An Interactive hunt:Semantics application over cloud

Nandini M. Ukey, Prof. P.P.Shelke

PG Scholar, Assistant Professor
Department of Computer Science and Engineering
Government College Of Engineering,Amravati,India.

Abstract **- It has become a major part of our life rather than carrying huge data files or even our data storage devices just attached those files on cloud applications and access from anywhere anytime. Then problem arises of not only data security also data accessibility. Keyword based search has become common practice for users but to get quick and accurate response semantics are applied. In this paper we build a cloud repository where people will upload their repository and store documents securely. Penetrating is done using semantics and even we use RAKE(rapid automatic keyword extraction) this technique provide key phrase extraction so as to quicker our response while examining. We need to explore and regain all the documents which are semantically interrelated to user commanded query along with appropriate storage and secured contact.**

The algorithms and embattled experiments of cloud security generating spontaneous results and future work will use more advanced algorithms with change in day to day technology.

Keywords: Semantic search, RAKE,cloud security.

I. INTRODUCTION

1.1 MOTIVATION

Cloud Computing enables cloud patrons to tenuously store their facts into the cloud so as to relish the on-request high quality offers and facilities from a cooperative group of configurable calculating returns [1]. The welfares brought by this new computing ideal include but are not partial to: relief of the load for filling administration, universal data access with self-regulating geographical locations, and evasion of capital outlay on hardware, software, and personnel maintenances, etc [2].

From years users are habitual to carry out keyword based search techniques ,to make users aware of semantic search by inputting keyword query as background processing is of not much importance to them. To adapt Keyword query to semantic search we need to overcome some difficulties:

Lex Gap:Web users didn't got any knowledge regarding background processing as the words of query are different which are understandable to human from those in ontology which are quite different.

Lack of relativity: relations between the concepts ,objects are required to be stated in query accurately as users to lack in its query formation.

Query ranking:As multiple queries can be formed using single keyword and their results will be different .Hence their ranking will be different the nearest possible ones vary each other with certain differences due to their content as semantics deal with meaning of query.

Then regarding mapping which is done semantically for which it utilizes general dictionaries to find semantically relevant words matched against the Knowledge base using different mapping techniques. To map directly using keyword is having higher value than that of synonym based matching. After mapping the words no longer tends to be lexical strings but become a meaningful terms and tends to what actual user wants. The mapped resources are splits into different query sets. This resources are constructed into graphs or hierarchy is formed.

Semantic search main focused is on:

1. Low Barrier for accessing to end users: To overcome the knowledge in the air and ensure that conjoint users can get valid results without having much knowledge about vocabulary or relation of ontologies or to learn any special query language.
2. To deal with complex queries and provide comprehensive means to handle .To provide quick results using semantic technology is prime benefit for users including accurate results .
3. Ordinary end users can understand the results without consulting backend semantic repositories or their underlying technologies which they too aren't aware of.

To conduct semantic analysis on the language expressions and achieve identicals between verbal expressions at semantic level. Semantic matching a technique for verifying information related to semantically related concepts like in 2-graph , 2-tree structures eg. XML schemas , taxonomies database, classifying ontologies etc like if nodes match each other then by describing through class diagram or entity relationship help to go through the intent somewhat similar. If in one file folder related to automobile or car represemts the same. Matching to query formation with term dependency , translation model , topic model , latent space model. Information is taken from linguistic resources like wordnet .

In recent years, searchable encryption (SE) techniques have been developed for safe contract out data search. Some further studies focus on exploration efficiency, multi-keyword search [11,12], and secure dynamic updating [13]. But they only support precise keyword hunt. To enhance the search obedience and usability, some inspection has been done on fuzzy keyword search [14-18]. These results delivery acceptance of minor typos and presentation variations, such as, search for "million" by carelessly typing "milion", or "datawarehouse" by typing "data-warehouse". These schemes mainly take the design of terms into reflection and use edit distance to estimate the likeness.

Semantic searching means:

1. Semantic Deepness: semantic penetration of the concept X, ( ) Penetration X , means the length of the path from the concept X to the root, and the semantic penetration of the root is demarcated as 0, which means ( ) 0 =Penetration root .
2. Semantic Intersection : Semantic Intersection means the sum of the same hyponymy perceptions of two ontology ideas, which shows the resemblance degree of the two ontology ideas.
3. Semantic Distance: Semantic Distance means the number of boundaries that link binary ontology notions in the domain ontology.

Data incorporation and association element summaries for problem arisen from the acquaintance gaining process, as varied sources bring out different evidence. Search services offer abstract sample of the functionality a semantic explore engine proposes that are predictable IR methods such as object and links. Lastly result is obtainable in chosen technique as graphical or symbolic view.

We are generating a cloud data storehouse where proprietors will upload there files and users will download securely by getting proper authentication. Users query will generate a trapdoor which provide access mechanism then the results are searched semantically over cloud and ranked output are sent, on top the expected file is there. Reference[15] provide proofs that semantic based search engine is more efficient than cloud based search engine by using TREC style average precision measure.

- Databases with coordination (Scenario 2): The databases have the knowledge of the set of the indices of the stored secret shares in other databases by coordination.

