Fast and Scalable Range Query Processing With

Strong Privacy Protection for Cloud Computing

ABSTRACT

Privacy has been the key road block to cloud computing as clouds may not be fully trusted. This paper is concerned with the problem of privacy-preserving range query processing on clouds. Prior schemes are weak in privacy protection as they cannot achieve index indistinguishability, and therefore allow the cloud to statistically estimate the values of data and queries using domain knowledge and history query results. In this paper, we propose the first range query processing scheme that achieves index indistinguishability under the indistinguishability against chosen keyword attack (IND-CKA). Our key idea is to organize indexing elements in a complete binary tree called PBtree, which satisfies structure indistinguishability(i.e., two sets of data items have the same PBtree structure if and only if the two sets have the same number of data items)and node indistinguishability(i.e.,thevalues of PBtree nodes are completely random and have no statistical meaning). We prove that our scheme is secure under the widely adopted IND-CKA security model. We propose two algorithms, namely PB tree traversal width minimization and PBtree traversal depth minimization, to improve query processing efficiency. We prove that the worst-case complexity of our query processing algorithm using PBtree is ,where is the total number of data items and is the set of data items in the query result.

EXISTING SYSTEM

In these paper, to leverage the computing and storage capability offered by clouds, we need to store data on clouds. On the other hand, due to many reasons, we may not fully trust the clouds for data privacy. First, clouds may have corrupted employees who do not follow data privacy policies. For example, in 2010, a Google engineer broke into the Gmail and Google Voice accounts of several children . Second, cloud computing systems may be vulnerable to external ma-licious attacks, and when intrusions happen, cloud customers may not be fully informed about the potential implications on the privacy of their data. Third, clouds may base services on facilities in some foreign countries where privacy regulations are difficult to enforce.

DISADVANTAGES

- However, even with such information, a pri-vacy-preserving range query scheme should not allow the cloud to infer additional information about the data based on past query results.
- Their communication cost is high because many data items in the query result do not satisfy the query.
- Reducing bucket sizes helps to reduce communication costs, but will worsen privacyprotection because the number of buckets becomes closer to that of data items.

PROPOSED SYSTEM

We propose the first privacy-preserving range query scheme that achieves index indistinguishability. Our key idea for achieving index indistinguishability is to organize all indexing elements in a complete binary tree where each node is represented using a Bloom filter, which we call a PBtree (where "P" stands for privacy and "B" stands for Bloom filter). PB-treesallow us to achieve index indistinguishability because they have two important properties. First, a PBtree has the property of structure indistinguishability, that is, two sets of data items have the same PBtree structure if and only if the two sets have the same number of data items. The structure of the PBtree of a set of data items is determined solely by the set cardinality, not the value of data items.

ADVANTAGES

- We propose two algorithms, namely PBtree traversal width minimization and PBtree traversal depth minimization, to improve query processing efficiency.
- We prove that the worst-case complexity of our query processing algorithm using PBtree is, where the total number of data items is and is the set of data items in the query result.
- We implemented and evaluated our scheme on a real-world dataset with 5 million items.

SYSTEM REQUIREMENTS:

HARDWARE REQUIREMENTS:

- System : Pentium IV 2.4 GHz.
- Hard Disk : 40 GB.
- Floppy Drive : 1.44 Mb.
- Monitor : 15 VGA Colour.
- Mouse : Logitech.
- Ram : 512 Mb.

SOFTWARE REQUIREMENTS:

- Operating system : Windows XP/7.
- Coding Language : ASP.net, C#.net
- Tool : Visual Studio 2010
- Database : SQL SERVER 2008