Fast Phrase Search for Encrypted CloudStorage

ABSTRACT:

Cloud computing has generated much interest in the research community in recent years for its many advantages, but hasalso raise security and privacy concerns. The storage and access of confidential documents have been identified as one of the centralproblems in the area. In particular, many researchers investigated solutions to search over encrypted documents stored on remotecloud servers. While many schemes have been proposed to perform conjunctive keyword search, less attention has been noted onmore specialized searching techniques. In this paper, we present a phrase search technique based on Bloom filters that is significantlyfaster than existing solutions, with similar or better storage and communication cost. Our technique uses a series of n-gram filters tosupport the functionality. The scheme exhibits a trade-off between storage and false positive rate, and is adaptable to defend againstinclusion-relation attacks. A design approach based on an application's target false positive rate is also described.

EXISTING SYSTEM:

- Boneh et al. proposed one of the earliest works onkeyword searching. Their scheme uses public key encryptionto allow keywords to be searchable without revealingdata content.
- Waters et al. investigated the problem forsearching over encrypted audit logs. Many of the earlyworks focused on single keyword searches.
- Recently, researchershave proposed solutions on conjunctive keywordsearch, which involves multiple keywords.

- Otherinteresting problems, such as the ranking of search resultsand searching with keywords that might containerrorstermed fuzzy keyword search, have alsobeen considered. The ability to search for phrases was alsorecently investigated.
- Some of the existing system hasexamined the security of the proposed solutions and, whereflaws were found, solutions were proposed

DISADVANTAGES OF EXISTING SYSTEM:

- The cloud can read any data it desired, providing no privacy to its users. The storage of private keys and encrypted data by the cloud provider is also problematic in case of data breach.
- By recognizing the almost exponential distribution of keywords, the entries in the keyword location tables are split into pairs to achieve normalization without the high cost of storing unused random data. However, the use of encrypted indexes and the need to perform client-side encryption and decryption may still be computationally expensive in certain applications.
- Its space-efficiency comes at the cost of requiring a brute force location verification during phrase search. Since all potential locations of the keywords must be verified, the amount of computation required grows proportionally to the file size. As a result, the scheme exhibits a high processing time.

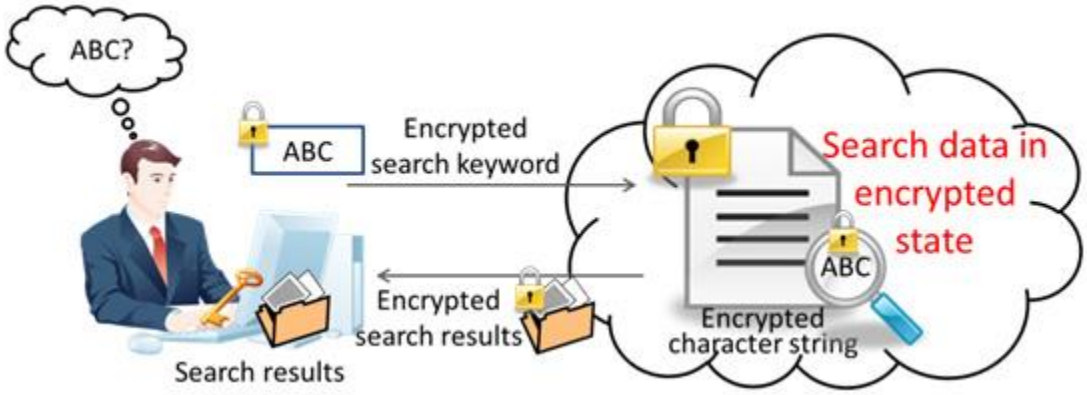
PROPOSED SYSTEM:

- In this paper, we present a phrase search scheme whichachieves a much faster response time than existing solutions.The scheme is also scalable, where documents can easily be removed and added to the corpus.We also describemodifications to the scheme to lower storage cost at a smallcost in response time and to defend against cloud providerswith statistical knowledge on stored data.
- Although phrase searches are processed independently usingour technique, they are typically a specialized functionin a keyword search scheme, where the primary function isto provide conjunctive keyword searches. Therefore, we describeboth the basic conjunctive keyword search algorithmand the basic phrase search algorithmalongwith design techniques.

ADVANTAGES OF PROPOSED SYSTEM:

- Our framework differs from some of the earlier works, where keywords generally consist of meta-datarather than content of the files and where a trusted keyescrow authority is used due to the use of Identity basedencryption.
- When compared to recent works, where an organizationwishes to outsource computing resources to a cloud storageprovider and enable search for its employees,where the aim is to return properly ranked files.Most other recent works related to search over encrypteddata have considered similar models such as, where theclient acts as both data owner and user.

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:

Hoi Ting Poon, Member, IEEE, and Ali Miri, Member, IEEE, “Fast Phrase Search for Encrypted CloudStorage”,**IEEETransactions on Cloud Computing, 2017.**