International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
Since 2012

http://www.ijcsjournal.com Volume 10, Issue 1, No 5, 2022 ISSN: 2348-6600
Reference ID: IJCS-408 PAGE NO: 2777-2780

DEVELOPMENT OF DATA LEAKAGE DETECTION

Jeevasuruthi S[#1], Dr. T. Velumani*[2]

[#]Student, M.Sc Computer Science,
Rathinam College of Arts and Science, Coimbatore,
Tamil Nadu, India – 641021. s.jeevasuruthi23@gmail.com

**Assistant professor, Department of computer science,*
Rathinam college of arts and science, Coimbatore,
Tamil Nadu, India – 641021. velumani.cs@rathinam.in

Abstract - **Our goal is to detect when the distributor's sensitive data has been leaked by interjected, and if possible, to spot the agent that leaked the info. an information distributor has given sensitive data to a collection of supposedly reliable agents. Sometimes data is leaked and located in unconstitutional place. as an example, a hospital may give patient records to researchers who will devise new treatments. Similarly, a corporation may have partnerships with other companies that need sharing customer data. Another enterprise may outsource its processing, so data could be given to varied other companies. The owner of the info is named as distributors and therefore the trusted third parties are called as agents. Data leakage happens a day when confidential business information like customer or patient data, company secrets, budget information etc. Are leaked out. When this information is leaked out, then the businesses are at serious risk. Most likely data are being leaked from agent's side. So, companies need to very careful while distributing such knowledge to agents. The Goal of Our project is to investigate "how the distributer can allocate the confidential data to the Agents in order that the leakage of information would be minimized to a Greater Extent by finding a guilty agent".**

Keywords – Patient Data, Sensitive Data, Agents, Data Leakage.

I. INTRODUCTION

Data Allocation Module:

The main focus of our project is the data allocation problem as how can the distributor intelligently" give data to agents in order to improve the chances of detecting a guilty agent, Admin can send the files to the authenticated user, users can edit their account details etc. Agent views the secret key details through mail. In order to increase the chances of detecting agents that leak data.

Fake Object Module:

The distributor creates and adds fake objects to the data that he/she distributes to agents. Fake objects are group give the wrong secret key to download the file, the duplicate file is opened, and that fake details also send the mail. Ex: The fake object details will display.

Optimization Module:

The Optimization Module is the distributor's data allocation to agents. The agent's constraint is to satisfy distributor's requests, by providing them with the number of objects they request or with all available objects that satisfy their conditions. They objective are to be able to detect an agent who leaks any portion of the data. User can able to lock and unlock the files for secure.

Data Distributor:

A data distributor has given sensitive data to a set of supposedly trusted agents (third parties). Some of the data is leaked and found in an unauthorized place (e.g., on the Merill Technologies group 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. Admin can able to view the file is leaking and fake user's details also.

II. SYSTEM STUDY

EXISTING SYSTEM

As data generation is far outpacing data storage it proves costly for small firms to frequently update their hardware whenever additional data is created. Also maintaining the storages can be a difficult task. In transmitting the file across the network to the client can consume heavy bandwidths. The problem is further complicated by the fact that the owner of the data may be a small device, like

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 10, Issue 1, No 5, 2022 ISSN: 2348-6600
Reference ID: IJCS-408 PAGE NO: 2777-2780

a PDA (personal digital assist) or a mobile phone, which have limited CPU power, battery power and communication bandwidth.

Disadvantages:

1. The main drawback of this scheme is the high resource costs required for the implementation.
2. Computing hash value for even a moderately large data file can be computationally burdensome for some clients (PDAs, mobile phones, etc).
3. Data encryption is large so the disadvantage is small users with limited computational power (PDAs, mobile phones etc.).

PROPOSED SYSTEM

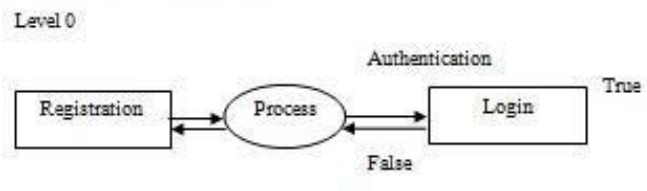
One of the important concerns that need to be addressed is to assure the customer of the integrity i.e. correctness of his data in the cloud. As the data is physically not accessible to the user the cloud should provide a way for the user to check if the integrity of his data is maintained or is compromised. In this paper we provide a scheme which gives a proof of data integrity in the cloud which the customer can employ to check the correctness of his data in the cloud. This proof can be agreed upon by both the cloud and the customer and can be incorporated in the Service level agreement (SLA). It is important to note that our proof of data integrity protocol just checks the integrity of data i.e. if the data has been illegally modified or deleted.

