduction about the existing searchable encryption framework, it is very inefficient to achieve efficient ranked search. Proposed efficient one-to many order preserving mapping function, which allows the effective RSSE to be designed. In additional to that proposed combination of concept based and keyword based searching techniques .This kind of techniques has the ability to categorize, and search large collections of unstructured text on a conceptual basis. This kind of searching technique is more reliable and efficient search method that is more likely to produce relevant results than traditional searches.

*5) Architecture*

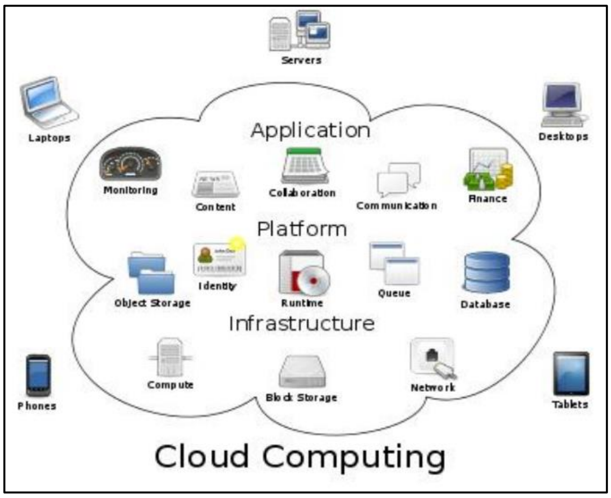

Fig. 3.1: Cloud Computing

*B. Enabling Efficient Fuzzy Keyword Search over Encrypted Data in Cloud (EEFK) Method*

Cloud Computing becomes prevalent, more and more sensitive information are being centralized into the cloud. For the protection of data privacy, sensitive data usually have to be encrypted before outsourcing, which makes effective data utilization a very challenging task. Although traditional searchable encryption schemes allow a user to securely search over encrypted data through keywords and selectively retrieve files of interest, these techniques support only exact keyword search. The first time formalize and solve the problem of effective fuzzy keyword search over encrypted cloud data while maintaining keyword privacy. Fuzzy keyword search greatly enhances system usability by returning the matching files when users' searching inputs exactly match the predefined keywords or the closest possible matching files based on keyword similarity semantics, when exact match fails. In our solution, exploit edit distance to quantify keywords similarity and develop two advanced techniques on constructing fuzzy keyword sets, which achieve optimized storage and representation overheads. Further propose a brand new symbol-based trie-traverse searching scheme, where a multi-way tree structure is built up using symbols transformed from the resulted fuzzy keyword sets.

Algorithm:

```
1) Step 1:   procedure CreateWildcardFuzzySet($w_i$,d)
2) Step 2:   if d >1 then
3) Step 3:   Call CreateWildcardFuzzySet($w_i$,d− 1);
4) Step 4:   end if
5) Step 5:   if d = 0 then
6) Step 6:   Set$S'_{w_i,d} = \{w_i\}$;
7) Step 7:   else
8) Step 8:   for $(k \leftarrow 1 \text{ to } |S'_{w_i,d-1}|)$ do
9) Step 9:   for j ← 1 to 2 ∗|$S_{w}$ ′ i,d$_{-1}$[k]| + 1 do
10) Step 10: if j is odd then
11) Step 11: Set fuzzyword as$S'_{w_i,d-1}[k]$;
12) Step 12: Insert ⋆at position [(j + 1)/2];
13) Step 13: else
```

```
14) Step 14:    Set fuzzyword as$S'_{w_i,d-1}[k]$;
15) Step 15:    Replace ⌊j/2⌋-th character with ⋆;
16) Step 16:    end if
17) Step 17:    if fuzzyword is not in$S'_{w_i,d-1}$ then
18) Step 18:    Set$S'_{w_i,d} = S'_{w_i,d} \cup$ {fuzzyword};
19) Step 19:    end if
20) Step 20:    end for
21) Step 21:    end for
22) Step 22:    end if
23) Step 23:    end procedure
24) Step 24:    end procedure
```

In the above straightforward approach, all the variants of the keywords have to be listed even if an operation is performed at the same position. Based on the above observation, proposed to use an wildcard to denote edit operations at the same position. The wildcard-based fuzzy set of w$_i$with edit distance d is denoted as$S_{w_i,d}$={$S'_{w_i,0}$, $S'_{w_i,1}$, $\cdots$, $S'_{w_i,d}$}, where$S'_{w_i,\tau}$ denotes the set of words$w'_i$ with τ wildcards. Note each wildcard represents an edit operation on w$_i$. The procedure for wildcard-based fuzzy set construction is shown in Algorithm 1. For example, for the keyword CASTLE with the pre-set edit distance 1, its wildcard based fuzzy keyword set can be constructed as SCASTLE,1 = {CASTLE, *CASTLE, *ASTLE, C*ASTLE, C*STLE, ⋯, CASTL*E, CASTL*, CASTLE*}.

