Ontology and Hyper Graph Based Dashboards in Data Warehousing Systems

Gitanjali.J [#1], C Ranichandra [#2], Meera Kuriakose [#3], Revathi Kuruba[#4]
[#] School of Information Technology and Engineering, VIT University
Vellore - 632014, TamilNadu, India
[1] gitanjalij@vit.ac.in,[2] cranichandra@vit.ac.in, [3] meerakuriakose@ymail.com

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. And we 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 paper 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.

Keyword- QBSE, QBVE, Ontology, Dashboard, Hyper graph, DBE

I. INTRODUCTION

The variety, complexity and diversity of analytical requirements lead to increased interest of on-demand DWH systems(data warehousing systems)[1],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 [2].Performance and flexibility in the dash board 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 [3]. Here hyper graph clustering and hyper graph partitioning techniques are used [4] 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. We are describing our method using a real-world based scenario and explaining the benefits regarding the solution's design flexibility.

II. RELATED WORKS

QBE has been very famous since its arrival and is a well-known[5] 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 [6]-[8]. E.g. QBSE (Query-by-semantic-example), QBVE (Query–by-visual example). Query by-example interface can be used by the dynamic dash boards 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.

III.HYPER GRAPH-BASED QBE LANGUAGE

With the availability of mete 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 Fig1

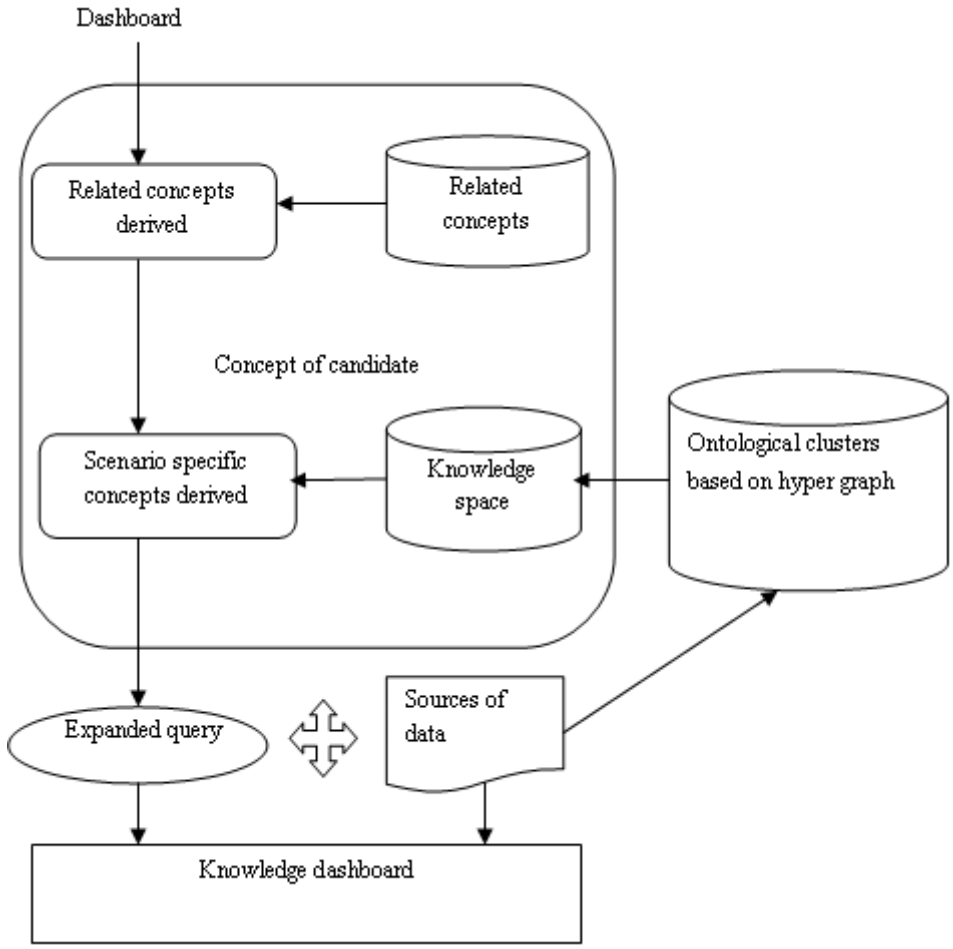

Fig1 Dashboard retrieval processes

A. Dashboard data model

The artifacts for the dashboard model [2] used in our approach are associated with data modeling, including the indicator models and the data. An abstract type presentation layer which includes the dashboard templates and navigation is defined by the dashboard model.

B. 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.

C. Logical Connectives

We 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.

