Two Level Signature Based Authorization Model for Secure Data Warehouse

Anjana Gosain[1] and Amar Arora[2](✉)

[1] USICT, Guru Gobind Singh Indraprastha University, Delhi, India
[2] National Informatics Centre, DeitY, Government of India, Delhi, India
amar.arora@nic.in

Abstract. Data Warehouse (DW) security has emerged as a crucial aspect since for the sake of high availability data warehouses started connected to internet. In order to comply with the security requirements, the authentication of legitimate users by verification of user credentials like username, password, etc. has become a standard. On successful verification, different variations of Role Based Access Control (RBAC) techniques are being used restricting user access to the facts and dimensions. But these RBAC's can only restrict the user access as per their respective roles and there is no check on the behavior pattern of the user access. In this paper, a two level signature based behavior analysis model has been introduced to keep a check on the user's access pattern. At the first level, the user provides its authentication credentials. On successful verification of these credentials, the user has been allowed to access elements as per its role. Once the user tries to access the DW elements his access pattern will be recorded to form usage access signature. Over the period of time user access profile is created which is used to match the signature of the user. If in case, the user's signature does not fit in the user access profile created over the period of time, the second level of verification will be performed in a form of secret question etc. The user query will be processed only on successful clearing of the second authentication level; else the current query will be suspended with regret message from the system. This further strengthens the security of the DW even on the compromise of the user's initial entry credentials.

Keywords: Data warehouse security · Signature based authorization · Two level authorization

1 Introduction

Data Warehouse is a collection of subject-oriented, time variant, integrated, non-volatile data which facilitates the management for making a knowledgeable decision (Inmon 1991). It needs special task of extracting, transforming and load procedures on historical data to help decision makers to improve their business process (Becker et al. 2008). Its global reach and web accessibility has made confidentiality as one of the major issue (Dhillon 2000). To deal with confidentiality and authentication, user credentials such as username, password, etc. are being used over the time. Once the users authenticate themselves by providing essential security information Audit Rules (AR) (Belén et al.

P. Mueller et al. (Eds.): SSCC 2016, CCIS 625, pp. 251–257, 2016.
DOI: 10.1007/978-981-10-2738-3_21

2012) are used to provide enhanced security features. It prevents misuse of authorization by logging the frustrated attempts over several multidimensional elements. But these audit logs are only used to identify the attacker who has tried to exploit the system by attempting to access the elements for which the user in not authorized. Here our intention is to develop extra security level for those cases, where the user is accessing its authorized areas but it is deviating from the user access pattern which has been developed over the period of time. It prevents the unauthorized access for most of the DW elements, even though the security parameters such as username, password for accessing the system stands compromised.

2 Background and Related Work

In this section the background and related work is organized into the following subsections: (1) DW Security Review, (2) Access Control in Data Warehouse.

2.1 DW Security Review

In an attempt to make data warehouse more secure, an adapted mandatory access control for OLAP-cubes based security approach is presented (Kirkgoze et al. 1997). The primary advantage of using this approach is its flexibility of assigning roles to different virtual sub-cubes. Metadata, which describes the contents of the data warehouse (Berson and Smith 1997) can also be used to data warehouse environment's security mechanism. Here, metadata is composed of access rules along with corresponding information about security subjects & objects. Over the period of time, steady growth in the number of OLAP necessitates the requirement of proper access control mechanisms, which ensures the confidentiality of the sensitive data (Santos et al. 2011). Some commercial systems (Cognos 1998; Oracle 1998; Chase et al. 1999; Microsoft 1999; MicroStrategy 2000) do provide mechanisms to cope with these requirements; but these approaches are highly proprietary. The solutions for confidentiality problems regarding DW's are also being discussed and an extension of UML for the secure DW is provided (Fernández-Medina 2004). It allows designers to specify main security aspects in the conceptual MD modeling, thus resulting in the design of a secure DW system. Various other approaches which uses UML extension for the DW Security (Fernandez-Medina et al. 2005; Villarroel et al. 2006; Eduardo et al. 2006; Eduardo et al. 2007; Emilio et al. 2009; Salem et al. 2012) have also been proposed. One of the solutions on DW security (Lopes et al. 2014) investigates the method for encrypting and querying a DW hosted in a cloud, but it leads to high amount of computation overhead at the server side.