The total number of variants on CASTLE constructed in this way is only 13 + 1, instead of 13 × 26 + 1 as in the above exhaustive enumeration approach when the edit distance is set to be 1. Generally, for a given keyword w$_i$with length ℓ, the size of S$_{wi,}$1 will be only 2ℓ + 1 + 1, as compared to (2ℓ + 1) × 26 + 1 obtained in the straightforward approach. The total number of variants on CASTLE constructed in this way is only 13 + 1, instead of 13 × 26 + 1 as in the above exhaustive enumeration approach when the edit distance is set to be 1. Generally, for a given keyword w$_i$with length ℓ, the size of S$_{wi,}$1 will be only 2ℓ + 1 + 1, as compared to (2ℓ + 1) × 26 + 1 obtained in the straightforward approach.

Architecture:

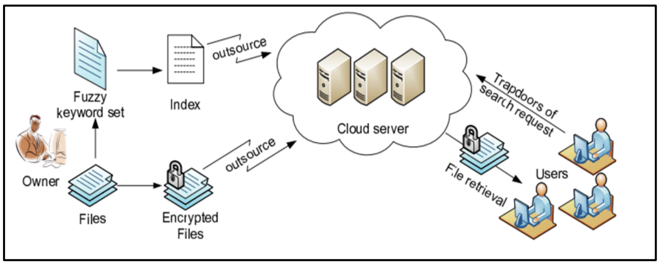

Fig. 3.2: Enabling Efficient Fuzzy Keyword Search

### *C. Efficient Ranked & Secure File Retrieval in Cloud Computing (ERSF) Method*

An important aspect of the cloud services is that user data are stored remotely in unknown machines in which users do not possess or manage. Since the data's are stored remotely, have to keep in mind that sensitive cloud data have to be encrypted before they are outsourced to the commercial public cloud, which makes efficient data utilization service. Searchable encryption file retrieval technique allows users to securely search over encrypted data through search word. Developing an automated system for both named and unnamed documents based on the clustering algorithms. Implement the ranking and searching algorithm to retrieve top k files. Also provides the mapping and encryption algorithm to protect the information. The resulting design is able to provide efficient ranking which will reduce the search time drastically and reduce the communication overhead. The mapping and encryption algorithms protect document against an outside attackers and prevent an untrusted cloud data provider from learning data.

Algorithm:

K-Means Clustering Algorithm:

1) Step 1. Decide K number of clusters.
2) Step 2. Decide K objects arbitrarily as the initial cluster center.
3) Step 3. Repeat

3.1) Allocate each object to their nearby cluster.
3.2) Compute new clusters.

4) Step 4. Until

4.1) No changes in the cluster centers or
4.2) No object changes its cluster.

The K-means clustering algorithm is the efficient algorithm for large data sets. This algorithm divide the set of objects, according to their attributes or features into K-clusters. K is the user defined constant. The main aim is to identify k centroids, one for every cluster. The centroid of a cluster is identified in such a way that it is strongly connected to all objects in the cluster.

This Algorithm takes one keyword at a time. For each keyword, it computes the partial rank/score for each document in the collection C. The final score/rank of each document in the collection C is computed when all the keywords are processed.

Architecture:

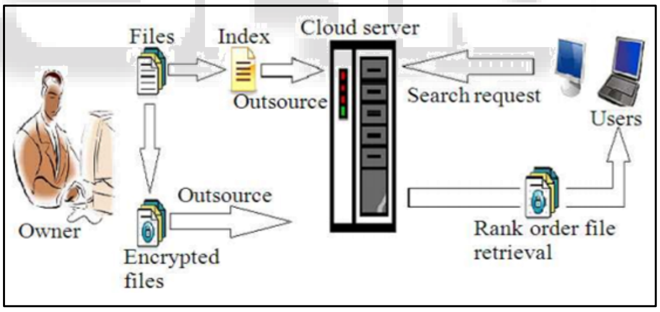

Fig. 3.3: Efficient Ranked & Secure File Retrieval

### *D. A Protected & Lively Multi-Keyword Hierarchical Search Arrangement over Encoded Cloud Data Method*

To enable secure, efficient, accurate and dynamic multi-keyword ranked search over outsourced encrypted cloud data under the models, our system has the following design goals.