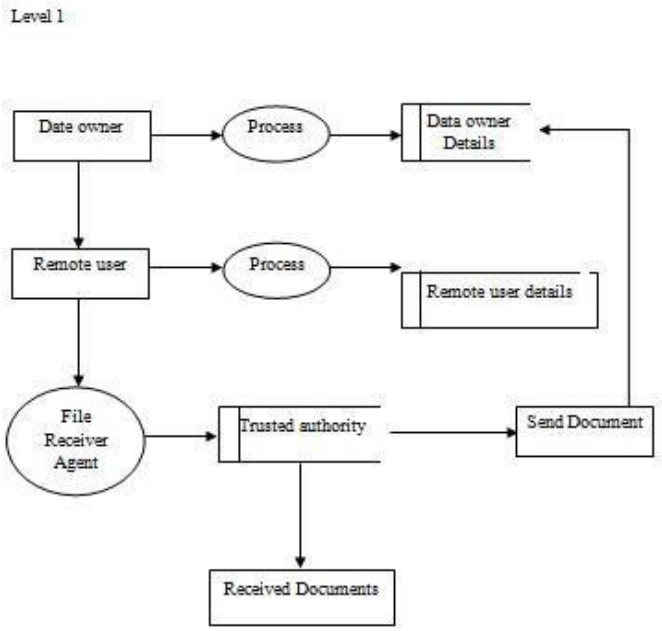
Advantages:

1. Apart from reduction in storage costs, data outsourcing to the cloud also helps in reducing the maintenance.
2. Avoiding local storage of data.
3. Reducing the costs of storage, maintenance and personal.
4. It Minimize the chance of losing data by hardware failures.
5. Owner of the data cannot be cheated.

III. PROPOSED WORK

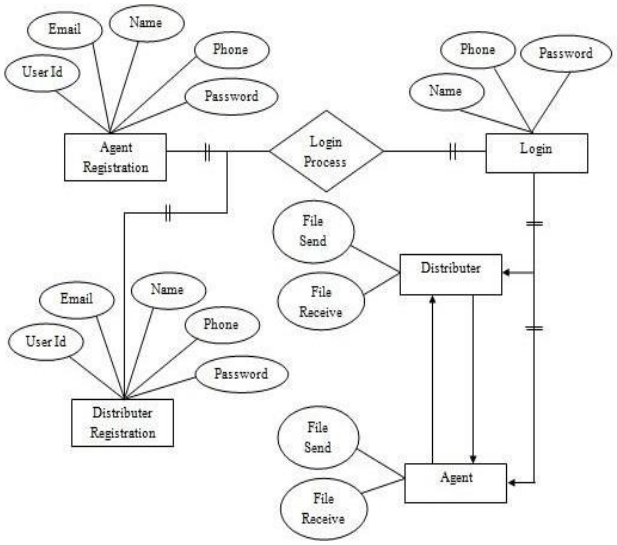
DATA FLOW DIAGRAM





International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 10, Issue 1, No 5, 2022 ISSN: 2348-6600
Reference ID: IJCS-408 PAGE NO: 2777-2780