II. LITERATURE SURVEY

References which provide details regarding single keyword search , then introduction of multiple keyword search in encrypted form [1][2]. Some of them [1] follows about multi-keyword ranked search over encrypted data in cloud computing against 2 threat model which is called MRSE. This supports vector space model (VSM) and secure inner product to release the high efficiency of search .References generates [1][2] search index with term frequency and vector space model and chooses cosine similarity to compare the source and the query which can help achieve more accurate search result . Then come how to return the ranked result through the frequency of keyword access.

Semantic searching use in searching the domain assets to find common requirement in large dataset [13]. Keyword query is inevitably decoded into official logic queries so that end users can use familiar keywords to perform a semantic search. But still is query scope as more keywords required to improve human interaction[14]. Usage of two cloud named one to be private and another public [15]as same our paper is having where keywords values are stored and encrypted index is stored in public cloud , but only the issue with relevancy count which is not done here.

A. *Searching issues*

Mayan [10] proposed hunt based on synonym queries. Identical data vector and query vector products is ended based on that inference are made ,but whose relevancy can't be said. Hang[4] familiarized semantic search process on conceptual graphs over encrypted outsourced data oncloud (SSCG) same encrypted document and query indexes are created. Conceptual graphs are created based on each file its three index vectors are made which are encrypted and directed over the cloud. Deliver more correct results than earlier search mechanism but intense time is taken.

Fu[5] proposed semantic hunt where prolonged concept hierarchy stored semantic relations ,between notions and search terms. Here protected exploration scheme based on asymmetric scalar product preserving encryption scheme(ASPE) is used where secure nearest neighbor end is regained.

B. *Privacy issues*

Rane[9] where keyword selection for queries as well as data fetched should be mysterious to server. Files are encrypted using symmetric key cryptographic algorithm are sent over the cloud and queries too are encrypted while sending and retrieving .While Mayan[10] encrypt using asymmetric encryption algorithm and for examining firmly using knn. References [11] introduce the concept of twofold server public key encryption with keyword search for secure cloud repository and guarding of data and again it was used while semantic searching[12].

III. PROPOSED METHODOLOGY

In this paper we had created a cloud document repository where a user can upload their text just like in Google Drive. Here the user will generate an account and then using login IDs form the repository, upload the documents in that repository. Even it could download the document from its own repository or from others repository for which Cloud Admin provide access mechanism by providing trapdoor key of respective document. Admin is the sole user who give authentication to repository created by user and then only it can be made to use for uploading of documents.

Dual server is created where on one server encrypted documents on other encrypted keywords and indexes are stored. To securely access the documents so as no intruder can enter and get query and document details so query is sent in encrypted by using Base-64 encoding mechanism and then again decrypted on document server where Rapid automatic keyword extraction(RAKE) technique is applied which extract keyphrases and that are stored on keyword server with its indexes .plain document get decrypted using AES encryption mechanism. The reverse mechanism is used on downloading side where user's query is matched and according to the relevancy score of document the desired results are shown . Here semantics are used in searching that how NLP techniques quickly provide responses.

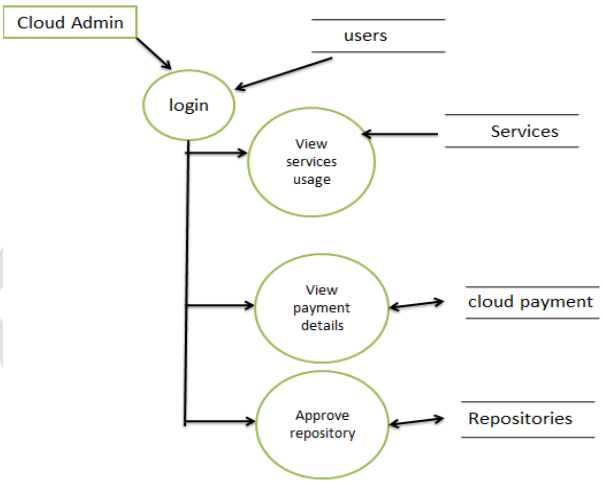

Figure 1.Overview

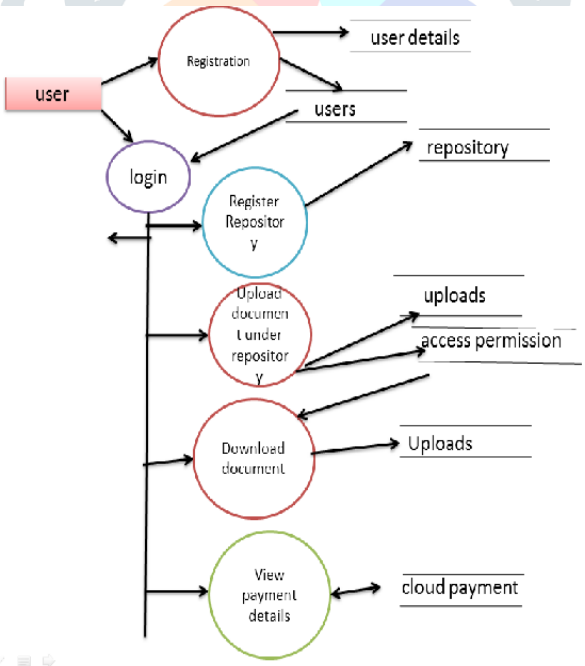

Figure 2. User Details

IV. RESULTS

Semantic hunt method which delivers user's explored query its relative meaning. This come from numerous societies of data retrieval semantic web, machine learning , natural language processing and so on. Addition of scope and advance quality of search. But here we not only increase its accessibility but also security while retrieving results from cloud. Usage of RAKE which improves our search mechanism . Not only searching but security too has been given focused. Here qualitative information is given importance and the relevant results are provided.

