FDSL Tool: An approach of Fuzzy Relational Database Management System

[1]Idris Mala, [2]Pervez Akhtar, [3]Abdul Rehman Memon, [4]Tariq Javid Ali

[1]PhD Research Scholar, FEST, Hamdard University, Karachi, Pakistan
[2]Professor, PNEC, National University of Science and Technology, Karachi, Pakistan
[3]Professor, Dean, FEST, Hamdard University, Karachi, Pakistan
[4]Assistant Professor, HITEC University, Taxila Cantt., Pakistan
E-mail: idrismala@yahoo.com

Abstract: Relational Database Management Systems (RDBMS) is commonly used in business, medical, environmental and industrial applications. The problem that is associated with RDBMS is that it gives precise values but since human understanding is based on vague set of values, we have selected Fuzzy RDBMS for our tool which is much easier for humans to interact. In this paper a Fuzzy Database Software Layer (FDSL) tool has been developed on top of a standard RDBMS to handle fuzzy SQL query giving results in fuzzy linguistic term. The implementation of FDSL tool is shown in this paper using heart database in the area of medical informatics.
[Idris Mala, Pervez Akhtar, Abdul Rehman Memon, Tariq Javid Ali. **FDSL Tool: An approach of Fuzzy Relational Database Management System.** *Life Sci J* 2013;10(2):1638-1643].(ISSN:1097-8135). http://www.lifescien cesite.com. 227

Keywords: Fuzzy RDBMS, Fuzzy SQL, FDSL Server, FMB, FRDB

1. Introduction

Structured query language (SQL) is used to communicate with relational database management system (RDBMS). SQL is a specialized language for updating, deleting and requesting information from databases. The tool which we have introduced is an interface that takes fuzzy SQL queries as inputs and converts in SQL queries. Further on these queries are used to calculate responses from the database and to generate results in human understandable format.

The fuzzy logic introduced by Zadeh [1], was further enhanced when Buckles and Petry combined the concept of fuzzy with SQL [2]. Then ER model was used with fuzzy logic by Zvieli and Chen [3]. The work on FRDBMS didn't end here. Medina et al. introduced Generalized Model of Fuzzy Relational Database Model (GEFRED) [4], which was implemented as Fuzzy Structured Query Language (FSQL) [5].

Fuzzy logic was applied to basic EER (Extended Entity-Relationship) concepts by Chen and Kerre [6]. A FRDBs (Fuzzy Relational Databases) design methodology has been proposed by Chaudhry et al. [7]. Fuzzy knowledge using relational database has been modelled called Fuzzy EER model [8]. An implementation of the model was done on the basis of FSQL – an SQL language with fuzzy capabilities.

To keep relational database separate and at the same time able to query in fuzzy terms, a tool kit was introduced named "FCQL" which is just an additional layer above RDBMS [9, 10, 11]. Another tool was introduced that automatically transforms fuzzy SQL query to SQL query based on fuzzy meta base (FMB) [12] in 2009.

In this paper, we propose a new FDSL Tool, a software layer on top of commercial database for implementation of FRDBMS. The tool takes its input in fuzzy query and the results generated are fuzzy results. The FDSL Tool comprises of two components, the FDSL Server and FP Editor / Algorithm. The FDSL Server works on top of a standard RDBMS, which is used as conversion layer to convert Fuzzy SQL to SQL and from SQL to fuzzy results (linguistic values). The FP Editor / Algorithm is used to define the Fuzzy Meta Knowledge Base (FMB) based on expert analysis or algorithm based on data. The FRDBMS which we have implemented in our tool is using heart ailment database for medical informatics.

The paper is organized as follows. A review of existing fuzzy tools for RDBMS is presented in Section 2. Section 3 discusses the proposed FDSL Tool. Section 4 presents implementation of FDSL Tool. The last section gives conclusion and suggests future research direction.

