UGC Approved Journal

IJIREEICE

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified

Vol. 5, Issue 8, August 2017

A Ranked Index-Tree Multi-keyword Search Method for Encoded Cloud Data

Soumyashree Malligeppagol[1], Prof. Vivekanandreddy[2]

PG Scholar, Computer Science and Engineering, Visvesvaraya Technological University, Belagavi, India[1]

Faculty, Computer Science and Engineering, Visvesvaraya Technological University, Belagavi, India[2]

Abstract: Cloud computing is becoming increasingly famous now a days because of its advantages like adaptability, accessibility, and on-demand access for network. These tempting qualities of cloud computing making organizations and people to outsource their sensitive and huge amount information to the cloud instead of maintaining external costly storage disks. In spite of the fact that there are many advantages there is always been security issue when out sourcing data in the form of plain text. This security issue actuates information proprietors to encode information and outsource to the cloud. By encoding information enhances the information security however the information effectiveness is reduced because of the fact that seeking on scrambled information is troublesome. To overcome from this issue, proposing a rated tree structured index-tree and allowing system to follow the Greedy depth first search over index tree for fetching files which are matched with given multi-keyword input query.

Keywords: Multi-keyword, Term frequency, Inverse document frequency.

I. INTRODUCTION

In spite of the rich benefits of cloud computing, for privacy concerns, individuals and enterprise users are hesitant to outsource their sensitive data, including emails, personal health records and organization’s confidential files, to the cloud. Because once delicate data are outsourced to a remote cloud, the corresponding data owners lose directly control of these data. Cloud service providers can have access to this sensitive data and misuse it. However, these techniques don't shield owner’s information from the CSP, since the CSP has full control of owner’s information, Cloud equipment and programming. Encoding of sensitive data before outsourcing can reserve information privacy against CSP. However, data encoding makes cloud data employment a very challenging task.

In this paper, we depict and tackle the issue of rated multi-keyword search encoded cloud information while preserving privacy in the cloud computing concept same as precise framework. Multi-keyword which select various matches as possible of data documents to the search query. We present a search scheme on encrypted cloud data which is secure is in the tree form, which allows user to play multi-keyword ranked search on encrypted file documents on cloud. For this vector space model is used where term frequency (TF) x inverse document frequency (IDF) together contributing in the formation of index and query to support multi-keyword ranked search.

We build an index based on the term frequency, and we use greedy-depth-first search (GDF search) which gives efficient multi-keyword ranked search over the index tree. We are encrypting this index tree and query vector by using KNN algorithm. To upgrade the secure multi-keyword search from basic multi-keyword ranked search, we proposing enhanced multi-keyword ranked search. And also proposing ways for owner to dynamic updates such as update and deletion of the file from the cloud.

II. RELATED WORK

K REN et al [1] In this paper the authors worked on very severe security challenges that may be faced when outsourcing sensitive data to the cloud such as computation outsourcing security, Data outsourcing security, Access control, Trustworthy service metering, Security overhead and many. By learning these a investigation can be done to surpass these problems. They listed some challenges can appear while outsourcing data to in reality are summarized below.

Mehmet Kuzu et al. [4]. In this paper, the author suggest a scheme for similarity seek over encrypted information. To accomplish that, we utilize a state-of- theart set of rules for immediate near neighbor seek in excessive dimensional spaces called locality touchy hashing. To make sure the confidentiality of the touchy statistics, we provide a rigorous security definition and prove the security of the proposed scheme beneath the supplied definition. Further, we provide a real global utility of the proposed scheme and affirm the theoretical results with empirical observations on a real dataset.

DOI 10.17148/IJIREEICE.2017.5808

wei Zhang et al. [9] For privacy issues, comfortable searches over encrypted cloud records encouraged several researches underneath the single owner model. However, maximum cloud servers in practice do no longer just serve one proprietor, instead, they assist more than one proprietors to share the benefits brought through cloud servers. In this paper, the authors suggested schemes to deal with secure multi-key-word ranked search in a multi-owner model. To enable cloud servers to perform secure seek without knowing the actual information of both key phrases and trapdoors. They systematically assemble a novel at ease search protocol. To rank the search results and hold the privacy of relevance scores among keywords and files, they advised a singular Additive Order and privacy preserving feature family. Vast experiments on actual-global datasets affirm the efficacy and performance of their proposed schemes.

