Building Confidential and Efficient Query Services in the Cloud with RASP Data Perturbation

Abstract

With the wide deployment of public cloud computing infrastructures, using clouds to host data query services has become an appealing solution for the advantages on scalability and cost-saving. However, some data might be sensitive that the data owner does not want to move to the cloud unless the data confidentiality and query privacy are guaranteed. On the other hand, a secured query service should still provide efficient query processing and significantly reduce the in-house workload to fully realize the benefits of cloud computing. We propose the random space perturbation (RASP) data perturbation method to provide secure and efficient range query and kNN query services for protected data in the cloud. The RASP data perturbation method combines order preserving encryption, dimensionality expansion, random noise injection, and random projection, to provide strong resilience to attacks on the perturbed data and queries. It also preserves multidimensional ranges, which allows existing indexing techniques to be applied to speedup range query processing. The kNN-R algorithm is designed to work with the RASP range query algorithm to process the kNN queries. We have carefully analyzed the attacks on data and queries under a precisely defined threat model and realistic security assumptions. Extensive experiments have been conducted to show the advantages of this approach on efficiency and security.

Existing System

HOSTING data-intensive query services in the cloud is increasingly popular because of the unique advantages in scalability and cost-saving. With the cloud infrastructures, the service owners can conveniently scale up or down the service and only pay for the hours of using the servers. This is an attractive feature because the workloads of query services are highly dynamic, and it will be expensive and inefficient to serve such dynamic workloads with in-house infrastructures. While new approaches are needed to preserve data confidentiality and query privacy, the efficiency of query services and the benefits of using the clouds should also be preserved. It will not be meaningful to provide slow query services as a result of security and privacy assurance. It is also not practical for the data owner to use a significant amount of in-house resources, because the purpose of using cloud resources is to reduce the need of maintaining scalable in-house infrastructures.
Disadvantages of Existing System

1. Service providers lose the control over the data in the cloud
2. Data confidentiality and query privacy have become the major concerns
3. Adversaries, such as curious service providers, can possibly make a copy of the database or eavesdrop users’ queries, which will be difficult to detect and prevent in the cloud infrastructures.

Proposed System

Based on the RASP perturbation method, we design the services for two types of queries: range query and kNN query. This section will dedicate to range query processing. We will first show that a range query in the original space can be transformed to a polyhedron query in the perturbed space, and then we develop a secure way to do the query transformation. Then, we will develop a two-stage query processing strategy for efficient range query processing.

1. Transforming Range Queries
2. Security Enhancement on Query Transformation

Advantages of Proposed System

1. Protecting Outsourced Data
2. Preserving Query Privacy

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.
- Coding Language : Java.
- Data Base : MY SQL