ISSN: 2321-7782 (Online)
Volume 2, Issue 2, February 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Paper / Case Study
Available online at: www.ijarcsms.com

A Technique for Avoiding Data Leakage and Misuse

Rashmi K. Bhat[1]
Department of Computer Science & Engg.
G.H. Raisoni College of Engineering
Nagpur – India

Nikita V. Mahajan[2]
Department of Computer Science & Engg.
G.H. Raisoni College of Engineering
Nagpur – India

Abstract: ***Data leakage, especially caused by insider threat is one of the important issue in information security research. This is because insider threats have potential to inflict severe damage to the organization's resources, financial assets and reputation. This paper reviews the different methods used for mitigating data leakage and misuse detection. Misuseability weight concept is very useful for data leakage and misuse detection since it measures the harm or the risk of damage that can be caused if the important data falls into the wrong hands. It assigns a score which estimates the sensitivity level of damage so that security measures can be applied in order to avoid data leakage and misuse detection. For this, user profile is created using the local knowledge base so that only required data can be given for avoiding data leakage and misuse.***

Keywords: data leakage, data misuse, insiders, misuseability weight, user profile.

I. INTRODUCTION

Organization's data is extremely important and proves as a main constituent in embodying the core of the organization's power. Organization definitely wants that to this power should be preserved and retained. On the other side, this data is required for daily working on different processes. Users within the organization (e.g., employees, subcontractors, or partners) perform different actions on this data and may be exposed to the important information while accessing the data[1],[16]. This may lead to data leakage and misuse.

Detection or avoidance of data leakage and misuse is a great challenging issue for organizations. Whether caused by the malicious intent or an inadvertent mistake, data leakage and misuse can diminish a company's brand and damage the company's reputation. This challenge becomes more difficult when trying to detect and/or prevent data leakage and misuse performed by an insider having legitimate permissions to access the organization's systems and its sensitive data. An insider is someone who has authorized access, privileges or knowledge of database and is motivated to violate the security policy of the system through authorized access[2]. While an insider attack or, as widely used in the literature, insider threat can be defined as a misuse of given privileges by a legitimate member of an organization for malicious purposes. With the widespread of Internet, the insider security threat is more serious. However, in a long period, researchers have different understanding of insider. According to latest 2010 Cyber security Watch Survey[3], on average 26% of cyber security events in the surveyed organizations were known to be caused by the insiders with another 24% coming from unknown sources. These insiders were the most damaging with 43% of the respondents reporting that their organization suffered data loss.

In short, the risk to data security from insiders is becoming more and more crucial because of the endless use of the computers and communication systems. Many mechanisms have been proposed for protecting data from outside attacks. Unfortunately, those mechanisms fail to protect data from authorized users who may misuse their privileges in carrying out malicious activities. So dealing with mechanisms that protect sensitive data against insiders has become a serious demand due to the amount of harm that can be caused by those malicious insiders.

Numerous attempts have been done to deal with the malicious insider scenario and many of these devised methods are generally based on behavioral profiles of the user that define normal user behavior and issue an alert whenever a user's behavior

significantly differs from the normal profile. For this, the most common approaches involves analyzing the SQL statements submitted by an application server to the database[12],[15] and analyzing the actual data exposed to the user[14].

But, still any of the approach does not prove to be effective and due to this fact the goal of the research is to use misuseability weight concept which will help in avoiding data leakage and misuse.

II. Related Work

In the recent years, several methods have been proposed to deal with the problem of data leakage and misuse in database systems, especially caused by the insider. A typical data leakage pattern of common software is shown in fig.1. It involves three typical kinds of data leakage patterns.

(1) Data is leaked to other process by process having accessed sensitive documents through inter-process data channel, then is written to non-security domain by other process. The channels involves shared memory, messaging channels and communication channels.

(2)Trusted process has to cross security domain and non-secure domain to complete its function. This leads to data leakage.

For example: In Microsoft WORD program, even if the temporary files are deleted in the end of process, they can also be recovered from the residual traces by disk recovery software, resulting in data leakage.

(3)Sensitive files and the network or removable storage devices resulting in data leakage are accessed by the processes.

