Privacy-Preserving Multikeyword SimilaritySearch Over Outsourced Cloud Data

ABSTRACT:

The amount of data generated by individuals andenterprises is rapidly increasing.With the emerging cloud computingparadigm, the data and corresponding complex managementtasks can be outsourced to the cloud for the management flexibilityand cost savings. Unfortunately, as the data could be sensitive,the direct data outsourcing would have the problem of privacyleakage. The encryption can be used, before the data outsourcing,with the concern that the operations can still be accomplished bythe cloud. We consider the multikeyword similarity search overoutsourced cloud data. In particular, with the consideration of thetext data only, multiple keywords are specified by the user. Thecloud returns the files containing more than a threshold numberof input keywords or similar keywords, where the similarity hereis defined according to the edit distance metric. We propose threesolutions, where blind signature provides the user access privacy,and a novel use of Bloom filter's bit pattern provides the speedupof search task at the cloud side. Our final design to achieve thesearch is secure against insider threats and efficient in terms of thesearch time at the cloud side. Performance evaluation and analysisare used to demonstrate the practicality of our proposed solutions.

EXISTING SYSTEM:

- Recently, privacy-preserving keyword search, suchas secure ranked search, where only the files with better matchingto the input keywords are returned, has been studied inthe settings of single keyword and multikeyword.
- Nonetheless, in the setting of secure ranked search, only thenumber of keyword matches is concerned, and the similaritybetween the input keywords and the actual words in text isnot taken into account.
- On a different front, privacy assuredsimilarity search, where the files containing exactly the samekeyword or containing similar keyword are returned, has alsobeen studied. However, in the setting of privacy assured similarityresearch, only single keyword is allowed, restricting thepractical use.

DISADVANTAGES OF EXISTING SYSTEM:

- Data encryption does not allow the cloud to answer the users' querieson the data.
- Practically infeasible because of the huge volume of the incurred bandwidth consumption.

PROPOSED SYSTEM:

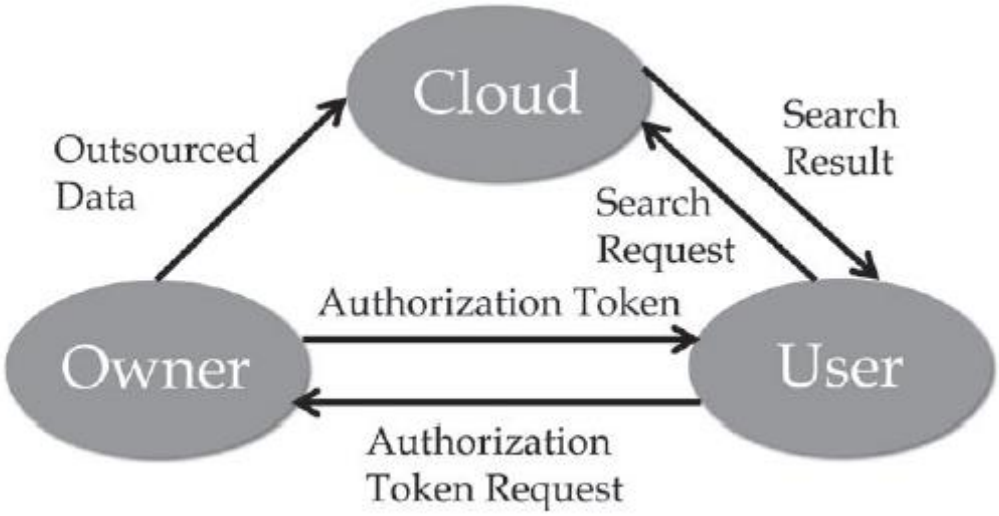
- In this paper, we focus on privacy-preserving multi-keywordsimilarity search (PPMKSS) over the outsourced cloud data.

- In the PPMKSS over the outsourced cloud data, the data areencrypted and then outsourced to the cloud. The user, aftergaining the authorization, sends keywords to the cloud, whichreturns to the user the files containing as many keywords ortheir variants as possible
- All of the files containing at least a threshold number of keywords similar to the input keyword specified by the user will be returned tothe user.
- The cloud cannot learn additional information from the outsourced encrypted data and thecorresponding index.

ADVANTAGES OF PROPOSED SYSTEM:

- PPMKSS over the outsourced cloud data is considered forthe first time in the literature.
- We propose a user authorization scheme with the guaranteeof user access privacy by using blind signature.
- By taking advantage of keyword suppressing techniqueand the Bloom filter, we propose three PPMKSS solutions,namely, PPMKSS-1, PPMKSS-2, and PPMKSS-3,to achieve PPMKSS.
- Simulation result and analysis are conducted to evaluatethe proposed methods and guarantee the performance.

SYSTEM ARCHITECTURE:



SYSTEM REQUIREMENTS:

HARDWARE REQUIREMENTS:

- System : Pentium Dual Core.
- Hard Disk : 120 GB.
- Monitor : 15'' LED
- Input Devices : Keyboard, Mouse
- Ram : 1 GB

SOFTWARE REQUIREMENTS:

- Operating system : Windows 7.
- Coding Language : JAVA/J2EE
- Tool : Netbeans 7.2.1
- Database : MYSQL

REFERENCE:

Chia-Mu Yu, Chi-Yuan Chen, and Han-Chieh Chao, *Senior Member, IEEE*, "Privacy-Preserving Multikeyword SimilaritySearch Over Outsourced Cloud Data" ,IEEE SYSTEMS JOURNAL, 2017.