Sub Table Based Access Security Model (STBAM) for OLAP Tools

Abid Sohail and Khalid Hassain

Abstract—**Data Warehouses are used to maintain and manage the historical information that is very vital and confidential. Keeping in view the significance of data various data security techniques have been introduced. Most of them consider data security requirements from the early developmental stage of Data Warehouse. In this paper, specific problems related to the existing data security techniques are discussed and a new data security model which is suitable for already developed data warehouses is presented. 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.**

Index Terms—**Decision support systems, computer-based decision aids, data warehouse, OLAP, access models, OLAP security.**

I. Introduction

Data of an organization is always very crucial for its survival. Minute disclosure can lead to disturbance of whole organizational informational structure. Because of its importance, data must be protected in favor of organization's unsurpassed management [1], [2]. During early stages, the main focus of data security was on the *Discretionary Access Control* (DAC) models. After that, the absorption moves to the *Mandatory Access Control* (MAC) models. For implementation of data security in *Data Warehouse*, a number of complications arose. Those are:,

The *Data Warehouse* repository has an open nature. A small inaccuracy in applying security constraints can hinder the process of analysis [1], [2]. The information in OLAP cubes is highly summarized. The aggregations contain the hierarchal levels. Due to these hierarchical levels data security complications are increased [3].

Data warehouse uses the relational structure for storage purpose. A question might arise that why don't we implement current DBMS security mechanisms? The answer is because the access privileges are not designed for the data warehouse structure [3], [4].

The security conditions are applied in terms of tables, rows and columns. It does not specify the security measures in data warehouse environment. We have to specify multidimensional security constrains.

Data warehouse collects data from all operational sources that are internal, external and online, and integrates data into a repository for the purpose of implementing the *Decision Support Systems* [1], [5], [6]. Decision Support Systems are used by the executives, analysts and high level users to take future business decisions. The complex analysis is made possible on historical data using different information delivery tools, such as OLAP tools [6]. OLAP tools use pre-aggregation mechanism to summarize data for fast evaluation of information. OLAP tools use DWH's data for summarization. OLAP tools summarize data structure called "data cube".

The data cubes are multidimensional in nature. Each data cube contains a huge amount of data. Different users are involved in accessing these data cubes. Security Levels are to be provided for each data cube which are facts, dimensions, dimensions attributes and cells. Security constraints are applied to whole data for all users. Increase in the security constraints will slow down the navigation and data analysis processes. By decreasing the number of access checks we can improve the retrieval of data during analysis session.

Our focus is on the fast access of information and data security for data warehouse users. We have developed a Sub-Table Based Access Model (STBAM) at the presentation layer of OLAP. After defining STBAM at presentation layer, data cube is created and structured by using MOLAP, ROLAP and HOLAP mechanisms. Finally the data cube is provided to a particular user which can easily be navigated by the user.

II. Existing Models and Their Limitations

Main problem with existing technique is that the application of security constraints becomes very complicated as the size of data cubes increases. Due to this, application time for security constraints increases. Hence, the applicability process will increase the administrative work for a security administrator [7]. Secondly, number of access checks will be increased. The large number of access checks will slow down the navigation and data retrieval processes. And finally they are not suitable for existing data warehouses [8].

The authors in [9] have proposed a data security model. This model is the inspiration of RBAC model and MAC Model. The security model is presented in steps. Each step is handled separately. The technique is based on an access control model and auditing model. The model is composed of authorization policies. Scheme contains a set of security

Manuscript received February 25, 2013; revised May 31, 2013.

The authors are with Department of Computer Science, COMSATS Institute of Information Technology, 1-Km Defense Road, Off Raiwand Road, Lahore, Pakistan (e-mail: abidbhutta@ciitlahore.edu.pk, khussain320@gmail.com)

DOI: 10.7763/IJCCE.2013.V2.266

policies. Each policy gives the specifications for subjects and objects. These security policies are deployed into *Multidimensional Model* (MD). After the development of conceptual design, the designer should specify the MD Modeling through UML technique. In this model for the authorization of Subjects there are two requirements. Firstly the reference of subject is required for efficient access to object. Secondly it accepts the reference of subject. To define access rules for subjects, Unified Modeling Language (UML) is used. In this paper [9] SIAR rules are used to specify the level base security information on each element of data warehouse's conceptual model. The values used in SIAR grammar are just tagged elements that are associated with the attributes of fact tables and dimensions. The SIAR are efficient when the access level of objects and subjects are relevant. If they are irrelevant then AUR are used. The ACA model considers the closed policy. The AUR specifies that which element is going to be accessed by a subject and for what purpose (action). At last, the model is translated into system independent code and integrated into the system.