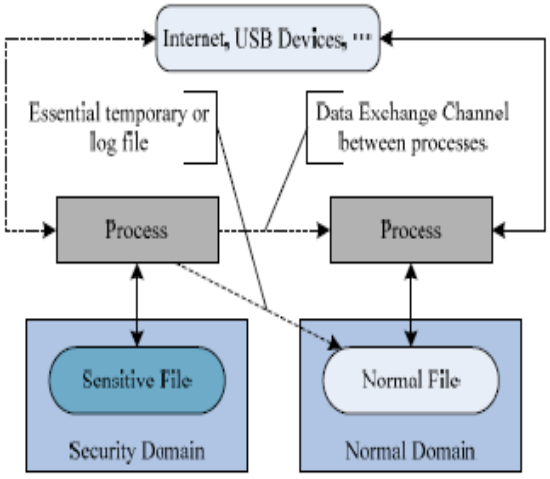

Fig1: Typical Data Leakage Patterns of Common Software

Jie, Zhou Jiangjiang Wu, Jun Ma, Songzhu Mei, Jiangchun Ren [3] presented an active data leakage prevention model which adds secure data container(SDC) to execute security prevention mechanism. By using SDC, the model can ensure that data is used in a trusted and controllable environment. of three dynamic isolation processes “read isolation”, “write isolation” and “communication isolation”, model ensures the availability of trusted process and achieves data leakage protection.

Celikel et al.[4] proposed a novel risk management system to provide enhanced security in Role Based Access Control(RBAC) employed distributed databases. Within the design, it considers risks imposed by illegitimate credentials, role misuse/abuse and system failure to the system. Another interesting work[5] given by Q. Yaseen and B. Panda investigates the problem of knowledge acquisition by an unauthorized insider using dependencies between objects in relational databases. Also, they introduces an algorithm to determine the knowledge base of an insider and explains how insiders can broaden their knowledge about various relational database objects to which they lack appropriate access privileges. For preventing insiders from misusing their privileges, Gates and Bishop[6] suggested the Group Based Access Control(GBAC) mechanism. This mechanism along with the user’s basic job description(i.e role)uses the user characteristics and behavioral attributes such as the time he/she normally comes to work or the customers with whom he/she usually interacts. A multi-perspective approach to insider threat detection given by Majid Raissi-Dehkordi and David Carr[7] focuses on detection of insider threats in

organizational networks and provides a logical framework for more accurate identification of malicious actions by legitimate users of the system. Combination of usage profiles defined for entities in the network beyond the traditional user profiles allows this system to achieve higher detection accuracies. To quantitatively limit the information leak Ning zhang, Ming Li, Wenjing Lou[8] introduced the model of data mining with differential privacy. This model proposes secure group differential private query(SDC) which combines techniques from differential privacy and secure multiparty computation. Also various issues related to private information leakage in social networks[10],[11] are addressed by different authors.

Although the existing work presented different techniques to deal with problem of data leakage and misuse detection, they do not prove to be much effective and therefore there is a need of the effective mechanism that will provide solution to the problem of data leakage and misuse detection.

III. Methodology For Data Leakage and Misuse Detection

A. *User Profile*:

Detection of an insider threat which is which mainly responsible for the data leakage and misuse, typically construct the normal user profile for each user by recording appropriate usage metrics. As shown in fig.2, user profile metrics can be typically collected from one of the two categories:

- Host or
- Network

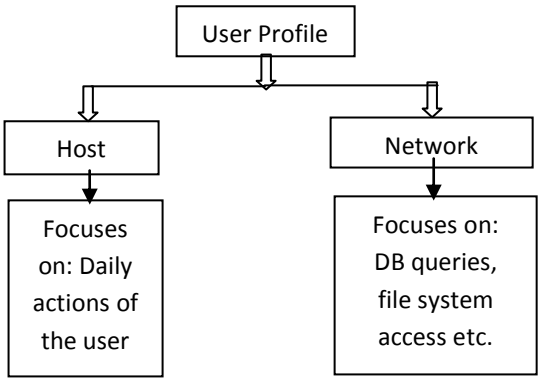

Fig.2 User Profile

In pure host monitoring, daily actions of the user on his/her machine are logged and the user behavior is modeled by using metrics such as the sequences of the OS commands, applications used by the user. Another way that is network monitoring focuses more on network aspect of user actions, such as database queries, file system access, printer access etc.

