

Privacy preserving Multi-keyword Top-k similarity search over Encrypted Data

Abstract:

Cloud computing provides individuals and enterprises massive computing power and scalable storage capacities to support a variety of big data applications in domains like health care and scientific research, therefore more and more data owners are involved to outsource their data on cloud servers for great convenience in data management and mining.

Existing system:

Cao et al. proposed the multi-keyword ranked search over encrypted data for the first time and built a searchable index based on the vector space model, and chosen "coordinate matching" to measure the similarity between queries and documents. However, in their schemes, the time complexity of search is $O(nm)$ (n is the number of keywords in dictionary, m is the size of the documents that stored in the cloud server), and the time complexity of trapdoor construction is also very high. Sun et al. proposed a tree-based index structure which is based on the vector space model and the TF_IDF model. This structure achieves sub-linear time complexity, but it is vulnerable in protecting data privacy.

Disadvantages:

1. There is time complexity for searching
2. Not sufficient for protecting the data privacy

Proposed System:

We propose the random traversal algorithm which can achieve that for two identical queries With different keys, the cloud server traverses different paths on the index, and the data user receives different results but with the same high level of query accuracies in the mean time. Then, in order to improve the search efficiency, we design the group multi-keyword top- k search

scheme, which divides the dictionary into multiple groups and only needs to store the top-*ck* documents of each word group when building index. Next, to protect the query unlinkability, we apply the random traversal algorithm.

Advantages:

1. Search efficiency has been improved here.
2. Time complexity is also reduced.

SYSTEM REQUIREMENTS

H/W System Configuration:-

- Processor - Pentium –III
- RAM - 256 MB (min)
- Hard Disk - 20 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - SVGA

S/W System Configuration:-

- Operating System : Windows95/98/2000/XP
- Application Server : Tomcat5.0/6.X
- Front End : HTML, Jsp
- Scripts : JavaScript.
- Server side Script : Java Server Pages.
- Database : MySQL 5.0
- Database Connectivity : JDBC