Comparative Analysis of Data warehouse Design Approaches from Security Perspectives

Shashank Saroop[#1], Manoj Kumar[*2]

[#] *M.Tech (Information Security), Department of Computer Science, GGSIP University*
Ambedkar Institute of Technology, Geeta Colony, New Delhi, India
[*] *Associate Professor, Department of Computer Science ,GGSIP University*

Ambedkar Institute of Technology, Geeta Colony, New Delhi, India

Abstract- **Data warehouse systems integrate data from heterogeneous sources and are used by decision makers to analyze the status and the development of an organization. Due to the sensitive data contain in the DW it is most important to specify a security, from early stages of DW design so that data must be secured. Traditionally, security was not considered as an important element in modeling of DW. Most of the author starts applying a security at conceptual level to implementation level, however they did not consider a "requirement level" which basically a main levels in DW design. In this paper we have surveyed the literature related to work done by various authors in last few years and Comparison of various approaches in data warehouse design from security perspective by using various parameters.**

Keywords - Data Warehouse, Security, Multidimensional modeling, UML extension, MDA, QVT, OLAP

I .INTRODUCTION

Data Warehouse (DW) systems are used by decision makers to analyze the status and the development of an organization, based on a large amount of data integrated from heterogeneous sources into multidimensional model (MD). Data Warehouses (DWs) manage business' historical information used to take strategic decisions and usually follow a multidimensional approach in which the information is organized in facts classified per subject called dimensions. In a typical DW architecture, ETL (extraction/transformation/load) processes extract data from heterogeneous data Sources and then transform and load this information into DW repository. Finally, this information is analyzed by Data base management systems (DBMS) and On-line Analytical Processing (OLAP) tools. Since data in DWs are crucial for enterprise, it is very important to avoid unauthorized access to information by considering in all layers and operations of the DW, from the early stages of development as a strong requirement to the final implementation in DBMS or OLAP tools (Thuraisingham, Kantarcioglu et al. 2007).

Several approaches for DWs modeling through a specific structural characteristic (Facts, Dimensions, Bases, Hierarchies, etc) exist, but only some of them include security aspects in their modeling [1]. However, these contributions deal with security problem in a static manner in which a set of security constraints basically establish what information will be shown to or hidden from the user, depending on his/her security profile.

In fact, as some authors have remarked [2,3] security of information is a serious requirement which must be carefully considered, not as isolated aspects, but as an element which turns up as an issues in all stages of the development lifecycle, from requirement analysis to implementation.

With regard to the modeling of DWs, we believe that the conceptual modeling phase has been widely recognized as an important step in the design of DWs, as the sooner we represent the main MD properties at the early stages of a DW project, the requirements of the final user. In recent years various approaches have been proposed for representing the main MD properties at conceptual levels (see section II). However none of these approaches for conceptual MD modeling considers security as important issues of its conceptual model, so they do not solve the problem of security within this kind of systems. Moreover, in most real world DW projects, security aspects are issues that usually rely on database management system (DBMS) administrators. .However design of these security aspects should be considered from requirement level of DWs from the early stages from DW projects.

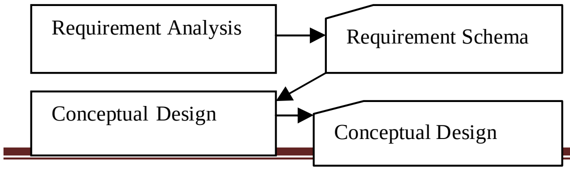
In this paper we have surveyed the literature related to the work done by various author in the recent past years and done the evaluation of various security approaches in DW Design by considering various parameters.

The reminder of this paper is organized as follows: Section 2 will present the related work to been done in past few years. Section 3 wills presents the modeling security in DW .Section 4 will present the evaluation of various approaches in DW design from security perspectives by using various parameters. Section 4 will presents our conclusion.

II. RELATED WORK

In the past years, there has been some good endeavor made to address security measures in Data Warehouse Design [4, 5]. Most of them focused on specific aspects related to multilevel security, federated databases and commercial applications [6]. However there are considerable major approaches which strived to achieve better security that is "*Classical Security Model*". Basically classical security model as shown in fig. 2 describes the whole design process of data warehouse. That starts from requirement analysis to the physical schema. Durning all this process and secure data warehouse schema can be developed. In [7] a security model (classical security model) is introduced that is fruitful if it is implemented in transactional databases, but as far as data warehouse is concerned this model may be unsuitable.

Data warehouse stores the statistical information not like a transactional data base that includes the rows and columns for the storage of records. So this classical security model may not be suitable to maintain the security in data warehouse.