Despite the differences among the above methods, some form of profile is constructed for each user and that profile is used as the baseline for monitoring data leakage and misuse.

B. *Overview Of Misuseability Weight Concept*:

In data leakage and misuse detection mechanism, misuseability weight concept determines the extent or risk of damage that the user can cause using the information he/she has obtained[1].

Assigning a misuseability to the given dataset is closely related to the way of presentation of data(e.g. structured or free text, tabular data). So, all types of data cannot be fitted be one measure of misuseability weight. Therefore four general dimensions of misuseability are given, which serves as guidelines while defining a misuseability weight measure. These dimensions are:

(1)Number of entities: This refers to the data size with respect to the different entities that involves in the data. Data about more entities obviously increases the potential damage.

(2)Anonymity level: It is defined as the effort that is required to fully identify the specific entity in the data. Anonymity level of data can decrease the misuseability weight.

(3)Number of properties: A variety of different properties can be included by data(e.g., employee income). As each additional property can increase the damage, number of different properties can affect the misuseability weight.

(4)Values of properties: Different properties of an entity can have different importance. Therefore property value of an entity can strongly affect misuseability level of data.

IV. PROPOSED PLAN OF WORK

From the previous discussion, it is observed that though many of the suggested methods contributed a lot towards the data leakage and misuse detection, there is a need of the system which will give more efficient and improved result. Thus proposed plan of work involves following steps which will try to resolve the problems occurred in the previous system. These steps can be given as follows:

(1) Making user profile using local knowledge base

(2) Analyze database for different dimensions of misuseability

(3) Calculate sensitivity level or risk of damage of the tabular data

(4) Altering the data to avoid data leakage and misuse

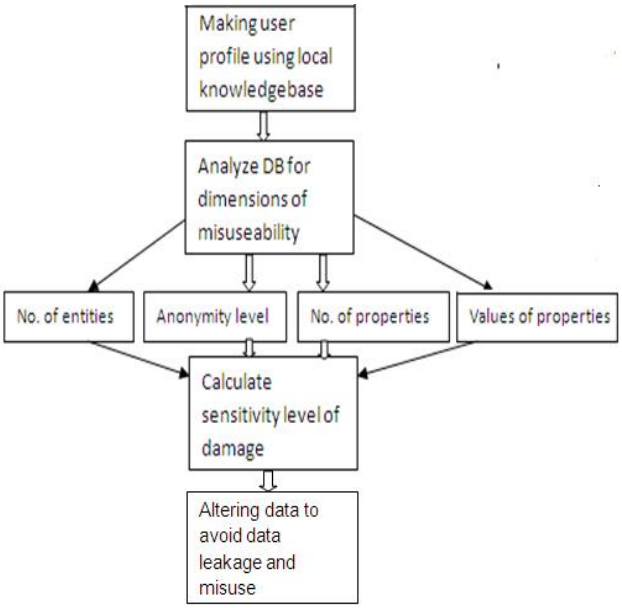

Fig.3 Steps of proposed work

As shown in fig.3, the very first step of the system is making a user profile which is based on the information submitted by the employee as well as the kind of work done by him. On the admin login, administrator is able to see the list as well as the complete information of all these employees. In the next step, tabular data or the database is analyzed for different dimensions of misuseability. This proves very helpful for calculating sensitivity level of the damage which is done in the next step. Measuring the sensitivity level of damage involves collection of information from the domain expert. Based on this, sensitive attributes are selected which help in calculating the sensitivity score as well as the misuseability score. For this, it mainly considers three factors which can be given as:

- Quality of data
- Quantity of data

- The distinguishing factor

Finally in the last step, security measures are applied to avoid data leakage and misuse. That is, here only the required or altered data will be sent to the user according to his profile in such a way that the data will be useful for the further work and at the same time data leakage will not happen.

V. CONCLUSION

This paper reviews the various techniques for data leakage and misuse detection. But still none of the techniques gives the sensitivity level of the damage caused to data while providing the data to the insider.