D. 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.

IV. KNOWLEDGE BASED DASH BOARD FRAMEWORK

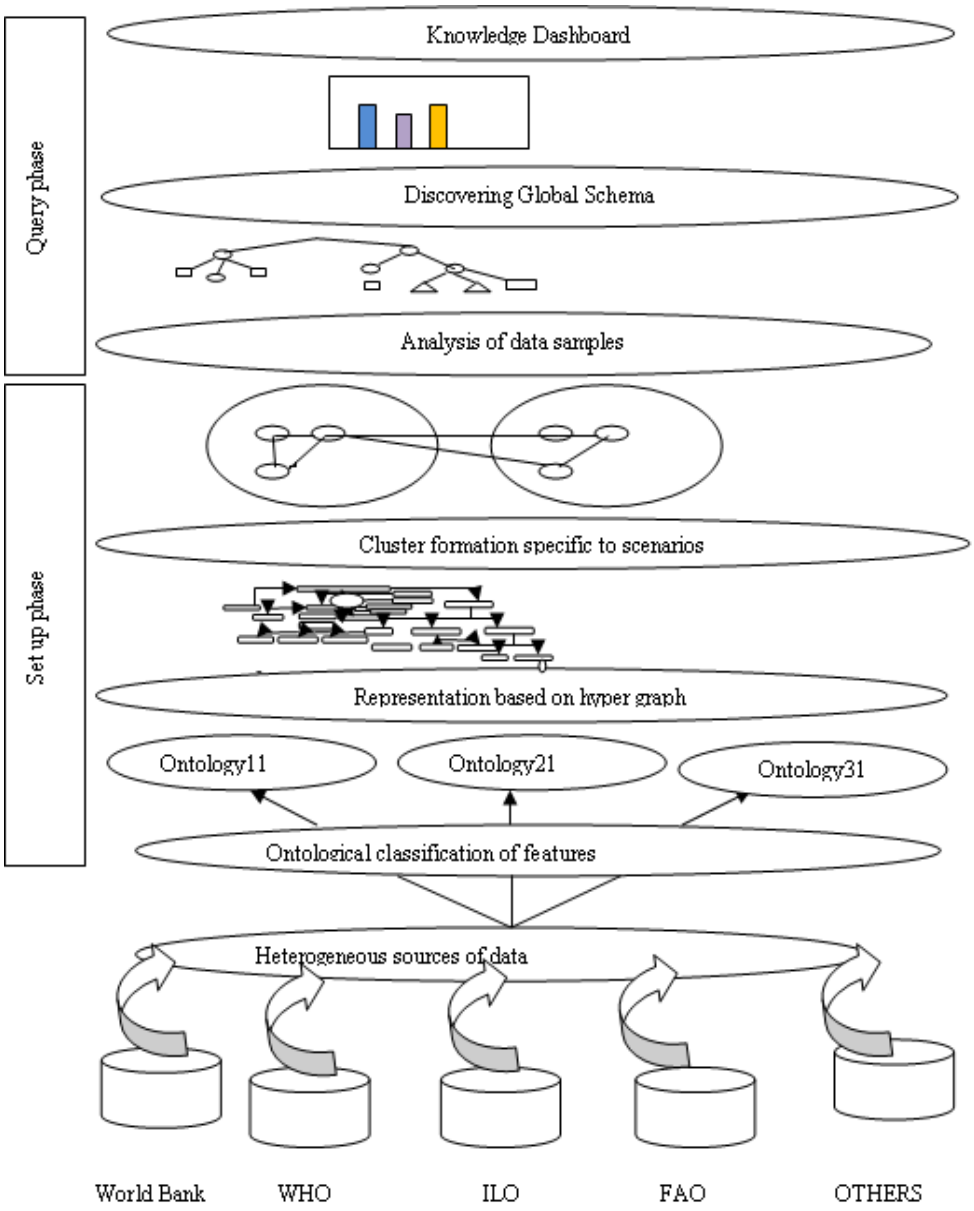

Fig 2. General Architecture of DBE Framework

DBE have an integrated view and use the dash board at different levels which helps user in obtaining different 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 is represented by means of hyper graph. Based on scenario different clusters are formed. Then data sampling is done and a global schema is discovered . And based on query suitable dash board design obtained.

V. 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 be satisfied based on the summary. Since the summery size is less than original graph, queries become 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 [9].

A. *Hyper Graph Builder*

It analyses the description based on ontology to build a labelled hyper graph which is directed and all synthesis the data dependencies.

