IOSR Journal of ComputerEngineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 22, Issue6, Ser. I (Nov. – Dec. 2020), PP 23-26
www.iosrjournals.org

Prevention of Data LosswithLabelsin Education

Cigdem Bakir
Yildiz Technical University, ComputerDepartment, 34015, Istanbul, Turkey

Abstract
Today, unauthorized access to data, propagation and modification of data bring many problems. Data leakage techniques are used to solve these problems. However, a model that will ensure data privacy by protecting the data of all patients in hospitals has not been fully developed. In our study, education data are defined as objects. Each object is labeled while being passed on to another user. Each object is allowed to be seen and processed by the users authorized by the user it owns. Co-owners of an object are authorized to operate on that object. Thus, each user performs the security management of his / her data in training.

Keys: *data leakageprevention, data access, authorization, data breach, privacy*

Date of Submission: 22-11-2020 Date of Acceptance: 07-12-2020

I. Introduction

Data leakage detection and prevention means data loss (DLP), data protection, information leakage prevention. Prevention of data leakage are techniques aimed at detecting data theft and protecting data by monitoring the access, use or transmission of data by unauthorized or unintentional persons. Briefly, it is the prevention of leakage of sensitive and valuable data from the transportation channels from the source to the target [1,2]. Thus, the movement of sensitive data on the network or in end-user systems is monitored and controlled [3,4].

Today, data breach, data propagation and data exposure pose a huge problem for many organizations. DLP techniques try to prevent attackers from breaching data. However, these methods cannot fully control the data traffic on the network. Data privacy, data integrity and a method that enables authorized users to access the system is required to control multiple nodes. Because organizations spend a lot of time and money to take security measures and to reduce risks. In addition, it is necessary to raise awareness of the users on this issue [5].

Each institution creates its own local security policies to prevent data leakage and data loss. However, it is very difficult for institutions to apply these policies to the system, to reduce the risk of data breaches, to improve compliance, to recognize malicious software, to optimize network bandwidth, to manage data, and to reduce time and costs. In particular, most security breaches are caused by intentional or unintentional behaviour by users within the organization. This necessitates the protection of personal information, that the data is not shared by unauthorized users, that it is not copied, and that the data are followed in communication paths. In addition, in case of loss of data, it must be securely backed up and stored.

II. Method

The intellectual property rights of companies and organizations, financial information, confidential information about patients in hospitals, information about diagnosis and treatment process, credit card information about customers in banks or other important information used in the industry constitute sensitive data. Leakage of this information to the outside by both people outside the organization and internal personnel brings some serious problems such as cost and time. For this reason, prevention of data leakage is of great importance in institutions and organizations. It is especially used in mobile devices, cloud computing, databases, and filing systems. Data leakage prevention techniques are shown in Table 1. [2].

Table 1: Data leakpreventiontechniques

Categories	UsedMethods
Defined DLP methods	Data in motion Data in use Data in rest
Access Control &Encryption	Device Control Encription Right Management Service
Advanced/Intelligent Security Metrics	Anomalydetection Activity basedverification
Standard Security Measures	Firewall Anti-viruses Intrusiondetectionsystems

1) Defined DLP methods: Techniques that prevent sensitive data from being sent, forwarded and copied to unauthorized persons, either intentionally or unintentionally. It shows that the objects (data in motion) on the broadcast node are in constant motion on the network. http, SMTP, P2P protocols, instant messaging, e-mail data protection are involved under this group [2]. The object used by the end user (data in use), means that the objects that are processed on the running node are in continuous use. The objects used are stored in the storage node (data in rest). Stored objects are stored in databases, file systems, and desktop computers as documents or files.

2) Access Control & Encryption: The text is encrypted with a key for unauthorized access to data. Data leakage can be prevented by decoding the encrypted text [3]. RMS (Right Management Systems) is used to protect sensitive file systems.

3) Advanced/Intelligent Security Metrics: Machine learning algorithms are used to detect abnormal behaviours in accessing data. Anomaly detection detects previously unseen attacks. It looks at events that are not considered normal. Users log into a system with their username and password. However, sometimes they choose passwords that are easily found or forget their passwords. In this case, users must be authenticated based on their behaviour and the actions they take. Users are authenticated with activity-based verification systems.

