International Journal of Scientific and Technical Advancements
ISSN: 2454-1532

Schema Free Data

Milina Salvi[1], Prajakta Satardekar[2], Sayali Mane[3], Chhaya Narvekar[4]
[1, 2, 3, 4]IT, Mumbai University, Mumbai, Maharashtra, India-400016
Email address: [1]milinasalvi@gmail.com, [2]prajakta.21081994@gmail.com, [3]sayalimanegb@gmail.com, [4]narverkarchhaya@gmail.com

Abstract— All real-world databases often have extremely complex schema s may be due to size or design of database. Database schema with lots of entities and their relationships, each with a multiple attributes. Schema free data can address this issue by allowing users with less or no knowledge of the schema to formulate database queries. The paper will show that most current Schema Free Query Interfaces provide a very limited degree of design independence the proposed method. In graphical database the data from the tables is store in the simple text file. So whenever we want to add any extra column will not required that much space which is utilize by the table form. As the time required to search the result is optimal in these method as compared with the tables. The only requirement to have this format is that to convert into graphical format we need to have the relational tables. The second feature is that keyword search on graph data. Keyword search provides a simple but user friendly interface for the information retrieval. Since many real life data sets to represent in graph format keyword search had become an attractive mechanism for data. The searching is done in bidirectional breadth first search.

Keywords— Data model; Schema free data model; DACR.

I. INTRODUCTION

A database is an organized collection of data. It is the collection of schemes, tables, queries, reports, views and other objects. The data is typically organized to model aspects of reality in a way that supports processes requiring information. It reduces the data redundancy, updating errors also it increase consistency and data security. But the databases may have very much complicated schemas may be due to its large size of database or due to complex design of database. Such database schemas with lots of entities and their relationships, whereas each entity contains multiple attributes. It is challenging for new database users to explore the data and formulate queries in order to extract information from real time databases. Schema free data (SFD) is introduced to address this issue by allowing users with less or no knowledge of the schema to formulate database queries.

II. SCOPE

An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it. The utility of this project is to all database users. The users who wants to search the data can make use of this project. This project will help in space utilization and will also help in the searching the result appropriately. Our project performs various operations in sequence to generate the answers to query in proper logical order.

1. Convert data to graph format which become easy for search using graph methods.
2. Find all possible CAs by LCA Method
3. Filter obtained results by Filtering Methods
4. Arrange the result in proper logical ranking using ranking methods.

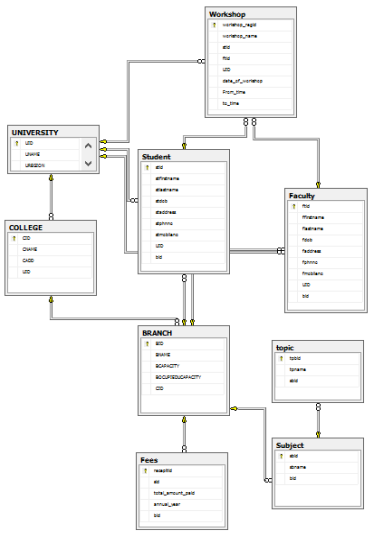
III. DISCUSSION AND EXPERIMENTAL RESULTS

We performed an extensive experimentation with the SFD system, which was implemented in C# technology with .NET framework. The experiments were carried out on various systems with following computer configurations,

1. *PC 1*: Pentium 4, CPU of 1.66 GHz and 1GB RAM
2. *PC 2*: Dual core, CPU of 2.4 GHz and 2GB RAM
3. *PC 3*: Intel i5 Processor, CPU of 3.66 GHz and 4GB RAM
4. *PC 4*: AMD Athlon, CPU of 2.00 GHz and 512MB RAM

We have implemented SFD using some research papers based on SFQI and evaluated the SFD system on two aspects:

1. *Search Quality*: Search Quality is evaluated using both a standard RDBMS benchmark and a heterogeneous data collection using XML data set.
2. *Search Performance*: We measure the overhead caused by evaluating schema-free query versus the schema-aware query. This is done by recording time and relativeness measures of query results.



Milina Salvi, Prajakta Satardekar, Sayali Mane, and Chhaya Narvekar, "Schema free data," *International Journal of Scientific and Technical Advancements*, Volume 2, Issue 2, pp. 65-66, 2016.

IV. Sample Database

The name of sample database is university which contains tables as follows: university, college, student, faculty, fees, workshop, branch, topic.