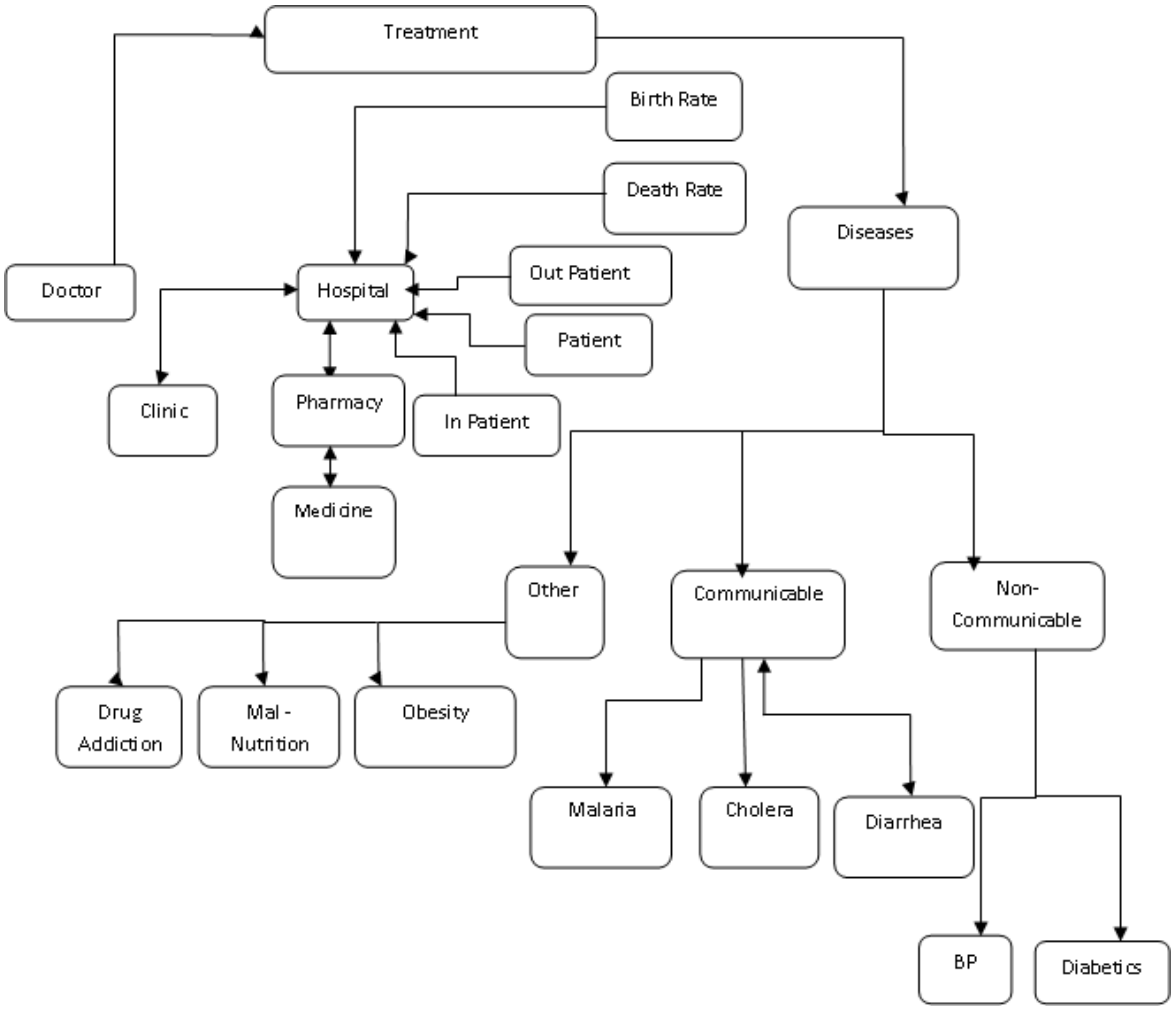

Fig 3.Hyper Graph for Health Care System

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 dash board 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, and diseases (etc).