This data security mechanism has a limitation that it is not suitable for existing data warehouse. As in this the authors suggest to integrate the security requirements at the conceptual level of data warehouse. Secondly, defining and integration of security model is very expensive procedure. As the conceptual model is not an ultimate model. Also, this paper suggests a huge number of security rules. This will slow down the retrieval process of data (as briefly discussed earlier).

The authors in [10] proposed a set of security policies which are based on the UML (Unified Modeling Language). According to the authors, UML (Unified Modeling Language) is not appropriate for modeling security constraints in data warehouse [11]-[13]. For that, authors made some extensions in the UML to make correspondence for multidimensional data models. The proposed approach is based on fundamentals of security semantics for objects. In this model security policies are implemented through the central security administration. The security requirements are considered at early stages (at conceptual level). The proposed approach presented a security constraint language that is used to secure the information at OLAP layer in data warehouse while delivering the information. In the proposed approach, access is given to different roles (subjects). After the association of security constraints, the conceptual model and security constraints models are transformed into logical model, finally, the logical model is transformed into physical.

But it is also not suitable for existing data warehouse because the authors have suggested integrating the security requirements at the conceptual level of data warehouse and a set of security constraints based on the operations of OLAP tools. For instances, each user will be restricted on the bases of operation.

III. Proposed Model (STBAM)

STBAM Model consists of seven steps. In this paper we apply this model on different case studies [11], [12], [14] and implemented it using web based application shown in Fig. 3. The Sub-Table Based Access Model (STBAM) provides a mechanism which reduces the number of access checks and thus enhances the efficiency of retrieval process. Firstly, this scheme enlists the security requirements and finally made association with their specified objects. The scheme (STBAM) considers three elements (S, O, -/+)

- S: *subjects* include the users and groups. These subjects are cataloged according to the Security Compartments (CS), Security Levels (SL) and Security Roles (SR). These users and groups would be analysts, managers and executives.
- O: *objects* are the elements that are made accessible. For example, (in our case) objects are considered as data cubes.
- **-/+:** the *sign* indicates that subject (user or group) have access to object or not. The minus sign indicates the deny property and plus sign indicates the allow property for the user.

The scheme is based on the *closed world policy*. The closed world means by default (when user is cataloged) user has no permissions on specified objects.

Scheme encompasses following steps

- Step 1: Collection and identification of Security requirements.
- Step 2: Identification of Subjects and Formulation of Subject Catalog.
- Step 3: Identification of the Objects and Creation of the Objects Catalog.
- Step 4: Identification of Association between Subjects and Objects.
- Step 5: Creation of tagged cubes.
- Step 6: Setting permission for Subjects.
- Step 7: Finalizing.

Step 1: Collection and identification of security requirements

The security requirements are the first one to be analyzed. The security requirements determine the subjects that are given permission for the objects. The security requirements are collected manually. For the applicability of security requirements, the subjects are classified according to the levels of sensitivity and the role played in an organization.

Step 2: Identification of subjects and formulation of the subject catalog

The user or user group is to be registered by a security manager or by a security administrator. An individual *catalog or profile* is maintained for each user or user group. The *profile* will contain the properties of user or user group. The properties will help in the selection of specific objects. The properties are defined and maintained for both subject and objects; by doing this we can match the security criteria of both entities. The essential properties for user catalog will be:

- An identifier of subject
- A name of subject
- Security Level (SL) of subject (top secret, secret, confidential, and unclassified)
- Security Role (SR) of subject (responsibility of a subject within an organization i.e. a teacher in a university, doctor in a hospital, manager in a bank)

- Security Compartment (SC) (classify subjects according to the geography)
- Group id that relates subject (creating a relationship between single subject and multiple subjects)
- Subject role (admin, owner and normal user)

Step 3: Identification of the objects and creation of the objects catalog

In this step the information that which fact tables, measures, dimensions and dimension attributes are to be protected is gathered. The names of fact tables, measures, dimensions and dimension attributes are inspected. Security administrator will select objects from object catalog. Objects catalog will contains the tagged values (identifiers) against each fact table, measure, dimension and dimension's attribute as shown in Fig. 1.