2. Existing Fuzzy Tools for RDBMS

In this section we will explain some exiting fuzzy tools that were used for RDBMS. It was Galindo et al. who proposed the Fuzzy SQL (FSQL) language which extended the SQL language to allow fuzzy queries [8, 13]. For example, SELECT command is extended to add flexible queries in FSQL and therefore a query which is valid in SQL is also valid in FSQL. Also each condition statement in the query is controlled by THOLD command for fulfillment of threshold. It also comprises of DML statements (INSERT, DELETE and UPDATE) and DDL statements (CREATE, DROP and ALTER). The

limitation of FSQL is that it only allows the flexible querying of FRDB and has been implemented for Oracle database only [12]. Another extension of SQL language by adding a layer to RDBMS was defined by Bosc at el. who contributed SQLf language, which is a method to acquire imperfect information in relational databases on flexible queries [15, 16, 17].

A fuzzy Classification Query Language (fCQL) for customer relationship management was proposed by Meier et al. [9]. The fCQL toolkit is an additional layer above the relational database system, making fCQL independent of underlying database systems. This approach enables fCQL to operate with every commercial product [11], which gives great flexibility in the design of FRDBMS. Another advantage of this approach is that the user can always query the database with standard SQL commands. Information is retrieved on the basis of fuzzy classification where objects are handled on the basis of classes, and not as attribute like in FSQL and SQLf. A drawback associated with fCQL is that it does not work on individual tuples but is based on classification of relations within the table.

Later on Touzi et al. presented a tool called FSQL_TO_SQL [12], and shown its implementation of fuzzy relational databases in FSQL proposed by Galindo [8, 13]. The tool automatically transforms FSQL script to equivalent SQL based on defined FMB. It handles both DDL and DML and has been implemented in C++. This tool does not perform automatic mapping of existing relational database to FRDB. A solution to fuzzy RDBMS was shown through a priority fuzzy logic enriched SQL language - PFSQL in [24, 25, 26].

Now in next section we will be discussing our proposed FDSL tool which is developed in open source, java programming language which can be implemented with any commercial RDBMS by connecting through ODBC (Open Database Connectivity).

3. Proposed FDSL Tool

Our proposed FDSL Tool is a software layer, which works as middle layer on top of commercial database for implementation of FRDBMS. On top of this FDSL tool is user/application interface and expert interface. This tool handles fuzzy query, which is input from user interface. The fuzzy query is converted to SQL query and is passed on to RDBMS. The RDBMS returns SQL query results to FDSL tool which are crisp value results. The FDSL tool then generates results into linguistic value from crisp value, which are interpretable.

3.1 FDSL Tool Architecture

The architecture of our proposed FDSL Tool is shown in figure 1. The tool works above the RDBMS and comprises of FDSL Sever and FP Editor/ Algorithm. The FDSL Tool works as middle layer, on top of which we have User / Application Interface and Expert Interface. The FDSL Tool works with the standard commercial RDBMS which stores database and fuzzy meta knowledge base (FMB). The FDSL Server is connected to User or Application layer interface to communicate with user or application. Expert interface is used to update the FMB based on expert experience or some FP algorithm.

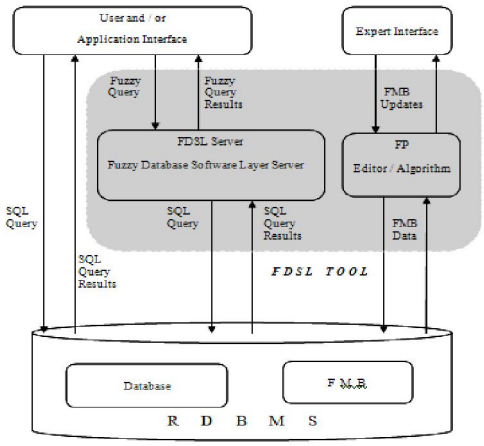

Figure 1. The architecture of developed FDSL tool

The components used in the above architecture are further described as follows:

1. *User and / or Application Interface:* Fuzzy query is automatically generated by this graphical user interface when the user selects options for the query. This query is passed to FDSL Server which converts fuzzy query to standard SQL and passes the query to standard RDBMS for query results. Instead of graphical user interface, the system can be connected to an application for answering fuzzy queries.
2. *Expert Interface:* This is an interface for experts to enter the FMB in the standard RDBMS. Fuzzy partitions for different attributes are defined by the expert. This data is entered through this interface in FMB. If expert is not available, then an algorithm designed to make fuzzy partitions for different attributes is run and results generated by this algorithm is stored in fuzzy meta knowledge base (FMB).
3. *Fuzzy Database Software Layer (FDSL) Server:* It uses FSQL_to_SQL algorithm to convert the fuzzy query to standard RDBMS query. It uses lexical, syntactic and semantic analysis converting fuzzy query into standard SQL. Another algorithm called Crisp_to_Linguistic is used to convert selected attributes to linguistic term for more understandability.

4. *Fuzzy Partitioning (FP) Editor / Algorithm:* It consists of fuzzy partitioning editor so that expert enters the fuzzy meta knowledge base (FMB). In case if expert is not available then an algorithm designed is used to generate fuzzy portioning to develop appropriate partition of data. For example in case of blood pressure we may have Low, Medium and High, as three linguistic values and its ranges are defined by this algorithm.
5. *Database:* It is the standard data, stored in the relational database format by the standard commercial RDBMS. The application data is stored in crisp values, which are converted by FDSL Server to fuzzy linguistic values based on stored fuzzy partitioning in FMB.
6. *Fuzzy Meta Knowledge Base (FMB):* FMB stores information about fuzzy relational database in the relational structure. It stores different linguistic terms for the attributes, such as name and range values.

3.2 Database used for FDSL Tool Implementation

The implementation of FDSL Tool is demonstrated using heart database of UCI Repository [18]. Heart ailment database comprises of 270 records, and is stored in standard RDBMS. The main attributes of this database are blood pressure, cholesterol, age, and heart rate. All four attributes membership function is defined based on the work carried out by Ali at el. [19] for heart ailment database.

3.3 Fuzzy Meta Knowledge Base (FMB)

It is the necessary knowledge required for conversion of crisp data to fuzzy and fuzzy data to crisp values. This information is stored in the relational format, showing fuzzy linguistic values with ranges from LOW to HIGH.

The FMB relation is based on the following:

OBJECT : defines the object identifier or the relational table referring to Field : attribute of the relation
FLV : fuzzy linguistic Value
LOW : Value range lower side
HIGH : Value range higher side

Figure 2 shows the example of heart data object with four attributes which are bp (blood pressure), having fuzzy linguistic values of Low, Medium, High and VeryHigh. Similarly, attributes are hr (heart rate), ch (cholesterol), and age having their own fuzzy linguistic values. Fuzzy Meta Knowledge Base (FMB) is necessary relation in RDBMS for conversion of crisp values to fuzzy linguistic values and is referenced by FDSL Server for mapping. This FMB data is either given by an expert of the domain or is generated by FP Algorithm.

Figure 2. Example Fuzzy Meta Knowledge Base.

OBJECT	Field	FLV	LOW	HIGH
heart	bp	Low	0	134
heart	bp	Medium	127	153
heart	bp	High	142	172
heart	bp	VeryHigh	154	999
heart	hr	Low	0	141
heart	hr	Medium	111	194
heart	hr	High	152	999
heart	ch	Low	0	197
heart	ch	Medium	188	250
heart	ch	High	217	307
heart	ch	VeryHigh	281	999
heart	age	Young	0	38
heart	age	Mild	33	45
heart	age	Old	40	58
heart	age	VeryOld	52	999

3.4 User / Application Interface

The user interface is a complete graphical interface for user to select the criteria of fuzzy query, so that the system itself generates the query from the selected criteria. The interface is shown in figure 3. The user enters the appropriate field required like medium blood pressure or high heart rate, etc. An example is shown in figure 3 for selecting medium blood pressure. A fuzzy query is generated by this user interface which is as follows:

Select * from heart where bp is medium

The above query generated is passed on to the FDSL Tool by the user interface. The user interface can generate a complex query comprising of four conditions as there are four attributes in the user interface.

