Data Leakage Detection Through Fake objects

Mrs. Anita Dixit, Ms. Priyanka Dudgikar, Ms. Nireeksha Nayak , Ms. Ranjana Khyade
SDM College of Engineering and Technology

ABSTRACT:

In today's business world, the owner of the data are called as distributors and the trusted third parties are called as agent. A data distributor of the any company has given sensitive data about their work or business to one or more authorized person. If that data are shared by these agents who does not authority to share this data to any employee and that data are leaked and found in an unauthorized place. In this project we are implementing the system for detection of leaked data and possibly the agent who is responsible for leakage of data. The distributor must access the leaked data came from one or more agents. We use some duplicate data which is not known to the agent, that will help in identifying leakages. In some cases, we can also use "realistic but fake" data records to further improve our chances of detecting leakage and identifying the unauthorized person who leaked data.

***Key Words*: Agent, Distributor,Third Party, Alert,Sensitive Data.**

Introduction:

A data distributor or head of the any company, etc. has given sensitive data that means the important data or information about their work or business to one or more agents or the authorized person or employee (third parties)[3][4]. It should not be handling without authority and it should not interfere by any unauthorized person.

The idea of our project is to find guilty agent who leaked the sensitive or confidential data of the company. And give alert message to the guilty agent if he/ she are break rule again and again then take a legal action. The distributor can register their name and information then he/ she has authority of distributor. The client or agent can register their information and send request to the distributor. After request is receive the distributor provide unique username and password to the agent. Agent has login by entering that username and password. The agent has send request to the distributor for data the request is explicit or sample. Then distributor can check the request send by the agent is our agent? Then distributor checks the type of request i.e. explicit or sample request. It can collect data from database and add fake object. After that the distributor check whether the data is already sends that agent or not. If data is already sends then it will send message sorry data is already sent. If data is already not sends then it can send that data to the particular agent. If the third party or any client send request to the agent and the agent send them sensitive data that has does not authority to send or share data. The alert message is send

to the distributor that the agent is guilty. Then distributors sends warning message to that agent don't do this again if it happens again by that agent then distributor take action against that agent.

Literature Survey

Traditionally, leakage detection is handled by watermarking, e.g., a unique code is embedded in each distributed copy. If that copy is later discovered in the hands of an unauthorized party, the leaker can be identified. Watermarks can be very useful in some cases, but again, involve some modification of the original data. Furthermore, watermarks can sometimes be destroyed if the data recipient is malicious. *E.g*. A hospital may give patient records to researchers who will devise new treatments. Similarly, a company may have partnerships with other companies that require sharing customer data. Another enterprise may outsource its data processing, so data must be given to various other companies. We call the owner of the data the distributor and the supposedly trusted third parties the agents.[]

Goals and Objectives:

To detect the agent who leaked the confidential data and send alert message to the distributor.
The objectives of the "Data Leakage Detection" are as follows:

- Detection of guilty agent
- Send message or email to the distributor with identification of guilty agent
- Send alert message to the guilty agent
- Take legal action on agent when he/she break rule after the alert message

Methodology:

Our goal is to detect when the distributor's sensitive data has been leaked by agents, and if possible to identify the agent that leaked the data. Perturbation is a very useful technique where the .0 data is modified and made "less sensitive" before being handed to agents. we develop *unobtrusive* techniques for detecting leakage of a set of objects or records.

Proposed System

In this section we develop a model for assessing the "guilt" of agents. We also present algorithms for distributing objects to agents, in a way that improves our chances of identifying a leaker. Finally, we also consider the option of adding "fake" objects to the distributed set. Such objects do not correspond to real entities but appear realistic to the agents. In a sense, the fake objects acts as a type of watermark for the entire set, without modifying any individual members. If it turns out an agent was given one or more fake objects that were leaked, then the distributor can be more confident that agent was guilty.Fig1 shows the over all architecture of the proposed system.