III. SYSTEM DESIGN AND IMPLEMEMNTATION

The Architectural model entails 3 extraordinary entities: Data owner, cloud server and data user, as illustrated in Fig 1.

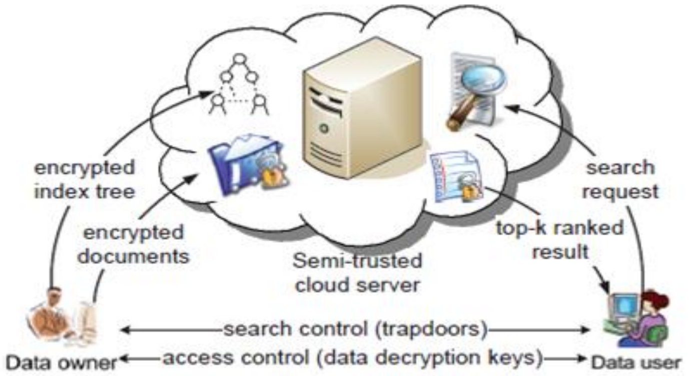


Fig. 1 A ranked index-tree multi-keyword query search scheme over encrypted data

A. Modules description:

DATA OWNER

Owner of the data has a collection of documents F = {f1; f2;…..; fn} that he wants to outsource to the cloud server in encrypted form at the same time as still maintaining the functionality to go looking on them for effective usage. In our scheme, the owner builds a cozy searchable tree index I from report series F, after which generates an encrypted file series C for F. Afterwards, he outsources the encrypted collection C and the relaxed index I to the cloud server, and securely distributes the important thing records of trapdoor generation (including keyword IDF values) and document decryption to the authorized users of data. Beyond this, the owner of the data is answerable for the update operation of his documents stored inside the cloud server. For updating, he generates the replace information regionally and sends it to the server. The owner of the data is also allowed to remove any files belongs to him locally, we call it as deletion of file.

DATA USER

Data users are registered and legal ones to have access to the documents of owner of the data. The authorized user make query to the server by t key-words. With these, authorized person can generate a trapdoor TD according to go looking control mechanisms to fetch k encrypted documents from cloud server. Then, the data user can decrypt the documents with the shared secret key and view the documents.

CLOUD SERVER

Cloud server is able to accumulate the encrypted document collection C and the encrypted searchable tree index I for owner of the data. Upon receiving the trapdoor TD from the statistics person, the cloud server executes search over the index tree I, and finally returns the corresponding collection of top-k ranked encrypted files. Except, upon receiving the replace facts from the information owner, the server wishes to update the index I and record collection C consistent with the received statistics.

B. PROPOSED SCHEME

1.GENERATION OF INDEX TREE

On this phase, we talk about the scheme primarily based on unencrypted and dynamic ranked multi-keyword search (UADRMS) scheme, two cozy search schemes are built further which are enhanced form of this scheme.

DOI 10.17148/IJIREEICE.2017.5808

Firstly we generate a node of tree for every document within the collection. These nodes are made as leaf nodes of the index tree. Then, the inner tree nodes are created based on those leaf nodes. The formal generation process of the index is depicted in algorithm Build-Index algorithm used in earlier schemes.

a. Enhanced dynamic ranked multi-keyword search (EDRMS)
The safety evaluation above shows that the BDRMS scheme can protect the Index Confidentiality and question Confidentiality in the recognized cipher text model. However, the cloud server is able to hyperlink the identical seek requests by way of tracking direction of visited nodes. In addition, in the regarded history version, it is miles possible for the cloud server to discover a key-word as the normalized TF distribution of the key-word can be precisely acquired from the very last calculated relevance rankings. The primary reason is that the relevance rating calculated from Iu and TD is precisely same to that from Du and Q. A heuristic technique to in addition improve the safety is to interrupt such genuine equality. As a result, we can introduce some tunable randomness to disturb the relevance score calculation. In addition, to in shape one-of-a-kind customers' alternatives for better correct ranked results or better protected keyword privacy, the randomness are set adjustable