4)Standard Security Measures: Firewalls, intrusion detection systems and anti-viruses fall into this group. Firewall checks incoming and outgoing packets over the network, such as IP filtering, content filtering. Intrusion detection systems, on the other hand, examine the status of the system, detect an attack or data security problem, and work to eliminate this problem.

Application Example

Iftheowners of a data labeledwith L set multiplepolicies, onlyreaders at theintersection set of allreadersetsreadthosepolicies.

K: total number of policies

i: anypolicies (1≤i≤K including)

oK_i :

the set of data owners of policy i

rK_i : the set of data readers of policy i

oK: the set of data owners of allpolicies

rK: Letallpoliciesrefertothe set of readers.

Inthisexample, wewanttoconveyinformationabouttheacademictaskamongtheusers in thefacultygroup.

Users={ Yrd.Doç_X, Doç.Dr_Y, Prof.Dr_Z, Arş.Gör_A, Doç.Dr_B, Prof.Dr_C, Arş.Gör_D, Prof.Dr_E, Doç.Dr_F, Yrd.Doç_H} Show alltheacademics in thedepartment.

L= { Prof.Dr_Z: Yrd.Doç_X , Doç.Dr_Y, Yrd.Doç_H, Prof.Dr_E, Doç.Dr_B;
Prof.Dr_C: Yrd.Doç_X, Arş.Gör_A, Doç.Dr_Y, Doç.Dr_B, Prof.Dr_E;
Doç.Dr_F: Doç.Dr_Y, Prof.Dr_E, Arş.Gör_D, Doç.Dr_B}

All data owners set of the L labelareshown in equation 1.

$$oK = \coprod_{i=1}^{K} oK_i = \{\text{Prof. Dr\_Z}, \text{Prof. Dr\_C}, \text{Doç. Dr\_F}\} \quad (1)$$

K_1policyflow set(X_1)

Prof.Dr_Zeducation dataProf.Dr_Z

Prof.Dr_Zeducation dataYrd.Doç_X

Prof.Dr_Zeducation dataYrd.Doç_Y

Prof.Dr_Zeducation dataYrd.Doç_H

Prof.Dr_Zeducation dataProf.Dr_E

Prof.Dr_Zeducation dataDoç.Dr_B

rK_1={ Prof.Dr_Z , Yrd.Doç_X , Doç.Dr_Y, Yrd.Doç_H, Prof.Dr_E, Doç.Dr_B}

K_2policyflow set (X_2)

Prof.Dr_Ceducation dataProf.Dr_C

Prof.Dr_Ceducation dataYrd.Doç_X

Prof.Dr_Ceducation dataArş.Gör_A

Prof.Dr_Ceducation dataDoç.Dr_Y

Prof.Dr_Ceducation dataDoç.Dr_B

Prof.Dr_Ceducation dataProf.Dr_E

rK_2={ Prof.Dr_C, Yrd.Doç_X, Arş.Gör_A, Doç.Dr_Y, Doç.Dr_B, Prof.Dr_E }
K_3policyflow set (X_3)
Doç.Dr_FeducationdataDoç.Dr_F

Doç.Dr_Feducation dataDoç.Dr_Y

Doç.Dr_Feducation dataProf.Dr_E

Doç.Dr_Feducation dataArş.Gör_D

Doç.Dr_Feducation dataDoç.Dr_B

rK_3={ Doç.Dr_F, Doç.Dr_Y, Prof.Dr_E, Arş.Gör_D, Doç.Dr_B }
Theall set of readers of the L label is shown in equation 2.

$$rK = \coprod_{i=1}^{K} rK_i = \{\text{Prof. Dr\_Z , Yrd. Doç\_X , Doç. Dr\_Y, Yrd. Doç\_H, Prof. Dr\_E, Doç. Dr\_B, Prof. Dr\_C , Arş. Gör\_A, Doç. Dr\_F, Arş. Gör\_D}\} \quad (2)$$