ER DIAGRAM

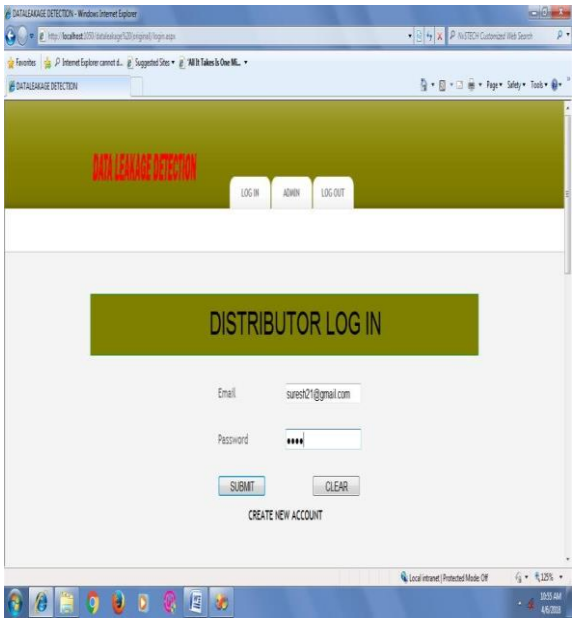


Distributor Details

IV. OUTPUT SCREENS

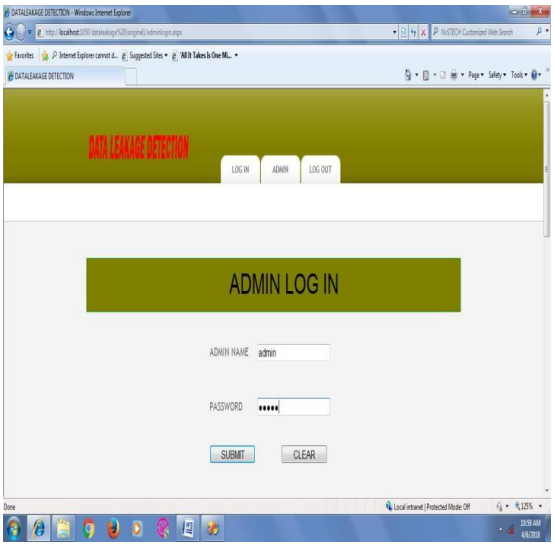
Distributer Registration

Distributor Login

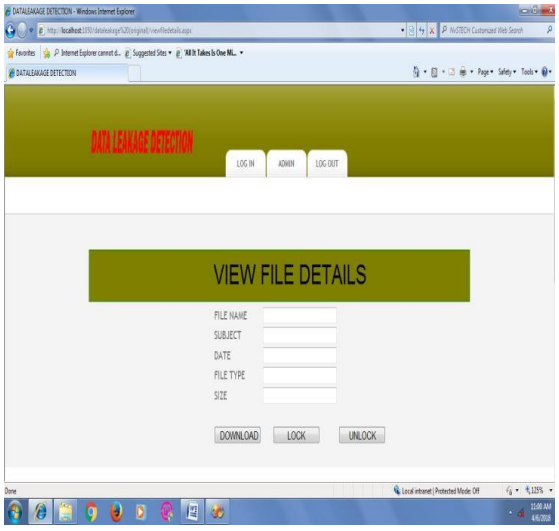


International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 10, Issue 1, No 5, 2022 ISSN: 2348-6600
Reference ID: IJCS-408 PAGE NO: 2777-2780

Admin Login



View File Details



V. CONCLUSION

In this paper we have worked to facilitate the client in getting a proof of integrity of the data which he wishes to store in the cloud storage servers with bare minimum costs and efforts. Our scheme was developed to reduce the computational and storage overhead of the client as well as to minimize the computational overhead of the cloud storage server. The size of the proof of data integrity is minimized to reduce the network bandwidth consumption. Many of the schemes proposed earlier require the archive to perform tasks that need a lot of computational power to generate the proof of data integrity. But in our scheme the archive just needs to fetch and send few bits of data to the client to provide proof of data integrity. Our future work is the extension of our allocation strategies so that they can handle agent requests in an online fashion (the presented strategies assume that there is a fixed set of agents with requests known in advance). Any application does not end with a single version. It can be improved to include new features. Our application is no different from this. In this, we proposed a model for assessing the "guilt" of agents. An algorithm for distributing objects to agents is proposed, in such a way that improves our chances of identifying a leaker. Finally, considering the option of adding "fake" objects to the distributed set. Such objects do not correspond to real entities but appear realistic to the agents. So, if it turns out that an agent was given one or more fake objects, that were leaked, then the distributor can be more confident that agent was guilty.

REFERENCES

[1] X. Shu, et al., "Fast Detection of Transformed Data Leaks," in IEEE Transactions on Information Forensics and Security, vol. 11, no. 3, pp. 528-542, 2016.

[2] X. Shu, et al., "Privacy-preserving detection of sensitive data exposure," IEEE Trans. Inf. Forensics Security, vol. 10, no. 5, pp. 1092-1103, 2015.

[3] F. Liu, et al., "Privacy-preserving scanning of big content for sensitive data exposure with MapReduce," in Proc. 5th ACM Conf. Data Appl. Secur. Privacy (CODASPY), pp. 195-206, 2015.

[4] Nadkarni and W. Enck, "Preventing accidental data disclosure in modern operating systems," in Proc. 20th ACM Conf. Comput. Commun. Secur., pp. 1029-1042, 2013.

[5] R. Hoyle, et al., "Attire: Conveying information exposure through avatar apparel," in Proc. Conf. Comput. Supported Cooperat. Work Companion (CSCW), pp. 19-22, 2013.

[6] H. A. Kholidy, et al., "DDSGA: A data-driven semi-global alignment approach for detecting masquerade attacks," IEEE Trans. Dependable Secure Comput., vol. 12, no. 2, pp. 164-178, 2015.