Fuzzy Keyword Search over Encrypted Data using Search Scheme in Cloud Computing

[1] K.Rajesh,[2] K.Venkatreddy,[3]G.RaviTeja
[1,2 ,3]M.Tech students K.L.University
Guntur,Andhra Pradesh,India

ABSTRACT

Cloud computing is a technology that uses the internet and middle isolated servers to uphold data and applications. Cloud computing allow consumers and businesses to use applications with no fitting and access their personal files at any computer with internet access. This technology allows for much more proficient computing by centralize storage, memory, processing and bandwidth. Perhaps the biggest concerns about cloud computing are security and privacy. If a client can log in from any location to access data and applications, it's possible the client's privacy could be compromised. In existing technique we retrieve the files from the cloud, by searching the keywords on the encrypted data. There are many searching technique which were implemented in the cloud these technique supports only exact keyword search. Typical users searching behaviors are happen very frequently these are the drawbacks with the existing system which are not suitable for cloud computing environment and which effects system usability. Using fuzzy search the exact keywords are displayed along with similarity keywords, which solve the problems faced by the cloud users. This paper concentrates on solving the problems of the user who search the data with the help of fuzzy keyword on cloud.

Keywords– **Edit distance, fuzzy search, trapdoor**

1. INTRODUCTION

Cloud computing, the new term for the long dreamed vision of computing as a utility, enables convenient, on-demand network access to a centralized pool of configurable computing resource that can be rapidly deployed with great efficiency and minimal management overhead . The amazing advantages of Cloud Computing include: On demand self-service, ubiquitous network access, location independent resource pooling, rapid resource elasticity, usage-based pricing, and transference of risk. Thus, Cloud Computing could easily benefit its users in avoiding large capital outlays in the deployment and management of both software and hardware. Undoubtedly, Cloud Computing brings unprecedented paradigm shifting and benefits in the history of IT. However, data encryption makes effective data utilization a very challenging task given that there could be a large amount of outsourced data files. Moreover, in Cloud Computing, data owners may share their outsourced data with a large number of users. The individual users might want to only retrieve certain specific data files they are interested in during a given session. One of the most popular ways is to selectively retrieve files through keyword-based search instead of retrieving all the encrypted files back which is completely impractical in cloud computing scenarios. Such keyword-based search technique allows users to selectively retrieve files of interest and has been widely applied in plaintext search scenarios, such as Google search [1]. Unfortunately, data encryption restricts user's ability to perform keyword search and thus makes the traditional plaintext search methods unsuitable for Cloud Computing. Besides this, data encryption also demands the protection of keyword privacy since keywords usually contain important information related to the data files. Although encryption of keywords can protect keyword privacy, it further renders the traditional plaintext search techniques useless in this scenario.

While maintaining keyword privacy. Fuzzy keyword search greatly enhances system usability by returning the matching files when user's searching inputs exactly match. The predefined keyword or the closest possible matching files based on keyword similarity semantics, when exact match fails. More specifically, we use edit distance to quantify keywords similarity and develop a novel technique, i.e., a wildcard based technique, for the construction of fuzzy keyword sets.. Based on the constructed fuzzy keyword sets, we propose an efficient fuzzy keyword search scheme. Through rigorous security analysis, we show that the proposed solution is secure and privacy-preserving, while correctly realizing the goal of fuzzy keyword search.

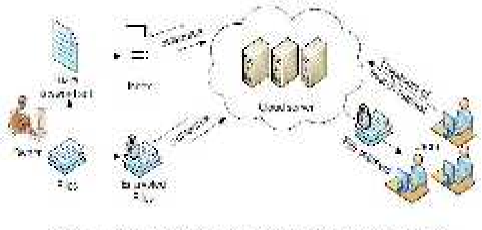

Fig. 1: Architecture of the fuzzy keyword search

2. RELATED WORK

2.1. Fuzzy Set Theory:

Fuzzy sets are sets whose elements have degrees of membership. Fuzzy sets were introduced simultaneously by Lotfi A. Zadeh and Dieter Klaua in 1965 as an extension of the classical notion of set. In classical set theory, the membership of elements in a set is assessed in binary terms according to a bivalent condition — an element either belongs or does not belong to the set. By contrast, fuzzy set theory permits the gradual assessment of the membership of elements in a set; this is described with the aid of a membership function valued in the real unit interval [0, 1]. Fuzzy sets generalize classical sets, since the indicator functions of classical sets are special cases of the membership functions of fuzzy sets, if the latter only take values 0 or 1. In fuzzy set theory, classical bivalent sets are usually called crisp sets. The fuzzy set theory can be used in a wide range of domains in which information is incomplete or imprecise, such as bioinformatics. Fuzzy sets can be applied, for example, to the field of genealogical research. When an individual is searching in vital records such as birth records for possible ancestors, the researcher must contend with a number of issues that could be encapsulated in a membership function. Looking for an ancestor named John Henry Pittman, who you think was born in (probably eastern) Tennessee circa 1853 (based on statements of his age in later censuses, and a marriage record in Knoxville), what is the likelihood that a particular birth record for "John Pittman" is your John Pittman? What about a record in a different part of Tennessee for "J.H. Pittman" in 1851? (It has been suggested by Thayer Watkins that Zadeh's ethnicity is an example of a fuzzy set).