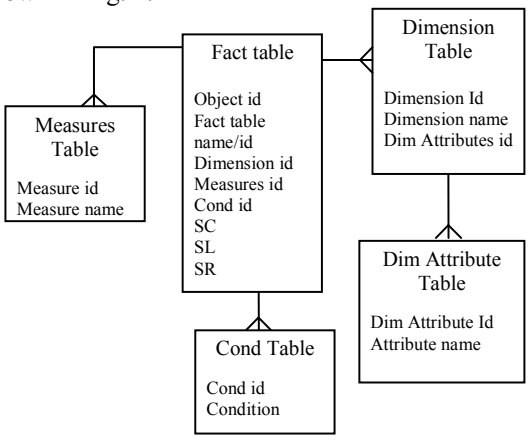

Fig. 1. Maintain a catalog for objects

Step 4: Identification of association between subjects and objects

For association between subjects and objects a catalog is required as shown in Fig. 2. *Subject id* indicates the specified subject. *Object id* indicates the related objects and *sign* indicates subject have access or not.

For association between subjects and groups a *group* table is required as shown in Fig. 2. *Group id* indicates the specified subject and *Subject id* indicates the related subjects.

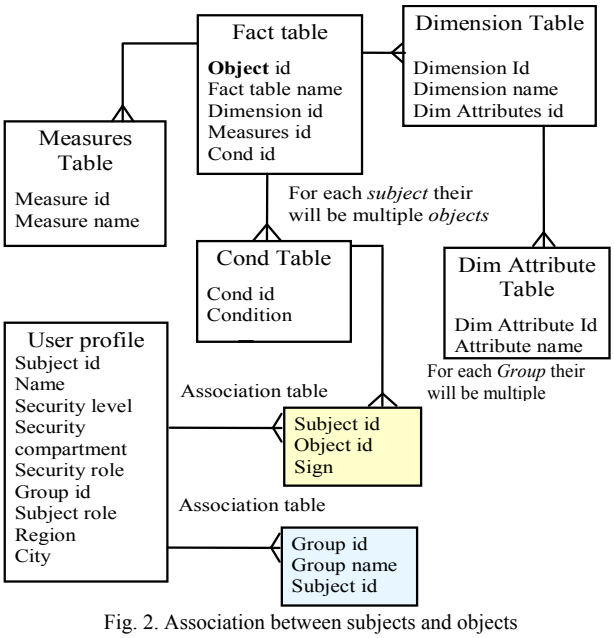

Fig. 2. Association between subjects and objects

Step 5: Creating of cubes with tags

The tagged values that are stored in above indicated tables as shown in the above Fig. 2 can now be used to create and structure a data cube into MOLAP, ROLAP or HOLAP cubes. These values are to be fetched at cube creation time. A *Cube* creation algorithm in STBAM is as follow

Let us consider we have to create a α cube Identified Facts δ
Identified Dimension γ
Identified Hierarchy Level ζ

Let c belong to α
{
Fact_Name = δ
Dimension_Name= γ
Hierarchy_Name=ζ
Create γ
Do until all dimensions created
{
Find the Level of Hierarchy for ζ from the granularity of the set containing {Year, Quarter, Month}
Set Time as Unique Key,
}
MEASURE [measure1] FUNCTION SUM
FORMAT 'format name',
Repeat until All MEASURED Created
}
where *Condition Expression*

WHERE Condition Expression can be written as, *Condition Expression* = "Fact Table.Object id = *ID* AND Measure Table. Measure id = *ID* AND Dimension Table.Dimension id = *ID* AND Dim Attribute Table.Dimension Attribute id = *ID* AND Cond Table.Cond id = *ID* ".

Using above cube creation statement, tagged values (as shown in Fig. 1 and Fig. 2) can be fetched using *WHERE* clause.