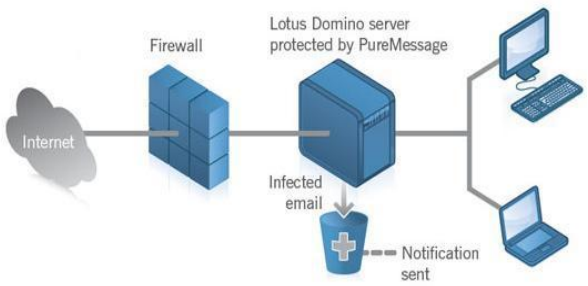

Fig1: Overview of the system

ADVANTAGES OF PROPOSED SYSTEM:

• If the distributor sees "enough evidence" that an agent leaked data, he may stop doing business with him, or may initiate legal proceedings.

• In this project we develop a model for assessing the "guilt" of agents.

• We also present algorithms for distributing objects to agents, in a way that improves our chances of identifying a leaker.

• Finally, we also consider the option of adding "fake" objects to the distributed set. Such objects do not correspond to real entities but appear.

• If it turns out an agent was given one or more fake objects that were leaked, then the distributor can be more confident that agent was guilty.

Problem Setup & Notation

A distributor owns a set $T=\{t_{1,\ldots,}t_m\}$of valuable data objects. The distributor wants to share some of the objects with a set of agents $U_1,U_2,\ldots U_n$, but does not wish the objects be leaked to other third parties. The objects in T could be of any type and size, e.g., they could be tuples in a relation, or relations in a database. An agent Ui receives a subset of objects, determined either by a sample request or an explicit request:

1. *Sample request*
2. *Explicit request*

Guilt Model Analysis:

Our model parameters interact and to check if the interactions match our intuition, in this section we study two simple scenarios as **Impact of Probability** p and **Impact of Overlap between** Ri **and** S. In each scenario we have a target that has obtained all the distributor's objects, i.e., T = S.

Algorithms:

1. Evaluation of Explicit Data Request Algorithms

In the first place, the goal of these experiments was to see whether fake objects in the distributed data sets yield significant improvement in our chances of detecting a guilty agent. In the second place, we wanted to evaluate our e-optimal algorithm relative to a random allocation.

2. Evaluation of Sample Data Request Algorithms

With sample data requests agents are not interested in particular objects. Hence, object sharing is not explicitly defined by their requests. The distributor is "forced" to allocate certain objects to multiple agents only if the number of requested objects exceeds the number of objects in set T. The more data objects the agents request in total, the more recipients on average an object has; and the more objects are shared among different agents, the more difficult it is to detect a guilty agent.

Conclusion:

Thus modules successfully work according to IEEE paper. It can successfully login distributor to the system and register the new agent request and show confirmation message for registration. In our work the distributor can check the list of registration request for new agent and the agent and distributor also updates its information successfully.

In this project in next modules we can implement following idea: User ID and password send to the agent for login to system. The agent should send data request to the distributor and distributor check the request and send data to the agent by adding fake object in data allocation module. In agent guilt module we can check send alert message to the distributor when the agent has share any confidential data. This is main goal of our Project.

References

1.International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 03 Issue: 03 | Mar-2016 www.irjet.net p-ISSN: 2395-0072

2.DATA LEAKAGE DETECTION USING DATA ALLOCATION STRATEGIES Jaymala Chavan, Priyanka Desai Thakur College of Engg. Tech, Mumbai, MH, India. International Journal of Advances in Engineering Technology, Nov 2013.

3.*International Journal of Advanced Engineering Research and Science (IJAERS) [Vol-3, Issue-1, Jan-2016]*
ISSN: 2349-6495

*4.*International Advanced Research Journal in Science, Engineering and Technology National Conference on Innovative Applications and Research in Computer Science and Engineering (NCIARCSE-2017) AGTI's Dr. Daulatrao Aher College Engineering, Vidyanagar Extension, Karad Vol. 4, Special Issue 4, January 2017

5.American Journal of Engineering Research (AJER) e-ISSN: 2320-0847 p-ISSN : 2320-0936 Volume-5, Issue-2, pp-82-87 www.ajer.org