The improved EDRMS scheme is almost the same as BDRMS scheme besides that:

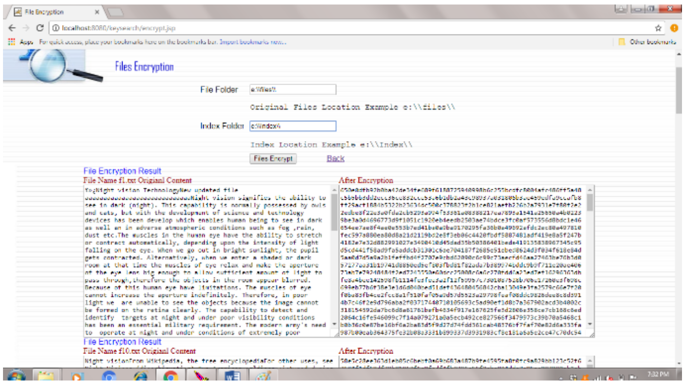
- SK ← Setup() in this set of rules, we set the secret vector S as a m-bit vector, and set M1 and M2 are (m + m′) × (m + m′) invertible matrices, where m′ is the range of phantom terms.
- I ← Gen_Index(F; SK) before encrypting the index vector Du, we make bigger the vector Du to be a (m+m′)-dimensional vector. Each prolonged element Du[m+j], j = 1; :::;m′, is ready as a random number "j .
- TD ← Gen_Trapdoor(Wq; SK) The query vector Q is prolonged to be a (m + m′)-dimensional vector. Many of the extended elements, some of m′′ elements are randomly chosen to set as 1, and the rest are set as zero.
- Relevance_Score ← SRScore(Iu;TD) After the execution of relevance assessment with the aid of cloud server, the very last relevance score for index vector Iu equals to Du · Q +Σ∈v, wherein v ∈ Q[m + j] = 1.

2. SEARCHING FOR FILE USING KEYWORDS
The search process of the UADRMS scheme is a recursive process upon the index tree, called as Greedy Depth first search (GDFS) algorithm. We assemble a result listing denoted as RList, whose element is defined as ⟨RScore; FID⟩. here, the RScore is the relevance score of the report fFID to the query, that's calculated in step with formulation (1). The RList shops the k accessed files with the most important relevance rankings to the query. The factors of the listing are ranked in descending order in step with the RScore, and may be updated well timed for the duration of the hunt technique.

Following are a few different notations, and the GDFS set of rules is defined below.

- l_child – the child node of a tree node with decrease relevance score. for the reason that feasible biggest relevance rating of documents rooted with the aid of the node u can be predicted, best part of the nodes in the tree are accessed for the duration of the search process.
- R_Score(Du;Q) – The feature to calculate the relevance score for query vector Q and index vector Du saved in node u, that's defined in formula (1).
- kth_score – The smallest relevance score in current RList, that is initialized as 0.
- h_child – the kid node of a tree node with higher relevance rating.

IV. EXPERIMENTAL RESULTS


File Encryption
localhost:8080/keysearch/encrypt.jsp
Apps
For quick access, place your bookmarks here on the bookmarks bar. Import bookmarks now...
Other bookmarks

Files Encryption

File Folder e:\\files\\

Original Files Location Example e:\\files\\

Index Folder e:\\index\\

Index Location Example e:\\Index\\

Files Encrypt   Back

File Encryption Result
File Name f1.txt Origianl Content

```
To;Night vision TechnologyNew updated file
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaNight vision signifies the ability to
see in dark (night). This capability is normally possessed by owls
and cats, but with the development of science and technology
devices has been develop which enables human being to see in dark
as well as in adverse atmospheric conditions such as fog ,rain,
dust etc.The muscles in the human eye have the ability to stretch
or contract automatically, depending upon the intensity of light
falling on the eye. When we go out in bright sunlight, the pupil
gets contracted. Alternatively, when we enter a shaded or dark
room at that time the muscles of eye relax and make the aperture
of the eye lens big enough to allow sufficient amount of light to
pass through,therefore the objects in the room appear blurred.
Because of this human eye have limitations. The muscles of eye
cannot increase the aperture indefinitely. Therefore, in poor
light we  are unable to see the objects because the image cannot
be formed on the retina clearly. The capability to detect and
identify  targets at night and under poor visibility conditions
has been an essential military requirement. The modern army's need
to  operate at night and under conditions of extremely poor
```