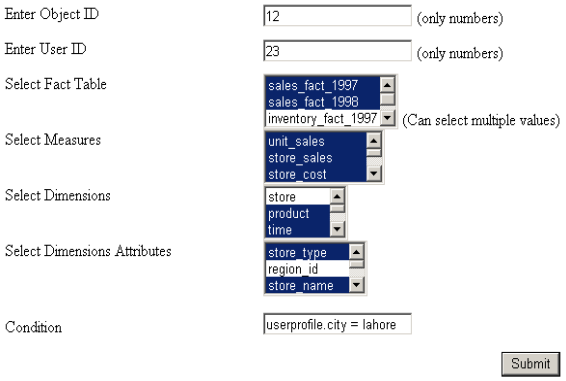

Fig. 3. Selection of security requirements in STBAM

Step 6: Setting permissions for subjects

At this stage, each subject has their own set of cubes. These cubes are now only accessible by their corresponding user or user groups. So the work load on security manager or security administrator decreases.

Step 7: Finalizing

Once the cubes are designed in MOLAP, ROLAP or

HOLAP then user can perform any operation on these cubes, easily without any restriction using the OLAP tools. Operations will be drill through, drill down, drill up and pivoting etc. In exiting OLAP tools there are restriction on each operation.

We apply this security model on Mondrian Food Mart case study [14] shown in Fig. 4. Identified object are shown in each table and condition table is attached with fact table which show that this model provide secure cube level security.

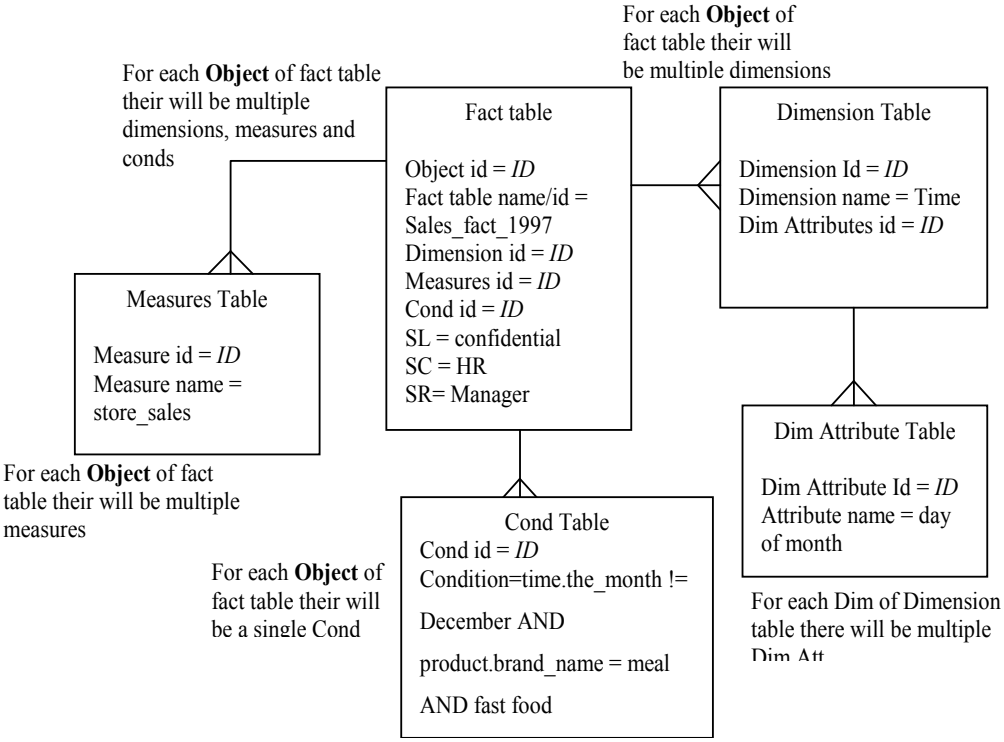

Fig. 4. Meta data of DWH for objects using case study [15]

IV. Conclusion and Future Work

Overcoming the Grant Read problem i.e. when *subject x* gives permission to *subject y* but security administrator deny *subject y* as a consequence permission conflict arises. STBAM provides a systematic way to prevent conflict as shown in Fig. 5. When subject X or any Subject (i.e. owner of data) want to give permission it is first checked whether the subject in question is permitted by security administrator and also by all other subjects, if answer is yes then that subject (in our case Subject X) can give permission to subject Y to access object Z.

