ABSTRACT

Databases hold data that represent properties of real-world objects. In most new commercial business software applications, the data is systematically stored in the database layers and precisely retrieved using Relational Database Management System (RDBMS). Structured Query Language (SQL) is the standard language used by RDBMS to communicate with relational databases. Though, SQL is a very successful and powerful entity-oriented query language for data manipulation using relational algebra, often it fall short of addressing some of the critical needs of searching 'near to' or 'close to' or 'just about' or 'approximate' values during decision making processes. To address these needs, this research work proposes certain improvement on standard SQL using fuzzy concepts called 'Enhanced SQL' (ESQL). It is also proposed to develop an automated query generator tool for the above ESQL to utilize it to the full capacity.

Human have the ability to arrive quick estimation based on "approximations". Whenever interacting with the database, human try to create nonstandard complex queries to arrive fast conclusions. The complexity in the query normally arises in the form of ambiguity from uncertainty. The conventional SQL is crisp, deterministic and firm in nature. Here firmness indicates that the structures and parameters of the model is to be undeniably known. But in real situations, human queries are not always crisp and deterministic and therefore, cannot be described precisely. Standard SQL is not capable of addressing such complex and vague queries. Standard SQL has the following limitations:

- SQL outputs crisp selection as it uses crisp logic in querying process. The issue in this is that the record would not be selected even if it is extremely close to the intent of the query criterion.

- SQL is not capable of satisfying the needs for data selection based on the degrees of truth and linguistic expressions.

The techniques based on the fuzzy set theory are very much useful while modeling the uncertainties especially, when the uncertainties are non-random in nature. In some earlier research works, exclusive Fuzzy Relational Databases were developed to address this problem. It is an extension of relational database which comprise fuzzy predicates or fuzzy conditions under shapes of linguistic expressions. But such Fuzzy data needs more storage space than crisp data does; it takes more I/O time to transfer data between main memory and secondary memory than crisp data. It takes more CPU time to evaluate a fuzzy query condition than crisp query condition. Hence, the present work considers improvement only on the classical query language viz. SQL using Fuzzy sets without affecting or changing the traditional Relational Database.

There is a problem in defining the data requirement by simply applying the literal query condition in the classical SQL. The traditional querying techniques in SQL have been tremendously improved by adding Fuzzy features in the proposed work to deal with the uncertain or imprecise information. This research proposes a new query system 'ESQL' to help the user, for defining their requirement for data retrieval through ESQL query with the help of fuzzy concept.

The SQL still face another issue Viz. 'nested query' because of expressing complex queries. The nested query is an important mechanism to ease the pain of expressing complex queries. However, execution of a nested query may incur heavy performance penalty due to compound query structure. Thus, the study of unnesting techniques is taken in this work. A technique is proposed to transform the nested query into an equivalent flat

query and then to execute the flat query. This unnesting technique has been studied extensively in the context of conventional relational database systems. The key to the accomplishment of the proposed work is that the improvised merge join algorithms used to execute the unnested queries are much more competent than the nested-loop algorithm used to execute the nested queries.

In general, SQL is a powerful mechanism for advanced searchers and experienced developers; however, majority of the users are not familiar with them and are unable to fully utilize the SQL capability. Hence an automatic query generation tool will be helpful for everybody to fully utilize the SQL capability. Under that consideration an automatic query generation tool have been developed. The proposed tool apart from improving traditional querying capability of standard SQL augments the generation of ESQL query as well.

This research work focuses on developing ESQL query that helps to extract the data from the database efficiently. The evaluation of the proposed work has been carried out using Java as programming language and SQL server as database management system with help of fuzzy database library that permits to extract the information from traditional relational database in an accurate and time efficient manner. The experimental evaluation has been carried out through the dataset which is collected from the site http://archive.ics.uci.edu/ml/. The experimental results of proposed approach have been evaluated using performance metrics like execution time, accuracy, cost and Memory. The proposed system will reduce the heavy performance penalty of nested queries by using the unnested queries and obtained less execution time and less cost than existing one.