2.2 Access Control in Data Warehouse

The Role Based Access Control (RBAC) (Sandhu et al. 1996) is one of the best ways to control the access of DW entities among the variety of users and the privileges associated with them. Once all the user-roles are populated into the database, the formulation of role-based rules are performed, followed by implementation of workflow engine modules. Then through these elements, role-based privileges can be quickly entered and updated across multiple systems, platforms, applications and geographic locations. Here RBAC provides companywide control process for managing data and resources (Iyer et al. 2007). Although RBAC is widely used and is capable of handling the entire system, there are some issues like the unclear definition of groups and user, and no mention of duties along with roles. As a solution an Extended RBAC for secure data warehousing has been proposed (Iyer et al. 2007). Another extension to the RBAC has been proposed as Temporal RBAC (TRBAC) which allows temporal restrictions on roles themselves, user-permission assignments (UA), permission-role assignments (PA) and role hierarchies (RH) (Uzun et al. 2014). In recent work another Role Based Access Technique is introduced which is applied to summarize data (Ali et al. 2014). The restrictions on the basis of summarized data classify the user access on the basis of a level of summarization they are authorized to. In all the above research studies RBAC is in place to make sure that user is accessing only those sections of the DW for which it has been authorized. But there is no check on its behavior once the initial verification has been performed and the user accesses DW elements as per its role. So, if any intruder enters the system with an authorized user credentials, RBAC will not be able to restrict the user access. In order to overcome this issue, we propose a two-level authorization system where the user access pattern will be tracked and monitored over a period of time once the initial verification of authentication has been done. When the user deviates from its normal access behavior, then the second level of authorization will be activated to verify its authentication by some different method like secret question etc. In a recent proposal, an intrusion detection mechanism for data warehouse has been proposed (dos Santos 2014). In this research work DW-DIDS (Data Warehouse Database Intrusion Detection system based on the analysis of user actions at the SQL command level etc., but in this article also there is no double check of the authenticity of the user if it's actions deviates from role profile.

3 Motivating Example

To illustrate our work let us consider an example, where the authentication details of the Managing Director of a banking DW which is authorized to access most of the facts and dimension at all the summarization levels has been compromised. Now this intruder who enters the system through compromised details will be able to access the entire banking DW elements including all financial transactions throughout the time span. Here, the RBAC can have very little impact as the user's privilege enjoys the maximum access. So the intruder can access most of the financial secrets of the bank irrespective of the existing RBAC protection. This creates a different kind of threat to the entire DW; it becomes more crucial when the cost of leakage of information is too high. Thus, if the user access signature has been maintained over a time interval and analyzed, any suspicious behavior can be detected and the user can be asked to provide the second level of authentication in the form of a secret question. So the user credentials which stand compromised would not be a threat to entire DW information for which the user is authorized and damage will only be limited to the DW elements which matches the user's accessed signature profile. The proposed system prevents most of the role wise accessible information, even if some user credential has been compromised. The second level of authorization also helps in reducing the false positive thus providing more robustness to the role profile created over the time (Fig. 1).

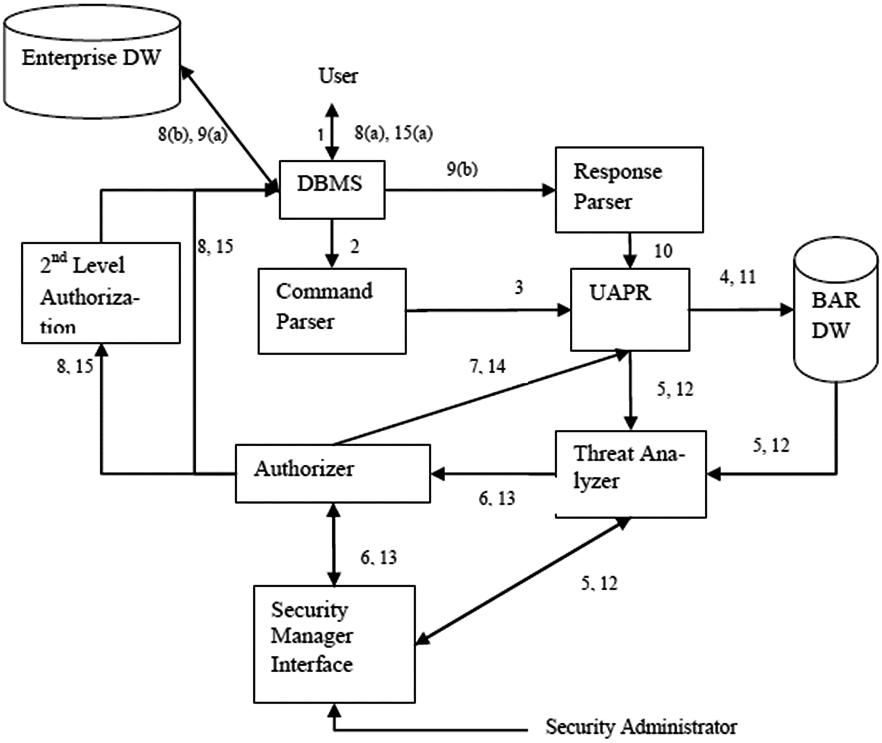

Fig. 1. Two Level Authorization Model for Accessing Data Warehouse

4 Authorization Model

Here two-level authorization model has been proposed which allows system to ask for second level authorization whenever a suspicious activity is detected. The working of the model is explained below:

1. At the first instance, user initiates a DW query to the DBMS after verification of its initial authentication details.
2. Once the DBMS has authorized the initial entry into the system, DBMS passes query to the command parser for converting it into the query tokens.
3. The details of the query and the details like username, CPU time, response size, processed rows, processed columns, etc. forms a signature in UAPR (User Access Pattern Recorder).

4. The UAPR updates the signature in the BAR (Behavioral Analysis Repository) to form a user access behavior profile which gets refreshed by completion of every request and response.
5. Then UAPR passes the control to threat analyzer to analyze the current request on the basis of updated signature available in BAR DW and some external inputs if any from security manager via security manager interface. The calculated similarity between the current and user access profile classifies query as authorized or unauthorized.
6. The threat analyzer sends its output to the authorizer for the decision whether to initiate the 2nd level authorization or not, which is decided on the basis of threshold on similarity set by the security administrator.
7. On the basis of input from threat analyzer, Authorizer decides whether to initiate the 2nd level authorization or allow the same command to continue. This decision has also been logged in BAR with the help of UAPR. This decision has been taken on the basis of input from threat analyzer and also the policy rules formulated by security administrator via security manager interface.
8. The DBMS asks for another authorization in the form of some secret question on the basis of the input from the authorizer before moving further for results, if the same has been initiated by the authorizer. Otherwise, the control has been passed to the DBMS without it.
   a. If 2nd level authorization has been initiated, then the user would be asked to prove its identity by answering the additional secret question. If user unable to answer it correctly then the current request would be suspended and the result would be sent to response parser for updating the user profile through UAPR, else the appropriate result sets would be sent to response parser.
   b. Once the previous stage has been passed the DBMS communicates Enterprise DW to provide the response to the DBMS accordingly.
9. The Enterprise DW processes the request from DBMS and provides the result back to the DBMS.
   a. The response containing the result set passes to the DBMS for further processing.
   b. Response from the user passes to the response parser for conversion of response data into the response tokens.
10. The response parser passes its tokens to UAPR which updates the user signature on the basis of response tokens to update the user profile in BAR. The response details like number of processed rows, processed columns, etc. further strengthens the user signature in UAPR (User Access Pattern Recorder). The response details also contain the details of the response of user on the basis of answer of 2nd level authorization, if the same has been initiated by the authorizer.
11. Then the same cycle will follow similar to the input query from steps 11 to 15 in order to finally take a decision to show the results to the user.

Here, 2nd level of authorization further strengthens the authorization process. Whenever the user access deviates from its long historical signatures maintained over the time, this mechanism provides a 2nd layer of security to ensure the authenticity of the user by some security question etc. This mechanism allows another chance to user to prove its authenticity in case of deviation. It also helps in the creation of more robust signature base over a period of time which in turn creates better user profiles.

5 Conclusion and Future Work

The mechanism of two level authorizations strengthens the security of the system in the following two ways:

1. The 2nd level of authorization itself bring robustness in the signature and user profile as some of the odd signatures which don't come in the acceptance ambit can be included in the accepted signature list by the verification of 2nd level of authorization.
2. It also act as a 2nd level challenge to the user, if the user's action doesn't match the user historical profile created over a period of time. So it also acts as an additional security level even if the initial security credentials stand compromised.

In future, a comparative study would be performed to support the model on the basis of various parameters such as no. of false positive, no. of true negative etc.