After Encryption

```
650e8dfb92b0ba42de34fe089f6180725940898b6c255bcdfc8004afc406ff5a40
c5b5bb5dd2ecc36cc852ccc55e6b1db2a45c98937a031805b53ec459edfa9cca5b8
ff29acf1084b5322b23634dc508c378823f2b1ce821aefb226b2a7931e7f80f2e2
2edbe8f22e3a0fda2cb5295a924f5556b1a0858821 7ea7895a1541a25b50a4b0223
9be3a0d4696773d9f1051c1920eb4eedb2503ae74bdce3fc0af573556d8bdc1e46
654ee7ae8f4ee0e553b7ed41ba0a0ba0170295fa36b0a49892efdc2ec80a497810
fec597a880ea80d8a21d2119bd2e8f3eb06c4420fbdf5887481adf419e8a5f247b
4182e7e32d882991027e3400410d45dad35b50386401beda41913583896 7345c95
d5cd441f58ad9fa5addcb31381c6ae704187f2685e51ebed8624d3f034f618e84d
5aa8d7d5a0a2b1feffbd4f2707e0cbd02090c6c99c73aecfd46aa27463be70b3d0
57277aa31b19741d8850ed5ef803fbd81f82ada7b389974bddc9b9f711c20ce406
73eb7e79240484f2ed724355 0e60dcc25000c6a6c270fdd6a21ed7ef162961 63db
fe8a4bee142598fb1114fe5fee5a2f12f5595 7c7b0108 7528b70bc1 7 200e5fb08c
699eb77b6f38e3e1d6d040bed51def436480456842cba13049e3fa2579c66e7f20
f0ba85fb4ce2fcc8a1f510fafe5a0d57d5523a29798feaf08ddc9828dee8c8d391
4b7c46f2e9d796aba2f03717440710105693c5ad90ef1d8b7a367902acd3b4002c
318154592da7bc8d8a81781bafb4834f917e167625fe5d2806a358ca7cb188c8ed
2064c16fe546099c7f14a079271a0a5ecb492ce827566f3479973c39b70a5468c1
b0b36c0e87be16bf6a2ba83d5f9d27d74fdd361cab48176bf7faf70e82d6a333fa
987b08eab364375fe32b08a3331b899337d3931983cf8e181a6a5e2ce47c70dc54
```

File Encryption Result
File Name f10.txt Origianl Content

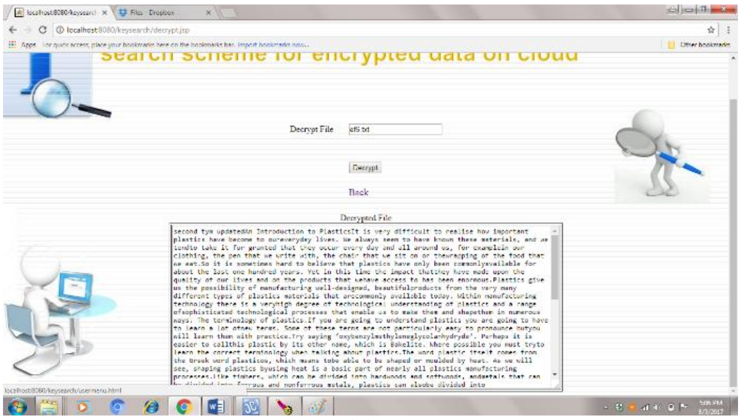
```
Night visionFrom Wikipedia, the free encyclopediaFor other uses, see
```