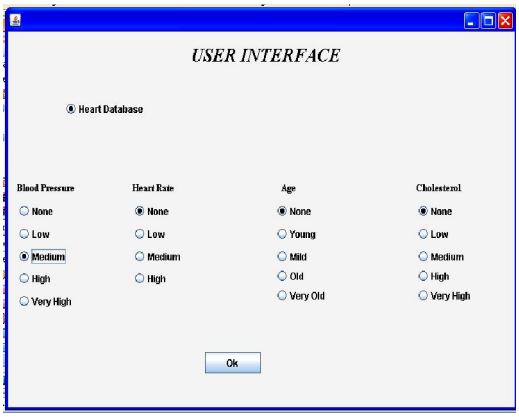

Figure 3. User Interface for fuzzy query generation

A complex query is generated, if all attributes are selected, as shown below:

Select * from heart where bp is medium and hr is medium and age is young and ch is low

3.5 Fuzzy Database Software Layer (FDSL) Server

The FDSL Server is middle ware layer (software layer) handles fuzzy query input and generates linguistic value results as output. FDSL tool is a standalone application implemented in Java programming language. FDSL Tool comprises of FDSL Server and FP Editor / Algorithm, where FDSL

Server uses lexical, syntactic and semantic analysis converting fuzzy query into standard SQL. It then converts SQL query results to fuzzy linguistic term which is the output of FDSL Server.

Any fuzzy SQL query once generated by input interface is passed to FSQL Server of FSQL tool, which performs as follows:

1. Syntax parsing is done on the fuzzy SQL query to verify it is correct.
2. Checks number of query conditions and generate multiple queries to find LOW and HIGH of each linguistic value.
3. Each condition will be converted into two conditions, one greater than equal to and second less than equal to.
4. Write the standard SQL query by now creating appropriate conditions.
5. Send standard SQL query to the (commercial) RDBMS for query generation.
6. Get results from RDBMS and return to FDSL Server for proper result generation.
7. Convert attributes, crisp values to linguistic values by referencing FMB.
8. Show results in linguistic value for interpretability.

In the first step, the FDSL Server converts any fuzzy query into the standard RDBMS query, by performing syntax analysis. The query first generates token for parsing, then it analyse each token to convert correct form of standard SQL query. Consider for example a fuzzy generated by user interface is

Select * from heart where bp is low and ch is high

The FDSL Server first creates tokens, from where the name of the table to be referred is checked, then it checks for the number of where clause attributes. The FSQL_to_SQL Algorithm runs query for each where condition to generate its appropriate range of result to be placed in the standard query like in case of above query, the first query will be generated for the blood pressure for Low and High bp range and second query is executed for cholesterol for Low and High ch range. The Low and High values of attributes are referenced from FMB. Once all the conditional values are generated, the algorithm generates standard query as shown in figure 4.

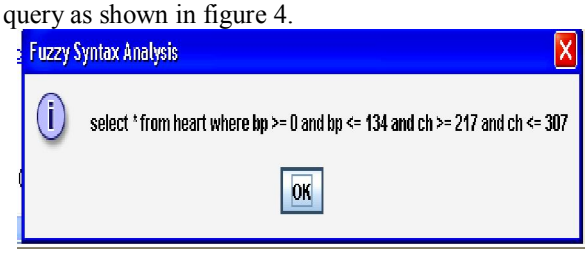

Figure 4. Standard SQL query converted by FDSL Server

The first algorithm that converts Fuzzy SQL Query to SQL Query called FSQL_to_SQL Algorithm is as follows:

FSQL_to_SQL Algorithm

INPUT : Fuzzy Query from User Interface
OUTPUT : Resultant SQL Query to be transferred to RDBMS
Begin
Load driver of jdbc : odbc
Make multiple connections
Get fuzzy query
Count number of tokens in the fuzzy query
Get the name of the table
Find all where clause attributes
Loop for all attributes
Generate query for each attribute
Execute each query to get range of value for each linguistic term
End loop
Write standard query
Convert each where clause condition with appropriate
Values of low and high range value
Place appropriate position with proper syntax
Generate query and display
Store query in public variable to transfer to FRDMS

Close connections
Close database
End

AGE	Blood Pressure	Cholestrol	Heart Rate
Old VeryOld	Low	High	Low Medium
VeryOld	Low Medium	High	Low
VeryOld	Low	High	Low Medium
Old VeryOld	Low Medium	High	Medium
VeryOld	Low	Medium High	Medium
VeryOld	Low Medium	Medium High	Medium
Old VeryOld	Low Medium	High VeryHigh	Medium High
VeryOld	Low	Medium High	Low Medium
Old	Low Medium	Medium High	Medium High
Mild Old	Low	High VeryHigh	Medium High

Figure 5. Results generated by Crisp_to_Linguistic algorithm

The FSQL_to_SQL algorithm generates standard SQL query from fuzzy SQL query. The Crisp_to_Linguistic Algorithm then gets results from RDBMS based on standard SQL query and converts it to linguistic values. All the four fields of the heart database are converted back to its linguistic term.

The query shown in figure 4 was passed on to standard RDBMS for execution of query, and results

were again retransferred to its linguistic values. Figure 5 shows the results generated by FDSL Server based on the Crisp_to_Linguistic Algorithm.

The Crisp_to_Linguistic Algorithm; used by FDSL Server to generate fuzzy linguistic terms from the standard SQL query results is as follows:

Crisp_to_Linguistic Algorithm

INPUT : SQL Query from RDBMS
OUTPUT : Results in Fuzzy Linguistic Variable term
Begin
Load driver of jdbc : odbc
Make multiple connections
Display headings of selected attributes
Get public variable having sql query converted by FSA
Execute query
Get result in a result set
Loop until end of the result set
Convert the result set variables
Get each variable
Run query of each variable to get its fuzzy linguistic variable
Store in a variable (range of values)
Display linguistic variable
Next line
End loop
Close all connections
Close database

3.6 Fuzzy Partitioning (FP) Editor / Algorithm

This fuzzy partitioning editor is used by expert to enter the fuzzy meta knowledge base in FMB of the RDBMS. In case if expert is not available then an algorithm is used to generate fuzzy partitioning to develop appropriate partition of data like for blood pressure we may have Low, Medium and High, as three linguistic values and its ranges are defined by this algorithm.

4. Implementation of the FDSL Tool

For testing the complete FDSL tool, we have used heart ailment database of UCI repository [18] and have taken four attributes out of thirteen attributes, as suggested by medical practitioner of heart disease namely blood pressure, age, cholesterol and heart rate.

The structure of fuzzy query is based on

SELECT attribute1, attribute 2, ...
FROM relation
WHERE condition 1, condition 2, ...

Where the CONDITION is defined as attribute in terms of some linguistic value like

bp is low and ch is high OR
hr is medium or age is old

where low, medium, high are linguistic values of the attributes

The results are based on fuzzy SQL query with

- no conditions
- one fuzzy condition
- two fuzzy conditions
- three fuzzy conditions
- four fuzzy conditions

The FDSL Tool is developed using java programming language with standard RDBMS engine. FDSL Server comprises of FSQL_to_SQL algorithm and Crisp_to_Linguistic algorithm to generate required fuzzy linguistic terms. Figure 6 shows fuzzy query and its results with four fuzzy conditions.