REFERENCES

[1] Ning Cao, Member, IEEE, Cong Wang, Member, IEEE, Ming Li, Member, IEEE, Kui Ren, Senior Member, IEEE, and Wenjing Lou, Senior Member, IEEE "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data".
[2] Zhangjie Fu, Member, IEEE, Xingming Sun, Senior Member, IEEE, Nigel Linge, Lu Zhou "Achieving Effective Cloud Search Services: Multi-keyword Ranked Search over Encrypted Cloud Data Supporting Synonym Query".
[3] Wei Zhang, Student Member, IEEE, Yaping Lin, Member, IEEE, Sheng Xiao, Member, IEEE, Jie Wu, Fellow, IEEE, and Siwang Zhou "Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing". DOI 10.1109/TC.2015.2448099, IEEE Transactions on Computers JOURNAL OF L ATEX CLASS FILES, VOL. 6, NO. 1, JANUARY 2015.
[4] Zhangjie Fu, Fengxiao Huang, Xingming Sun, Athanasios V. Vasilakos, and Ching-Nung Yang "Enabling Semantic Search based on Conceptual Graphs over Encrypted Outsourced Data" DOI 10.1109/TSC.2016.2622697, IEEE Transactions on Services Computing.
[5] Zhangjie Fu1,2, Xingming Sun1, Sai Ji1, Guowu Xie3"Towards Efficient Content-aware Search over Encrypted Outsourced Data in Cloud" IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications.
[6] Zhangjie Fu, Member, IEEE, Fengxiao Huang, Kui Ren, Fellow, IEEE, Jian Weng, and Cong Wang, Member, IEEE "Privacy-Preserving Smart Semantic Search Based on Conceptual Graphs Over Encrypted Outsourced Data". IEEE Transactions On Information Forensics And Security, Vol. 12, No. 8, August 2017.
[7] Zhangjie Fu, Member, IEEE, Xinle Wu, Qian Wang, Member, IEEE, Kui Ren, Fellow, IEEE " Enabling Central Keyword-based Semantic Extension Search over Encrypted Outsourced Data" DOI 10.1109/TIFS.2017.2730365, IEEE Transactions on Information Forensics and SecurityIEEE Transactions On Information Forensics And Security.
[8] Zhihua Xia, Member, IEEE, Xinhui Wang, Xingming Sun, Senior Member, IEEE"A Secure and Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data" .
[9] Miss Deepali D.Rane1 Dr.V.R.Ghorpade2 "Multi-User Multi-Keyword Privacy Preserving Ranked Based Search Over Encrypted Cloud Data".
[10] Albert Mayan.J1 , Shaik Arifa2, Pavithra.R3 "Semantic based Multi lexical Ranking Technique for an Effective Search in Protected Cloud "2016 International Conference on Control, Instrumentation, Communication and Computational Technologies.
[11] Rongmao Chen*, Yi Mu*, Senior Member, IEEE, Guomin Yang, Member, IEEE, Fuchun Guo and Xiaofen Wang"Dual-Server Public-Key Encryption with Keyword Search for Secure Cloud Storage".
[12] Zhangjie Fu Lili Xia Xingming Sun Alex X. Liu Guowu Xie" Semantic-aware Searching over Encrypted Data for Cloud Computing"IEEE Transactions on Information Forensics and Security.
[13] Wei Guo, Jian Wang*, Keqing He, Dunhui Yu, Pengfei Du State Key Lab of Software Engineering, Wuhan University, Wuhan 430072, China "Towards Searching Domain Assets Based on Semantic Similarity".
[14] Qi Zhou, Chong Wang, Miao Xiong, Haofen Wang and Yong Yu "SPARK: Adapting Keyword Query to Semantic Search".
[15] "Privacy-preserving Indexing and Query Processing for Secure Dynamic Cloud Storage ",Minxin Du, Qian Wang, Member, IEEE, Meiqi He, and Jian Weng.
[16] Jason W. Woodworth1∗and Mohsen Amini Salehi2∗ 2 High Performance Cloud Computing (HPCC) Laboratory 12 School of Computing and Informatics University of Louisiana at Lafayette, Louisiana, USA" S3BD: Secure Semantic Search over Encrypted Big Data in the Cloud".
[17] "Efficient semantic search over Encrypted Data in Cloud Computing" Teng-Sheng Moh, Department of Computer Science, San Jose State University.