IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 3, Issue 4, Aug-Sept, 2015
ISSN: 2320 – 8791 (Impact Factor: 2.317)
www.ijreat.org

Application Of Enhanced Security In Mondrian OLAP To Secure Big Data

Fernandez raj D[1], Gunasekaran G[2]

[1]Research Scholar Department of Computer Science Engineering, St. Peter's University, Avadi, Chennai, Tamil Nadu, India.

[2]Professor and Principal, Meenakshi College of Engineering, K.K. Nagar, Chennai, Tamil Nadu, India.

Abstract

Security policies in Mondrian OLAP system is designed to protect sensitive data from unauthorized access while, at the same time, ensuring that authorized requests can be consistently satisfied. Hence such policies are allowing administrators to define rules, restrictions and exceptions which can be associated with the components of the ROLAP data model. Mondrian memory protection supports different access permissions for individual words rather than setting access on a page-level as traditional memory protection scheme and it is not enough to implement required complex security techniques.

In this paper, specific problems such as the data security techniques in practice and a new data security model which is suitable for already developed data warehouses are discussed. We implemented this model by using web based application and verified it through different case studies. Key advantages of our approach are reduction of number of access checks that leads to improved data retrieval and reduced analysis time in a secure environment.

.

***Keywords**: Decision support systems, data warehouse, OLAP, access models, OLAP security, Big-Data, Mondrian OLAP security.*

1. Introduction

One of the most important features of any OLAP system is the protection of data against unauthorized disclosure (privacy) , while at the same time ensuring accessibility by authorized users whenever needed (availability). Considerable effort has been devoted to addressing various aspects of privacy and availability. The two main objectives are considered in this context. The first is the identification of specification of suitable security policies. The second is the development of suitable access control mechanism implementing the stated polices.

The Big Data processing is considered as a complex task and it is usually performed in parallel software systems and infrastructure that can traverse through the huge amount of data without much difficulty as a personal desktop computer. The process of collecting data in digital form is believed to improve an organization's development proportionally. The analyzing and retrieving of data will be much simpler and easier if the data are stored digitally. The challenge being faced in Big Data is not mainly about the storage, the real problem arises when the data in the dataset is manipulated or retrieved from the Big Data

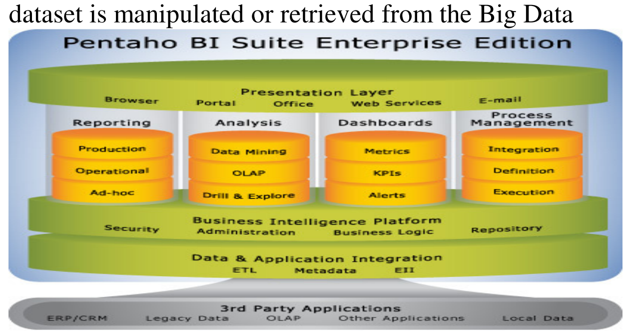


Fig: OLAP architecture

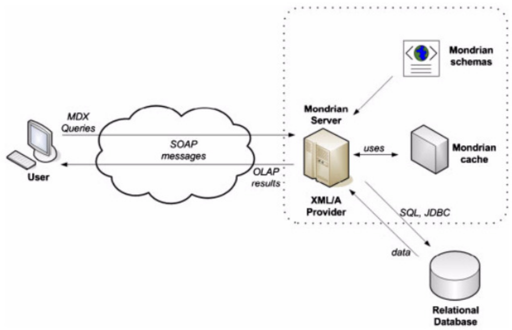
OLAP – On Line Analytical Programming

The entire IT system can be split into OLTP and OLAP used for data transaction and data analysis respectively. OLAP tool is mainly used for analyzing the huge set of business data from the data warehouse. Figure-1 depicts the relationship between the OLAP and the data warehouse.

2. Various security processes in Mondrian

Security mechanisms in general provide means to define which security subjects (users, groups, roles) may or may not access certain security objects (i.e. sensitive data) applying a particular access type (usually read access in OLAP). The closure assumption specifies whether everything is forbidden unless explicitly allowed (closed world) or vice versa (open world). Due to the usually high-level users of OLAP applications an open world policy might be appropriate.

IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 3, Issue 4, Aug-Sept, 2015
ISSN: 2320 – 8791 (Impact Factor: 2.317)
www.ijreat.org



Mondrian operation.