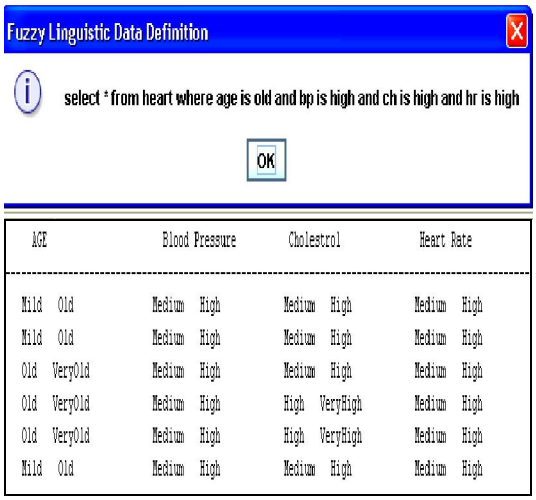

Figure 6. Four conditions fuzzy query and its linguistic value results

The tool worked excellent and produced results of fuzzy query into fuzzy results. Results shown in figure 5 and 6 of the heart database comprises of attributes with fuzzy linguistic values instead of crisp values which are comparable with results shown by Bedi et al. in student database [20], Gupta et al. in employee database [21], Galindo et al. in team database [22] and Mala et al in similar heart database [23].

5. Conclusion

We have shown fuzzy RDBMS implementation using our developed FDSL Tool, which is a software layer, using heart database in Java Programming Language. Input to this tool is fuzzy SQL query by the user interface. The fuzzy query is converted to SQL query by the FSQL_to_SQL algorithm and the query is passed to standard RDBMS which generates results from the query. The results are converted back to fuzzy linguistic value by Crisp_to_Linguistic algorithm, where both these algorithms are part of FDSL tool. The output is fuzzy linguistic value results.

The FDSL Tool can be used with many fuzzy expert system applications. This tool provides interpretability in the field of medical informatics,

where it is easier for the medical practitioner to query and get results both in fuzzy linguistic term. Another advantage is that results generated by this FDSL tool can be stored as fuzzy RDBMS storage and the stored fuzzy relational database can be used to generate Fuzzy Knowledge Base System, for any Fuzzy Expert System.

The application in this paper is shown for a medical informatics, but is not restricted only to medical domain, and can be used in any application related to other domains like business, education, engineering etc.

References

[1] L. A. Zadeh, (1965)"Fuzzy sets". Information and Control, Vol. 8, 1965, pp. 338-353.

[2] B. P. Buckles, F. E. Petry, (1982) "A Fuzzy Representation of Data for Relational Databases", Fuzzy Sets and Systems, Vol. 7, Issue 3, pp. 213-226, 1982.

[3] A. Zvieli, P. Chen , (1986) "ER modeling and fuzzy databases". In Proceedings of the Second International Conference on Data Engineering. pages 320-327, LA, USA.

[4] Medina M., Pons O., Vila A., (1994) "GEFRED: A Generalized Model of Fuzzy Relational Databases". Information Science, No. 76, Pages 87-109.

[5] Medina M., Pons O., and Vila A., (1994) "An Elemental Processor of Fuzzy SQL". Computer Journal of Math Ware and Soft Computing, vol. 1, no. 3, pp. 285-295.

[6] Chen, G., Kerre, E., (1998) "Extending ER/EER concepts towards fuzzy conceptual data modeling". In: Proc. IEEE Intl. Conf. Fuzzy Syst. (Anchorage, AK), pp. 1320-1325.

[7] Chaudhry, N., Moyne, J., Rundensteiner, E.,(1994) "A design methodology for databases with uncertain data". In: Proc. 7th International Working Conference on Scientific Statistical Database Management .Charlottesville, VA., pp. 32 - 41.

[8] Galindo, J., Urrutia, A., Piattini, M., (2006) "Fuzzy Databases: Modelling Design and Implementation". IDEA Group, Hershey, PA.

