RESEARCH ARTICLE OPEN ACCESS

Ranked Keyword Search Technique for Privacy Preserving

Sandhya Pradip Mohite
(Computer Department , Pune University,MIT Academy of engineering, Alandi devachi ,Pune)

Abstract:

Privacy preserving is the essential aspect for cloud. In privacy preserving ranked keyword search, data owner outsource the document in an encrypted form. So for privacy purpose, Reliable ciphertext search technique is essentially. To design ciphertext search technique which provide encrypted document is critical task. In this paper, hierarchical clustering method is designed for semantic search and fast ciphertext search within a big data environment. In hierarchical clustering, documents are based on maximum relevance score and clusters are divided into sub-clusters.

Keywords — **hirarchical clustering, ciphertext search technique, relevance score, security.**

Literature survey-

Information leakage is problem in big data environment. Encryption of data is common method to reduce information leakage searching encrypted documents on the server side is big challenging task. Many cryptographic techniques are developed in past, but these techniques are much complex and time consuming.

In this paper, vector space model is used, every document is represented by vector. Relationship between different documents are classified into several categories. Due to desired document categories, document search time is reduced. Due to the small number of documents, cluster can be categorized into sub-categories. Cloud server first search document in cluster. Cloud server will select the desired k-document. The value of k is previously decided by user and send to server. If document cannot find in nearest cluster, it goes for sub-clusters. Further k-document is not satisfied then, cloud server will trace back to the parent node and select the desired document. This process repeated recursively until respected k-document get satisfied.

Proposed System-

In this paper, multi-keyword ranked search over encrypted data based on hierarchical clustering index i.e (MRSE-HCI).By computing the relevance score between document, every document will classified into specific clusters. Relevance score between document can be used for calculating relationship. The constraints on the cluster may broken due to newly added documents to the clusters. If any new document is added on the cluster, then new cluster center can be added. So new document choose temporal cluster center. So all documents are reassigned and cluster centers will be released. Therefore, number of clusters depend on number of documents. If cluster size exceeds the limitation, cluster will be divided into sub-clusters. Rank privacy can be improved due to hierarchical clustering. At first level of search phase, cloud server compute the relevance score between the query and cluster center and searches nearest cluster. This process iterate until smallest cluster is found. If smallest cluster is not found, which satisfies document, then cloud server goes to parent cluster.

System model-

System model consists of three entities i.e. data user, data owner and the cloud server. Data owner can collect documents, outsource them into encrypted format to cloud server. Data user cannot access document from cloud Server without authorization from data owner.

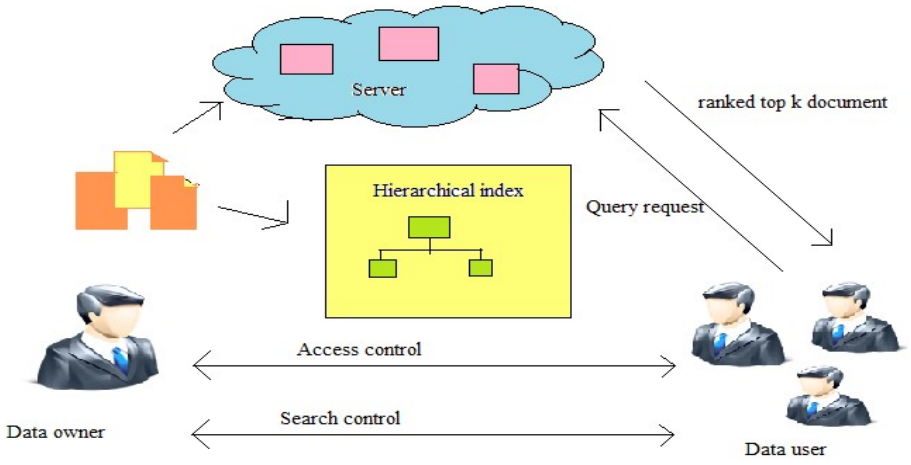

Fig1: System Architecture

Data user should authenticate from data owner. Then data user can send request for encrypted document to cloud server. If authenticated data user is requesting for document, then cloud server searches requested document in dataset and sends top k-document which matches with query. This helps to protect the information leakage.

Conclusion-

In this paper, MRSE-HCI architecture is proposed for semantic relationship between different plain document over related encrypted document. Due to relevance score measure between query and the document, search technique is improved. So ciphertext search technique can be improved.

Reference-

[1]S. Grzonkowski, P. M. Corcoran, and T. Coughlin, "Security analysis of authentication protocols for next-generation mobile and CE cloud services," in Proc. IEEE Int. Conf. Consumer Electron.,2011, Berlin, Germany, 2011, pp. 83–87.

[2]D. X. D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. IEEE Symp. Security Priv,BERKELEY, CA, 2000, pp. 44–55.

[3]D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. EUROCRYPT,Interlaken, SWITZERLAND, 2004, pp. 506–522.

[4]Y. C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in Proc. 3rd Int.Conf. Applied Cryptography Netw. Security, New York, NY, 2005,pp. 442–455.

[5]R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable Symmetric encryption:Improved definitions and efficient constructions," in Proc. 13th ACM Conf. Comput. Commun. Security, Alexandria, Virginia, 2006, pp. 79–88.

[6]M. Bellare, A. Boldyreva, and A. O'Neill, "Deterministic and efficiently searchable encryption," in Proc. 27th Annu. Int. Cryptol.Conf. Adv. Cryptol., Santa Barbara, CA, 2007, pp. 535–552.

[7]D. Boneh and B. Waters, "Conjunctive, subset, and range queries on encrypted data," in Proc. 4th Conf. Theory Cryptography,Amsterdam, NETHERLANDS, 2007, pp. 535–554.