V. Data Conversion and Ranking

This section describes our method for converting a relational database into a graph database. Data needs to be converted from relational data into graph modeled data as, many organizations required to make their information available to all, usually stored in relational databases, on the Web using RDF or graph. For this reason, several solutions have been proposed to support the translation of relational data into RDF. Some of them focus on mapping the source schema into an ontology and rely on a naive transformation technique in which every relational attribute becomes an RDF predicate and every relational values becomes an RDF literal. In place of RDF we are presenting our information in for of graph which will be far better for searching process.

The formula used for ranking is as follows:

$$\text{Rank}(t, Q) = F_0 * K_0 + F_1 * K_1 + F_2 * K_2 + F_3 * K_3$$

Where,

F_0 =TDIF Distance based ranking

F_1 = Result base ranking

F_2 = Relation Importance

F_3 = Entity Based ranking

Distance based ranking

Distance based ranking consider distance between the two query searched.

Result Importance Weight

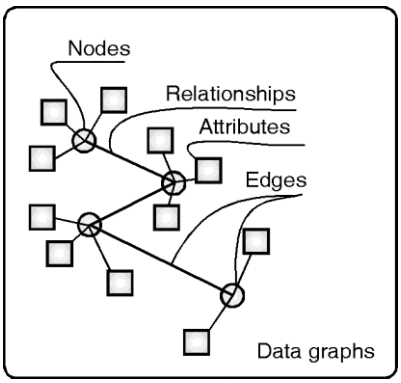
It makes use of traditional Distance Based Method to compute importance of search pattern in entire dataset.

Relation importance weight

To find out importance of how closely data is related in dataset.

Entity importance

The total Relation Importance Weight is given by average of importance weight for all relations in a given result pattern.



Representation of data graph

VI. Conclusion

In this project, we have studied the requirement and concept of Schema free data techniques. System of SFD should also include the efficient data search and ranking results in meaningful order in the presence of keyword search ambiguities. This is done even for those users with less knowledge of database schema. We can formalized this approach with help of many guidelines, which are analyzed and compared with many design independence concepts of current keyword search and schema free query interfaces.

We utilize various searching statistics to infer query answers and used DACR technique to rank the query results. This technique can enable end users to gain lot of advantage over query techniques like XQuery for querying data precisely and efficiently without requiring full knowledge of the database schema. We have shown that it is possible to express a wide variety of queries in a schema free manner and have them return correct results over a broad diversity of schema. In this paper we proposed the approach for obtaining high precision result by removing ambiguity present in query answers. At last, we have proposed the new approach for ranking results using DAR-CR method is better than the technique DA-CR which is used by some SFQI as it also considers the importance of relationships.

Acknowledgment

We would like to thank our Project Guide Prof. Chhaya Narvekar, who has been very supportive throughout the entire project tenure. We thank her for providing her valuable feedback at all times and always encouraging us to do the best. She has definitely extracted the best performance out of us. A special thanks to our college Principal Dr. Y. D. Venkatesh and college Director Fr. Francis De'Melo for providing us the best possible environment to work and complete our project. We would also like to thank the entire faculty of the IT Department, who have been very supportive towards us and helped us in every possible way. We would also like to thank our friends for helping us in the project documentation. Lastly we would like to thank those who have helped either directly or indirectly in any manner and those whose name we forgot to mention.

References

[1] M. Rammohanrao, M. Brahmaiah, E. Ramesh, and V. Rajesh, "Schema and design free keyword search interfaces for XML databases," *International Journal of Engineering Research and Development*, vol. 5, issue 9, 2013.

[2] A. Termehchy, M. Winslett, Y. Chopathumwan, and A. Gibbons, "Design independent query interfaces," *IEEE transaction on knowledge and data engineering*, vol. 24, no. 10, 2012.

[3] L. Han, T. Finin, and A. Joshi, "Schema-free structured querying of DBpedia data," in *CIKM*, 2012.

[4] A. Termehchy, M. Winslett, and Y. Chopathumwan, "How schema independent are schema free query interfaces?," *ICDE '11 Proceedings of IEEE 27th International Conference on Data Engineering*, 2011.

Milina Salvi, Prajakta Satardekar, Sayali Mane, and Chhaya Narvekar, "Schema free data," *International Journal of Scientific and Technical Advancements*, Volume 2, Issue 2, pp. 65-66, 2016.