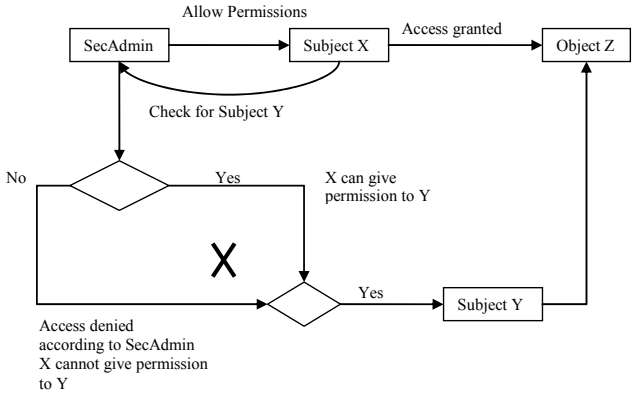

Fig. 5. Overcoming grant read problem in STBAM

STBAM Provides a flexible security model through which security rules can be modified without effecting whole system, this approach provides easy implementation method which is suitable for both newly creating DWH and already existing system. Comparison of STBAM with other models is shown in Fig. 6. STBAM analyzes the security requirements, maintain a user profile and then by using security requirements, objects are associated with subject. It presents a single sign in policy which only checks the access privilege at the time of sign in.

Technique	Security rules update problem	Suitable for existing DWH's	difficult process to apply security rules	only read operation is considered	Grant read permission problem	View base security
Access Control and Audit Model	No	Yes	No	Yes	Yes	No
Pragmatic approach to secure data warehouse	No	No	Yes	Yes	Yes	No
An Authorization model for DWH and OLAP	Yes	No	Yes	Yes	No	No
View Security as Data Warehouse Security approach	Yes	No	Yes	No	Yes	Yes
STBAM	No	Yes	No	No	No	No

Fig. 6. Comparison of STBAM with other approaches

Subjects are free to operate any operation on their related objects. In conclusion STBAM is more efficient and proficient as it (i) reduces the applicability process (ii)

improves the retrieval process (iii) reduces administrative work. Also it is indicated by this research that how administrator can applying the security checks in a simple and quick way.

This approach can be extended to present STBAM as a well maintained plug-in for widely used OLAP Tools.

REFERENCES

[1] R. Kirkgoze, N. Katic, M. Stolda, and A. M. Tjoa, "A security concept for OLAP," in *Proc. 8th International Workshop on Database and Expert System Applications*, pp. 619-626, 1997.

[2] T. Priebe and G. Pernul, "Towards OLAP security design - survey and research issues," presented at 2nd International Workshop on Design and Management of Data Warehouse (DMDW'00), Sweden, 2000.

[3] C. Blanco, E. F. Medina, J. Trujillo, and M. Piattini, "Implementing multidimensional security into OLAP tools," in *Proc. the Third International Conference on Availability, Reliability and Security*, pp. 1248 – 1253, 2008.

[4] E. Soler, V. Stefanov, J. Norberto, J. Trujillo, F. Medina, and M. Piattini, "Towards comprehensive requirement analysis for data warehouses: Considering security requirements," in *Proc. the Third International Conference on Availability, Reliability and Security*, pp. 104-111, 2008.

[5] H. Khalid, "Sub Table Based Access Model (STBAM) for OLAP Tools," MSCS thesis, CIIT, Lahore, Pakistan, June 2009.

[6] 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.

[7] 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.

[8] 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.

[9] E. F. Medina, J. Trujillo, R. Villarroel, and M. Piattini, "Access control and audit model for the multidimensional modeling of data warehouses," *Decision Support Systems* (*DSS*), vol. 42, pp. 1270-1289, IEEE, 2006.

[10] T. Priebe and G. Pernul, "A pragmatic approach to conceptual modeling of OLAP security," in *Proc. 20th Int. Conference on Conceptual Modeling*, Springer-Verlag, Yokohama, Japan, 2001.

[11] *SQL Server Books Online*, Chapter Analysis Services, Case Study "FoodMart2000," SQL Server Books Online, 2000.

[12] E. F. Medina *et al.*, "Developing secure data warehouses with a UML extension," *Information Systems*, vol. 32, no. 6, pp. 826-856, 2007.

[13] 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.

[14] Case study. Mondrian Food Mart. [Online]. Available: http://www.sourceforge.net.

[15] 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.

Abid Sohail graduated in computer science in 2008 from COMSATS Institute of Information Technology and currently pursuing his Ph.D. from the University Technology PETRONAS (Malaysia), His current interests of research are Data warehouse and its applications.

Khalid Hassain graduated in computer science in 2010 from COMSATS Institute of Information Technolgy His current interests of research are Data warehouse development.