

MODELING CONFLICT OF INTEREST IN THE DESIGN OF SECURE DATA WAREHOUSES

Salah Triki, Hanene Ben-Abdallah, Jamel Feki

Laboratoire Mir@cl, Faculté des Sciences Economiques et de Gestion de Sfax,

Route de l'Aéroport Km 4 – 3018 Sfax, BP. 1088

Salah.Triki@fsegs.rnu.tn, Hanene.BenAbdallah@fsegs.rnu.tn, Jamel.Feki@fsegs.rnu.tn

Nouria Harbi

Laboratoire ERIC, Université Lyon 2, 5 avenue P. Mendès France 69676 Bron, Cedex

Nouria.Harbi@univ-lyon2.fr

Keywords: Data Warehouse, Security, Conflict of interest, Uml profile

Abstract: Security is very important in a data warehouse that often contains data confidential to an enterprise (like turnover) and/or data private to individuals (like health information). Furthermore, while providing access to some particular data in isolation could be safe, their combination could leak confidential information; such leak is known as conflict of interest. In this paper, we extend an existing UML 2.0 profile for the design of secure data warehouse with concepts to model conflict of interest. We illustrate the extended profile through a case study.

1 INTRODUCTION

A data warehouse is a special type of databases that collects and integrates data from multiple, transactional data sources in order to assist the decision maker. An improper disclosure of such data is a threat to the competitive power and at times even the survival of the enterprise. Similar to the case of transactional databases, countering this threat involves establishing access control mechanisms.

In the case of transactional databases, there are two main access control models, adopted as *de facto* standards: mandatory access control (MAC) (USDoD, 1985) and role-based access control (RBAC) (Sandhu, R. S., Coyne E.J., Feinstein H.L. and Youman C.E., 1996). In an MAC, information objects are classified into different levels and subjects are cleared for particular levels. On the other hand, RBAC simplifies permission management through the concept of *roles*: a role represents a set of access privileges to information objects (read/write/update operations) and *users* are assigned appropriate roles. In addition, RBAC defines separation of duties *constraints* used to

prevent conflicts of interest that arise when a role allows a user to access data out of his/her privileges; in other words, this data allows the user to get confidential data.

Inspired from the MAC and RBAC models, several researchers proposed design models for secure data warehouses, *cf.*, (Torsten P., Gunther P., 2000, Edgar W., Oscar M., Wolfgang E., Franz L., Werner W., 2001, Torsten P., Günther P., 2001, Rodolfo V., Eduardo F., Mario P., Juan T. 2006, Soler, E., V. Stefanov, J.-N. Mazón, Trujillo J., Fernández-Medina E., et Piattini M. 2007 et 2008). Among these propositions, we have the UML 2.0 profile SECDW (Secure Data Warehouses) (Rodolfo V., Eduardo F., Mario P., Juan T. 2006). Overall, the proposed models agree upon the concepts to secure in terms of objects (cubes, slices, dices, measures) and subjects (users, roles). They differ in supporting or not security levels and in the granularity of objects. However, none of them provides for the specification of conflict of interest. This paper extends the UML 2.0 profile SECDW (Rodolfo V., Eduardo F., Mario P., Juan T. 2006) by adding static separation of duties as proposed in RBAC. The static

separation of duties prevents conflicts of interests that arise when permissions are associated to the same role; in opposition, dynamic separation of duties places restrictions on the roles that cannot be activated during the same user session.

The remainder of this paper is structured as follows. Section 2 overviews current security models for data warehouses. Section 3 presents the newly extended UML 2.0 profile. Section 4 summarizes the contributions and outlines future work.

2 SECURITY MODELS FOR DATA WAREHOUSES

Inspired by the security models for databases, various approaches have been proposed for modeling secure DW. We next overview the most complete propositions.

In (Torsten P., Günther P., 2001), the authors propose a multidimensional security constraint language (MDSCL) that is based on MDX, a language for manipulating OLAP data. With MDSCL, the designer can define authorization constraints based on the requirements defined in (Torsten P., Gunther P., 2000). The authorization constraints use objects (*i.e.* fact, dimension, hierarchy level in a dimension) and security roles to allow or deny access to (*i.e.*, reading) the corresponding data.

In (Soler, E., V. Stefanov, J.-N. Mazón, Trujillo J., Fernández-Medina E., et Piattini M. 2008), the authors propose a profile based on i* (Eric S. K. Y. 1997) to model the security requirements at business level of MDA (Model Driven Approach) approach (Soler, E., V. Stefanov, J.-N. Mazón, Trujillo J., Fernández-Medina E., et Piattini M. 2007).

For a precise interpretation of a security model, in (Edgar W., Oscar M., Wolfgang E., Franz L., Werner W., 2001), the authors propose an authorization model based on a formal notation. It supports the following concepts: OLAP operations (*i.e.*, read, drill-down, roll-up, and slice), subjects (*i.e.*, end user), and the multidimensional objects of facts and dimensions.

