International Journal of Research in Advent Technology, Vol.2, No.2, February 2014

E-ISSN: 2321-9637

A Review: Probabilistic Scheme for IQP and XML Query Construction by Keyword Search

Deepika Joge[1], Minal Kamble[2], Prof. P. S. Chhaware[3]

Department of Computer Science & Engineering [1, 2, 3], G.H.Raisoni Institute of Engineering & Technology for Women, Nagpur[1, 2,] Priyadarshani College of Engineering & Technology, Nagpur[3]

Email: deepikajoge@gmail.com[1] , minalkamble327@gmail.com

ABSTRACT - A query can be constructed in various forms and using various algorithms. Incremental query construction is one approach, to construct a query. A probabilistic method is given to find out the query interpretation by incremental query construction algorithm to construct a query in SQL when data is searched using any keyword. In this paper we are presenting the review for two different technologies. An Incremental Query Construction Probabilistic Scheme (IQP) for constructing a query having an incremental probability of occurrence of the data for input keyword. The other one is XSeek, a semantic search engine using keyword. The IQP method uses a probabilistic approach for formulation of structured queries. Here SQL is used for the construction of structured queries. Formation of the structured queries is the error-prone task; IQP minimizes this. XSeek uses the VLCA node technology, which represents the data in XML structured query format.

Keywords – probabilistic, keyword, XML.

1. INTRODUCTION

The current generation of the search engines ranks the keyword search. An informational retrieval (IR) system returns the sets of data which satisfy the information need expressed by a user's question. The purpose of informational retrieval is to retrieve all the relevant documents, while filtering out non-relevant document. In search engines user may face the difficulty to precisely formulate their queries. The problem is to illustrate about the schema complexity. Second, the partial or misspelled attribute values. Third, we would like the user to issue queries that are meaningful in terms of result size a query listing all options related to keyword. Ranked keyword search has been quite successful in the past; we would like to minimize the user's efforts providing XML query formats. However, keyword search has its common limits and the search engine XSeek; the next generation search engine will be more semantic in one way.

In this paper, we present the novel scheme IQ^P, which bridges the gap between usability of keyword search and expressiveness of the database queries. IQ^P allows a user to start with a random keyword query, and then incrementally refines the keywords into the desired structured query through an interactive interface. IQ^P enables users to construct structured queries efficiently; without necessarily knowing the database schema or mastering a query language. IQ^P translates the searched query into ranked list of structured queries, which give different interpretations keywords. Simultaneously, IQ^P generates a set of queries construction options.

2. RELATED WORK

The [1] has suggested a system for keyword-based search over relational database. They named it as "DBXplorer" [1]. The proposed methodology states the concept of relational database, where to search some data you have to enter a keyword. From that keyword the relational queries will get fired and the informational needs will be satisfied by the algorithm.

In [2] the methodology for Identifying Meaningful return Information for XML Keyword Search, in which XSeek, is introduced that addresses an open problem of inferring desirable return nodes without elicitation of user preferences and has achieved problem results.

[3] Has presented IQ^P: Incremental Query Construction, a Probabilistic Approach. In this paper, authors state the probabilistic approach for construction of queries in incremental order, where IQP enables users to construct structured queries efficiently and there is no necessity of knowing database schema.

In 2012, a probabilistic scheme for keyword-based Incremental Query Construction has been implemented in which the same probabilistic approach for incremental query construction of structured queries from keywords is proposed. In this methodology the keyword search technology is used

for satisfying the informational needs which will be represented by structured queries.

3. PROBLEM DEFINITION

***3.1 Schema of XML data*:** An XML document can optimally have a schema and a DTD i.e. Documented Type Definition. A DTD is a commonly used method to describe the structure of an XML document.

3.2 Keyword Search: The user input is a set of keywords, each of which may match value node in the XML tree.

Example: If we search a keyword like 'Taj', then the resulted queries will be shown as 'TajMahal (Movie)', 'Hotel Taj (Hotel Name)', 'MoomTaj (name of person)' and so on. A tree for keyword Taj will be structured. The probable incremental query will be constructed in XML data format using XML schema. The schema may be in DTD. The constructed tree may have some nodes, which includes some probable options above.

The above example contains another possible search as mentioned above. The result will show the most accessed or most important search at most priority. Hence, the user can get the result in structured form and the most possible results for a same keyword. Most importantly users don't need to know what XML is.

This implementation is most significant as the numbers of optional results are shown by our system and user get to know what are other relevant information is present over there.

4. FRAMEWORK

The existing methodologies consist of the following framework which helps in query construction and query processing.

I. Analyzing XML Data Structure – XML data structure can be analyzed by using the entity-relationship model.

II. Analyzing Keyword Match patterns – keyword which are entered by user are to be analyzed and the keyword patterns are to be matched in database.

III. Processing query structures – Query processing structures processes the query for incremental query construction and to provide a priority to search results.

IV. Generating search results – The result panel will display the result in module with probable answers.

Here are using some greedy algorithms,

a. Query Construction Algorithm and
b. Query Processing Algorithm.

5. CONCLUSION

In this paper, we present IQP – a system to search a keyword for information retrieval, and aXSeek, a system which executes a result in XML query representation form. These both systems are executed separately. We are aiming to execute a new system which will be a combination of IQP and XSeek.

6. REFERENCES

[1] S. Agrawal, S. Chaudhuri, and G. Das, (2002): DBXplorer: A system for Keyword-Based Search over Relational Database, Proc. Int'l Conf. Data Eng. (IDCE).

[2] Ziyang Liu, Yi Chen, (2007): Identifying Meaningful Return Information for XML Keyword Search, ACM, *SIGMOD*..

[3] Elena Demidova, Xuan Zhou*, Wolfgang Nejdl, (2010): IQ^P: Incremental Query Construction, a Probabilistic Approach, IEEE, ICDE Conference.

[4] Elena Demidova, Xuan Zhou, and Wolfgang Nejdl, (2012): A Probabilistic Scheme for Keyword-Based Incremental Query Construction, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 24, NO. 3.

[5] Youjin Chang, Minkoo Kim, Vijay V. Raghavan, (2006): Construction of query concepts based on feature clustering of documents, Springer Science+Business Media.

[6] HolgerBast, AlexandruChitea, Fabian Suchanek, Ingmar Weber, (2007): ESTER: Efficient Search on Text, Entities, and Relations, ACM, SIGIR'07, Amsterdam, The Netherlands.

[7] Yi LuoXuemin Lin Wei Wang, Xiaofang Zhou, (2007): SPARK: Top-k Keyword Query in Relational Databases, ACM, SIGMOD.

[8] Arnab Nandi, H. V. Jagadish, (2007): Assisted Querying using Instant-response Interface, ACM, SIGMOD.

[9] Ziyang Liu, Jeffrey Walker, Ya Chen, (2007): XSeek: A Semantic Search Engine Using Keyword, ACM, VLDB.