DATA LEAKAGE DETECTION

Apoorva Kulkarni [1], Apeksha Ganeshwar [2], Bhumika Bhange [3], Pooja Tonge[4], Atharva Vyavahare[5]

[1] *Student, Computer Technology, Rajiv Gandhi College of Engineering and Research, Nagpur*
[2] *Student, Computer Technology, Rajiv Gandhi College of Engineering and Research, Nagpur*
[3] *Student Computer Technology, Rajiv Gandhi College of Engineering and Research, Nagpur*
[4] *Student, Computer Technology, Rajiv Gandhi College of Engineering and Research, Nagpur*
[5] *Student, Computer Technology, Rajiv Gandhi College of Engineering and Research, Nagpur*
Name of Guide: Prof. Sandeep Kamble

ABSTRACT

In this techno-savvy world from sending letters to photos, transmission takes place digitally. Everything from text to images and videos are clustered in category called 'data'. Moreover organizations have sensitive data that is given to a data distributor. There is quite a possibility of leakage of the data by the agents and can be distributed in unauthorized place. This project describes the ways in which data leaked data can be detected and some ways to prevent the leakage of data where fake data can be injected in the system that appear to be realistic. The unauthorized user when try to access or download the document, the fake data file is downloaded instead of actual data file that seems to be realistic to the user. Algorithms describes the encryption and decryption of data using shared keys. The system also propose the ways to secure the data when the keys are mismatched or when wrong key is used to download the data. The organizations can implement these techniques to secure their sensitive data from the agents often the trusted third parties who share the data in unauthorized way.

Keyword - *Data, Fake data, Data leakage.*

---

1. INTRODUCTION

In organizations the data must be handed to agents often called as third parties. Companies may share their data with other companies or data processing may be outsourced. Data leakage defines the unauthorized, sometimes unintentional transfer of information from alleged third parties to unauthorized users. Once the sensitive data is leaked, it pose threat to the functional integrity of an organization. The data exposed in any way cause serious damage to the organization resulting in direct or indirect loss. Direct loss can be measured in terms of cost, indirect loss on the other hand cannot be measured quantitatively that has broader effect on customer relationship with the organizations. Thus, detection of leaked data is principle measure that prevents further threats to an organization.

2. PROPOSED SYSTEM

The proposed system for data leakage detects when the distributor's sensitive data has been leaked by agents. Data can be modified and made less sensitive before giving to agents. There is an option of adding fake objects to the distributed set of data called as 'fake data'. This data is not the actual data that is being viewed by the unauthorised user or agents but it appear realistic to them. The fake data acts as a type of watermark for the data set. If an agent was given one or more fake objects that were leaked, then the distributor can be more confident that agent was guilty.

The admin will upload the original data on cloud and data is in encrypted form and the fake data (predefined) is also uploaded on the cloud. Admin is authorised data owner of the company and has the right to distribute the company's sensitive data to its employees. Admin is authorized to register the new employees, upload the data to server or send the data to the respective employees of the company.

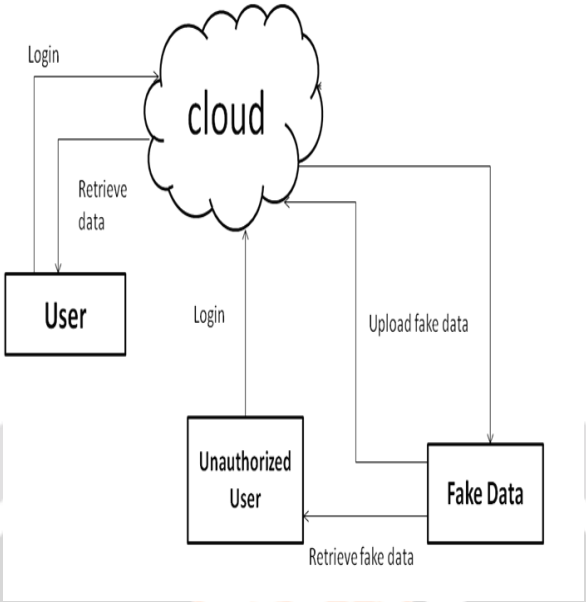

Fig 3.1: System Architecture

- TPA uploads the data on the cloud.
- Employee of the company will login using its credential and retrieve the data from the cloud.
- Predefined fake data along with the original data is uploaded on cloud.
- Unauthorised user using any employees credential will try to login and leak the data but he will receive fake data as he enters the invalid key.

4. MODULES

Admin here act as TPA is authorised data owner of the company and has the right to distribute the company's sensitive data to its employees. Admin is authorized to register the new employees, upload the data to server or send the data to the respective employees of the company.

Admin module contains following functions:-

1. Manage User
2. Upload Data
3. Employee Registration
4. Data Leakage
5. File Details

I. Manage User: - Admin has the authority to maintain the record of the users of its company. Also he has the right to manage the employee records, perform operation on the user details like editing and delete operation.

II. Upload Data: -In this part, the admin has the authority to browse the system and select the data (may be critical or sensitive data of the organization) and upload it on the cloud server. If necessary he has the

option to select the users to whom admin wants to share the data with and will generate the key. The key generated will be forwarded to the selected users through mail.

III. **Employee Registration: -** Admin at this part will have the functionality of new user registration. New employee in the company can get registered through the admin panel.

IV. **File Details: -**Admin will keep a complete record of the uploaded data at the cloud server and the data forwarded by him to the users name of the data which is uploaded will be explicitly mentioned.

5. CONCLUSIONS

The proposed model suggest various data distribution techniques to ensure the detection of the leaked data. The fake data (initially present) deceives the unauthorized user when he tries to access or download the data. In this case admin who handles the transaction of the data is notified when someone try to download the document with the incorrect key. Thus the proposed system helps to detect the leakage of the data set.

6. REFERENCES

1. Panagiotis Papadimitriou, and Hector Garcia-Molina, "Data Leakage Detection", IEEE Transactions on Knowledge and Data Engineering, Vol. 23, NO.1, January 2011.

2. Data Leakage: Affordable Data Leakage Risk Management by Joseph A. Rivela Senior Security Consultant P.P (4-6) Data Leakage Prevention: A news letter for IT Professionals Issue 5 P.P (1-3)

3. Data Leakage: What You Need to Know by Faith M. Heikkila, Pivot Group Information Security Consultant. P.P (1-3)

4. International Journal of Computer Applications in Engineering Sciences **[**VOL I, ISSUE II, JUNE 2011**] [**ISSN: 2231-4946**]** P.P (1, 4) Development of Data leakage Detection Using Data Allocation Strategies Rudragouda G Patil *Dept of CSE,* The Oxford College of Engg, Bangalore.

5. Mr.V.Malsoru, Naresh Bollam/ International Journal of Engineering Research and Applications (IJERA) ISSN:2248-9622 www.ijera.com Vol. 1, Issue 3, pp.1088-1091 1088 | P a g e REVIEW ON DATA LEAKAGE DETECTION.