References

Inmon, W.H.: Building the Data Warehouse. Wiley, Hoboken (1991)

Becker, B., Kimball, R., Mundy, J., Ross, M., Thorthwaite, W.: The Data Warehousing Lifecycle Toolkit. Wiley, Hoboken (2008)

Kirkgoze, R., Katic, N., Stolba, M., Tjoa, A.: A security concept for OLAP. In: Proceedings Eighth International Workshop on Database and Expert Systems Applications (DEXA). IEEE (1997)

Berson, A., Smith, J.S.: Data Warehousing Data Mining & OLAP. Series on Data Warehousing and Data Management. McGraw-Hill, New York (1997)

Santos, R., Bernardino, J., Vieira, M.: A survey on data security in data warehousing: issues, challenges and opportunities. In: EUROCON - International Conference on Computer as a Tool (EUROCON), pp. 1–4. IEEE (2011)

Cognos Incorporated: Schrittweise Anleitungen for Transformer. Cognos Power-Play Version 6.0 (1998)

Microsoft Corporation: Microsoft SQL Server OLAP Services Cell-level. Security White-paper (1999)

MicroStrategy Incorporated: MicroStrategy. 7 Administrator Guide (2000)

Oracle Corporation: Oracle Express Database Administration Guide. Release 6.2, Part No. A59962-01 (1998)

Chase, D., Spofford, G., Thomsen, E.: Microsoft OLAP Solutions. Wiley, New York (1999)

Fernández-Medina, E., Trujillo, J., Villarroel, R., Piattini, M.: Extending UML for designing secure data warehouses. In: Atzeni, P., Chu, W., Lu, H., Zhou, S., Ling, T.-W. (eds.) ER 2004. LNCS, vol. 3288, pp. 217–230. Springer, Heidelberg (2004)

Fernandez-Medina, E., Piattini, M., Trujillo, J., Villarroel, R.: A UML profile for designing secure data warehouses. Latin Am. Trans. **3**(1), 40–48 (2005). IEEE

Villarroel, R., Soler, E., Fernández-Medina, E., Trujillo, J., Piattini, M.: Using UML packages for designing secure data warehouses. In: Gavrilova, M.L., Gervasi, O., Kumar, V., Tan, C., Taniar, D., Laganá, A., Mun, Y., Choo, H. (eds.) ICCSA 2006. LNCS, vol. 3982, pp. 1024–1034. Springer, Heidelberg (2006)

Eduardo, F., Juan, T., Rodolfo, V.: A UML 2.0/OCL extension for designing secure data warehouses. J. Res. Pract. Inf. Technol. **38**(1), 31–44 (2006)

Eduardo, F., Juan, T., Rodolfo, V., Mario, P.: Developing secure data warehouses with a UML extension. Inf. Syst. **32**(6), 826–856 (2007). Elsevier

Emilio, S., Eduardo, F., Juan, T., Mario, P.: A UML 2.0 profile to define security requirements for Data Warehouses. Comput. Stand. Interfaces **31**(5), 969–983 (2009). Elsevier

Salem, A., Triki, S., Ben-Abdallah, H., Harbi, N., Boussaid, O.: Verification of security coherence in data warehouse designs. In: Fischer-Hübner, S., Katsikas, S., Quirchmayr, G. (eds.) TrustBus 2012. LNCS, vol. 7449, pp. 207–213. Springer, Heidelberg (2012)

Dhillon, G.: Information Security Management: Global Challenges in the New Millennium. IGI Global, Hershey (2000)

Iyer, S., Kantarcioglu, M., Thuraisingham, B.: Extended RBAC-based design and implementation for a secure data warehouse. Int. J. Bus. Intell. Data Min. (IJBIDM) **2**(4), 367–382 (2007)

Belén, V., Carlos, B., Eduardo, F., Esperanza, M.: A practical application of our MDD approach for modeling secure XML data warehouses. Decis. Support Syst. **52**(4), 899–925 (2012). Elsevier

Lopes, C.C., Times, V.C., Matwin, S., Ciferri, R.R., Ciferri, C.: Processing OLAP queries over an encrypted data warehouse stored in the cloud. In: Bellatreche, L., Mohania, M.K. (eds.) DaWaK 2014. LNCS, vol. 8646, pp. 195–207. Springer, Heidelberg (2014)

Ali, S., Rauf, A., Khusro, S., Zubair, M., Farman, H., Ullah, S.: An authorization model to access the summarized data of data warehouse. Life Sci. J. **11**(6 s) (2014)

Sandhu, R.S., Coyne, E.J., Feinstein, H.L., Youman, C.E.: Role-based access control models. IEEE Comput. **29**(2), 38–47 (1996)

Uzun, E., Atluri, V., Vaidya, J., Sural, S., Ferrara, A.L., Parlato, G.: Security analysis for temporal role based access control. J. Comput. Secur. **22**, 961–996 (2014)

dos Santos, R.J.R.: Enhancing data security in data warehousing. Ph.D. thesis submitted at Department of Informatics Engineering, Faculty of Sciences and Technology, University of Coimbra (2014)