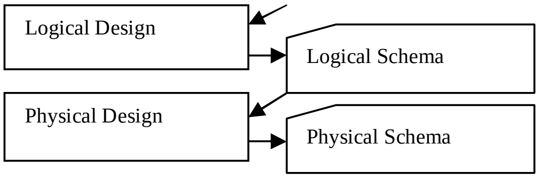


Figure 2. Development Process of Data Warehouse

However [8, 9] propose a several approaches have been proposed for representing the main multidimensional (MD) properties at the conceptual level. Nevertheless, none of these approaches for MD modeling considers security to be an important issue in their conceptual models, so they do not solve the problems arising from this question in these kinds of systems. Although [10] propose an access and audit control model integrated with an Unified Modeling Language (UML) extension, this allowing the development of secure multidimensional models at conceptual level. In [11] defines a UML 2.0 extensibility mechanism which is being used to secure a data warehouse at conceptual level. Although after it same author [12] proposes an extension of relational model considering security and audit measures. This proposal is promising, but still it does cover all the stages of a DW development cycle.

It is true that**,** in the relevant literature, we can find several initiatives for the inclusion of security in data warehouses [6, 9, and 10]. However, none of them considers security aspects which incorporate all stages of the system development cycle, nor the introduction of security into MD design.

A new standard that addresses the complete life cycle of developing applications by using models in software development is arising: Model Driven Architecture (MDA) [13]. In the MDA technology the new standard for defining transformations is Model Object Facility (MOF) 2.O Query/Views/Transformations. (QVT).

Current specialized literature comprises several proposals to integrate security with the MDA technology, but all of them are related with information systems, access control, security services and secure distributed applications, so none of them, is related with the design of secure DW.

However by using models in software development that is (MDA), it will further used in data warehouse design process to secure the data at various levels. In [14] author propose a framework for the design of secure data warehouse in which they defines two meta models to represents a security and audit measures at conceptual and logical levels In [15] same author proposes a two kinds of requirement for data warehousing as information and quality-of-services and combines them in a comprehensive approach based on MDA. In [16] author proposes a MDA architecture which is being used to secure a data warehouse which basically defines a security models at various abstraction levels. After this approach [17] proposes a MDA which basically defines a security issues and provided an automatic transformation between models.

III MODELING SECURITY IN DW

Information security is a basic requirement for a wide range of applications. In the case of DWs, among the different aspects of security, confidentiality (i.e., ensuring that users can only access the information they have privileges for) is particularly relevant, because business information is a very sensitive and can be discovered by executing a simple query.

One of the main concerns in DW design is data security, which is usually seen as a non-functional requirement. The security modeling for DWs comprises several initiatives to include security in the DW design. In [18] the authors describe a prototype model for DW security based on metadata, whose main goal is to reduce user queries to only those data which are to be seen by that user. However, this does not permit the specification of complex restrictions of confidentiality such as deny-allow access to a special user combining groups and security constraints. Rosenthal and Sciore [19] extend SQL grants and create a mechanism of inferences through which to establish DW security, which derives permission on tables and views of the system, thus establishing easy administration. A further attempt is that of the architecture for both Federated Information Systems (FIS) and DWs which preserves MultiLevel security integration between FIS and DWs [20].The authorization of the DW scheme built takes into account the security policy of the federation itself. Kirkgöze et al. [21] defines a model based on the Discretionary Access Model (DAC) which propose a security concept for OLAP, a role based security model for DWs. According to these security rules, a derived data cube is defined for each role. Essmayr et al. [22] shows how access privileges for DWs and OLAP can be expressed more intuitively than by using SQL's grant statements. This access control model focuses specifically on expressiveness and usability. These approaches are attractive but only focus on practical issues such as acquisition, storage and access control on the OLAP side. None of them examine the representation of security at the early stages of the DW design.

More elaborate initiatives which propose authorization models for DWs design also exist. For example, "Priebe and Pernul" [23] propose a security design methodology similar to the classical database methodology (requirement analysis, conceptual, logical, and physical design) which covers requirements and concrete implementations in commercial systems.

Though most conceptual models for DW in the literature does not address security, lately some interesting proposals were devised which define specific authorization models (21). However these proposals mainly deal with OLAP operations accomplished with OLAP tools, thus they are unsuitable for integration in multidimensional modeling as part of DW design

However In [24] authors extend the "ADAPTed" UML (which uses ADAPT symbols as UML stereotypes) model for the aforementioned conceptual phase and specify a methodology and an MD security constraint language for the conceptual modeling of OLAP security. These approaches offer security models at the conceptual level by means of security constraints, but basically deal with OLAP operations. In short, these works implement the security rules considered in their conceptual approach to commercial database systems.

The proposed methodology focuses solely on the conceptual stage of the DW life cycle, other stages are not taken into account and no method exists with which to perform it.

None of the various approaches concerns security from requirement level which completes the whole DW development lifecycle. However these various approaches start deals with providing the security from conceptual model to the implementation of Data warehouse.

IV. COMPARTIVE ANALYSIS OF VARIOUS SECURITY BASED APPROACHES IN DW DESIGN:

In Data Warehouse various numbers of approaches have been introduced in designing a data warehouse. However these design approaches did not deals with a security. After, 2004 few good endeavors will introduce a security in the design of data warehouse i.e. (“Emilio Soler”, “Juan Trujillo”, “Fernandez-Medina”, “Mario Piattini”). These authors introduce a security in data warehouse designing by the use “UML” which basically defines security at conceptual level not in logical and physical levels. Some author introduces a security in data warehouse by the use of models in software development i.e. “MDA” (Model Driven Architecture). This standard of technology basically defines the transformations Model Object Facility (MOF) 2.O Query/Views/Transformations. (QVT).

According to authors propose work regarding to secure data warehouse we have done the evaluations of various security based approaches in the designing of data warehouse by various parameters.

TABLE: SHOWS THE COMPARSION OF DW DESIGN APPROACHES FROM SECURITY PERSPECTIVES:

Year	**Methodologies Proposed by:**	**Security Specific at levels**				**Security based approaches**	**Parameters**		
		Requ.	**Concep.**	**Logical**	**Physical**		**Extensibility**	**Easy to implement**	**Traceability**
2005	Fernandez-Medina, Juan Trujillo, Mario Piattini **(6)**		Yes			Security model (ACA) and extend UML by this by model.	Yes (Extend in 2006,2007)	No (Only be defines security at conceptual level by the used of security model)	No (not give complete information at each level but provided a security)
2006	Fernandez-Medina, Juan Trujillo, Mario Piattini **(25)**		Yes			Using UML 2.0 Extensibility Mechanism approaches	Yes(Extend in 2007 in which UML extension was being introduced)	No (Only be defines security at conceptual level but not to define how it can be implement at this level)	No (not give complete information at each level) but provided a security)
2006	Emilio Solar , Rodolfo Villarreal, Juan Trujillo, Fernandez-Medina **(11)**		Yes	Yes		Defines Extension of Relational Model to consider a security and audit measures	Yes (Extend in 2008 in which security and audit measures being defined)	No(Only be defines security at conceptual and logical level but not define how it can be implement at this level)	No (not give complete information at each level but provided a security))
2007	Fernandez-Medina, Juan Trujillo, Rodolfo Villarreal **(10)**		Yes			Defines Extension of UML Profile for specific security constraints	No (after this approach it cannot be extend because this approach is	No(Only be defines security at conceptual level but not define how to implement at this level)	No (not give complete information at each level but provided a security)

							feasible only at conceptual level)		
2008	Emilio Soler, Fernandez-Medina, Juan Trujillo, Mario Piattini **(14)**	Yes				Defines a requirement based security models used as modeling security requirement	No (because this approach is only feasible at requirement level)	No (Only be defines security at Requirement level so, there is no implement work at this level)	No (not give complete information at each level but provided a security)
2009	Carlos Blanco, Ignacio Garcia-Rodriguez de Guzman, Juan Trujillo **(15)**		Yes	Yes	Yes	Define MDA Approaches defines security models.	Yes (this security based approach can be extend in 2009)	Yes (Because DW implement at physical level that is: by the use DBMS tools)	Yes (give complete information at each level and also provided a security)
2009	Carlos Blanco, Ricardo Perez-Castillo, Juan Trujillo **(16)**		Yes	Yes	Yes	Using MDA Approach defines security issues and provided automatic transformaions b/w models.	No (after this approach it cannot be extend by an author)	Yes (Because DW implement at physical level that is: by the use DBMS tools)	Yes (give complete information at each level and also provided a security)

V. CONCLUSION

In this paper we have seen that several approaches have been proposed by various authors to secure a data warehouse but most of the approaches basically start from conceptual level to final implementation of DW. However the approaches that starts from the conceptual level gives the complete information in data warehouse design from security perspectives. But none of the approaches will be start to secure a data warehouse design from requirement level to final implementation. However In this paper we surveyed the literature related to work done by various authors in past few years and evaluation of various security based approaches in data warehouse design by considering various parameters.

REFERENCES

[1] F. Saltor, M. Oliva, A. Abelló, J. Samos, Building secure data warehouse schemas from federated information systems, in: H. Bestougeff, J.E. Dubois, B. Thuraisingham (Eds.), Heterogeneous Information Exchange and Organizational Hubs, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002, pp. 123–134. ISBN: 1-4020-0649–7

[2] P. Devanbu, S. Stubblebine, Software engineering for security: a roadmap, in: A. Finkelstein (Ed.), The Future of Software Engineering, ACM Press, New York, 2000, pp. 227–239.

[3] E. Ferrari, B. Thuraisingham, Secure database systems, in: M. Piattini, O. Dı´az (Eds.), Advanced Databases: Technology Design, Artech House, 2000.

[4]. R. Kirkgöze, N. Katic, M. Stolda, and A. M. Tjoa, "A Security Concept for OLAP. (DEXA'97)," Toulouse, France

[5]. Rosenthal and E. Sciore, "View Security as the Basic for DW Security," (DMDW'00), Sweden, 2000

[6]. Emilio Soler, Juan Trujillo, Eduardo Fernandez-Medina and Mario Piattini, "A Framework for the Development of Secure Data Warehouses based on MDA and QVT" Second International conference on availability, reliability and security 2007.

[7]Diego Calvanese, Data Integration in Data Warehousing, International Journal of Cooperative Information Systems Vol. 10, No. 3 (2001) 237–271

[8] Golfarelli, M., Maio, D., and Rizzi, S., *The Dimensional Fact Model: A Conceptual Model for Data Warehouses.* Int. Journal of Cooperative Information Systems (IJCIS),
(1998) (2-3): 215-247.

[9]. Sapia, C., Blaschka, M., Höfling, G., and Dinter, B.: *Extending the E/R Model for the Multidimensional Paradigm.* in *1st Int. Workshop on Data Warehouse and Data Mining.*

[10**]** Fernandez-Medina, E., et al., *Access Control and Audit Model for the Multidimensional Modeling of Data* Decision Support Systems, 2006. **42**: p. 1270-1289.

[11] Fernandez-Medina, Juan Trujillo, Rodolfo Villarroel “Developing a secure data warehouses with a UML extension” published in science direct 0306-4376.

[12] Emilio Soler , Rodolfo Villarroel, Juan Trujillo” Representing security and audit rules in Data warehouse at the logical levels by using the common warehouse met model” Proceedings of the First International Conference on Availability, Reliability and Security (ARES'O6)0-7695-2567-9106 $20.00 2006 IEEE.

[13] **J**. Miller and J. Mukerji, “MDA guide version 1.0.1, “2003”.

[15] Emilio Soler, Juan Trujillo, Fernandez-Medina, “Towards comprehensive requirement analysis for DW: Considering security requirement Published in IEEE Conference in 2008.

[16] Carlos Blanco, Ignacio Garcia-Rodriguez de Guzman,” Automatic generation of secure multidimensional code for data warehouse by using QVT transformation: Proceeding in international conference in 2009

[17] Carlos Blanco, Ricardo Perez-Castillo,“ Towards a Modernization Process for secure Data warehouse” Published DAWAK 2009, LNCS 5691, pp. 24–35, 2009. © Springer-Verlag Berlin Heidelberg 2009

[18] K. Katic, G. Quirchmayr, J. Schiefer, M. Stolba, A.M. Tjoa, A prototype model for Data Warehouse security based on metadata, in: Proceedings of the 9th International Workshop on Database and Expert Systems Applications (DEXA’98), Vienna, Austria, 1998, pp. 300–309.

[19] A. Rosenthal, E. Sciore, View security as the basic for Data Warehouse security, in: Proceedings of the Second International Workshop on Design and Management of Data Warehouses (DMDW’00), Stockholm, Sweden, 2000, pp. 1–8.

[20] F. Saltor, M. Oliva, A. Abelló, J. Samos, Building secure data warehouse schemas from federated information systems, in: H. Bestougeff, J.E. Dubois, B. Thuraisingham (Eds.), Heterogeneous Information Exchange and Organizational Hubs, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002, pp. 123–134. ISBN: 1-4020-0649–7.

[21] R. Kirkgöze, N. Katic, M. Stolba, A.M. Tjoa, A security concept for OLAP, in: *Proceedings of the 8th International Workshop on Database and Expert System Applications (DEXA’97), Toulouse, France, 1997,* pp. 619–626.

[22] W. Essmayr, E. Weippl, F. Lichtenberger, W. Winiwarter, O. Mangisengi, An authorization model for Data Warehouses and OLAP, in: *Workshop on Security in Distributed Data Warehousing, in Conjunction with 20th IEEE Symposium on Reliable Distributed Systems (SRDS’2001), USA,* 2001, pp. 9–13.

[23] T. Priebe, G. Pernul, Towards OLAP security design – survey and research issues, in: Proceedings *of the 3rd ACM International Workshop on Data Warehousing and OLAP (DOLAP’00), VA, USA, 2000,* pp. 33–40.

[24] T. Priebe, G. Pernul, A pragmatic approach to conceptual modeling of OLAP security, in*: Proceedings of the 20th International Conference on Conceptual Modeling (ER’01)*, Yokohama, Japan, LCCS, vol. 2224, 2001, pp. 311–324.

[25] Eduardo Fernandez, Juan trujillio, Mario Piattini: UML/2.0 Extension used to secure a DW” *International Journal published in 2006*