*1) Dynamic*

The proposed scheme is designed to provide not only multi-keyword query and accurate result ranking, but also dynamic update on document collections.

*2) Search Efficiency*

The scheme aims to achieve sub- linear search efficiency by exploring a special tree-based index and an efficient search algorithm.

*3) Privacy-preserving*

The scheme is designed to pre- vent the cloud server from learning additional information about the document collection, the index tree, and the query.

The search process of the UDMRS scheme is a recursive procedure upon the tree, named as "Greedy Depth-first Search (GDFS)" algorithm. Construct a result list denoted as RList, whose element is defined as ⟨RScore,FID⟩. Here, the RScore is the relevance score of the document fFID to the query, which is calculated according to Formula. The RList stores the k accessed documents with the largest relevance scores to the query. The elements of the list are ranked in descending order according to the RScore, and will be updated timely during the search process. The "Greedy Depth-first Search" algorithm to provide efficient multi-keyword ranked search.

#### 4) *Algorithm*

Greedy Depth-first Search Algorithm:

```
Step 1: if the node u is not a leaf node then
Step 2: if$RScore(D_u,Q) > k^{th}$score then
Step 3: GDFS(u.hchild);
Step 4: GDFS(u.lchild);
Step 5: else
Step 6: return
Step 7: end if
Step 8: else
Step 9: if$RScore(D_u,Q) > k^{th}$ score then
Step 10: Delete the element with the smallest relevancescore from RList;
Step 11: Insert a new element$\langle RScore(D_u,Q),u.FID\rangle$and sort all the elements of RList;
Step 12: end if
Step 13: return
Step 14: end if.
```

Document Ranking Algorithm:

Require: Query q and document collection C.

```
Step 1: Score [N] = 0:0
Step 2: For all term term t in query q do
Step 3: for all document di in collection C do
Step 4: $w_{it} = tf_{it}$: LogN/$N_t$
Step 5: Score[i] = Score[i] + wit
Step 6: end for
Step 7: End for
Step 8: R = SORT(C; Score [])
Step 9: return R
```

$RScore(D_u,Q)$ – The function to calculate the relevance score for query vector Q and index vector Du stored in node u, which is defined in Formula. $k^{th}$score – The smallest relevance score in current RList, which is initialized as 0. hchild – The child node of a tree node with higher relevance score.lchild – The child node of a tree node with lower relevance score. Since the possible largest relevance score of documents rooted by the node u can be predicted, only a part of the nodes in the tree are accessed during the search process. Fig. 3 shows an example of search process with the document collection F = {fi|i = 1,...,6}, cardinality of the dictionary m = 4, and query vector Q = (0,0.92,0,0.38).

### E. *Architecture*

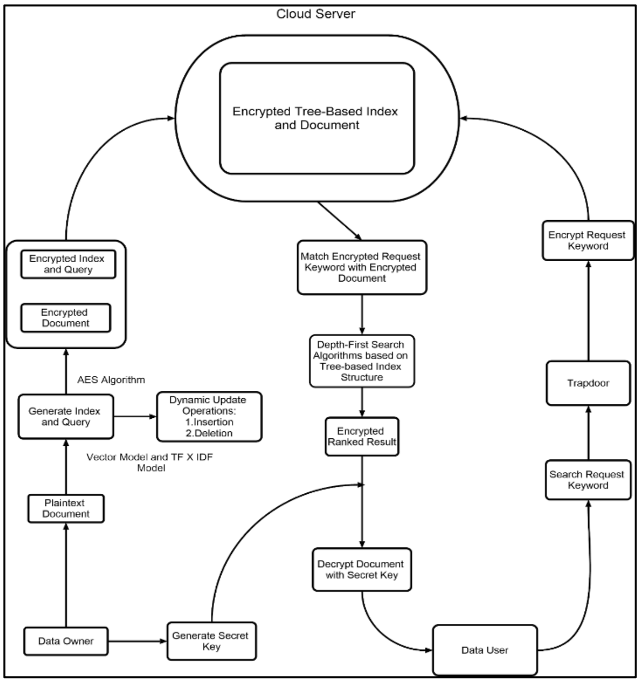
Cloud Server
Encrypted Tree-Based Index and Document
Encrypted Index and Query
Encrypted Document
Match Encrypted Request Keyword with Encrypted Document
Encrypt Request Keyword
Depth-First Search Algorithms based on Tree-based Index Structure
AES Algorithm
Trapdoor
Generate Index and Query
Dynamic Update Operations: 1.Insertion 2.Deletion
Encrypted Ranked Result
Vector Model and TF X IDF Model
Plaintext Document
Search Request Keyword
Decrypt Document with Secret Key
Data Owner
Generate Secret Key
Data User

Fig. 3.4: Unencrypted Dynamic Multi-keyword Ranked Search

## IV. RESULTS & DISCUSSION

### A. *Problem Definition*

In order to address the above problem, the research have designed some general-purpose solutions with fully-homomorphic encryption or oblivious RAMs. 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.