Thus, this paper introduces a system which will measure the risk of damage that can be caused when data is exposed to the insider. This involves collecting knowledge from the domain expert as well as use of the risk measuring algorithm. Measuring risk before data exposure will help administrator to take proper action to prevent or minimize the damage. Also, this system will alter the data in such a way that the risk will be reduced and at the same time modified data will be useful for performing desired task.

References

1. Amir Harel, Asab Shabtai, Lior Rokach and Yuval Elovici, "Mscore:A misuseability Weight Measure" , may/june 2012 , IEEE Transactions on dependable and secure computing, vol. 9, no. 3.
2. Jiangjiang Wu, Jie Zhou, Jun Ma, Songzhu Mei, Jiangchun Ren, 2011, "An Active Data Leakage Prevention Model for Insider Threat", IEEE Computer Society.
3. 2010 CyberSecurity Watch Survey, 2012, http://www.cert.org/archive/pdf/ecrimesummary10.pdf
4. E. Celikel et al., "A Risk Management Approach to RBAC ", 2009, Risk and Decision Analysis, vol. 1, no. 2, pp. 21-33.
5. B. Panda and Q. Yaseen, "Knowledge Acquisition and Insider Threat Prediction in Relational Database Systems", 2009, Proc. Int'l Conf. Computational Science and Eng., pp. 450-455.
6. C. Gates and M. Bishop, "Defining the Insider Threat,", 2008, Proc. Ann. Workshop Cyber Security and Information Intelligence Research, pp. 1-3.
7. Majid Raissi-Dehkordi and David Carr,"A Multi-Perspective Approach to Insider Threat Detection", 2011, Military Communications Conference, IEEE.
8. Ning Zhang, Ming Li, Wenjing Lou, "Distributed Data Mining with Differential Privacy ", 2011, IEEE ICC.
9. K. Wang, C. M. Fung R. Chen, and P.S. Yu, "Privacy-Preserving Data Publishing: A Survey on Recent Developments," , 2010, ACM Computing Surveys, vol. 42, no. 4, pp. 1-53.
10. A. Friedman and A. Schuster, "Data Mining with Differential Privacy," , 2010, Proc. 16th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining, pp. 493-502.
11. C. Dwork, "Differential Privacy: A Survey of Results" , 2008, Proc. Fifth Int'l Conf. Theory and Applications of Models of Computation, pp. 1-19.
12. E. Terzi, A. Kamra, and E. Bertino, "Detecting Anomalous Access Patterns in Relational Databases" , 2008, Int'l J. Very Large Databases, vol. 17, no. 5, pp. 1063-1077.
13. F. Duran, S. H. Conrad, G. N. Conrad, D. P. Duggan, and E. B. Held,"Building a system for insider security" , Nov.-Dec 2009, IEEE Security & Privacy,Vol. 7, Issue 6, pp 30-38.
14. S. Mathew, M. Petropoulos, H.Q. Ngo, and S. Upadhyaya, "Data-Centric Approach to Insider Attack Detection in Database Systems", 2010, Proc. 13th Conf. Recent Advances in Intrusion Detection.
15. Elisa Costante, Sokratis Vavilis, Sandro Etalle, Jerry den Hartog, Milan Petkovi´c and Nicola, "Database Anomalous Activities: Detection and Quantification",2007.
16. Predd, J. et al. "Insiders Behaving Badly" ,July-Aug. 2008 , IEEE Security & Privacy, vol. 6, no. 4, pp. 66-70.

AUTHOR(S) PROFILE

Rashmi K. Bhat received the Bachelors degree in Information Technology from KDK College of Engineering in 2011. She has 3 months of teaching experience from Government Polytechnique College, Nagpur. Her main area of interest includes Data mining and Information Security. She is now pursuing Masters in Technology in Computer science and Engineering from Raisoni College of Engineering, Nagpur.

Asst.Prof. Nikita V. Mahajan is currently working as assistant professor in G.H. Raisoni College of Engineering. She has completed Masters in Technology in Computer science and Engineering from Raisoni College of Engineering, Nagpur. She has an experience of teaching for the 3 years. She has many national and international publications to her credit.