IJARES/May 2015/Volume-3/Issue-5/1831-1835 **ISSN 2347- 9337**

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN ENGINEERING AND SCIENCE

ASSESSMENT OF SIMILARITY FOR CAPTURING IMPORTANCE OF DATA DOCUMENTS

Suraneni Lavanya[1], Dr.B.Vijayakumar[2]

[1]**M.Tech Student, Dept of CSE, Malla Reddy Engineering College for Women, Hyderabad, T.S, India**

[2]**Professor, Dept of CSE, Malla Reddy Engineering College for Women, Hyderabad, T.S, India**

ABSTRACT:

Searching for privacy preserving and useful search service over encrypted cloud information is of vital importance. Scheming of a well-organized encrypted data search method that supports multi-keyword semantics without privacy violations turns out to be a challenging trouble. Preserving problem of multi-keyword ranked searching on encrypted data was worked out while maintaining strict system wise confidentiality in cloud computing concept. Among a variety of multi-keyword semantics, we select the proficient similarity measure of coordinate matching, specifically as many matches as feasible, to confine the significance concerning data documents. We put forward two multi-keyword ranked search over encrypted information schemes on basis of similarity assess of coordinate matching while meeting several privacy needs in two various threat representatiosn such as privacy-preserving system in known ciphertext representation and in known background representation. To attain multi-keyword ranked search, we suggest utilizing inner product similarity to quantitatively assess competent similarity measure coordinate matching.

Keywords: Privacy preserving, Multi-keyword ranked search, Threat models, Coordinate matching, Cloud computing.

1. INTRODUCTION:

For protecting data privacy in cloud as well as beyond, sensitive data may have to be encrypted by means of data owners earlier than outsourcing to commercial public cloud. This obsoletes established data

utilization service on basis of plaintext keyword search [1]. Searchable encryption is a supportive method allows a user to strongly search all the way through a single keyword and recover documents of importance. Direct application of these approaches towards huge scale cloud data exploitation system would not be essentially suitable, since they are cannot put up high service-level needs. Ranked search eliminates redundant network traffic via sending back most appropriate data, which is highly enviable in pay-as-you-use cloud idea. For protection of privacy, such ranking process, should not disclose any keyword associated information. Coordinate matching is a well-organized resemblance measure among multi-keyword semantics to improve result relevance, and was extensively used in plaintext information recovery [2][3]. Scheming of a well-organized encrypted data search method that supports multi-keyword semantics without privacy violations turns out to be a challenging trouble. To attain multi-keyword ranked search, we suggest utilizing inner product similarity to quantitatively assess competent similarity measure coordinate matching. Preserving problem of multi-keyword ranked searching on encrypted data was worked out while maintaining strict system wise confidentiality in cloud computing concept. The adapted secure inner product computation proposal is not superior enough for MRSE systems.

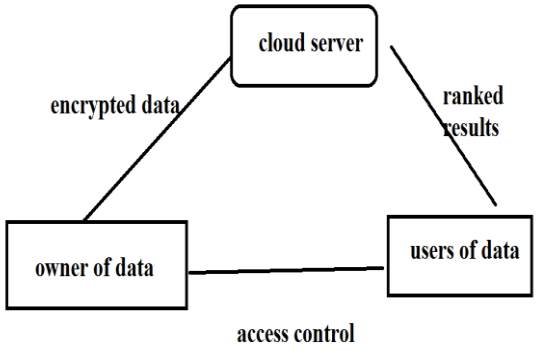

Fig1: Cloud data hosting service.

2. STRUCTURE OF MULTI-KEYWORD RANKED SEARCH OVER ENCRYPTED INFORMATION:

Searching for privacy preserving and useful search service over encrypted cloud information is of vital importance. We put forward two multi-keyword ranked search over encrypted information schemes on basis of similarity assess of coordinate matching while meeting several privacy needs in two various threat models. The cloud server is measured as honest-but-curious in our representation, which is reliable with associated works on cloud security. Fig1 considers cloud data hosting service. Among a variety of multi-keyword