### B. *Existing Method*

Traditional searchable encryption schemes allow a user to securely search over encrypted data through keywords without first decrypting it, these techniques support only conventional Boolean keyword search, without capturing any relevance of the files in the search result. When directly applied in large collaborative data outsourcing environment, they may suffer from the following two main drawbacks. On the one hand, for each search request, users without pre-knowledge of the encrypted data have to go through every retrieved file in order to find ones most matching their interest, which demands possibly large amount of post processing overhead On the other hand, invariably sending back all files solely based on presence/absence of the keyword further incurs large unnecessary network traffic, which is absolutely undesirable in today's pay-as-you-use network paradigm. To mention a few, some of the risks are that the server may leak data information to unauthorized entities or sometimes even the server might be hacked. Therefore, to prevent data privacy and to combat unsolicited accesses, it is necessary that sensitive data have to be encrypted prior to outsourcing. However, such data encryptions render effective data utilization a very challenging task due to the basic reason that there could be a large amount of outsourced data files. Therefore, to prevent data privacy and to combat unsolicited accesses, it is necessary that sensitive data have to be encrypted prior to outsourcing. However, such data encryptions render effective data utilization a very challenging task due to the basic reason that there could be a large amount of outsourced data files.

In the existing system, To improve feasibility and save on the expense in the paradigm, it is preferred to get the retrieval result with the most relevant files that match users' interest instead of all the files, which indicates that the files should be ranked in the order of relevance by users' interest and only the files with the highest relevance are sent back to users. A series of searchable symmetric encryption schemes have been proposed to enable search on .Traditional SSE schemes enable users to securely retrieve the cipher text, but these schemes support only Boolean keyword search, i.e., whether a keyword exists in a file or not, without considering the difference of relevance with the queried keyword of these files in the result. Preventing the network from involving in ranking and entrusting all the work to the user is a natural way to avoid information leakage. However, the limited computational power on the user side and the high computational overhead precludes information security.

### C. Disadvantages

To improve security without sacrificing efficiency, schemes presented in show that they support top-k multi keyword retrieval under various scenarios.

- Authors of made attempts to solve the problem of top-k multi-keyword over encrypted cloud data.
- These schemes, however, suffer from two problems - Boolean representation and how to strike a balance between security and efficiency.

In the former, files are ranked only by the number of retrieved keywords, which impairs search accuracy. In the latter, security is implicitly compromised to tradeoff for efficiency, which is particularly undesirable in security-oriented applications.

### D. Proposed Method

The research propose 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, construct a tree-based index structure and propose a "Greedy Depth-first Search" 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. 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.

### E. Advantages

- The proposed search scheme can flexibly achieve sub-linear search time and deal with the deletion and insertion of documents.
- Provide accurate relevance score calculation between encrypted index and query vectors.
- The proposed "Greedy Depth-first Search" algorithm to obtain better efficiency than linear search.
- The parallel search can be flexibly performed to further reduce the time cost of search process.

## V. Conclusion & Future Scope

The research, 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. 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 (BDMRS and EDMRS) by using the secure kNN algorithm. 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. Experimental results demonstrate the efficiency of our proposed scheme.

### A. Future Scope

The future works, will try to improve the SE scheme to handle these challenge problems. There are still many challenge problems in symmetric SE schemes. 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 secure keys to the unauthorized ones.

## References

[1] Sonam.K.Darda, Prof. (Mrs). Manasi.K.Kulkarni, "Multiuser Multi-Keyword Ranked Search over Encrypted Cloud Using MHR and KP-ABE", International Journal of Computer Science Trends and Technology (IJCST) – Volume 4 Issue 4, Jul - Aug 2016.

[2] Raghavendra S, Geeta C M, "DRSMS: Domain and Range Specific Multi-Keyword Search over Encrypted

Cloud Data", International Journal of Computer Science and Information Security (IJCSIS), Vol. 14, No. 5, May 2016

[3] M.Gomathi, Mr. D.Seenivasan, "Dynamic multi-keyword rank scheme using Top key over encrypted cloud data", International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056, Volume: 03 Issue: 04 | April-2016.