[9] Meier A., Werro N., Albrecht M., Sarakinos M., (2005). "Using a Fuzzy Classification Query Language for Customer Relationship Management". Proceedings 31st International Conference on Very Large Data Bases (VLDB), Trondheim, Norway, pp. 1089-1096.

[10] Werro N., Stormer H. and Meier A., (2006) "A Hierarchical Fuzzy Classification of Online Customers". Proceedings of the IEEE International Conference on e-Business Engineering (ICEBE), Shanghai, China, pp. 256-263.

[11] Meier A., Werro N., (2007) "A Fuzzy Classification Model for Online Customers". Informatica, Vol. 31, No. 1, pp. 175-182.

[12] Touzi A., Hassine M., (2009) "New Architecture of Fuzzy Database Management Systems", The International Arab Journal of Information Technology, Vol 6, No. 2.

[13] J. Galindo, J. M. Medina, O. Pons, J. C. Cubero. (1998) "A Server for Fuzzy SQL Queries". In T. Andreasen, H. Christiansen, and H. L. Larsen, editors, Flexible Query Answering Systems, volume 1495 of Lecture Notes in Artificial Intelligence. Springer.Galindo J., (2008) "Le Serveur FSQL and FQ", http://www.lcc.uma.es/personal/ppgg/FSQL.html.

[14] Bosc P. , Pivert O., (1995) "SQLf: A Relational Database Language for Fuzzy Querying," Computer Journal of IEEE Transactions on Fuzzy Systems, vol. 3, no. 1, pp. 80-89,

[15] Bosc P., Lietard L., Pivert O., (1998) "Bases de Donnees et Flexibilite: Les Requetes Graduelles," Computer Journal of Techniques et Sciences Informatiques, vol. 7, no. 3, pp. 355-378.

[16] Bosc P. , Pivert O., (2000) "SQLf Query Functionality on Top of a Regular Relational Database Management," in Proceedings of Knowledge Management in Fuzzy Databases, Heidelberg, pp. 171-190.

[17] Statlog Heart Dataset, UCI Dataset Repository, V. A. Medical Center, Long Beach and Clevelan Clinic Foundation. http:// www.archive. ics.uci. edu/ml/dataset/Heart+Disease.

[18] Ali A., Mehdi M., (2010). "A Fuzzy Expert System for Heart Disease Diagnosis", Proceedings of the International Multi Conference of Engineers and Computer Scientists, Vol I, IMECS 2010.

[19] Bedi P., Harmeet K., Ankit M., (2002) "Fuzzy dimension to databases". In proceedings 37th Annual Convention of Computer Society of India, October 29-31, 2002, Bangalore.

[20] Gupta P., Rishi R., Mittal H., (2011)"Database Design for Storage of Fuzzy Information in Traditional Databases". International Journal of Computer Applications vol 15 No. 2.

[21] Galindo J., Medina J., Garido M., (2001)"Fuzzy division in fuzzy relational databases: an approach", Fuzzy Sets and Systems, vol 121, 471-490.

[22] Mala I., Akhtar P., Zia S., Mirza S., (2011)"Application of Fuzzy Relational Databases in Medical Informatics", Proceedings of the 14th IEEE Multi-topic conference (INMIC). Pages 41-44.

[23] Skrbic. S., Rackovic. M., (2009) "PFSQL: a fuzzy SQL language with priorities". In: Proc. PSU-UNS Inter. Conf. on Engineering Technolgies, (ICET 2009) Novi Sad, Serbia, pp. 58 -63.

[24] Skrbic. S., Rackovic. M., Takaci. A., (2011) "Towards the methodology for development of fuzzy relational database applications". Computer Science and Information System Vol. 8, Issue 1, pp 27-40.

[25] Skrbic. S., Takaci. A., (2009). "An interpreter for priority fuzzy logic enriched SQL". In: Proc. 4th Balkan Conf. Informatics.Thessaloniki, Greece. pp. 96-100.

5/25/2013