2.1 Declaring roles in the Mondrian schema:

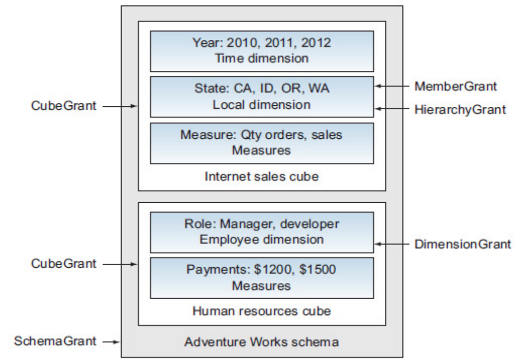
Mondrian allows you to define role-based restrictions for security. For example, if a user is assigned to the role of sales manager, that person can only see data that a sales manager is allowed to see. If sales managers aren't allowed to see information about customers, anyone assigned to this role couldn't see customer information unless they're assigned to a second role that gives access to customer information.

- one-to-one role mapper
- lookup-map role mapper
- user-session role mapper
- your very own role mapper

2.2 Security Grants:

Mondrian security grants can be thought of as a set of filters on the data, and the role can only see what their filters let through. At each level in the schema, the user can have data explicitly blocked or shown. The nesting of the security grants matches the general nesting of the schema design.

- SchemaGrants that can limit access to entire schemas
- CubeGrants that can limit access to specific cubes
- DimensionGrants that can limit access to entire dimensions
- HierarchyGrants that can limit access to dimensional hierarchies
- MemberGrants that can limit access to specific members within a hierarchy level
- Measuregrants, a special case of the MemberGrant that limits access to measures.



Security grants within security grants

2.3 Dynamic Security:

We need a way to determine which data a user is allowed to see. The approach we'll use iinvolves setting session attributes for the users when they log in and then checking the values of these attributes when queries are made to Mondrian.

- Creating an action sequence
- Restricting data using a dynamic schema processor
- Restricting data using dynamic role modification

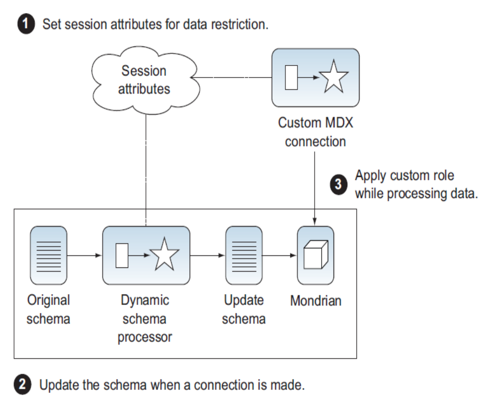


Fig: Dynamic security process in Mondrian

2.4 Multidimensional K-Anonymity in Mondrian:

K-Anonymity has been proposed as a mechanism for protecting privacy in micro data publishing, and numerous recoding "models" have been considered for achieving k-anonymity.

IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 3, Issue 4, Aug-Sept, 2015
ISSN: 2320 – 8791 (Impact Factor: 2.317)
www.ijreat.org

K-anonymity has been proposed to reduce the risk of this type of attack. The primary goal of achieving the K-anonymity is to protect the privacy of the individuals to whom the data pertains. However, subject to this constraint, it is important that the released data remain as "useful" as possible. Numerous recoding models have been proposed in the literature for k-anonymity , and
often the "quality" of the published data is dictated by the model that is used.

3. Related work:

Early work in the area of policy specification tended to focus on networked and distributed environments. The ponder model ,for example, targets networked domains and represents policies as entries in a table consisting of multiple attributes. This model was extended to fully distributed environments with ponder2, an xml-based language that specifies security and management policies
In a subject-action-target format. These approaches are not well-suited to olap domains as they are often fragmented, dependent on infrastructure and lack any native understanding of olap's multidimensional data model.

