iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 2 **Issue:** XI **Month of publication:** November 2014

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Effective Multi-keyword Semantics Search Over Encrypted Cloud Data by Using Hash Function

J Kirubakaran[1], P Venkatesan[2], M Yesudoss Winster[3]
[1, 2, 3] Department of CSE, Vel Tech Multitech Dr RR & Dr SR Engg College,Chennai,India

Abstract— ***Today's buzz world cloud computing plays a vital role in storage of the end user data. So we can store large amount of user data such as personal health records, banking information, email data, government documents, personal multimedia data. The major issue in cloud computing is security. So in this paper we use a new technique known as encrypt the data before store it in the cloud. More sensitive information is encrypted before outsourcing to the cloud [1]. In the public cloud has no encryption technique to upload files. In the past few years' personal information are stored in the encrypted form. But accessing those encrypted files will take more time. In order to quick index to required data we use semantic search on encrypted data. Ranking also embed in to the semantic search for quick retrieval of needed information because it will give more relevant data for query data. Hash function also used here for finding similar records in the cloud storage. We assign multiple keywords to single data to index it.***
Keywords— Cloud Computing, Semantics search Sensitive Information, Ranking, and Encryption Technique.

I INTRODUCTION

Cloud computing is a new technology that use a technique virtualization on the single hardware it leads to platform independent on the single hardware storage. We can run multiple operating system on a single memory space. In the past recent years usage of cloud storage is rapidly increased in IT industries because of reduction in cost, efficiency, flexibility and free cost for storage and retrieving. Cloud has some cloud models such as infrastructure as a service, platform as a service, software as a service. Based on user requirement types of cloud can be defined such as

- *Private cloud (*Cloud infrastructure is operated solely of an organization it may be managed by organization third party also)
- *Public cloud (*Public cloud can accesses any user without using any encryption techniques.)
- *Community cloud (*Cloud infrastructure can shared more than two organizations and they have an common cloud requirements)
- *Hybrid cloud (*Combinations of all clouds like private, public cloud and community cloud) [2].

Cloud can also provide many services. Based on the user needs cloud computing provide many services such as

A. Infrastructure As a Service (IASS)

Infrastructure as a service provides physical or virtual machines a hypervisor such as xeon, oracle virtual box. By using VMM (virtual machine monitor) can run multiple operating systems on the same hardware. Hardware cost can be reduced at the same time resource utilisation can be high..one of the good example of IAAS is Amazon web service(AWS).

B. Platform As a Service (PASS)

Platform as a service allows consumer to create software applications using tools supplied by the provider. Google app is one of the platforms as a service. PASS offers operating system, database management system, storage, network access, server software.

C. Software As a Service (SASS)

In this service user can directly view their documents in on-line without usage of the software that is software is given to the user as a service. This can be achieved by using internet [3].

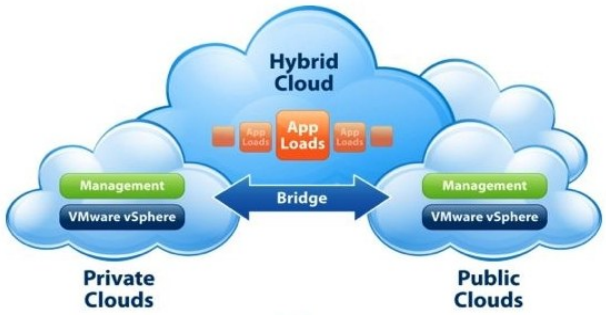


Fig. 1 Depict the different types of cloud services and their tools.

II. PROPOSED SYSTEM

A. Secure multi-keyword ranked search over cloud data

Developing a private cloud is very expensive. Storing of sensitive data in public cloud is very risky. To make it as a possible we should have to avoid unauthorized access to storing the data in encrypted format. In this paper we propose an multi-keyword ranked search by returning matching files from searchable index before upload the file to cloud server. The statistical approach information retrieval (IR) and text mining analyses the text word by word.

B. Rank function

In information retrieval a ranking function is usually used evaluate relevant scores of matching files to a request. There are lot of ranking functions are available among those ranking functions "TF*IDF" rule is most widely used where TF (term frequency) denotes the occurrence of the terms appearing in the document and IDF (inverse document frequency) is often obtained by dividing the total number of documents by the no files containing the term .That the TF represents the importance of the term in the document and IDF indicates the importance or degree of distinction in the whole document [4].

C. Indexing

Purpose of indexing is to optimize the code and performance in finding an relevant documents for a given search query without indexing would scan all document in the cloud. We can also use inverted indexing structure to index data because it is highly used and very fast retrievable indexing technique. We can also use multi-dimensional technique to index it.

There are many multi-dimensional indexing techniques are available. The existing popular multi-dimensional indexing techniques are Bucketing algorithm, k-d tree, priority kd-tree, quad-tree, K-D-B tree, hB-tree, R-tree and its variants R^{+}-tree and R^{*}-tree. Among those techniques BA-KD-tree gave the best performance

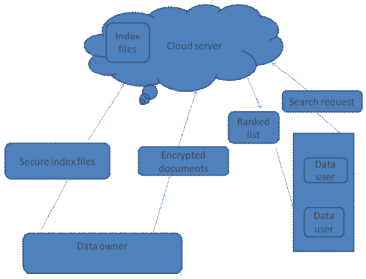


Fig. 2 Frame work of cloud server.