2.2. Plaintext fuzzy keyword search.

Recently, the importance of fuzzy search has received attention in the context of plaintext searching in information retrieval community they addressed this problem in the traditional information access paradigm by allowing user to search without using try-and-see approach for finding relevant information based on approximate string matching. At the first glance, it seems possible for one to directly apply these string matching algorithms to the context of searchable encryption by computing the trapdoors on a character base within an alphabet. However, this trivial construction suffers from the dictionary and statistics attacks and fails to achieve the search privacy.

2.3. Searchable encryption.

Searchable encryption schemes provide an important mechanism to cryptographically protect data while keeping it available to be searched and accessed. In a common approach for their construction, the encrypting Entity chooses one or several keywords that describe the content of each encrypted record of data. To perform a search, a user obtains a trapdoor for a keyword of his/her interests and uses this trapdoor to find all the described by this keyword We present a searchable encryption scheme that allows users to privately search by keywords on encrypted data in public key setting and decrypt the search result. To this end, we define and implements two primitives: public key encryption(PEOKS) and oblivious keyword search and committed blind anonymous identity-based encryption. PEOKS is an extension of public key encryption with keyword search in which users can obtain trapdoor from the secret key holder without revealing the keywords. PEOKS scheme is used to build public key encrypted database that permits private searches i.e.; neither the keyword nor the search result are revealed.

3.Problem Definition

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. More specifically, we use edit distance to quantify keywords similarity and develop a novel technique, i.e., wildcard-based technique and grambased technique for the construction of fuzzy keyword sets. This technique eliminates the need for enumerating all the fuzzy keywords and the resulted size of the fuzzy keyword sets is significantly reduced. Based on the constructed fuzzy keyword sets, we propose an efficient fuzzy keyword search scheme. Through rigorous security analysis, we show that the proposed solution is secure and privacy preserving, while correctly realizing the goal of fuzzy keyword search.

4 CONCLUSION & FUTURE WORK

In this project, for the first time we formalize and solve the problem of supporting efficient yet privacypreserving fuzzy search for achieving effective utilization of remotely stored encrypted data in Cloud Computing. We design an advanced technique (i.e., wild card-based technique) to construct the storageefficient fuzzy keyword sets by exploiting a significant observation on the similarity metric of edit distance. Based on the constructed fuzzy keyword sets, we further propose an efficient fuzzy keyword search scheme. Through rigorous security analysis, we show that our proposed solution is secure and privacy-preserving, while correctly realizing the goal of fuzzy keyword search. Future work is on security mechanisms that support search semantics that takes into consideration conjunction of keywords, sequence of keywords, and even the complex natural language semantics to produce highly relevant search results and search ranking that sorts the searching results according to the relevance criteria.

5.References:

[1] Google, “Britney spears spelling correction,” Referenced online at http://www.google.com/jobs/britney.html, June 2009.

[2] M. Bellare, A. Boldyreva, and A. O’Neill, “Deterministic and efficiently searchable encryption,” in Proceedings of Crypto 2007, volume 4622 of LNCS. Springer-Verlag, 2007.

[3] D. Song, D. Wagner, and A. Perrig, “Practical techniques for searches on encrypted data,” in Proc. Of IEEE Symposium on Security and Privacy’00, 2000.

[4] E.-J. Goh, “Secure indexes,” Cryptology ePrint Archive, Report 2003/216, 2003, http://eprint.iacr.org/.

[5] D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, “Public key encryption with keyword search,” in Proc. of EUROCRYP’04, 2004.

[6] B. Waters, D. Balfanz, G. Durfee, and D. Smetters, “Building an encrypted and searchable audit log,” in Proc. of 11th Annual Network and Distributed System, 2004.

[7] Y.-C. Chang and M. Mitzenmacher, “Privacy preserving keyword searches on remote encrypted data,” in Proc. of ACNS’05, 2005.

[8] R. Curtmola, J. A. Garay, S. Kamara, and R. Ostrovsky, “Searchable symmetric encryption: improved definitions and efficient constructions,”in Proc. of ACM CCS’06, 2006.

[9] D. Boneh and B. Waters, “Conjunctive, subset, and range queries on encrypted data,” in Proc. of TCC’07, 2007, pp. 535–554.