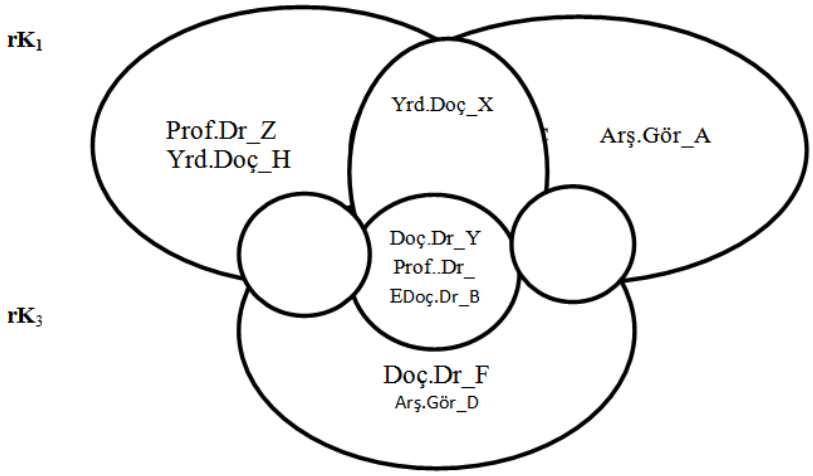

Figure 1 rK_1, rK_2 ve rK_3intersection of reader sets

Figure1showstherK$_1$, rK$_2$ ve rK$_3$readerclusters of the data labeledwiththe L. Accordingtothiscluster, readers in theintersection set of allreadersetswithrK$_1$ηrK$_2$ηrK$_3$={ Doç.Dr_Y, Prof.Dr_E, Doç.Dr_B}readthistraining data.

III. Conclusion

Privacy, security and confidentiality of personal data is to prevent the patient's information from being viewed other than authorized persons. Our work aims to protect personal sensitive data by providing security and privacy.

Data should be monitored, audited and recorded against cyber attacks. In short, it is necessary to ensure confidentiality, security, data integrity, traceability, control of data, access by authorized users. It is important to determine which users with which authorizations the data will be given and their access rights. Access to confidential data should be prevented by both internal personnel and external users. Data transmission, sharing, access to, viewing, use of authorized users, protection against cyber attacks, ensuring confidentiality, integrity and confidentiality, and performing risk analysis are the most important problems. In our study, a common

consent management was provided that could protect the data of each user. With the labelling method, the patient determines their local policies for confidentiality and integrity. Thus, each user performs his/her own consent management.

References

[1]. Prathaben K., "Data LossPrevention", *SansInstituteInfosec Reading Room, 2008.*
[2]. Asaf S., Yuval E. AdnLion R., "A Survey of Data LeakageDetectionandPrevention Solutions", 2012.
[3]. Jorge B., Julio C. andJuan E.T., "BypassingInfprmationLeakageProtectionwithtrustedapplications", *Computer& Security,* pp.557-568, 2012.
[4]. "Data LeakPrevention", Isaca White Paper, 2010.
[5]. PrathabenKanagasingham,"Data LossPrevention", *SansInstituteInfoSec Reading Room,* 2008.
[6]. Olca E., Can Ö., "Türkiye'de Elektronik Sağlık Kaydı Bağlamında Gizlilik ve Güvenlik Üzerine Teknolojiler", *3rd International Symposium on DigitalForensicsand Security",* 2015.
[7]. Öğütçü G., Köybaşı S.C, "Elektronik Sağlık Kayıtlarının İçeriği, Hassasiyeti ve Erişim Kontrollerine Yönelik Farkındalık ve Beklentilerin Değerlendirilmesi", pp.88-97, 2015.
[8]. İzgi M.C., "Mahremiyet Kavramı Bağlamında Kişisel Sağlık Verileri", *Türkiye Biyoetik Dergisi,* vol.1, no.1, pp.25-37, 2014.
[9]. T.Pasquier, J.SinghandD.Eyers, "Information FlowAuditforPaaSclouds", *IEEE International Conference on CloudEngineering (IC2E),* 2016.
[10]. T.PasquierandD.Eyers, "Information FlowAuditTransparencyandCompliance in the Handling of Personal Data", *In IC2E International Workshop on Legal Technical andScience,* 2016.
[11]. Turgut N., Karaarslan E., Ergin A., Kılıç Ö., "Elektronik Sağlık Kayıtlarının Gizlilik ve Mahremiyeti", 2015.
[12]. Kişisel Verilerin Korunması Kanunu, 2016.

Cigdem Bakir. "Prevention of Data LosswithLabelsinEducation." *IOSR Journal of ComputerEngineering (IOSR-JCE),* 22(6), 2020, pp. 23-26.