Abstract

In the virtual and widely distributed network, the process of handover sensitive data from the distributor to the trusted third parties always occurs regularly in this modern world. It needs to safeguard the security and durability of service based on the demand of users. A data distributor has given sensitive data to a set of supposedly trusted agents (third parties). Some of the data are leaked and found in an unauthorized place (e.g., on the web or somebody's laptop). The distributor must assess the likelihood that the leaked data came from one or more agents, as opposed to having been independently gathered by other means. We propose data allocation strategies (across the agents) that improve the probability of identifying leakages. These methods do not rely on alterations of the released data (e.g., watermarks). In some cases, we can also inject "realistic but fake" data records to further improve our chances of detecting leakage and identifying the guilty party. The idea of modifying the data itself to detect the leakage is not a new approach. Generally, the sensitive data are leaked by the agents, and the specific agent is responsible for the leaked data should always be detected at an early stage. Thus, the detection of data from the distributor to agents is mandatory. This project presents a data leakage detection system using various allocation strategies and which
The Guilt Detection Approach in Data Leakage Detection

assess the likelihood that the leaked data came from one or more agents. For secure transactions, allowing only authorized users to access sensitive data through access control policies shall prevent data leakage by sharing information only with trusted parties and also the data should be detected from leaking by means of adding fake record's in the data set and which improves probability of identifying leakages in the system. Then, finally it is decided to implement this mechanism on a cloud server.

References

- Rakesh Agrawal, Jerry Kiernan. Watermarking Relational Databases// IBM Almaden Research Center.

Index Terms

Computer Science Distributed Systems
Keywords
Cloud environment  data leakage  data security  fake records