semantics, we select the proficient similarity measure of coordinate matching, specifically as many matches as feasible, to confine the importance of data documents to the search query. Improved Multi-keyword ranked search over encrypted information schemes attains a variety of stringent privacy needs in two threat models with improved attack capabilities [4]. To meet challenge of supporting multi keyword semantic lacking privacy breaches, we put forward a basic thought for MRSE by means of secure inner product computation, adapted from safe k-nearest neighbour method. We put forward two multi-keyword ranked search over encrypted information schemes on basis of similarity assess of coordinate matching while meeting several privacy needs in two various threat models. Multi-keyword ranked searching on encrypted data consists of four algorithms such as Setup which takes a security parameter m as input, and outputs a symmetric key as EK. Build Index: on basis of data set J, a searchable index S was built by data owner which is encrypted by symmetric key EK and subsequently outsourced towards cloud server. Subsequent to index building, document gathering is encrypted as well as outsourced separately. Trapdoor: by k keywords of importance in D as input, algorithm generates an equivalent trapdoor. Query: When cloud server obtains a query request it carries out ranked search on index S with trapdoor, and at last returns ranked id list concerning top-k documents that are sorted by similarity [5]. Based on information known by cloud server two threat models were considered with several attack capabilities as follows Known background representation: in which the cloud server is thought to hold additional knowledge than what can be accessed in recognized ciphertext model. Such information might comprise correlation relationship of specified search requests, in addition to data set associated statistical information. Known ciphertext representation: In this representation the cloud server is believed to only recognize encrypted data set as well as searchable index that are outsourced from data owner.

3: SCHEMES OF MULTI-KEYWORD RANKED SEARCH OVER ENCRYPTED INFORMATION:

To attain multi-keyword ranked search, we suggest utilizing inner product similarity to quantitatively assess competent similarity measure coordinate matching. We suggest two multi-keyword ranked search over

encrypted information schemes on basis of similarity assess of coordinate matching while meeting several privacy needs in two various threat models. Privacy-Preserving system in Known Ciphertext representation: The adapted secure inner product computation proposal is not superior enough for MRSE systems. In our more superior design, rather than just removing extended dimension in query vector as we plan to carry out at initial glance, dimension extending process but assigning a new random number to the extended dimension within each query vector was preserved. Such a newly added randomness is usual to enhance complexity for cloud server to find out association between received trapdoors. Randomness should be cautiously calibrated in search result to conceal document frequency and reduce the chances in support of re-identification of keywords. Introducing randomness in concluding similarity score is an effectual way toward what we imagine here. Unlike the randomness concerned in query vector, a dummy keyword was introduced into each data vector and allocates a random value to it. Privacy-Preserving system in Known Background representation: When cloud server has information of some background information on outsourced data set, for instance, the correlation association of two specified trapdoors, certain keyword confidentiality might not be assured anymore by privacy-preserving system in known ciphertext representation [6]. This is feasible in known background representation since cloud server can utilize scale analysis to work out the keyword particular information, for instance, document frequency, which is combined with background information to recognize keyword within a query at high likelihood.

4. CONCLUSION:

Searchable encryption is a supportive method allows a user to strongly search all the way through a single keyword and recover documents of importance. Ranked search eliminates redundant network traffic via sending back most appropriate data, which is highly enviable in pay-as-you-use cloud idea. Coordinate matching is a well-organized resemblance measure among multi-keyword semantics to improve result relevance, and was extensively used in plaintext information recovery. Preserving problem of multi-keyword ranked search on encrypted datawas worked out while maintaining strict system wise

confidentiality in cloud computing concept. We put forward two multi-keyword ranked search over encrypted information schemes on basis of similarity assess of coordinate matching while meeting several privacy needs in two various threat models such as privacy-preserving system in known ciphertext representation and in known background representation. Among a variety of multi-keyword semantics, we select the proficient similarity measure of coordinate matching, specifically as many matches as feasible, to confine the importance of data documents to the search query. The multi-keyword ranked searching on encrypted data consists of four algorithms. Improved Multi-keyword ranked search over encrypted information schemes attains a variety of stringent privacy needs in two threat models with improved attack capabilities.

REFERENCES

[1] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy Keyword Search Over Encrypted Data in Cloud Computing," Proc. IEEE INFOCOM, Mar. 2010.

[2] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W.E.S. III, "Public Key Encryption That Allows PIR Queries," Proc. 27th Ann. Int'l Cryptology Conf. Advances in Cryptology (CRYPTO '07), 2007.

[3] P. Golle, J. Staddon, and B. Waters, "Secure Conjunctive Keyword Search over Encrypted Data," Proc. Applied Cryptography and Network Security, pp. 31-45, 2004.

[4] L. Ballard, S. Kamara, and F. Monrose, "Achieving Efficient Conjunctive Keyword Searches over Encrypted Data," Proc. Seventh Int'l Conf. Information and Comm. Security (ICICS '05), 2005.

[5] D. Boneh and B. Waters, "Conjunctive, Subset, and Range Queries on Encrypted Data," Proc. Fourth Conf. Theory Cryptography (TCC), pp. 535-554, 2007.

[6] R. Brinkman, "Searching in Encrypted Data," PhD thesis, Univ. Of Twente, 2007.