After Encryption

```
50e5c28ee363d1eb05c6bebf0a69b683a487b9fe4595fa0f0fc9a829bb123c52f6
```

Fig. 2 Encryption of files

DOI 10.17148/IJIREEICE.2017.5808



```
second tym updatedAn Introduction to PlasticsIt is very difficult to realise how important
plastics have become to oureveryday lives. We always seem to have known these materials, and we
tendto take it for granted that they occur every day and all around us, for examplein our
clothing, the pen that we write with, the chair that we sit on or thewrapping of the food that
we eat.So it is sometimes hard to believe that plastics have only been commonlyavailable for
about the last one hundred years. Yet in this time the impact thatthey have made upon the
quality of our lives and on the products that wehave access to has been enormous.Plastics give
us the possibility of manufacturing well-designed, beautifulproducts from the very many
different types of plastics materials that arecommonly available today. Within manufacturing
technology there is a veryhigh degree of technological understanding of plastics and a range
ofsophisticated technological processes that enable us to make them and shapethem in numerous
ways. The terminology of plastics.If you are going to understand plastics you are going to have
to learn a lot ofnew terms. Some of these terms are not particularly easy to pronounce butyou
will learn them with practice.Try saying 'oxybenzylmethyleneglycolanhydryde'. Perhaps it is
easier to callthis plastic by its other name, which is Bakelite. Where possible you must tryto
learn the correct terminology when talking about plastics.The word plastic itself comes from
the Greek word plasticos, which means tobe able to be shaped or moulded by heat. As we will
see, shaping plastics byusing heat is a basic part of nearly all plastics manufacturing
processes.Like timbers, which can be divided into hardwoods and softwoods, andmetals that can
be divided into ferrous and nonferrous metals, plastics can alsobe divided into
```

Fig. 2 Decryption of files and view content

Before outsourcing or downloading the files the user has to be registered. The owner after completing pre-processing things. The files containing sensitive data to be outsourced and index for these files are encoded as shown in fig.2 and outsourced to the cloud. The authorized user first requests for trapdoor information, and query is generated by searching for keyword. At the cloud server, the top k files matching for query keyword are fetched and made available to user. Now the user uses the trapdoor information to decode the files and view information as shown in fig.3.

V. CONCLUSION AND FUTURE SCOPE

In this project, gives a secure way to outsource the sensitive data to the cloud, and a multi-keyword search scheme securely and successfully searches for the top k files containing multi-keywords given by the user and returns the files. Which also performs the dynamic operations such as update and delete of the documents from the cloud.For this work we construct a keyword based ranked index tree and endorse a "greedy depth-first seek" algorithm to achieve higher performance than linear search.

The outsourcing sensitive files are encrypted to be secure. Presently the proposed scheme works on storing and retrieving the encrypted data to/from single cloud. This can be extended to two/many clouds that the replica of original data resides on multiple clouds hence the security level increases.

REFERENCES

[1] K. Ren, C.Wang, Q.Wang et al., "Security challenges for the public cloud," IEEE Internet Computing, vol. 16, no. 1, pp. 69–73, 2012.

[2] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Advances in Cryptology-Eurocrypt 2004. Springer, 2004, pp. 506–522.

[3] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," in INFOCOM, 2010 Proceedings IEEE. IEEE, 2010, pp. 1–5.

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

[5] B. Wang, S. Yu, W. Lou, and Y. T. Hou, "Privacy-preserving multikeyword fuzzy search over encrypted data in the cloud," in IEEE INFOCOM, 2014.

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

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

[8] C. Orencik, M. Kantarcioglu, and E. Savas, "A practical and secure multi-keyword search method over encrypted cloud data," in Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013, pp. 390–397.

[9] W. Zhang, S. Xiao, Y. Lin, T. Zhou, and S. Zhou, "Secure ranked multi-keyword search for multiple data owners in cloud computing," in Dependable Systems and Networks (DSN), 2014 44th Annual IEEE/IFIP International Conference on. IEEE, 2014, pp. 276–286.

DOI 10.17148/IJIREEICE.2017.5808