Visualization of Query Processing on Data Warehouse with UML

Ugale Shrikant, Godse Gaurav, Kale Onkar

Abstract- Data transformations are the main subject of visual modeling concerning data warehousing dynamics. A data warehouse integrates several data sources and delivers the processed data to many analytical tools to be used by decision makers. Therefore, these data transformations are everywhere: from data sources to the corporate data warehouse by means of the ETL processes, from the corporate repository to the departmental data marts, and finally from data marts to the analytical applications Data warehousing involves complex processes that transform source data through several stages to deliver suitable information ready to be analyzed. Usually Database designer process and extract the data for their use or according to the business needs. Anyone from the organization can specify the query and get the data related to it from the data warehouse. There may be the case when the database designers or business analysts need to view query flow. In other words, if they want to analyze the flow of query in data warehouse that is how query actually flows from one table to another, they cannot easily visualize it. Though many techniques for visual modeling of data warehouses from the static point of view have been devised, only few attempts have been made to model the data flows involved in a data warehousing process. Besides, each attempt was mainly aimed at a specific application, such as ETL, OLAP, what-if analysis, data mining. Data flows are typically very complex in this domain; for this reason, designers would greatly benefit from a technique for uniformly modeling data warehousing flows for all applications. The visualization of query flow is interactive to the designer and analyst. If the designer have the model to view the internal flow then it will be easy for them to visualize the important data and attribute. UML has an activity diagram which shows the complete and accurate flow actions. So if we have activity diagram to specify the flow. It will be easy to the designers to understand the query.

Index Terms- OLTP, OLAP, ETL, data warehouse, data mining

I. INTRODUCTION

Visual Modeling of Data Warehousing Flows with UML is done to make the analyst easier to view the complete flow of query. The visualization of query flow is interactive to the designer and analyst. If the designer have the model to view the internal flow then it will be easy for them to visualize the important data and attribute. UML has an activity diagram which shows the complete and accurate flow of actions. So if we have activity diagram to specify the flow, it will be easy to the designers to understand the query.

Manuscript Received on September, 2013.

Ugale Shrikant, B.E. student of Vishwabharti Academy's College of Engg.,Department of Computer Engg Ahmednagar.State-Maharashtra, India.

Godse Gaurav, B.E.student of Vishwabharti Academy's College of Engg.,Department of Computer Engg Ahmednagar.State-Maharashtra, India.

Kale Onkar , B.E. student of Vishwabharti Academy's College of Engg., Department of Computer Engg Ahmednagar.State-Maharashtra, India.

The project and all its details are processed swiftly. When user of the system fires the query on data warehouse, the framework represents the flow of the query in the form of activity diagram of UML. With the help of class diagram, the designers would get to know the attributes and operations of the table. With the help of the query fired on database and the class diagram, XML file is generated. And then the activity diagram is displayed.

Data warehousing commonly implies complex data flows, either because of the large number of steps data transformations may consist of, or of the different types of data they carry. These issues raise interesting challenges concerning design-oriented modeling of data warehousing flows. In particular, the thorough visualization of these models has a deep impact on the current trends for data warehousing design, where the so-called model-driven technologies promote diagrams as the main, tentatively unique, design artifacts managed by software engineers. Cognitive aspects, such as diagrams readability, are thus related to the productivity of the whole development process. Many techniques for visual modeling of data warehouses from the static point of view have been devised, only few attempts have been made to model the data flows involved in a data warehousing process. Besides, each attempt was mainly aimed at a specific application, such as ETL, OLAP, what-if analysis, data mining. To resolve the problem, visual modeling framework is better way to overcome the inefficiency of the decision makers. In the proposed framework, designers get the internal flow of the query. When user of the system fires the query on data warehouse, the framework represents the flow of the query in the form of activity diagram of UML.

Intended audience:

- Business Analyst
- Decision makers
- Database Handler

II. LITERATURE SURVEY

Data warehousing commonly implies complex data flows, either because of the large number of steps data transformations may consist of, or of the different types of data they carry. These issues raise interesting challenges concerning design-oriented modeling of data warehousing flows. In particular, the thorough visualization of these models has a deep impact on the current trends for data warehousing design, where the so-called model-driven technologies promote diagrams as the main, tentatively unique, design artifacts managed by software engineers. Cognitive aspects, such as diagrams readability, are thus related to the productivity of the whole development process. Nevertheless, the main research efforts made so far have concerned the static modeling of

the data warehouse repository and, even when data warehousing flows were considered, it was done within specific business intelligence applications (OLAP, data mining and so on). While these efforts were addressed at designing individual modeling frameworks, all of them characterized nothing but data transformations.

Many techniques for visual modeling of data warehouses from the static point of view have been devised, only few attempts have been made to model the data flows involved in a data warehousing process. Besides, each attempt was mainly aimed at a specific application, such as ETL, OLAP, what-if analysis, data mining.

III. SYSTEM MODEL

Visual modeling framework is better way to overcome the inefficiency of the decision makers. In the proposed framework, designers get the internal flow of the query.

When user of the system fires the query on data warehouse, the framework represents the flow of the query in the form of activity diagram of UML. A database design that is based on central detail fact table link to surrounding dimensional tables. They allow access to data using business terms and perspectives. Business intelligence is broad category of technology that allows for gathering, storing accessing and analyzing the data to help business, users and make better decisions.

In this framework, there is a class diagram which represents the name of the table, attributes and operations which are related to that query. With the help of class diagram, the designers would get to know the attributes and operations of the table. With the help of the query fired on database and the class diagram, XML file is generated. And then the activity diagram is displayed. Various types of queries can be fired on data warehouse such as group by, having, nested etc. If the query format specified by the designer is not according to the format and even if the query is incomplete, the query does not get executed. So here query checking is also performed.

The framework is linked with data warehouse. Database should be selected by the designer before firing the query. Only the authenticated user should have access to the system

Objective 1: To provide query input for activity diagram generation specified by user. The system will allow users to provide username and password and authenticate them.

Objective 2: To provide query authentication.

System will accept input query and validate the query.

Goal 2: Activity diagram generation Objective 1: Provide output.

The system will enable user to view the activity diagram. The diagram is generated with the help of XML file which is generated by query processing.

Objective 3: Save diagram

Output will be displayed on screen and can be saved to a file. The system will enable the user to view the activity diagram of an input query in future therefore the diagram is saved in jpg format. The generated diagram can be saved in user defined path or default path of system.

The following block diagram shows the major components of the system, interconnections, and the external interfaces.

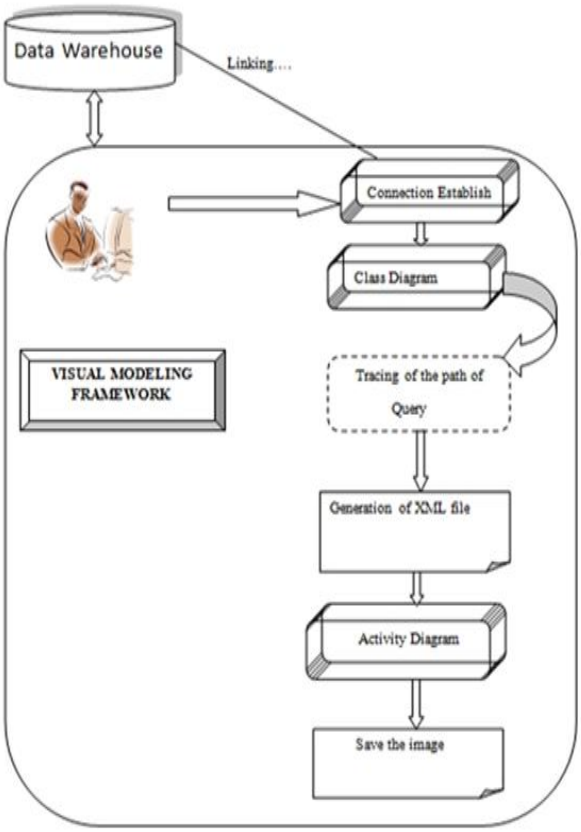

Figure 1:Block Diagram of VMDW

IV.PROPOSED WORK

To resolve the problem, visual modeling framework is better way to overcome the inefficiency of the decision makers. In the proposed framework, designers get the internal flow of the query.

When user of the system fires the query on data warehouse, the framework represents the flow of the query in the form of activity diagram of UML.

A database design that is based on central detail fact table link to surrounding dimensional tables. They allow access to data using business terms and perspectives. Business intelligence is broad category of technology that allows for gathering, storing accessing and analyzing the data to help business, users and make better decisions.

In this framework, there is a class diagram which represents the name of the table, attributes and operations which are related to that query. With the help of class diagram, the designers would get to know the attributes and operations of the table. With the help of the query fired on database and the class diagram, XML file is generated. And then the activity diagram is displayed.

VI. TECHNICAL SPECIFICATION

A. Advantages

- Designers would greatly benefit from a technique for uniformly modeling data warehousing flows for all applications.
- End user readability will be enhanced.
- Data warehouse implies complex data flows.

- So developing design-oriented modeling of data warehousing flows would be interesting challenge.
- Business users want to tease out the interesting and relevant aspect of the data.
- Business users want analytic applications to help them perform analysis and gain understanding.
- It will reduce the time consumed in understanding the flow of data in database.
- If you are taking long time in understanding the flow of diagram, it will surely be affecting your performance. It will help in taking quick decision by database developers.

B. Applications

- It will be helpful to new database developer as well as experience persons for dealing with complex database queries.
- IT data wants to ensure access, security and performance tuning and archiving of the data warehouse remains appropriate for the overall organization needs.

VII.CONCLUSION

The state-of-the-art for visual modeling of fw's comprises a wide range of techniques, each taking into account specific aspects of application domains, but overlooking their common foundational concepts. In this work project identified two challenging issues concerning design oriented fw visual modeling: how to handle complex data structures, and how to specify the semantics of the involved data transformations in a formal and straightforward mode. For this reason, we devised an fw visual modeling framework where two kinds of diagrams are provided by using UML as scaffolding. Their suitability to visually manage the complexity involved in fw's is shown by applying them to an example scenario relying on the Net Beans platform. The results of this work have interesting implications for data warehouse practitioners. Regarding the integrated vision of fw's, the current modeling tools that were conceived for a specific kind of fw may be reused for the others. This fact sets a bridge between current visual modeling techniques.

ACKNOWLEDGMENT

My sincere thanks go to Vishwabharti Academy's College of Engineering for providing me a strong platform to develop my skill and capabilities. I would like to thanks to my friends & family for their constant support and motivation for me. We would like to thank Prof. Shital Aher for their support. Last but not least, I would like to thank all those who directly or indirectly assist me in presenting the paper.

REFERENCES

1. Jes´us Pardillo1, Matteo Golfarelli2, Stefano Rizzi2, and Juan Trujillo1,Visual Modelling of Data Warehousing Flows with UML Profiles_, 2009.
2. OMG: Unified Modeling Language (UML) Superstructure, version 2.1.2.(November2007).
3. Luj´an-Mora, S., Trujillo, J., Song, I.Y.: A UML profile for multidimensional modeling in data warehouses.Data Knowl. Eng. 59(3), 725-769 (2006).
4. A UML based Approach for Modeling ETLProcesses in Data Warehouses_ Juan Trujillo and Sergio Luj´an-Mora.
5. Data Mapping Diagrams for Data Warehouse Design eith UML Sergio Lujan-Mora, Panos Vassiliadis and Juan Trujillo.
6. Data warehouse design with uml Sergio Luján-MoraAdvisor:Juan Trujillo.
7. Jia Han, Micheline Kamber., "Data Mining concepts and techniques", second edition 2006, Elsevier Inc.
8. Cabibbo, L., Torlone, R.: From a Procedural to a Visual Query Language forOLAP. In: Proc. SSDBM, pp. 74–83 (1998).