DATA THEFT DETECTION

Ujjwala Chavan, Bharati Pawar, Prof. H. A. Chavan

Bharati Vidyapeeth College of Engineering, Navi Mumbai

Sector-7, C.B.D, Belpada, Navi Mumbai-400614, India
ujjwalachavan04@gmail.com

Abstract

In the course of doing business, we come across data that is supposed to be confidential and need to be taken care of not being leaked. As this data need to be circulated among the employees, the possibility of data being leaked by these employees of the organization increases. So, as per the system of data theft detection, it will try to find out leakage of data and prevent the leakage in future. System basically consists of admin and cluster of users working in the organization. System contains sensitive data (video) which admin does not want to be leaked so, for the security of system data allocation strategies have been implemented on the admin's side. Only registered users can work with admin and they are authorized to download data i.e. video. If any of the user leaks the data on some other site, the data allocation strategies can be able to track them and thus admin gets the notifications on his/her corresponding email address. After that immediately the admin blocks the particular user. Implementation of the project is done in JSP. To generate a unique key, RSA algorithm is used. The key is embedded in the data by stenography process.

Keywords: *Data Leakage, Key Embedding, Theft Detection.*

I. INTRODUCTION

When it comes to work in large organisation, over there sharing of data becomes a need . For example, 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. In this project, the owner of the data is the distributor and the data is accessed by the users. The project's goal is to detect the user who leaked sensitive data.

II. EXISTING SYSTEM

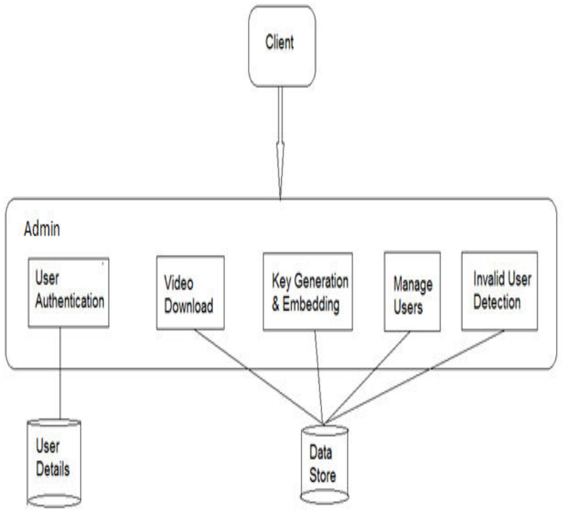
As per the literature survey made, existing system is working on digital watermarks. These watermarks are embedded in the sensitive data by admin. So whenever admin finds the sensitive data on un trusted sites/machines, he can check the watermark and decides whether the data belongs to his site or not.
Demerit:

- The watermarks are removable and can be manipulated. Also the admin can understand, whether the data belongs to his site or not, but fails to find the person responsible for the leakage.

III. PROPOSED SYSTEM

Now-a-days many systems are used for detecting the leaked data by using a number of different software's. In the current system users download videos from the server after registering into the system. Some of these users misuse the data and then distribute the content to other users. By such activates the data falls into high risk. The software's used till now only detect the leaked data. The current system fails to track the user who has downloaded and leaked the video to other unauthorized user. The proposed system works with an exception of embedding dynamically generated key with the video. Thus the user leaking the data is tracked. And thus, this proposed system provides an authority to the admin so that the admin can block the suspected user.
Advantages:
The proposed system empowers admin to track the source user who has leaked the video and take appropriate action upon him

BLOCK DIAGRAM:



The above system architecture consists of various modules. Whenever user enters into the system, he has to register first. If he is already registered user then, he can login directly into the system. The first module works basically on the authentication of the users. So this is how only registered users remain linked with system. There is another part in the system i.e. admin module. Admin have authorization to upload the video and even to manage the data. The users of the system are authorized to download the video. Whenever the user downloads the video, the data allocation strategies which have been implemented in the admin's site runs parallel. A key generation algorithm runs simultaneously and the key is embedded dynamically. The key so formed is random and unique in nature. The user which have downloaded the video along the key information is stored in the admin's repository. So, a subsystem, which basically helps the admin to find the data of his possession. If he founds the data which he thinks belongs to his site, then he decodes the video, extracts the embedded key, and thus finds the match with the data stored in his repository. If match found, the admin easily investigate the un trusted person working in his system, and thus can block the suspected user permanently. This is how, system is able to detect leakage of data and also prevent it.

IV. CONCLUSION

This system proves beneficial in detecting the exact user who leaked the data, and thus overcomes this drawback of existing system i.e. watermarking. Over model is relatively simple which actually overlaps the data present at the admin's site with the data found in the unauthorized place.

REFERENCES

Book,

[1] R. Agrawal and J. Kiernan, Watermarking Relational Databases, Proc. 28th Intl Conf. Very Large Data Bases (VLDB 02), VLDB Endowment, pp. 155-166, 2002.

[2] F. Hartung and B. Girod, Watermarking of Uncompressed and Compressed Video, Signal Processing, vol. 66, no. 3, pp. 283-301, 1998.

Journal Paper,

[3]Panagiotis Papadimitriou, Student Member, IEEE, and Hector Garcia-Molina, Member, IEEE.

Proceeding paper,

[4]J.J.K.O.Ruanaidh,W.J.Dowling,and F.M. Boland,"Watermarking Digital Images for Copyright Protection,"IEEEProc.Vision,Signal and Image Processing, vol. 143 ,no. 4, pp.250-256, 1996.