Abstract

The owner of large amount of data chooses cloud facility to outsource their precious data. For the purpose of privacy the data owners encrypted their data before outsourcing. This encryption hides the connection between the documents. So this makes the ciphertext search critical. Caused by software/hardware failure, data search results returning to the users may carry injured data. Thus, a verifiable method should be provided for users to verify the accuracy of the search results. This paper proposes a hierarchical clustering method in order to get a better clustering result. This method is based on k-means clustering algorithm. Every document will be clustered and hashed; this hash result will be used as the representative of the document. This paper gives a new search technique to adopt the backtrack algorithm on the above clustering method. With the increment of the data volume, the usefulness of the proposed method in rank privacy tends to be more evident. By appealing the Merkle hash tree and cryptographic signature to authenticated tree anatomy, we give a verification mechanism to persuade the accuracy and perfectness of query results. By using this method, the search time increased only sequentially not exponentially. The exploratory result demonstrates that the proposed system
solves multi keyword search problems and also brings the relevance between retrieved documents and raises the search efficiency.

References

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

Index Terms

Computer Science  Distributed Systems

Keywords

Cloud computing, Ciphertext, Multi-keyword search, Hierarchical Clustering, Security, Query, Symmetric Key encryption.