Table 1. Security feature comparison of the evaluated products

	Product	ROLAP based products	Microsoft SQL Server 2000	MicroStrategy 7	Cognos PowerPlay	Oracle Express
Info	Evaluated release	N/A	8.0 BETA	7.0 BETA	6.0	6.2
	Supported security feature(s)	SQL views	Cell-level and dimension security	Access control list and security filters	User class and dimension views	Permission programs
	Security enforcing architecture component	DBMS	OLAP server or OLAP front-end	OLAP server or OLAP front-end	OLAP front-end	OLAP server
	General approach	View	Hybrid[3]	View	View	Rule
	Security policy	Closed world	Open world	Open world	Open world	Open world
	Security administration	Ownership	Administrator	Ownership	Administrator	Administrator
Requirements	Hide whole cubes	●	●	●	●	●
	Hide certain measures	●	●	●	●	●[3]
	Hide slices of a cube	●	●	●	●	●[3]
	Hide levels of detail	●	●	●	●	●[3]
	Hide levels of detail in certain slices of same dimension	●[4]	●	○[6]	●	●
	Hide certain measures in certain slices	●[5]	●	○[6]	–	●
	Hide complex slices (dices) of a cube	●[5]	●	●[5]	–	●
	Hide levels of detail in certain slices of a different dimension	○[5,7]	●	○	–	●
	Dynamic/data driven constraints	○[5,8]	○[8]	○[5,8]	–	○[8]
	Inference control	–	–	–	–	–

3.1 Demerits of existing Techniques:

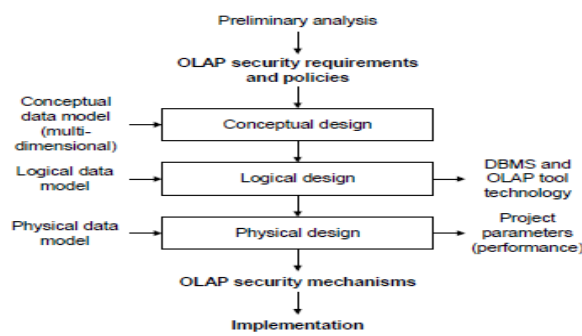
a) It is very complicated to implement the security constraints as the size of data cubes increases.

b) Application time for security constraints increases.

4. Proposed Enhanced Security Model:

Our Enhanced security model for Mondrian OLAP is developed on the assumption of a CSP (central security policy). The access privileges are defined as authorization constraints making the identification of security objects & subjects are necessary and we assume that the notion of (non-overlapping, non-hierarchical) roles as security subjects. Therefore, in addition to the elements cubes, dimensions, etc. the element role is introduced. Authorization constraints can either be positive (explicit grants) or negative (explicit denials). We base the security model on an open world policy (i.e. access to data is allowed unless explicitly denied) with negative authorization constraints. This corresponds to the open nature of OLAP systems. Additionally, we limit our security model to read access. The typical queries in OLAP systems are read-only. There are some systems that support write-back mechanisms (e.g. for future plan data), but these will not be considered in our model.

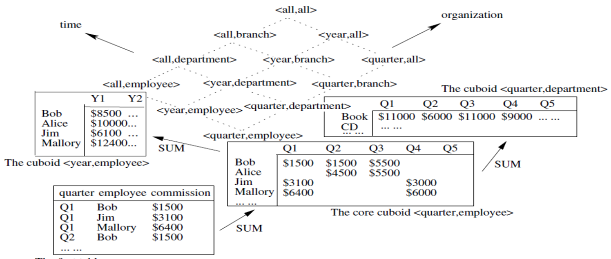
The authorizations are depicted as rules to the OLAP. We use discretionary access controls (DAC) model that are based on a collection of concepts , including a set of security subjects (S), a set of access types (A) , and a set of security objects (O) . In general , a security rule is quadruple, (*s,a,o,p*), where subject *s* has the access type *a* to access security object *o* within the range of predicate *p*.



OLAP security design methodology

In order to grant any access right to a role a, any combination of described rule definition types can be used. The derived data cube of the role has its own dimensions and dimension hierarchies which are also a sub-set of the not restricted dimensions and dimensional hierarchies.

The sub-cube Ci (Ci Ɛ V) defined on OLAP data represents the area of the data cube to which corresponding role has access. If F (Ci ∩ Cj) be a fragment then F is a data cube area which two roles have access in common. If F=Ci\Cj the F is only accessible by roles having sub-cube Ci as their interface to the data cube.



An Example of Data Cubes

5. Conclusions

We have presented a methodology for conceptual modeling of OLAP security. Even though we used a rather pragmatic approach (limiting our model in several ways), it should be applicable not only for the GOAL prototype applications, but also for most other real-life projects. The aim was not to create an exhaustive model that would be able to cover all special requirements that might come up in rare occasions, but rather to present an approach that would be applicable in practice. Also throughout the design of the methodology we had the idea of a possible tool support in mind.

