# MULTI LEVEL AUTHENTICATION AND ACCESS CONTROL BASED SECURITY ENHANCEMENTS IN DATA WAREHOUSE

*Thangaraju. G [1], Agnise Kala Rani. X[2]*

**Abstract**

Business Intelligence environment the data warehousing is occupy a major role, the DWH consists the major steps to processes or handle the data's in data warehouse environment. The data can be retrieved from the different sources like ORACLE, SQL Server, Excel, MS-Accesses, Terradata etc.., the each source can have their own security mechanism. In data warehouse processing the first stage the data's can be retrieved from the sources based on our requirements. The ETL processing stage the user can connect to the data sources via the with our connection utility based on the software, on that time the user can connect to the OLTP databases, In that situation the Source database can have access privileges' with options of either in Operating System based authentication or database based authentication or third party authentication. The access time of data and accessibility of data based on the authentications chosen by the user. In this paper to provide the mechanisms to access the data from the OLTP and other sources in efficient manner. The remaining part of this paper explains the Introduction about the authentications level and methodology to implement the above requirements.

*Index Terms* : Multi level Authentication, Data Warehousing, Security Management, Access Control, User.

## I. Introduction

In current scenario of the enterprises are taking the decisions for the feature development based on the existing data or information available in the enterprise or company. In up-to-date information or data available on the company databases, the size of the data should be update on every day activities of the firm. In olden days the data's are maintained at the relational model but the huge amount of data has to be stored in well organized manner so that it could be retrieved easily. The datawarehounig is the processes of retrieving data from the sources and storing data on to the database organizing data among the stored data and take decision for feature plan.

The Enterprise contains the huge information related to the different departments like purchase, sales, production, marketing, and administration , the each department can have large amount of data with the latest updating information . In such large organization, the security is the major issue which restricts the unauthorized user from accessing the information stored. In particular the user in purchase department can access only the purchase department details, he can not allowed to the sales department data, like role and access rights are vary from

[1]Research Scholar, Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore - 641 021.

[2]Professor, Department of Computer Applications, Karpagam Academy of Higher Education, Coimbatore-641 021. Email : agneskala72@gmail.com.

user to user. Providing such a restriction in the data warehousing is essential and named as security management.

The Integrity management is the process of maintaining security to the data from the DWH according to the information given. In the data warehouse there exist a number of relational data objects and some of them would keep standard information and some of them keep some unique information about the organization Such information is more responsive and secret which must be secured from unofficial access. To keep the information more secure and to restrict the access in proficient manner there are a number of approaches which have been discussed in the literature but each of them has their own merits and demerits. This paper focuses on improving the effectiveness of integrity management and how the security can be enforced in a proficient manner.

Authentication and Access control in the data warehousing environment can be imposed in many ways. The simple password based approach is not enough for the management of large warehouse where exist a large number of users. During the data access from the other sources like oracle, excel, Sql server and MYSQL can have the own authentication or support of operating system authentication, the user can select the source from the list the security based on the source database. After connecting the OLTP database the connected user's access control rights are verified after that based on the accesses rights the data will be provided to the users. The source profile can the meta data about the sources authentication type and authentications rights and user profile is the Meta data about the access rights of different users and using such Meta data the system can identify whether the user has the rights to read the data object. Here verify the authentication type of the database with the source profile and the authentication type of database whether it is match or not, after that users available in the respective sources can verify with the user profile and also can check with the access rights of the individual users based on the access rights the data or information accessed by the user. This paper focus the methodology of the authentication type of the source database, user selection and access rights of the users.

For any given request from the user, the method can identify a authentication type and user selection and access rights of the users on database . Using the source profile and the list of users needed to access, we can compute the level .

## II. Related Works

There are a number of techniques discussed for the improvement of integrity management and we discuss some of the methods are discussed here in this section.

Data integration on data warehousing proposed by Diego Calvanese for secure data warehouse data ie. Is called Classical Security Model. It describes the security in design process it covers requirements analysis and Conceptual, Logical, Physical design. It deals with the operational data of the industry. In this method only suitable for traditional databases. So this model is not suitable for data warehouse security [1].

A Survey on current security strategies in Data Warehouses, IJEST International Journal of Engineering Sciences and Technology, Anshunman Kumar et al. et al

discussed about the Classical Security model , Applying an MDA-based approach for develop secure data warehouses , Building a secure star schema in data warehouse , data filtration and data encryption before storage in data warehouse and A hybrid approach for securing data warehouse after the study of the above security solution methods he concludes some of the authors describe the security about design level and some authors describe the security in access level and came to the conclusion drawbacks available the above discussed methods[2].

A Prototype Model for Data Warehouse Security Based on Metadata, is a method proposed by N.Katic et al. in this method based on the metadata security it consist the further security levels structural metadata security, access metadata security and discussed about the R-OLAP layer, the M-View environment, Secure Query Management layer and modes of operation of M-View and finally he conclude the data or information accessed by the user or limit ie. Other than the limit of the user can not view by the users [3].

A Survey on Current Security Perspective in Data Warehouses , International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE) by G.Thangaraju et al he was survey the nearly 13 papers related to the security perspectives of data warehouse environment and he came to the conclusion security implementation in data warehouses are in different perspectives and he suggest the Integrity management security is in poor finally he proposed a method called Multi Attribute integrity management,

An Improved Security Framework for Data Warehouse: A Hybrid Approach, Proposed by Sajjad Ahmad et al.., Proposed a hybrid model for data warehouse security it encomposses three stage security first stage encryption filter is used in between the OLTP and Data warehouses, the internal security is in DWH and the security in DWH to destination named as encryption filter, data warehouse Internal security, User levels. Finally he concludes the DWH is typically developed as DSS, so this attempt of identifying the level of users can be a major help for the top level management, this approach is useful for hide the information about low level and middle level management [5].

A Schematic Technique Using Data type Preserving Encryption to boost Data Warehouse Security, proposed by M.Sreedhar Reddy et al. discuss the security features in the data type preservation, some of code not supported to the encryption and decryption, the proposed solution is support to the encryption and decryption operations , the methodology encompasses the mod operations , and the enhanced encryptions uses the techniques DES with key values[6].

Data warehouse Security Using Log Based Analysis: A Review, Raj et al. data masking are major security measure, the classification of normal and malicious user. The main focus is the data vulnerability over the data transmission on the network. The conclusion of this paper prevention of malicious user is the major consideration [7].

Using Secret Sharing Algorithm for Improving Security in Cloud Computing, the strongest cryptography algorithm named Shamir's secrete sharing algorithm, which has number of features including security, client-side

aggregation. It claims that security is maintained even when k or more servers collude. We have found that much investigate has been done to ensure the security of the single cloud and cloud storage whereas multi-clouds have received less attention in the area of security. We validate the moving to multi-clouds due to its ability to decrease security risk that affect the cloud computing user. The key conclusion is that proposed work provides privacy, data integrity, improved availability and capacity to handle multiple requests at a time [8].

Enabling Privacy Preservation Technique to Protect Sensitive Data with Access Control Mechanism using Anonymity, Barkha kasab et al., proposed a new methodology for ensuring the data integrity with the access control mechanism, he stated the some methodologies first one is General System Architecture access Control model, Access control for relational data, Anonymization, permission and imprecision bound, these algorithms are have the minimum security in dwh users, K-Anonymization algorithm and L-Diversity Algorithm uses the Generalization and suppression these two techniques are used prevent the data from un authorized users[9].

BYOD with Multi-Factor Authentication, Surabhi Shukla and Neelam Joshi , BYOD stands for Bring Your Own Device, this concept is role play in cloud environment, The security in the cloud computing is in different levels , here the author quoted in Multi-Factor Authentication , like KBA, OTP, and TPA based on the services the security ie Identity-Access Management is implemented on the MFA[10].

Multi Attribute Schematic Relational Mapping Based Integrity Management for Security Enhancement of Data Warehouse, G.Thangaraju et al. proposed a method MASRM, this method provides the Integrity management in data warehouse via the Relational Rule Generation, Rule Mapping and Schematic Score computation, both of the methods generate the relational rule based on the sensitive and nonsensitive attributes, and compared to the roles of the attribute, so the integrity will be maintained by the MASRM technique is efficient[11].

Performance in case of data intrusion and service All the methods discussed above have the problem of ensuring security and restricting access in data warehouse and produce higher false classification.

## III. Methodology

Multi Level Authentication and Access Control Based Security Enhancements in Data Warehouse:

The multi level authentication and access control based security enhancements in data warehouse based authentication approach reads the source connection , select the any one sources among the list, then verify the authentication type of the source , based on the type select the source database object and no.of users authenticated to the database object , the user or the source connection authenticator accesses rights are verified if having the read permission , then the table or data objects accssed by the datawarehouse and assigned to the target databases , if the target database have the same no.of attributes and data types , the data or records are stored to the target database. During this process the no.of levels can be verified to identify the exact users are calculated by the access level computation methods. The

entire process can be split into different stages namely Source profile verification, User Access tree generation, Generation, multi authentication and access control, access depthness measure computation and Validation. This section discusses about each of the stages in detail.

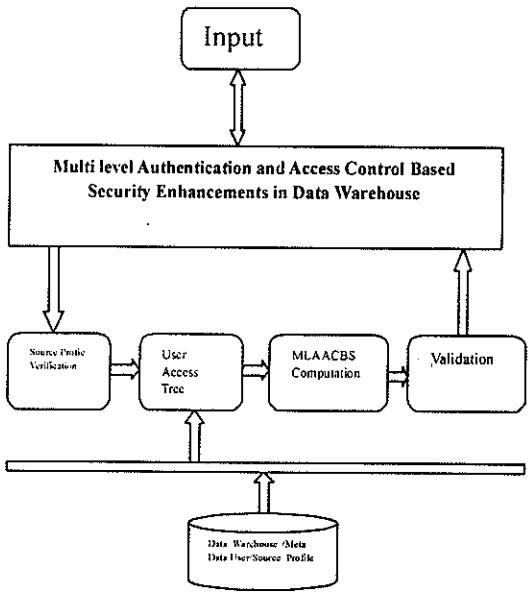



Figure 1 : Structural design of the proposed method

Figure 1, shows the Structural design of the proposed approach and shows the functional components.

**A). Source Profile Verification**

The method takes the input query from the user given the source connection parameter and performs parsing of input query. From the parsed result, the method identifies the set of sources available in the meta data tables mentioned in the query. Also the method identifies the authentication types of the sources chosen by the user. Also list the users available in the database or operating system. Also the method verifies the username and if the username password is matched then the connection will be established and call a User Access Tree generation procedure for further process. All these information are passed to the multi level authentication and access control based security enhancements in data warehouse, attribute access depthness measure computation. Based on the result from the MLAACBS value, the method decides the access result of the user query.

**Algorithm :**

Input: Data warehouse Metadata Med, Source Profile Verification Query SPVQ

Output: Result Res.

Start

Read metadata Med.

Read input query SPVQ.

Initialize Source Name set Sns.

Initialize Authentication Type set Ats.

Initialize user names set Un.

Initialize passwords set Pw.

Key set Ts = $\int Parse(SPVQ,'')$

For each term Ti from Ts

Identify the data Source name.

If $\int_{i=1}^{size\ (med)} Med(i).sname == Ti$ then

Add to Source Name set Ats = $\sum Sourcenames \in (Ss) + Ti$

End

$\int_{i=1}^{size\ (Aut\ hentication\ )} Authentication.sname == Ti$ Then

Add to Authentication set Ats = $\sum Authentication(Med) + Ti$

Identify list of user names and passwords accessed.

Uns set un= $\int_{i=1}^{size\ (Uns\ (Ats\ (i)))} \sum Usernames\ \ Required$

Pws set Pw= $\int_{i=1}^{size\ (Pws\ (Ats\ (i)))} \sum Passwords\ Required$

End

MLAACBS set Mlaacbs = Compute Multi Authentication access control Username password Access Depthness Measure (Sns, Ats, Un,Pw).

For each level l

If Maacbss(i)>Un then

Process the query and send result

Else

Drop the query

End

Stop.

The algorithm displayed above performs SPV identifies the set of sources available in the metadata required and set of authentication types of all the sources accessed and then identifies list of usernames and passwords. Using all the above the method computes the multi Authentication access control user name password access depthness measure and based on the value of MlAACBS value the method restrict the user from accessing the system.

**B). User *Access Tree Generation***

In this stage the function generates the user access tree from the Meta information available. The function maintains different level of access modes by placing the sources at different levels. First the functions creates the root node and the first level contains leaf of source nodes which has open access. According to the number of sources which has public access the functions creates a number of nodes and each assigned with different sources of different data tables. Similarly according to the Metadata information the functions creates a number of levels according to the importance of sources. If the functions classifies the sources as 3 levels then there will be three levels in the tree. The number of levels can be extended up to any level.

Pseudo Code of User Access Tree Generation:

Input: Meta Data Med

Output: User Access Tree UAt.

Start

Create Tree Tre.

Create Root node Ron.

Add Rn to T.

'For each source table Sti from Dm

Identify all sources.

Sns= $\int_{i=1}^{size\ (Dm)} \sum Souce \in Dm(i)$

For each source Si from Sns

Identify level l.

L = $\int_{k=1}^{size\ (level)} Med(k) \in Si$

Create Node N.

Initialize N with Si.

Add to tree Tr.

T = $\int \sum (Nodes \in Med(k)) \cup N$

End

End

Stop

The above discussed algorithm generates the access tree which will be used to compute the MLAACBS value.

**C). MLAACBS Computation :**

The function reads the user access tree and the source profile with the Authentication type set Ats, user name set, password set and attribute set. Using all this, for each level the functions computes the access depthness measure based on a number of sources in any level has been required and a number of sources in any level is

given. The lookup is performed from the source profile and computes the access depthness measure for each level. The computed value will be given for query processing.

Pseudo Code for MLAACBS Computation:

Input: Source Profile Sp, user access Tree UAt, Authentication type Ats.

Output: MLAACBS set Mls.

Start

For each level l from tree UAt

Extract total Sources.

$$LS = \int_{i=1}^{size\,(nodes)} \sum Node.Source \in UAt(i)$$

Compute number of sources.

NS = size(LS).

Identify number of sources required.

$$NU = \int_{i=1}^{NS} \sum Source(As) \in LS$$

Compute Multi Attribute Access depthness measure MLAACBS.

$$MLAACBS = \frac{NS}{NU}$$

Add to Mls.

End

Stop

The algorithm discussed above computes the multi level authentication and access control depthness measure which will be used to restrict the user from unauthorized access.

**D).** ***Validation:***

At this stage, the functions performs all the operations mentioned above by using the procedures. First, the functions performs query preprocessing and then generates user access tree using the Meta data. Second, the functions MLAACBS computes depthness, and based on the value returned, the functions executes the query for the user and returns the results to the user. The results to the user are fully based on the values of Authentication type sets, Source sets, obtained on the query submitted which are computed depending on the source profile and depthness of the user query.

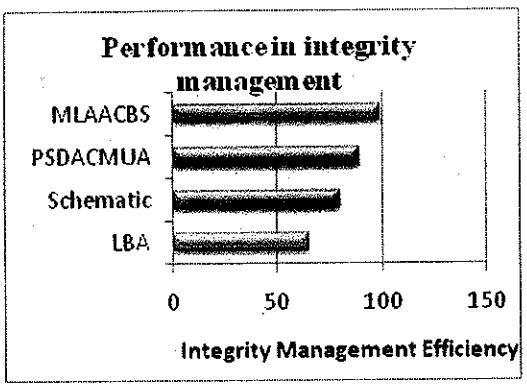
## IV. Results And Discussion

The proposed functions has been designed and implemented using the SQL data base which has number of relational database and other sources like flat file. The warehouse has been created with large no.of relational database and has been evaluated from the user query in lacks. The details of evaluation have been listed below :

Table 1: Details of model parameters

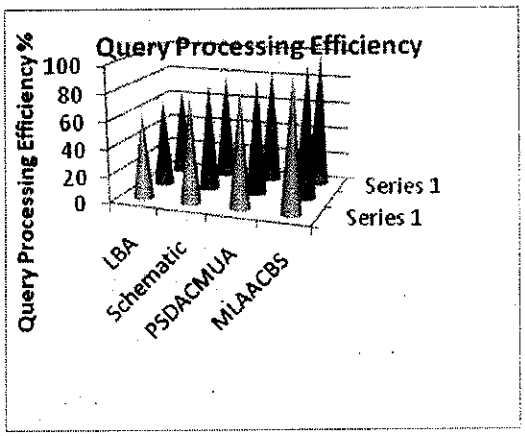
| **Parameter** | **Name** |
|---|---|
| Data warehouse Tool | SQL |
| Number of database | 7500 |
| Number of Source dbs | 5 |
| Number of queries | 100000 |
| Number of users | 575 |

Table 1 shows the details of model parameters being used to perform evaluation of the proposed approach.




Graph 1: Comparison of integrity management efficiency

Graph 1, shows the efficiency of integrity management produced by different functions and it shows clearly that the proposed functions has produced efficient integrity management than other functions.




Graph 2 : Comparison of query processing efficiency

Graph 2, shows the comparative analysis of query processing produced by different functions and it shows clearly that the proposed method has produced more efficiency than other functions.

## V. Conclusion

In this paper, an intelligent source profile authentication type user access control function is proposed for the integrity management of data warehouse systems. First the function identifies the key terms and sources and functions from the input query. Then from the inferred result, the function identifies a set of all sources of different users require must to process the query. Using this information and the source profile and Meta information the function builds the user access tree and using the access tree with the information extracted the method computes the access depthness measure for each level of the tree. Then for each level the method verifies the measure and the profile value to proceed with the query processing. The proposed method improves the performance of the query processing and improves efficiency also.

## References

[1] Diego Calvanese, Data Integration in data warehousing, International Journal of Cooperative Information Systems, Vol.10, No.3 (2001)237-271.

[2] Anshunman Kumar Saurabh, Bharti Nagpal, A Survey on current security strategies in Data Warehouses, IJEST International Journal of Engineering Sciences and Technology, Vol. 3, No 4 April 2011.

[3] N.Katic and G.Quirchmayr, *"A Prototype Model for Data Warehouse Security based on Metadata"*.

[4] G.Thangaraju and Dr.X.Agnes Kala Rani: A Survey on Current Security Perspectives in Data warehouses, International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE) ISSN: 0976-1353 Volume 19 Issue 2 – JANUARY 2016.

[5] Sajjad Ahmad, Rohiza Ahmad, Rohiza Ahmad, An Improved Security Framework for Data Warehouse: A Hybrid Approach, IEEE Conference, 978-1-4244-6716-7/10@2010.

[6] M.Sreedhar Reddy , Prof.M.Rajitha Reddy , Proff R.Viswanath, Prof.G.V.Chalam et al., A Schematic Technique Using Data type Preserving Encryption to Boost Data Warehouse Security, International Journal of Computer Science Issues, Vol. 8, Issue 1, January 2011.

[7] Raj Rani, Data warehouse Security Using Log Based Analysis: A Review, International Journal of Advanced Research in Computer Science and Software Engineering, April-2014, pp.447-449.

[8] Swapnila S Mirajkar, Santoshkumar Biradar, Using Secret Sharing Algorithm for Improving Security in CloudComputing, International Journal of Advanced Research in Computer Science & Technology (IJARCST 2014)

[9] Barkha Kasab, vinayak pottigar and Swapnaja Ubale, Enabling Privacy Preservation Technique to Protect Sensitive Data with Access Control Mechanism using Anonymity, International Journal of Computer Science and Engineering, Volume-3,issue-10.

[10] Surabi Shukla and Neelam Joshi , BYOD with Multi-Factor Authentication, International Journal of Computer Science and Engineering, Volume-3,issue-10.

[11] G.Thangaraju and Dr. X.Agnise Kala Rani, Multi Attribute Schematic Relational Mapping Based Integrity Management or Security Enhancement Of Data Warehouse.

## Author's Biography

**Mr.G.Thangaraju**,received his M.Phil in computer science from Mononmaniyam University, Thirunalvelli, India in 2004. Pursuing Ph.D in computer science at Karpagam Academy of Higher Education. Currently he is working as Assistant Professor in the Department of Computer Applications, Thanthai Hans Roever College. His current research interests Data warehousing, Distributed Data Base and software metrics.

**Dr.X.Agnise Kala Rani** is a Professor in the Department of Computer Applications at Karpagam Academy of Higher Education. She received her Ph.D. in 2011 from Mother Teresa University. She holds the Master of Engineering degree from VMKV University and Master of Computer Applications degree from Madras University. She has several publications including scientific journals and top-tier networking conferences to her credit.