[4] S.SaravanaKumar, C.Periyanayaki, "A Secure and Dynamic Multi-Keyword Ranked Search Scheme Over Encrypted Cloud Data", International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST) Vol. 2, Special Issue 10, March 2016.

[5] K.S.Karthick, P.Deepa, "A Secure and Dynamic Multi-keyword Ranking Search on Encrypted Cloud Data using GDFS", International Journal on Advanced Computer Theory and Engineering (IJACTE), ISSN (Print): 2319-2526, Volume -5, Issue -2, 2016.

[6] Ms. Pradnya H. Unde& Ms. ArtiMohanpurkar, "Fastened Multi keyword Search over Encrypted Cloud Data", Imperial Journal of Interdisciplinary Research (IJIR) Vol-2, Issue-8, 2016.

[7] VeerrajuGampala, Sreelatha Malempati, "An efficient Multi-Keyword Synonym Ranked Query over Encrypted Cloud Data using BMS Tree", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, No 1, 2016.

[8] A.Raghavendra Praveen Kumar, K.Tarakesh, "A Secure and Dynamic Multi Keyword Ranked Search Scheme over encrypted", International Journal of Computer Engineering In Research Trends Volume 2, Issue 12, December-2015.

[9] Jaikishan Tindwani, Aruna Gupta, "A Survey on Multi-Keyword Ranked Query Search over Encrypted Cloud Storage", International Journal of Science and Research (IJSR), Volume 4 Issue 11, November 2015.

[10] Miss. SuvarnaDandekar, Miss. SunitaKhamkar, "Privacy-Preserving Of Encrypted Cloud Data through Dynamic Multi-Keyword Search", IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 10, October 2015.

[11] Singhal, "Modern Information Retrieval A Brief Overview" , IEEE Data Eng. Bull., vol. 24, no. 4, pp. 35-43, Mar. 2001.

[12] I.H. Witten, A. Moffat, and T.C. Bell, "Managing Gigabytes Compressing and Indexing Documents and Images", Morgan Kaufmann Publishing, May 1999.

[13] D. Song, D. Wagner, and A. Perrig, "Practical Techniques for Searches on Encrypted Data", Proc. IEEE SympSecurity and Privacy, 2000.

[14] E.-J. Goh, "Secure Indexes Cryptology reprint Archive", IEEE international conference. 2003.

[15] Y.-C. Chang and M. Mitzenmacher, "Privacy Preserving Keyword Searches on Remote Encrypted Data", Proc. Third Int'l Conf. Applied Cryptography and Network Security, 2005.

[16] R. Curtmola, J.A. Garay, S. Kamara, and R. Ostrovsky, "Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions", Proc. 13th ACM Conf. Computer and Comm. Security (CCS '06), 2006.

[17] D. Boneh, G.D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public Key Encryption with Keyword Search" , Proc. Int'l Conf. Theory and Applications of Cryptographic Techniques (EUROCRYPT), 2004.

[18] M. Bellare, A. Boldyreva, and A. ONeill, "Deterministic and Efficiently Searchable Encryption", Proc. 27th Ann. Int'l Cryptology Conf. Advances in Cryptology (CRYPTO '07), 2007.

[19] M. Abdalla, M. Bellare, D. Catalano, E. Kiltz, "Searchable Encryption Revisited: Consistency Properties, Relation to Anonymous Abe, and Extensions", vol. 21, no. 3, pp. 350-391, 2008.

[20] 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.

[21] L.M. Vaquero, L. Rodero-Merino, J. Caceres, and M.Lindner, "A Break in the Clouds: Towards a Cloud Definition", ACM SIGCOMM Computer Common. Rev., vol. 39, no. 1, pp. 50-55, 2009.

[22] N. Cao, S. Yu, Z. Yang, W. Lou, and Y. Hou, "LT Codes-Based Secure and Reliable Cloud Storage Service" , Proc. IEEE INFOCOM, pp. 693- 701, 2012.

[23] S. Kamara and K. Lauter, "Cryptographic Cloud Storage" , Proc. 14th Int'l Conf. Financial Cryptography and Data Security, Jan. 2010.

[24] J. Katz, A. Sahai, and B. Waters, "Predicate encryption supporting disjunctions, polynomial equations, and inner products", in Proc. of EUROCRYPT, 2008.

[25] Lewko, T. Okamoto, A. Sahai, K. Takashima, and B. Waters, "Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption", in Proc. of EUROCRYPT, 2010.