References

1. Bishop, M., Computer Security: Art and Science, Addison-Wesley, 2003.
2. Tsyrklevich, E. and Yee, B., "Dynamic detection and prevention of race conditions in file accesses," Usenix Security Symposium, 2003.
3. Castano, S., Fugini, M., Martella, G., Samarati, P.: Database Security. ACM Press, 1994.
4. [2] Chaudhuri, S., Dayal, U.: An Overview of Data Warehousing and OLAP Technology. 1996.
5. Essmayer, W., Wagner, R., Kapsammer, E., Tjoa, A.M.:\Meta-Data for Enterprise-Wide Security Administration. In Proceedings of the Third IEEE Computer Society Metadata Conference; NIH Campus, Bethesda, Maryland, April 6-7, 1999.
6. Oliveira R, Bernardino J. Building OLAP tools over large databases. Proceedings of IADIS Virtual Multi Conference , 2006.
7. S. Rizzi, A. Abelló, J. Lechtenbörger, and J. Trujillo, "Research in data warehouse modeling and design: Dead or alive?" in Proc. 9th ACM International Workshop on Data Warehousing and OLAP (DOLAP'06), pp. 3-10, 2006.
8. M. R. Villarroel, E. Soler, E. F. Medina, J. Trujillo, and M. Piattini, "Representing levels of abstraction to facilitate the Secure Multidimensional," in Proc. The First International Conference on Availability, Reliability and Security, 2006.
9. E. Weippl, O. Mangisengi, W. Essmayr, F. Lichtenberger, and W. Winiwarter, "An authorization model for data warehouses and OLAP," in Proc. Workshop on Security in Distributed Data Warehousing, 2001.
10. E. F. Medina, J. Trujillo, R. Villarroel, and M. Piattini, "Access control and audit model for the multidimensional modeling of datawarehouses," Decision Support Systems (DSS), vol. 42, pp. 1270-1289, IEEE, 2006.
11. T. Priebe and G. Pernul, "A pragmatic approach to conceptualmodeling of OLAP security," in Proc. 20th Int. Conference on Conceptual Modeling, Springer-Verlag, Yokohama, Japan, 2001. [12] E. F. Medina et al., "Developing secure data warehouses with a UML extension," Information Systems, vol. 32, no. 6, pp. 826-856, 2007.
12. K. Shazad and A. Sohail, "A systematic approach for transformation of ER schema to dimensional schema," in Proc. the 7th International Conference on Frontiers of Information Technology, CIIT, Abbottabad, Pakistan, 2009.
13. Case study. Mondrian Food Mart. [Online]. Available: http://www.sourceforge.net.
14. E. Soler, R. Villarroel, J. Trujillo, E. Fernández- Medina, and M. Piattini, "Representing security and audit rules for data warehouses at the logical level by using the common warehouse metamodel," presented at First International Conference on Availability, Reliability and Security (ARES'06), Vienna, Austria, 2006.

AUTHORS BIOGRAPHY

D. Fernandez raj was born in Sirkali, Tamilnadu, India, in 1973. He received the **B.Sc.,** Bachelor degree in Mathematics (1995) , the **B.Ed.,** Bachelor degree in Education (1996), and the **M.A.,** Master degree in English (2006) from the Annamalai University, Chidambaram, Tamilnadu. He received **M.C.A.,** professional Master degree in Computer Science Applications (2002) from the Bharadhidasan University,Tiruchirapalli, Tamilnadu and the **M.B.A.,** Master degree in Business Administration in Computer systems (2010) from the University of Madras , Chennai, Tamilnadu . He is currently pursuing the **Ph.D.,** degree with the Department of Computer Science and Engineering, St. Peter's University, Chennai. His research interests include Data Mining concepts, Big Data analytics and web data mining.

Dr. G. Gunasekaran, Principal, Meenakshi College of Engineering was born in Tamilnadu, India in 1965. He received the Bachelor degree in Computer Science and Engineering from the Madurai Kamaraj University, in 1989 and the Master degree in Computer Science and Engineering from Jadavpur University, Kolkata, in 2001. He got his **Ph.D**. degree from the Department of Computer Science and Engineering, Jadavpur University, Kolkata in 2009. His research interests include Data Mining, Bio informatics, Software Engineering, Graphics and Multimedia.