Volume 6, Issue 1, January 2016 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

An Improved Privacy Preserving Technique to Avoid Sensitive Data Leakage Using t-Closeness Technique: A Review

Pranjali A Meshram*, Ranjana Shende
Computer Science & Engineering, RTMNU, Nagpur,
Maharashtra, India

Abstract— In government organizations and in research institutions there is huge amount of data that need privacy .And data privacy is becoming the big issue day by day. In this paper we will study method which will provide privacy to sensitive data. The method called fuzzy fingerprints and DLD (Data leak Detection) are the techniques which provide privacy to sensitive data and leaks if any. t-closeness technique is used to make Fuzzy Fingerprints. Privacy preserving data-leak detection (DLD) solution is used to solve the issue of special set of sensitive data digests. It enables the owner of data to safely give the detection operation to a semi-honest provider without revealing the sensitive data to the provider.

Index Terms- Fuzzy Fingerprints, Data leak, Sensitive Data, Privacy

I. INTRODUCTION

According to various reports from Risk Based Security, the number of leaked sensitive data records has increased dramatically. Attacks, inadvertent leaks, and human mistakes lead to the data-leak incidents. Detecting and preventing data leaks requires total solutions, which may include data-leak detection [1].

Fuzzy fingerprint technique enhances data privacy during data-leak detection operations [5]. It enables the data owner to securely give the content-inspection task to DLD providers without exposing the sensitive data. In detection procedure, the data owner compute fuzzy fingerprints from the sensitive data and then discloses only a small amount of data to the DLD provider [3] who is semi-honest. The DLD provider computes fingerprints from network traffic and identifies potential leaks in them. Using detection method, the DLD provider, who is modeled as an honest-but-curious (semi-honest) provider, can only get limited knowledge about the sensitive data from either the released digests, or the content being inspected [2].

II. RELATED WORK

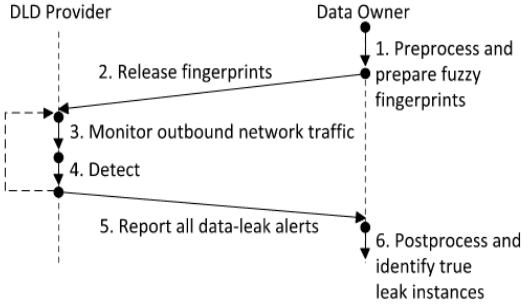
Now a day's privacy of sensitive information is big issue. In the previous research they used the different techniques and algorithms to give privacy to sensitive data, but it is still necessary to improve their performance in terms of accuracy rate because they got failed to provide good output.

Also, the previously proposed technique l-diversity does not support reverse engineering. While using l-diversity method there are more chances of leakage of sensitive data than t-closeness technique.

III. PROPOSED RESEARCH

In this paper, data leak detection solution is provided for the privacy of sensitive data. Fuzzy fingerprints is one of the most important term in data leak detection method. In the proposed methodology data is presented in tabular form. From the table sensitive data and non-sensitive data fields are differentiated. for e.g. an organisation having data of person having data fields name, age, address, account no. etc. In this name, age, and address are non-sensitive data fields and account no. is sensitive data field. After retrieving sensitive data field, we compute fuzzy fingerprints from sensitive data fields. These fuzzy fingerprints are computed using t-closeness. In previous method, fuzzy fingerprints were computed using l-diversity technique which has some disadvantages. t-closeness is successor of l-diversity technique. Fuzzy fingerprints of sensitive data organization can safely give to semi-honest provider who can find the data leaks if any.

The diagrammatic representation of the data leak detection method is given below



IV. PROPOSED MODEL AND OVERVIWE

In the privacy preserving data model there are two most important players

1. The Organization
2. Data leak detection (DLD) provider

1. The Organization

The Organisation is a data owner who has sensitive data which need privacy and give permission to the Data Leak Detection (DLD) provider to inspect the data contents but the organisation does not want to disclose the information to DLD providers.

2. DLD provider

DLD providers inspect the sensitive data and inform the organisation that if any data leak is present or not. While inspecting for the data leaks the DLD provider may try to get the sensitive (Private) information, to prevent this only the organisation make fuzzy fingerprints of the sensitive data which he want to inspected using t-closeness method. And if DLD provider found any data leak then it will send alerts to the organisation.

V. EXPECTED OUTCOME

1. t-closeness method will provide full privacy and security to sensitive data fields.
2. DLD server will send information about the data leaks if any.

VI. CONCLUSION

Fuzzy fingerprint and a privacy-preserving data-leak detection model provide privacy to data contents. Using special digests, the exposure of the sensitive data is kept to a minimum during the detection. Fuzzy fingerprints which are designed using t-closeness method provide more security and privacy to sensitive data than l-diversity method.

ACKNOWLEDGMENT

I thank to Ms. Ranjana Shende for her valuable guidance and also thank to IEEExplore.org for references. .

REFERENCES

[1] X. Shu , D Yao "Privacy Preserving Detection of Sensitive Data Exposure " in IEEE transaction on information forensics and security . VOL.10 N 5, may 2015

[2] B. Wang, S. Yu, W. Lou, and Y. T. Hou, "Privacy-preserving multikeyword fuzzy search over encrypted data in the cloud," in Proc. 33th IEEE Conf. Comput. Commun., Apr./May 2014, pp. 2112–2120.

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

[4] A. Kapravelos, Y. Shoshitaishvili, M. Cova, C. Kruegel, and G. Vigna, "Revolver: An automated approach to the detection of evasiveweb-based malware," in Proc. 22nd USENIX Secur. Symp., 2013, pp. 637–652.

[5] X. Shu and D. Yao, "Data leak detection as a service," in Proc. 8th Int. Conf. Secur. Privacy Commun. Netw., 2012, pp. 222– 240.

[6] Jordi Soria-Comas_ and Josep Domingo-Ferrer," Differential Privacy via t-Closeness in Data Publishing,"in proc 18th IEEE conf may 2011 pp.563-574

[7] X. Jiang, X. Wang, and D. Xu, "Stealthy malware detection and monitoring through VMM-based 'out-of-the-box' semantic view reconstruction," ACM Trans. Inf. Syst. Secur., vol. 13, no. 2, 2010, p. 12

[8] K. Borders, E. V. Weele, B. Lau, , "Protecting confidential data on personal computers with storage capsules," in Proc. 18th USENIX Secur. Symp., 2009, pp. 367–382

[9] H. Yin, D. Song, M. Egele, C. Kruegel, and E. Kirda, "Panorama: Capturing system-wide information flow for malware detection and analysis," in Proc. 14th ACM Conf. Comput. Commun. Secur., 2007, pp. 116–127.

[10] Ninghui Li, Tiancheng Li , Suresh Venkata subramanian," t-Closeness: Privacy Beyond k Anonymity and -Diversity " in Proc 14th ACM Conf. Comput. Commun Secur,2007, pp.255-270