In (Soler, E., V. Stefanov, J.-N. Mazón, Trujillo J., Fernández-Medina E., et Piattini M. 2007), the authors propose a MDA approach for developing a secure data warehouse. This approach exploits the language QVT (Query View Transformation) (OMG 2005) to automate the transition from the conceptual level to logical level.

Finally, in (Rodolfo V., Eduardo F., Mario P., Juan T. 2006), the authors present a UML 2.0 profile called SECure Data Warehouse (SECDW). This profile incorporates the MAC model and partially the RBAC model. Thus, it provides for the specification of security aspects such as security levels and user roles on the main elements of the multidimensional model such as facts, dimensions and hierarchies. While being the most complete proposition, one shortage of this profile is its lack of support for conflict of interests. Enriching this profile is the main contribution of this paper. We next review in detail the SECDW profile and present our extensions.

3 SECDW ENRICHED WITH CONFLICT OF INTEREST SUPPORT

UML (Unified Modeling Language) is a graphical language for modeling data and treatments. It is a standard developed by the OMG (Object Management Group). Profiles allow the adaption of UML to a particular application domain. They are defined through three UML extension mechanisms provided: stereotypes, tagged values and constraints.

Using these extension mechanisms, SECDW (Rodolfo V., Eduardo F., Mario P., Juan T. 2006) is a UML profile for the design of secure DW models. The first type of extensions in SECDW is meant to introduce the multidimensional concepts (fact class, dimension class, fact attribute, etc.); the second type of extensions is meant to introduce security concepts. These latter can be used to specify, for each multidimensional concept, its security information in terms of: a sequence of security levels, a set of user levels and a set of user roles.

We next overview the security concepts of SECDW and then we present our extensions and a set of well-formedness rules.

3.1 SECDW: an overview

In (Lujan-Mora, S., Trujillo, J. and Song, I. Y., 2002), the authors have defined a UML profile for the design a data warehouse. This profile extends UML in order to provide for the multidimensional concepts: fact, measure, dimension, parameter and hierarchy.

Based on this profile, the authors then defined their SECDW profile to provide for the specification of

certain security concepts in data warehouses. SECDW is able to specify security information for each element of a data warehouse. It takes into account RBAC and MAC model. Whereas SECDW allows designer to define which role needed to access data warehouse elements it doesn't deal with conflicts of interests issues. Elements in conflict should not be assigned to the same role. For example, if facts *Sale*, *Expenditure* and *Purchase* are assigned to the same role a user assigned to that role will deduce the earnings of the company. The detailed description of the stereotypes is presented in Table 1.

3.2 SECDW+: the extensions

Our SECDW enrichment consists of adding the concepts required for the specification of Static Separation of Duties (SSD). The resulting profile, called SECDW+, provides for the specification of both conflicts of interests (CoI) among multidimensional concepts, and conflicts depending on data.

For the CoI among multidimensional concepts, SECDW+ allows a designer to indicate conflicts among: facts; attributes; dimensions; hierarchies; a hierarchy level and a dimension; and measures and parameters of a dimension. This level of CoI is independent of data.

On the other hand, at the second CoI level, SECDW+ allows a designer to specify CoI that are data dependant. Here, for instance, the designer can specify that the address of a patient and his/her illness type are in conflict if the illness type is equal to *cancer*.

To provide for the specification of these two CoI levels, the SECDW+ profile defines two new stereotypes, new tagged values and one new constraint.

3.2.1 SECDW+ stereotypes

SECDW+ adds the following two stereotypes:

- *Conflict class* stereotype that contains tagged values associated to its attributes and a conflict list;
- *Conflict attribute* stereotype that specifies a conflict list through tagged values.

A conflict list contains the set of attributes and classes that are in conflict with the tagged element (class or attribute). The detailed description of the stereotypes is presented in Table 1.

Table 1: Stereotypes

Name	ConflictClass
Description	They represent classes in conflict with other classes. Classes in conflict must not have the same security role.
Name	CFact
Description	They represent facts within a multidimensional model. They are sub-class of ConflictClass.
Name	CDimension
Description	They represent dimensions within a multidimensional model. They are sub-class of ConflictClass.
Name	CBase
Description	They represent dimension hierarchy within a multidimensional model. They are sub-class of ConflictClass.
Name	ConflictAttribute
Description	They represent attributes in conflict with other attributes. Attributes in conflict must not have the same security role.
Name	CFactAttribute
Description	They represent fact attributes in conflict with other ConflictAttributes and/or ConflictClass. They are sub-class of ConflictAttribute.
Name	CDimensionAttribute
Description	They represent dimension attributes in conflict with other ConflictAttributes and/or ConflictClass. They are sub-class of ConflictAttribute.
Name	ConflictHierarchyLevel
Description	They represent dimension hierarchy level in conflict with other ConflictAttributes and/or ConflictClass. They are sub-class of ConflictAttribute.
Name	ConflictRule
Description	This type of rule defines the conflicts between ConflictClass and/or ConflictAttribute.

