

Ontology and Hyper Graph Based Dashboards in Data Warehousing Systems

J. Gitanjali, Meera Kuriakose and Revathi Kuruba
School of Information Technology and Engineering (SITE), VIT University,
632014 Vellore, Tamil Nadu, India

Abstract: Analytical and computational requirements are increasing in Data Warehousing Systems. So, it is very crucial to identify semantically equivalent data, linking them and discovering structural dependencies. This can be performed using dashboards. Dynamic dashboard will perform the customization of information for specific need of user. Researchers are using an improved method for dashboard development using hyper graph approach and ontology exploration. Existing dashboard design satisfies given user query and provides flexibility in changing user interactions. In this study, a quick specification of dashboard interfaces is obtained for detailed DB (Data Base) query language where semantics of user interaction is in a declarative manner.

Key words: Dashboard, hypergraph, ontology, DBE, query

INTRODUCTION

The variety, complexity and diversity of analytical requirements lead to increased interest of on-demand DWH Systems (Data Warehousing Systems) (Pedersen *et al.*, 2013) where qualitative difference in data monitoring, exploration and other tasks is obtained by interactive and analytical data processing. This maximized the popularity of dashboard applications which analyze and monitor flexible, rapid information for specific need of users (Palpanas *et al.*, 2007). Performance and flexibility in the dashboard development process is maintained by the QBE Method based on knowledge.

Ontological knowledge is utilized by DBE framework to formalize the context and content of heterogeneous sources of data. DBE query language defines queries on data source's base attributes, base data semantics which is ontology encoded and on the analyzed data and the ontology data connections. Ontology explorations associated with hyper graph transversals problem (Brogi *et al.*, 2006). Here, hyper graph clustering and hyper graph partitioning techniques are used (Papa and Markov, 2007) which provide rich information sources for data linkage and semantic browsing tasks.

Dashboard is an easily readable, real-time user interface can be viewed as a single page. It includes current status and organization's historical trends which are presented graphically thus instantaneous and important decisions can be made at a glance. In computer and information science, ontology is representing knowledge as concept sets within a domain and different

relationships among concept pairs. Researchers are describing the method using a real-world based scenario and explaining the benefits regarding the solution's design flexibility.

LITERATURE REVIEW

QBE has been very famous since its arrival and is a well-known (Zloof, 1975) and most of modern database related products use its variants. Most QBE related research works are focused on extension and enrichment of QBE as a query language and obtaining efficient query generation and processing methods (Song, 2009; Rasiwasia and Vasconcelos, 2008; Rasiwasia *et al.*, 2007), e.g., QBSE (Query By Semantic Example) and QBVE (Query By Visual Example). Query by-example interface can be used by the dynamic dashboards where users can draw the required sequence as a specific query, flexible criteria for search is specified and similar scenarios are searched in a user friendly and efficient manner.

HYPER GRAPH-BASED QBE LANGUAGE

With the availability of meta information about the source content and various source access mechanisms, all of the necessary data can be stored inside a data warehouse dashboard retrieval processes is described in Fig. 1.

Dashboard Data Model: The artifacts for the Dashboard Model (Palpanas *et al.*, 2007) used in the approach are associated with data modeling including the indicator

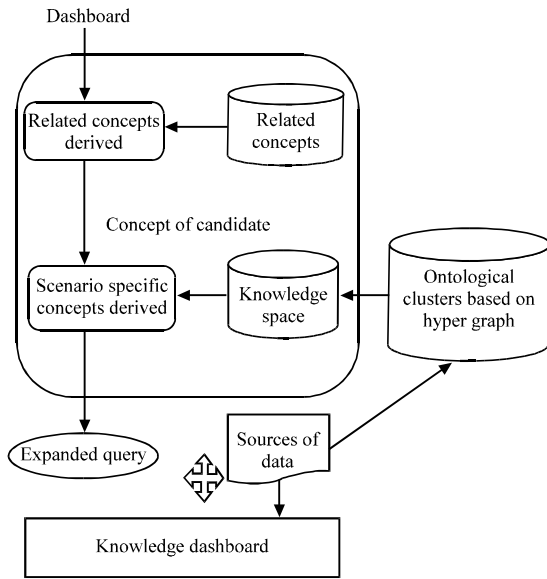


Fig. 1: Dashboard retrieval processes

models and the data. An abstract type presentation layer which includes the dashboard templates and navigation is defined by the Dashboard Model.

Dashboard Navigation Model: In a normal scenario, initially some scenario concepts are defined by the analyst and then these scenario concepts are associated with the indicators widgets. In the last step, a navigation element is introduced by DBE for capturing the navigation paths among various indicator widgets which results in the formation of the dashboard reports.

Logical connectives: Researchers can bind a tuples subset in the related indicators widgets when a suitable value for a widget is given. By partitioning the widgets in the dashboard hierarchically, decomposition can be done.

Interaction semantics: Different types of interactions including addition of a new value to the chosen value set by user are associated with on-demand dashboard designs. Thus, interaction semantics includes a consistent data view and allow the user traversal to other data view in a complete and correct manner. To explore the semantics of the query again the ontological clusters based on hyper graph can be applied.

KNOWLEDGE BASED DASHBOARD BY EXAMPLE FRAMEWORK

DBE have an integrated view and use the dashboard at different levels which helps user in obtaining different

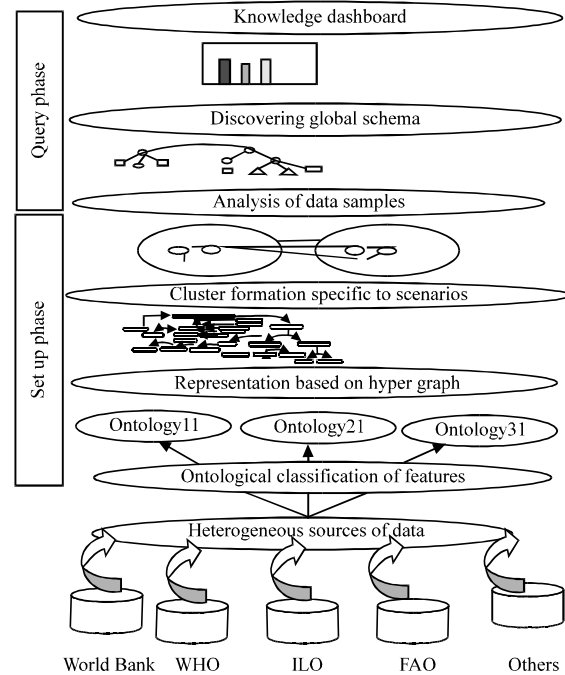


Fig. 2: General architecture of DBE framework

perspective of datasets. Knowledge of DBE includes abstract view of different sources of data depending on ontology. The architecture of dash board frame work (Fig. 2) includes the usage of knowledge representation based on scenario in ontological cluster form.

Semantic DBE framework include two phases: setup phase then query phase. Thus, initially acquisition of knowledge occur data from heterogeneous data sources is obtained. Then, the data classified based on ontology that it is represented by means of hyper graph. Based on scenario different clusters are formed. Then, data sampling is done and a global scheme is discovered. And based on query suitable dashboard design obtained.

KNOWLEDGE ACQUISITION BASED ON HYPER GRAPH

Ontology based hyper graph is used to summarize the ontology graph and hence the queries in the background can he satisfied based on the summary. Since, the summery size is less than original graph, queries becomes faster. The ontological hyper graph is obtained from original graph such that outgoing path nodes are grouped as one and similar incoming path nodes are grouped together (Hoang *et al.*, 2012).

Hyper graph builder: It analyzes the description based on ontology to build a labeled hyper graph which is directed and all synthesis the data dependencies.

Hyper graph development is purely query independent and it can be computed before answering of query in the offline mode. Hence, it does not hinder the dashboard design efficiency.

Hyper graph builder construct the hyper graph vertices based on ontological concepts and the hyper edges corresponding to these vertices include the ontological relationships. Initially ontological concept added to the hyper graph by the hyper graph builder. Next hyper edges are drawn which represent the dependency and relationship between the ontology concepts that are added. The hyper graph builder constructs the hyper graph for specific scenarios. For example health care system hyper graph includes nodes and edges as shown in Fig. 3. Thus, nodes include hospital, patient, inpatient, outpatient, medicine, diseases, etc.

Scenario specific views based on hyper graph: The semantic information flow in many dimensions is discovered by clustering process to obtain different ontological concepts. All related entities collected together to form cluster. The structural relationship

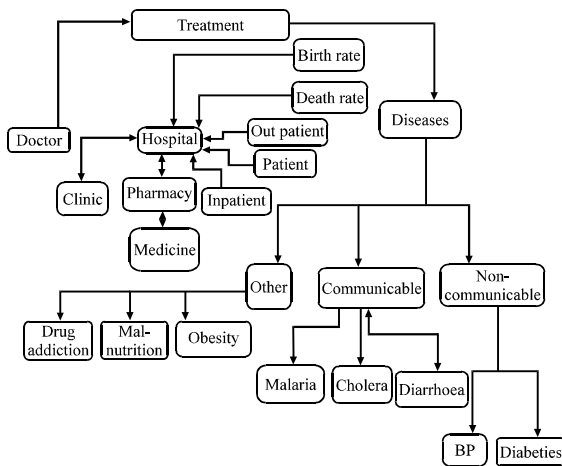


Fig. 3: Hyper graph for health care system

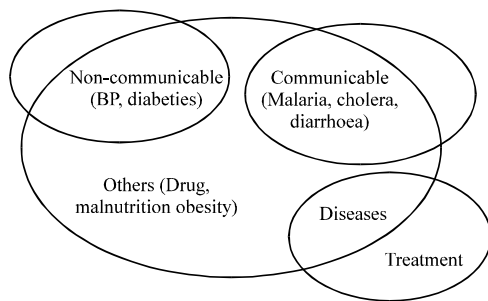


Fig. 4: Cluster border and indicator families

among ontological cluster is obtained by strategy of cluster mapping (Fig. 4). The meaningful data captured, semantic trees formed and cluster relations are discovered.

DBE EXAMPLE

Researchers are considering the healthcare systems. The disjoint and vast data in various hospital department results in a huge decision making bottleneck. There are many aspects associated with a hospital like clinical, financial and operational which determine overall performance of the hospital.

Researchers are also dealing with general health care which includes various health issues across the world like obesity, communicable diseases, diabetes, UV rays, pollution, etc., thus a dashboard is very useful in obtaining an idea about the health issues in a single view.

Also, the hospital administrators are also using dashboards on a large scale for gathering all critical information. Thus, the overall performance can be analyzed. For example different regions prone to malaria during different time period can be easily analyzed. Or obesity rate in different regions during various time span can be studied.

Various hospital related matters like no of outpatients, inpatients, treatment, hospital charges, salary, diseases, births, deaths, financial aspects, etc. can be analyzed. Thus, by referring dashboard, researchers will get all necessary information in a user friendly manner as shown in Fig. 5.

EXPERIMENTAL RESULTS

Aim is to generate dashboards for the various input data. Here, input is taken from user and saving it in

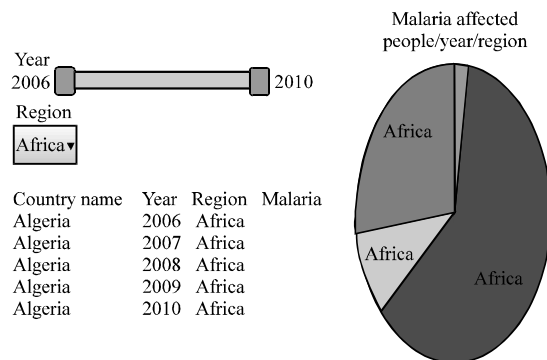


Fig. 5: Sample dashboard showing malaria affected region/year

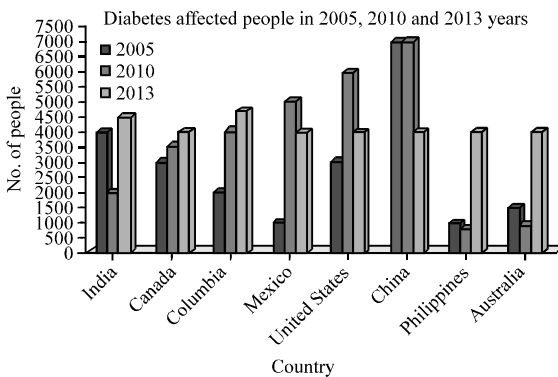


Fig. 6: Output dashboard obtained

the database. Then, based on the specified parameters dashboard is generated. Here, PHP 5.4.3 and WAMP server and MySQL 5.5.24 database is used for generating dashboard. The example considered is of diabetes affected people during specified years for different countries. The dashboard denoting required information is obtained (Fig. 6).

CONCLUSION

A hyper graph based dashboard development is presented which provide more flexibility in varying user requirements. Thus, complex data can be arranged as components of dashboard. All the details can be obtained as single view. Thus, analysis and comparison operations can be done excellently. The model can be extended to specific DBE query languages to model user interaction semantics in a declarative manner. Also, as a future research more dynamic dashboards for frequently changing data can be developed.

ACKNOWLEDGEMENTS

Thanks to the researchers who have contributed towards development of this study. That is Prof. Gitanjali. J (Asst. Prof. (Sr.) School of Information Technology and Engineering (SITE), VIT University), Meera Kuriakose (MTech Software Technology (2nd year School of Information Technology and Engineering (SITE), VIT University), Revathi Kuruba (MTech Software Technology (2nd year) School of Information Technology and Engineering (SITE), VIT University).

REFERENCES

- Brogi, A., S. Corfini and J. Aldana, 2006. Automated Discovery of Compositions of Services Described with Separate Ontologies. In: Service-Oriented Computing-ICSOOC 2006: 4th International Conference, Chicago, IL, USA, December 4-7, 2006, Proceedings, Dan, A. and W. Lamersdorf (Eds.). Springer, Berlin, Heidelberg, ISBN: 978-3-540-68147-2, pp: 509-514.
- Hoang, D.T.A., T.B. Nguyen and AM. Tjoa, 2012. Dashboard by-example: A hypergraph based approach to on-demand data warehousing systems. Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, October 14-17, 2012, Seoul, pp: 1853-1858.
- Palpanas, T., P. Chowdhary, G. Mihaila and F. Pinel, 2007. Integrated model driven dashboard development. Inform. Syst. Front., 9: 195-208.
- Papa, D. and I.L. Markov, 2007. Hyper Graph Partitioning and Clustering. In: Approximation Algorithms and Metaheuristics, Gonzalez, T.F. (Ed.). CRC Press, USA.
- Pedersen, T.B., D. Pedersen and K. Riis, 2013. On-demand multidimensional data integration: Toward a semantic foundation for cloud intelligence. J. Supercomput., 65: 217-257.
- Rasiwasia, N. and N. Vasconcelos, 2008. A study of query by semantic example. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, June 23-28, 2008, Anchorage, AK., pp: 1-8.
- Rasiwasia, N., P.J. Moreno and N. Vasconcelos, 2007. Bridging the gap: Query by semantic example. IEEE Trans. Multimedia, 9: 923-938.
- Song, I., 2009. Semantic query-by-example for RDF data. Proceedings of the 1st International Conference on Emerging Databases, August 27-28, 2009, Busan, Korea, pp: 8.
- Zloof, M.M., 1975. Query-by-example: The invocation and definition of tables and forms. Proceedings of the 1st International Conference on Very Large Data Bases, September 22-24, 1975, Framingham, Massachusetts, USA., pp: 1-24.