D. *Architecture has three entities*:

1) *Data owner:* Data owner having number files that he/she wants to upload into the cloud server in encrypted format, this will increase the effectiveness.
2) *Data user:* When the owner or user wants to search required files enters a keyword in a secret form.
3) *Cloud server:* In this cloud server have bulk amount of data are to stored previously user can search data files in the form of Boolean search and plaintext search method of downloading of all the files and then retrieving is not practical. In order to avoid this problem we introduce an effective multi-keyword semantics search here symmetric encryption is used to secure files .Same key is used for both encryption and decryption .Cloud data consisting of data owner, data user and cloud

server. Collections of encrypted data files C= (F1,F2, FN) stored in the cloud server. Set of keywords W= (W1, W2, W3) cloud service provider provide service to authorized user only.

Cloud server is responsible for matching searching request to a set of related files [5].

Step1 Data owner registered his details
Step 2 Enter into cloud server
Step 3 Meta data constructed for each file
Step 4 Upload user data into cloud server
Step 5 Search data using keyword
Step 6 Return semantically matched files

III. SEMANTICS SEARCH

In this paper we use a similar search solution based on user queries corresponding meta data constructed for each file are uploaded to the cloud server using cloud meta dataset cloud server build an secure inverted index. similarity can be measured y according to Robert and spark jones [6].

$$W = \log (r / R) / (n/N)$$

N=number of documents in the collections
R=number of relevant documents for query
n=number of documents having terms t
r=number of relevant documents having term t

A. Benefits of semantics search

Semantics search find the conceptual meaning of the query rather than find an literal meaning of the keywords several advantages are tenses and plural forms, synonyms with correct meaning, generalization, concept matching, knowledge matching [7].

IV. HASH FUNCTION

Hash functions allow us to verify the input data matches stored in the values. Return value of the hash function can be called as hash value, hash codes, hash sums. Hash function hash many advantages such as finding duplicate data in the records, protecting data, finding similar records and finding similar substrings in the records. We can use the hash function in many applications such as authentication, message integrity, message finger printing, data corruption detection and digital signature efficiency [8].

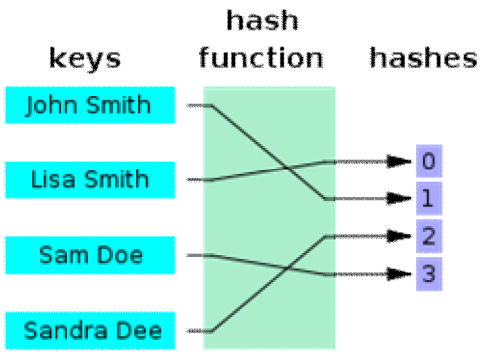


Fig 3 Show the perfect hash function used for indexing the encrypted data [9].

In this paper we are embed the perfect hash function to index the query data. Perfect hash function has the property of injective (i.e. one to one function mapping between domain and co-domain). Figure 3 describe the perfect hashing technique that one to one map between keys and hashes. With such a function one can directly locate the desired entry in a hash table, without any additional searching. The main advantage of the using perfect hashing is reducing the duplication of similar records. And it can directly index required data without searching of all the files in memory [10].