3.2.2 SECDW+ tagged values

SECDW+ the following tagged values:

- *ConflictListAttribute* that specifies a set of attributes in conflict with an attribute or a class.
- *ConflictListClass* that specifies a set of class in conflict with an attribute or a class.
- *ConflictStatus* that specifies whether there is a conflict between classes and attributes.
- *DerivationRule* that represents the derivation rule.
- *InvolvedClass* specifies the classes that have to be involved in conflict rule.
- *InvolvedAttribute* specifies the classes that have to be involved in conflict rule.

These tagged values can be used to define the conflict list among the different elements (fact, measure, dimension, etc.) of a multidimensional model.

3.3 Security well-formedness rules

To be used consistently, our extension must respect the following six well-formedness rules:

- If a dimension D has a parameter P that is in conflict with a parameter P' of another dimension D' , then all the parameters below P are in conflict with P' .
- If a class C (or attribute A) is in conflict with a class C' (attribute A'), then C' (A') is also in conflict with C (A).
- If a class C is in conflict with a class C' , then all the properties of C are in conflict with the properties of C' .
- Classes or attributes in conflict must have different security roles.
- The derived¹ attribute is in conflict with attributes and classes which are in conflict with the base attributes.
- When a dimension class is specialized by several base classes, if the dimension class is in conflict with an element of data warehouse then the sub-classes are also in conflict with them.

4 Conclusion

In this paper, we proposed an extension of the SECDW (Rodolfo V., Eduardo F., Mario P., Juan T. 2006) UML profile in order to model conflicts of interest. Hence, the proposed model, called SECDW+, provides for the specification of: multidimensional concepts needed for the design data warehouses, security concepts adapted from the standard MAC and RBAC models, and data-independent and data-dependent conflicts of interest. The security pertinent concepts provide for the specification of secure data warehouse designs.

We are currently developing a CASE toolset to support SECDW+. In addition, we are investigating how to formally analyze a SECDW+ model through its transformation a formal language such as Z

¹ A *derived* attribute in a class C is an attribute that is computed from attributes of other classes related to C . The latter attributes are called *base attributes*.

(Michael S. J., 1992). Such a transformation will allow us to prove several properties of a given design like its freedom of inconsistencies (*i.e.*, consistency).

REFERENCES

- Edgar W., Oscar M., Wolfgang E., Franz L., Werner W., 2001. An Authorization Model for Data Warehouses and OLAP. In : Workshop on Security in Distributed Data Warehousing
- Eric S. K. Y. 1997. Towards modelling and reasoning support for early-phase requirements engineering. In, the Third IEEE International Symposium on Requirements Engineering, Annapolis, MD, U.S.A
- Michael S. J., (1992). The Z Notation: A reference manual (2nd edition ed.). Prentice Hall International Series in Computer Science
- Lujan-Mora, S., Trujillo, J. and Song, I. Y., 2002. Extending the UML for multidimensional modeling. In: 5th International Conference on the Unified Modeling Language. Dresden, Germany, Springer-Verlag. LNCS 2460.
- OMG 2005. MOF QVT final adopted specification.
- Ravi S. Sandhu and Edward J. Coyne and Hal L. Feinstein and Charles E. Youman, 1996. Role-Based Access Control Models, pp 38--47. IEEE Computer Vol. 29 Num. 2
- Rodolfo V., Eduardo F., Mario P., Juan T. 2006. A UML 2.0/OCL Extension for Designing Secure Data Warehouses. Journal of Research and Practice in Information Technology, Vol. 38, Num. 1 pp. 31—43
- Sandhu, R. S., Coyne E.J., Feinstein H.L. and Youman C.E., 1996. Role-Based Access Control Models, IEEE Computer 29(2): 38-47, IEEE Press.
- Soler, E., V. Stefanov, J.-N. Mazón, Trujillo J., Fernández-Medina E., et Piattini M. 2008. Towards comprehensive requirement analysis for data warehouses : Considering security requirements. In : The Third International Conference on Availability, Reliability and Security. pp. 104–111. IEEE Computer Society. Barcelona, Spain.
- Soler, E., Trujillo J., Fernández-Medina E., et Piattini M. 2007. A framework for the development of secure data warehouses based on MDA and QVT. In The Second International Conference on Availability, Reliability and Security, ARES, pp. 294–300. IEEE Computer Society.
- Torsten P., Gunther P., 2000. Towards OLAP Security Design - Survey and Research Issues. In: International Workshop on Data Warehousing and OLAP
- Torsten P., Günther P., 2001. A Pragmatic Approach to Conceptual Modeling of OLAP Security. In : The 20th International Conference on Conceptual Modeling, pp 311--324. Springer-Verlag London, UK
- USDoD, 1985. Trusted Computer System Evaluation Criteria. United States Department of Defense. DoD Standard 5200.28-STD.