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

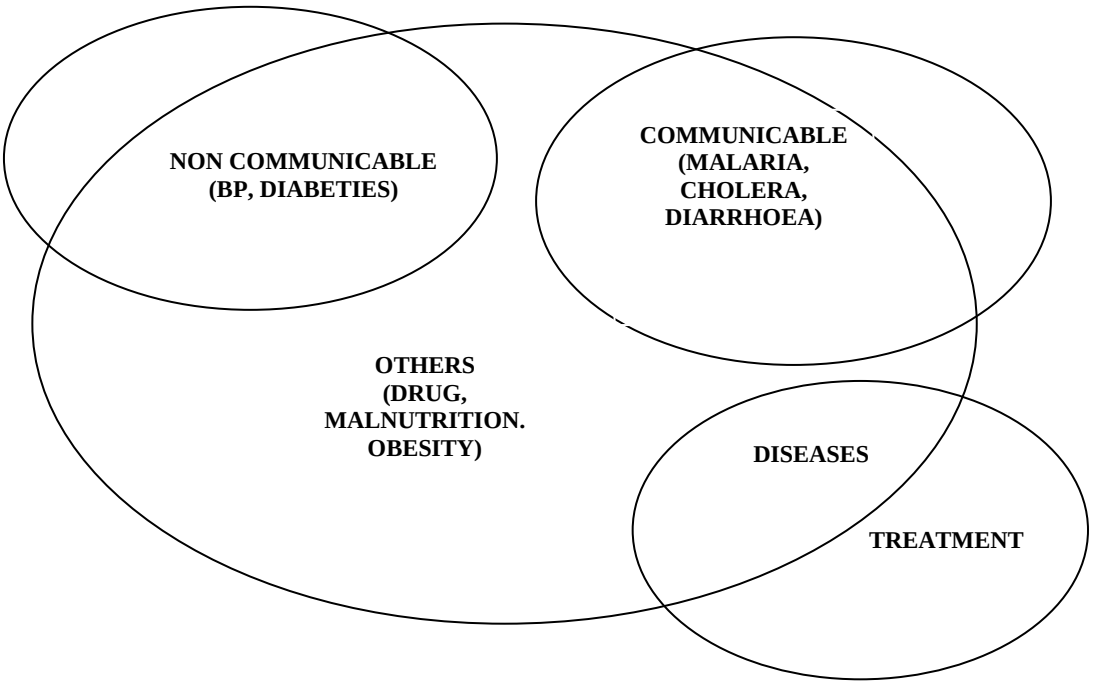

Fig 4.Cluster Border and Indicator Families

VI. DBE EXAMPLE

We are considering the healthcare systems here. 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. We are also dealing with general health care which includes various health issues across the world like obesity, communicable diseases, diabetes, UV rays, and 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 analysed. 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. And 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 we will get all necessary information in a user friendly manner as shown in Fig 5.