V. CONCLUSION

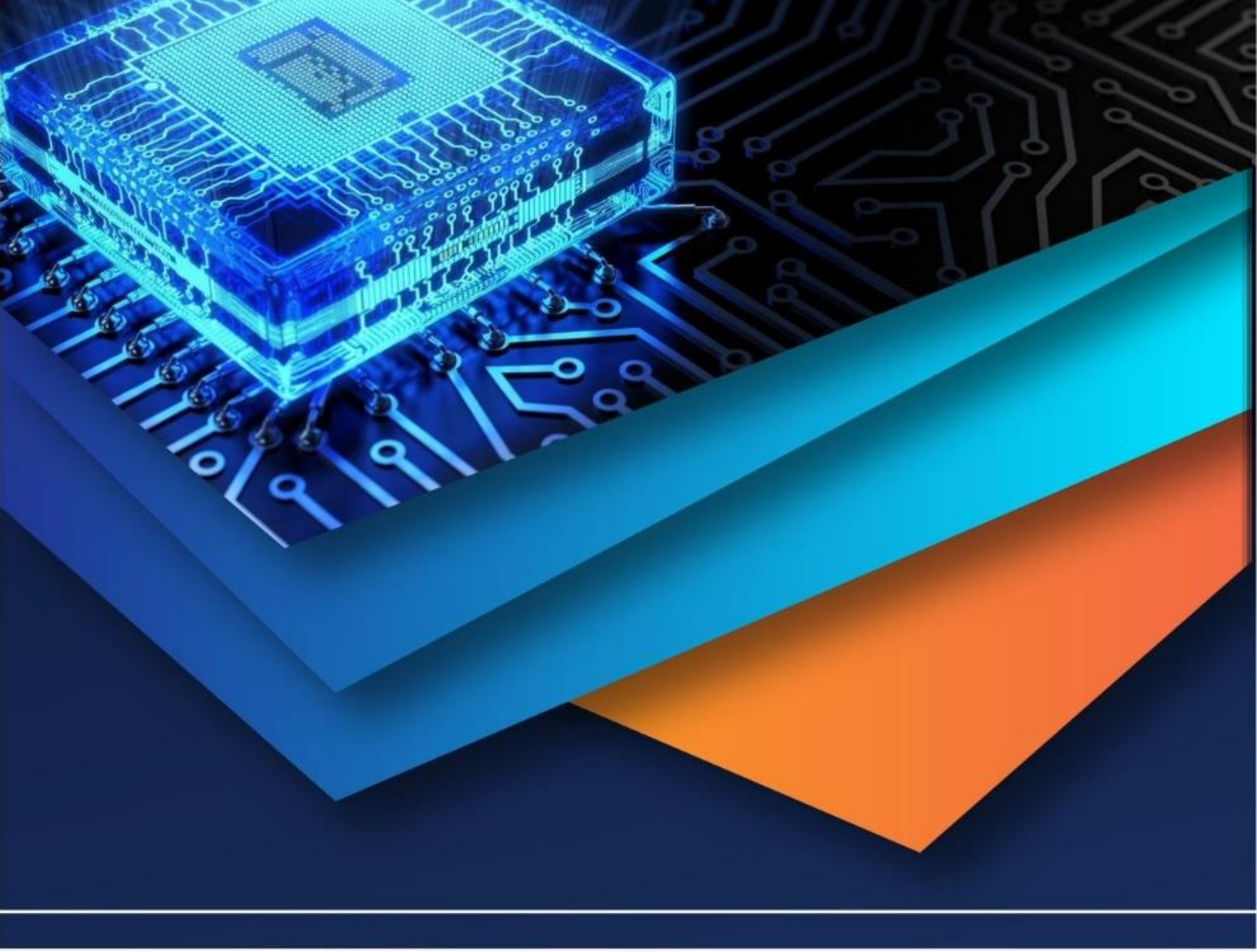
In this paper, we propose a searching method over the cloud data to improve the efficiency of multi-keyword ranked search by using symmetric encryption .Future work secure semantics search enhanced security and a single data may has many keywords so keyword collision will occur. Further work of this paper to minimizes the keyword collision to index the encrypted data. Our proposed system greatly improves the efficiency of multi-keyword ranked search.

ACKNOWLEDGMENT

I would like to thank my institution to allow me to work on this domain and my project guide Mr.P.Venkatesan M.Tech..,for giving the technical and moral support to do this work.

REFERENCES

[1] Z hangjie Fu, Xingming Sun, Nigel lingie and Lu Zhou Achieving effective cloud service:multikeyword ranked search over encrypted cloud data. IEEE Trans.Consumer electronics, vol. 60,No. 1,February 2014
[2] Ms Mayura R.Girme and prof.G.M. Bhandari. Efficient Secure Ranked Search Algorithms over outsourced Cloud data. IJETICS Vol 2, Issue 5, September- October 2013
[3] Mahesh Lanjeswar , Swapnali Ghadge, Sneha Mane, Priti Dalvi. Fuzzy Keyword Search Over Encrypted Data Using Cloud Computing Vol 2, Issue 2,Mar-Apr 2012,pp.870-874
[4] Ning Cao, Cong Wang, Ming Li, Kui Ren and Wenjing Lou . Privacy-preserving multi-keyword ranked search over encrypted cloud data. IEEE Trans on Parallel and Distributed Systems, vol.25, No 1, January 2014
[5] Jun Xu, Weiming Zhang, Ce Yang Jiajia Xu and Nenghai Yu. Two-step ranking secure multi-keyword search over encrypted cloud data. 2012 International Conference On Cloud Computing and Service Computing
[6] Fabrizio Lamberti, Andrea Sanna, and Claudio Demartini.A relation -based page rank algorithm for semantic web search engines.IEEE Trans on Knowledge and Data engineering , vol.21, NO.1, January 2009
[7] Peter Mika, Glovanni Tummarello. Web semantics in clouds. Pulished By IEEE computer society
[8] Cong Wang, Ning Cao,Jin Li,Kui Ren,and Wenjing Lou. secure ranked keyword search over encrypted cloud data. 2010 International computer on distributed computing systems
[9] Amith sheth.Knoe.sis,Semantics scale up beyond search in web 3.0. Published by IEEE computer society .November/December 2011
[10] Cong Wang, Ning Cao,Jin Li,Kui Ren,and Wenjing Lou Enabling secure and efficient ranked keyword search over outsourced cloud data. IEEE Trans on Parallel and Distributed Systems. Vol.23, No.8, August 2012



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY