International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 5 | ISSN : 2456-3307

A Survey on Semantic Search Encryption Using Cloud Computing

[1]Dr.R.Shankar, [2]R.Gowriprakash

[1]Assistant Professor, Department of Computer Science, Chikkanna Government Arts College, Tirupur, Tamilnadu, India

[2]M.Phil Research scholar, Department of Computer Science, Chikkanna Government Arts College, Tirupur, Tamilnadu, India

ABSTRACT

As Cloud Computing becomes widely, more and spread sensitive data are being centralized into the cloud. Data in outsourced to the public cloud for economic savings and ease of access. However, the privacy information has to be encrypted to guarantee the security. To implement efficient data utilization, search over encrypted cloud data has been a great challenge. The existing solutions depended entirely on the submitted keyword-based search schemes, and almost all of them depend on predefined keywords extracted in the phases of index construction and query. In view of the deficiency, as an attempt, we propose a semantic expansion based similar search solution over encrypted cloud data. Our solution could return not only the exactly matched files, but also the files including the terms semantically related to the query keyword. In the proposed scheme, a corresponding file metadata is constructed for each file. Then both the encrypted metadata set and file collection are uploaded to the cloud server. Further more, we a basic idea for significantly improved scheme to satisfy the security guarantee of searchable symmetric encryption (SSE).

Keywords : Searchable Symmetric Encryption, Cloud Computing, Web Ontology Language, Fuzzy Concept, GCP, RSL

I. INTRODUCTION

Semantic search should occur while a term is ambiguous, meaning it able to have numerous meanings foe a example , if one considers the lemma "bark", which many be understood as "the sound of a dog," "the skin of a tree," or "a three-masted sailing vessel"), In a structure of this type, every lexical concept coincides therefore with a semantic network node and is related to others with aide of unique semantic relationships in a hierarchical and hereditary structure. during this manner, each construct is enriched with the characteristics and that mean of the near nodes. Each node of the network (called Synset) groups a collection of synonyms that represent the similar lexical concept.

Semantic search used in a methods RDF: the useful resource Description Framework Semantic search seeks to enhance search accuracy by understanding the searcher's intent and therefore contextual meaning phrases as they a appear with in the searchable dataspace, whether on the Web or inside a closed system, to get larger relevant consequence. Semantic search structure consider varied points as well as context of search, location, intent, variation of words, synonyms, generalized and specialized queries, concept matching and natural language queries to provide applicable search results. Major web search engines like Google and Bing incorporate a many elements of semantic search. In vertical search, LinkedIn publishes their semantic search methodology to activity search by way of recognizing and standardizing entities in every queries and documents, e.g., companies, titles and skills, then constructing various entity-award capabilities supported on the entities. semantic search hard and fast techniques for retrieving data from primarily based facts sources like ontologies and XML as discovered on the Semantic Web. Such technologies modify the formal articulation of domain information at a high level of quality and may allow the user to specify their purpose semantic search based on disambiguation cloud be a perceive what a user is searching find, word sense disambiguation (RDF) may be a trendy framework for the way to describe any internet along with a web site and its content material. An RDF description (such descriptions are often time referred as metadata). include the authors of the resource, date of creation or change, the organization of the pages on a site (the sitemap), data that describes content in terms of audience or content rating, key words for search engine data collection, subject classes and so on.

Keywords to concept mapping: The concept map can assist you brainstorm your subject matter and see what principles or keywords to use as you search for statistics. It additionally helps you become response to what you understand regarding your topic, presents you with an chance to consider your subject matter in new strategies and identify gaps on your experience.

The Web Ontology Language(OWL): Cloud also be circle of relatives of data representation languages for authoring ontologies. Ontologies are correct methodology formal method measure to correct justify taxonomies and sort networks, primarily method the form of information for varied domains.

Fuzzy concept: A form of representation suitable for notion that cannot be defined precisely but which depend upon their contexts. Fuzzy examples of : “not clear, distinct (or) precise, blurred”. Such keyword-based search approach permits users to by selectively retrieve document of interest and has been typically implemented in plaintext search scenarios, like as Google search. Unfortunately information encryption restricts user’s capability to perform keyword search and therefore makes the traditional plain-text search techniques improper for Cloud Computing. except this, data encryption also demands the protection of keyword privacy account that key phases some times include essential statistics related to data files. though encryption of keywords will protect keyword privacy, it additionally renders the normal plaintext search techniques useless in though out state of affairs. To securely search over encrypted data, searchable encryption method were advanced developed in recent years [5–13]. Searchable encryption schemes generally build up an index for every keyword of typically an associate the index with the files that incorporate the keyword.

II. RELATED WORK

Proposed scheme are only appropriate for plaintext or for performing searches before encryption. They cannot recognize encrypted searches based on Conceptual graphs. As a result we have to first construct a search scheme in plaintext that can also be easily applied to encrypted information. Simplifying the sentences received in text summarization by Tregex and building the CGs consistent with the simplified sentences. Then we will describe an effective unencrypted semantic search based on conceptual graphs (USSCG) scheme as the basis of the subsequent encrypted scheme. Finally, based on the USSCG and referring to the encrypted scheme We apply order preserving symmetric encryption (OPSE) to our scheme to enhance security.

Another way of proposed scheme in Fuzzy keyword using semantic search encryptions are:

1. Wildcard-primarily based approach

A wildcard is used to edit the operations at the equal function. The edit distance may be calculated the usage of substitution, deletion and insertion.

2. Gram-primarily based approach

Right here the bushy set is built based on grams. The gram of a string is a substring and can be used for powerful approximate search. The order of the characters after the primitive operation is always kept the same before the operations.

3. Symbol-based trie-traversed

In this technique, a multi-way tree is constructed for storing the fuzzy keyword set over a finite symbol set. All the trapdoors sharing a common prefix have common nodes. The fuzzy keyword in the trie can be found by depth first search approach.

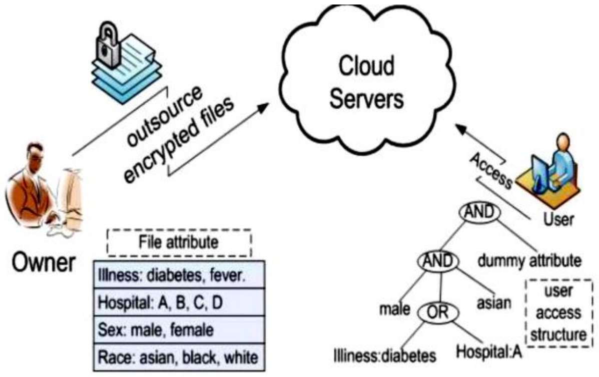

Fig1.Symbol-based trie-traversed scheme

We examine the problem of retaining the near dating between extraordinary undeniable files over an encrypted domain and encrypted a clustering method to solve this problem. We designed the MRSE-HCI structure to hustle up server-side searching phases. Accompanying with the exponential development of document series, the quest time is decreased to a linear time alternatively of exponential time. We layout a seek strategy to increase the rank privacy. Using searching strategy adopts the backtracking algorithm upon the above clustering technique. With the increase of the facts extent, the gain of the proposed approach in rank privacy has a tendency to be extra perceptible. through applying the Merkle hash tree and cryptographic signature to real tree shape, we provide a authentication mechanism to assure the correctness and completeness of search results.

III. SEMANTIC RELATIONSHIP LIABRAY (RSL)

Cloud Data store provides programmatic access to a variety of its information to support meta programming, implementing backend body functions, modification of consistent caching, and similar purposes; you may use it, for example, to create a custom data store viewer for your application. The metadata accessible includes information regarding the entity groups, namespaces, entity varieties, and properties your application uses, thus because of property representations for every property. The Cloud Data store Dashboard with in the GCP Console additional provides some metadata regarding your application, how ever the information displayed there differs in some necessary respects from that came back by these functions.

Freshness: Reading metadata using the API gets current information, whereas data with in the dashboard is updated just only once daily. **Contents:** Some metadata in the dashboard is not on the market via the APIs. the reverse is additionally true.
Speed: Metadata gets and queries are billed in the same way as data store gets and queries.

Meta lexical indicates information that isn't always contained in the lexicon, even though it could be statistics about words. the primary members of the family discussed here are exemplified as follows:
synonymy:sofa=couch=divan=davenport
antonymy:good/bad,existence/lossoflife,come/moe**contrast:**candy/sour/bitter/salty,strong/liquid/gasoil
hyponymy or classfunction inclusion:cat<mammal < animal **meronymy or the part-whole relation:** line<stanza<poem

The equals sign(=) is used to indicate synonymy. The curb (/) among participants of antonym or evaluation units indicates the semantic incompatibility of the contrasting phrases. Antonym is a subtype of assessment, in that it is comparison inside a binary paradigm. while the term antonym is every now and then reserved for more unique family members, it's miles used right here for any binary semantic assessment among lexical gadgets (whereas opposite is used greater extensively right here, not restricted to assessment between linguistic expressions. The 'less than' sign (<) inside the hyponymy and meronymy examples shows that those members of the family are hierarchical and asymmetrical. this is stanza is a meronym of poem, but poem is not a meronym of stanza The speak family members of hyponymy and holonymy can be represented by the 'greater than' signal (>), as a poem >stanza(i.e., 'poem is the holonym of stanza').

IV. META DATA QUREIES

A receiving a query request, the cloud server first finds out the keywords that are semantically related to the query keyword according to SRL. Then both the query keyword and the extensional words are used to retrieve the files. these kind will not conflict with others of the equal names which can exist already to your software. The means of querying on those unique kind, you may retrieve entities containing the desired metadata. The entities returned through metadata queries are generated dynamically, based on the current state of the data store. while you can create local entity object of kind **__namespace__**, **__kind__**, **or __property__**, any try to store them in the records store will fail.

Keyword Search over Encrypted Cloud Data

Unique data structure called Secure File Object (SFO) to modify keyword search over encrypted cloud data. Once a data owner desires to upload a data file to the cloud storage, the client side application will create and attach a SFO to the encrypted data file before uploading to the cloud storage. Every SFO include fact that describes the necessary fact that document. Throughout SFO creation, the client-side application extracts unique keyword from the uploading data file and encrypts them to create list of encrypted keywords that will be stored in the SFO. When a user desires to search for a specific keyword, the user will put up the keyword to statistics owner and the data owner will compute the search capability by encrypting the keyword with equal key that used to generate the listing of encrypted keywords in the SFO. The user will submit the comeback search capability from the data owner to the cloud server. The cloud server can come the encrypted data file if the list of encrypted keywords with in the SFO contains the search functionality.

V. SECURITY ANALYSIS

Secure and privacy preserving efficiency is shown through security analysis and effectiveness is demonstrated through evaluation of experiments. in the proposed used in the related symmetric encryption:-

Order Preserving Symmetric Encryption (OPSE) OPSE may be a cryptographic primitive that helps sort the ranking scores in an encrypted type. However, the distinctive OPSE forever maps the plaintext into the identical random-sized non overlapping interval bucket that result in to some information leakage. During this project, we focus on the change OPSE cites in they designed a one-to-many OPSE that appends files' id into the plaintext and build identical plaintext now not factor to the same cipher text deterministically. The modified OPSE effectively reduces the information leakage.

Ranked Search Symmetric Encryption(RSSE): Ranked searchable symmetric encryption for the subsequent security and performance assurance. specially, we've got the subsequent desires: **Ranked key-word seek**: To explore special mechanisms for designing effective ranked search schemes based on the prevailing searchable encryption framework.

Protection guarantee: To save you cloud server from learning the plaintext of both the information documents or the searched key phrases, and acquire the "as strong- as-viable" protection power as compared to current searchable encryption schemes.

Performance: Above dreams have to be finished with minimum verbal exchange and computation overhead

VI. CONCLUSION

Semantic expansion based similar search solution over encrypted cloud data used in a cloud based keyword search encryption analytics on encrypted with the aim to protect users' privacy. The metadata set, the cloud server builds the inverted index and constructs semantic relationship library (SRL) for the keywords set. Once receiving a query request, the cloud server first finds out the keywords that are semantically related to the query keyword in keeping with to SRL. Then each the query keyword and therefore the extensional words are used to retrieve the files. The result files are returned in order according to the total relevance score. Privacy preservation and security achieved regarding to the definition of (SSE searchable symmetric encryption). Useful of The effectiveness and efficiency are given a demo through evaluation

VII. REFERENCES

[1]. Ning Cao, Zhenyu Yang, Cong Wang, Kui Ren, and Wenjing Lou. Privacy-preserving query over encrypted graph structured data in cloud computing. In Distributed Computing Systems (ICDCS), 2011 31st International Conference on, pages 393–402. IEEE, 2011.

[2]. Khorshed, M.T., Ali, A.S., Wasimi, S.A. "Monitoring Insiders Activities in Cloud Computing Using Rule Based Learning*", Security and Privacy in Computing Communications* (TrustCom), 2011 IEEE 10th International Conference on, 2011, pp 757-764.

[3]. S.Miranda- Jimnez, A.Gelbukh, and G.Sidorov, "Summarizing conceptual graphs for automatic summarization task," Conceptual Structures for STEM Research and Education, Springer Berlin Heidelberg,pp. 245-253,2013.

[4]. R. Ferreira, L.de Souza Cabral, and R.D. Lins, "Assessing sentence scoring techniques for extractive text summarization," Expert systems with applications,vol.40,no.14,pp.5755-5764,2013.

[5]. Z.Fu, X.Sun,and Q.Liu,"Achieving Efficien tCloud Search Services: Multi keyword Ranked Search over Encrypted Cloud Data Supporting ParallelComputing,"IEICETransactionsonCom munications,vol.98,no.1,pp.190-200,2015.

[6]. S. Ji, G. Li, C. Li, and J. Feng, "Efficient interactive fuzzy keyword search," in Proc. of WWW'09,2009.

[7]. C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in Proc. of ICDCS'10,2010.

[8]. Wang, li et al [2014] - Privacy-Preserving Multi-keyword Ranked Search over Encrypted Cloud Data.

[9]. Wang, Yu et al [2014] Privacy-Preserving Multi-Keyword Fuzzy Search over Encrypted Data in the Cloud.

[10]. Y. Zhu, G. J. Ahn, H. Hu, S. S. Yau, H. G. An, and C. J. Hu, "Dynamic audit services for outsourced storages in cloud," IEEE Transactions on Services Computing. vol. 6, no. 2, pp.227-238, Apr.-Jun. 2013. 0018.