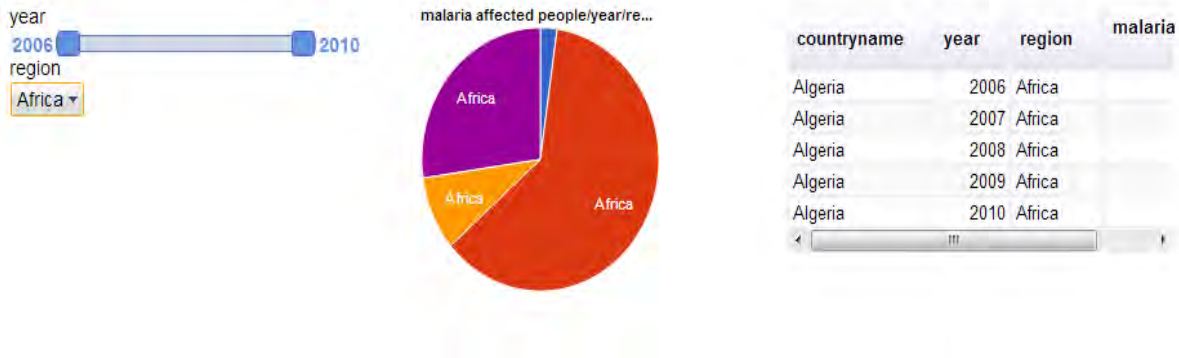

Fig 5: Sample Dashboard Showing Malaria Affected Region/Year

VII. EXPERIMENTAL RESULTS

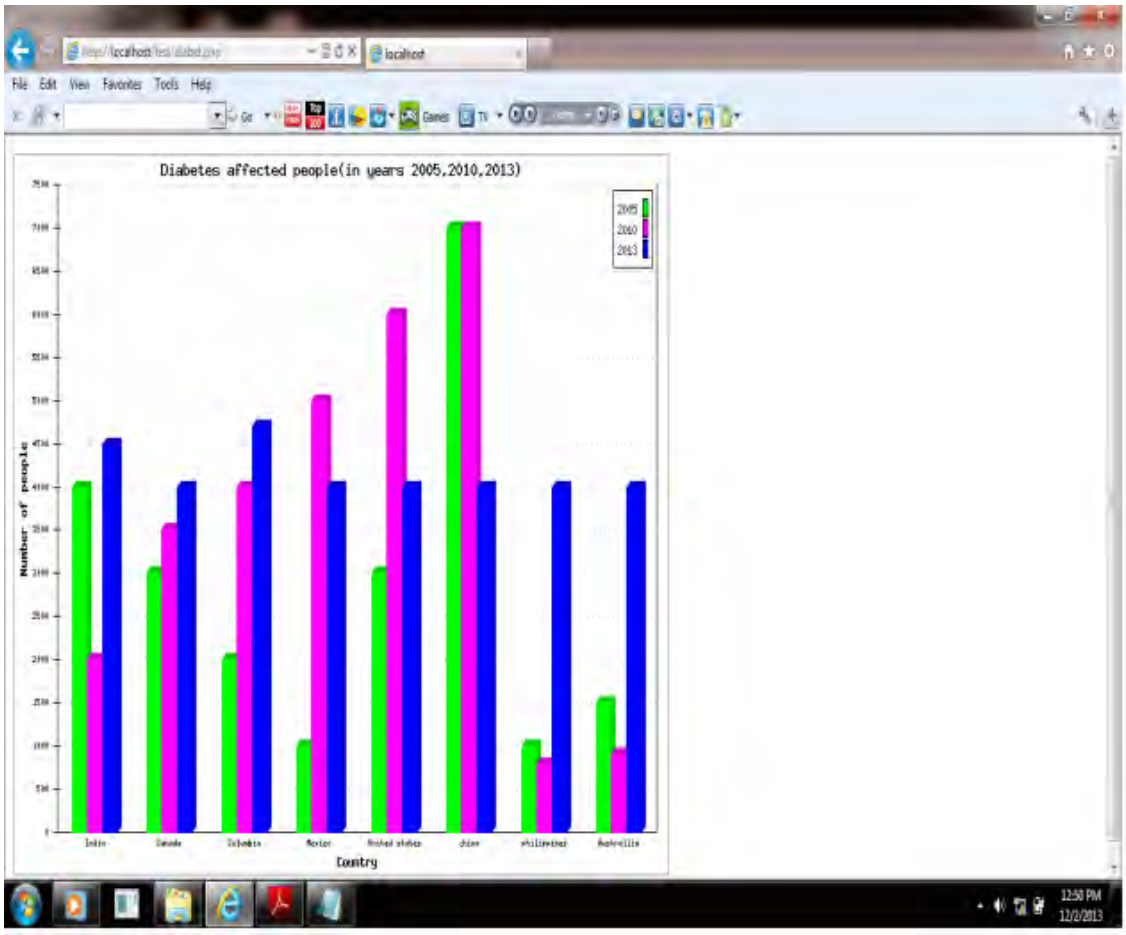

Fig 6.Output Dashboard Obtained

Finally we have generated dashboards for the various input data. Here input is taken from user and saving it in 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. And the dashboard denoting required information is obtained. (Fig.6).

VIII. 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. And 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 work more dynamic dashboards for frequently changing data can be developed.

REFERENCES

[1] T. B. Pedersen, D. Pedersen, and K. Riis, “On-demand multidimensional data integration: toward a semantic foundation for cloud intelligence, “The Journal of Supercomputing, Oct. 2011.
[2] T. Palpanas, P. Chowdhary, G. Mihaila, and F. Pinel, “Integrated model driven dashboard development,” Information Systems Frontiers, vol. 9, no. 2-3, pp. 195–208, May 2007.Available:http://linkinghub.elsevier.com/retrieve/pii/S1045926X85710087
[3] A. Brogi, S. Corfini, and J. Aldana, “Automated discovery of compositions of services described with separate ontologies,” Proceedings of the 4th international conference on Service-Oriented Computing ICSOC’06, pp. 509–514, 2006.
[4] D. Papa, “Hyper graph partitioning and clustering,” Approximation algorithms and metaheuristics, pp. 1–38, 2006.
[5] M. Zloof, “Query-by-example: the invocation and definition of tables and forms,” in Proceedings of the 1st International Conference on, 1975, pp. 1–24.
[6] I. Song, “Semantic Query-by-Example for RDF data,” in Proceedings of the First International Conference on Emerging Databases, 2009, p. 8.
[7] P. Moreno and N. Vasconcelos, “Bridging the Gap: Query by Semantic Example,” IEEE Transactions on Multimedia, vol. 9, no. 5, pp. 923–938, Aug. 2007.
[8] N. Rasiwasia and N. Vasconcelos, “A study of query by semantic example,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. IEEE, Jun. 2008, pp. 1–8.
[9] Duong Thi Anh Hoang, Thanh Binh Nguyen, A Min Tjoa”Dashboard by-Example: A Hypergraph based approach to On-demand Data warehousing systems” in 2012 IEEE International Conference on Systems, Man, and Cybernetics October 14-17, 2012.