Semantic-aware Searching over Encrypted Data for Cloud Computing

Abstract

- With the increasing adoption of cloud computing, a growing number of users outsource their datasets to cloud. To preserve the privacy, the datasets are usually encrypted before outsourcing.
- However, the common practice of encryption makes the effective utilization of the data difficult. For example, it is difficult to search the given keywords in encrypted datasets. Many schemes are proposed to make encrypted data searchable based on keywords.
- However, keyword-based search schemes ignore the semantic representation information of users retrieval, and cannot completely meet with users search intention. Therefore, how to design a content-based search scheme and make semantic search more effective and context-aware is a difficult challenge. In this paper, we propose ECSED, a novel semantic search scheme based on the concept hierarchy and the semantic relationship between concepts in the encrypted datasets

Contd..

- ECSED uses two cloud servers. One is used to store the outsourced datasets and return the ranked results to data users. The other one is used to compute the similarity scores between the documents and the query and send the scores to the first server.
- To further improve the search efficiency, we utilize a tree-based index structure to organize all the document index vectors. We employ the multi keyword ranked search over encrypted cloud data as our basic frame to propose two secure schemes.
- The experiment results based on the real world datasets show that the scheme is more efficient than previous schemes. We also prove that our schemes are secure under the known ciphertext model and the known background model.

Existing system

- many researchers have proposed a series of efficient search schemes over encrypted cloud data.
- The general process of search scheme can be divided into extracting document features, constructing a searchable index, generating search trapdoor, searching the index based on the trapdoor and returning the search results. These search schemes provide different query capabilities, including single keyword search

Disadvantages

- enterprise IT infrastructure that provides high quality applications and services .
- The cloud customers can outsource their local complex data system into the cloud to avoid the overhead of management and local storage. However, the security of outsourced data cannot be guaranteed, as the Cloud Service Provider (CSP) possesses whole control of the data.

Proposed system

- Many schemes are proposed to make encrypted data searchable based on keywords.
- However, keyword-based search schemes ignore the semantic representation information of users retrieval, and cannot completely meet with users search intention.
- Therefore, how to design a content-based search scheme and make semantic search more effective and context-aware is a difficult challenge.
- In this paper, we propose ECSED, a novel semantic search scheme based on the concept hierarchy and the semantic relationship between concepts in the encrypted datasets. ECSED uses two cloud servers. One is used to store the outsourced datasets and return the ranked results to data users

Advantages

- consider keywords as the document feature, do not take the semantic relations between words into consideration, both in the steps of extracting document features and generating search trapdoor.
- As we all know, the semantic relations between words are diverse, such as synonymy and domain correlation. Considering the potentially huge amount of outsourced data documents in the cloud, the search accuracy and search efficiency are influenced negatively if the semantic relations between words are not handled well.

Hardware Requirements

Processor	:Intel Pentium IV 1GHz
RAM	:256MB (Min)
Hard Drive	:5GB free space
Monitor	:1024 * 768, High Color inch
Mouse	:Scroll Mouse(Logitech)
Keyboard	:104 keys

Software requirements

OS	:	Windows XP/7/8
Front End	:	Visual Studio 2010/ netbeans 7.1
Back End	:	SQL Server 2005/ heidisql 3.2
Browser	:	Any Web Browser

conclusion

- In this paper, to address the problem of semantic retrieval, we propose effective schemes based on concept hierarchy. Our solutions use two cloud servers for encrypted retrieval and make contributions both on search accuracy and efficiency.
- To improve accuracy, we extend the concept hierarchy to expand the search conditions. In addition, a tree-based index structure is constructed to organize all the document index vectors, which are built based on the concept hierarchy for the aspect of search efficiency. The security analysis shows that the proposed scheme is secure in the threat models. Experiments on real world dataset illustrate that our scheme is efficient.

References

[1] L. M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, “A break in the clouds: towards a cloud definition,” ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 1, pp. 50–55, 2009.

[2] C.Wang, N. Cao, K. Ren, andW. Lou, “Enabling secure and efficient ranked keyword search over outsourced cloud data,” IEEE TPDS, vol. 23, no. 8, pp. 1467–1479, 2012.

[3] D. Song, D.Wagner, and A. Perrig, “Practical techniques for searches on encrypted data,” in Proc. of S&P, 2000.

[4] R. Curtmola, J. A. Garay, S. Kamara, and R. Ostrovsky, “Searchable symmetric encryption: improved definitions and efficient constructions,” In Proc. of ACM CCS, 2006, pp. 79–88.

[5] A. Swaminathan, Y. Mao, G.-M. Su, H. Gou, A. L.Varna, S. He, M. Wu, and D. W. Oard, “Confidentiality-preserving rank-ordered search,” in Proc. of the 2007 ACM Workshop on Storage Security and Survivability, 2007, pp. 7–12.