Valentin FILATOV[1], Natalia KASATKINA[2]

Kharkov University of Radioelectronics (1), Supreme Attestation Commission of Ukraine (2)

Fuzzy models presentation and realization by means of relational systems

Abstract. *The feature of the research is the proposed fuzzy models representation method by means of relational systems, which unlike the known approaches, can solve integrated data mining problems in databases and fuzzy systems.*
Streszczenie:

Keywords: information system, relational databases, intelligent systems, fuzzy model, data mining.
Słowa kluczowe:

Introduction

Computer technologies of intelligent algorithms are experiencing their heyday. This is due to the flow of new ideas coming from computer science, which was formed at the intersection of artificial intelligence, statistics and database theory. Many scientists research the problem of designing fuzzy databases, particularly on the basis of the relational model. Various versions of fuzzy relational database models have been designed. The results of the surveys, can be found in these works: Buckles B.P., Petty F.E., Shenoi S., Melton A., Vila M.A., Lipski W. Jr., Prade H.

Systems built by combining the databases and fuzzy logic can significantly extend the functionality and range of tasks of data mining. The theory of fuzzy databases is not yet complete from a mathematical point of view, and there are still many issues that require resolution.

Development of a relational model representation and storage of fuzzy data

To solve the problems of storing fuzzy data we will define a specialized type of relation. The scheme of such relation must satisfy two conditions: compatibility with the requirements of the classical relational data model and effective storage and representation of linguistic variable model [1]. Now let's look into it's graphical representation. Any line on a coordinate plane can be represented as a binary relation, where $Dom\ R$ is represented by the values of the x-axis, and $Im\ R$ - the y-axis. There are three indicators in the fuzzification problem that should be considered when forming relation. Define the fuzzy variable as a set (N, X, Y), where N - is the name of a variable, X - the area of research, Y - a fuzzy set on X. Using this definition, we define the three domains corresponding to elements of the variable.

Assume that $N = \{n_1, ..., n_m\}$, $Y = \{0, 0.1, ..., 1\}$, $X = \{x_0, ..., x_k\}$. The X and Y correspond to the selected scale sampling of the coordinate axes and represent the region belonging to the parameter N. For this particular case we define the respective domains in order to present the values of fuzzy variable. $D_1 = \{n_1, n_2, n_3, ..., n_m\}$; $D_2 = \{x_0, ..., x_8, x_9, ..., x_{14}, ..., x_{19}, ..., x_{29}, ..., x_{30}, ..., x_m\}$; $D_3 = \{0, 0.1, ..., 1\}$. We define a set of domain names and mapping to formulate a set of attributes: for a set of names $A = \{A_1, A_2, A_3\}$, - mapping $\rho : (A_1 \to D_1; A_2 \to D_2; A_3 \to D_3)$ determines the set attributes $A = \{A_1, A_2, A_3\}$, which corresponds to the relation scheme $S(A_1, A_2, A_3)$. Thus, in general, we can talk about a universal relation, which includes the full set of tuple resulted by domains Cartesian join $D_1 \times D_2 \times D_3$ [2].

Developing a model of integration fuzzy relation and relational database

Let's look at the problem in general. Let $U\ (R_1, ..., R_n)$ – database that stores basic data about the studied domain, $R^1(A_1, A_2, A_3)$ – fuzzy relation. The problem will make sense if there is an attribute in U database, fuzzification is based on

In order to provide simultaneous work of two databases U and R^1 we need to formalize the process of integration, based on the gradual normalization technique. The structure of U relation is based on basic functional dependencies $F = \{M_i \to N_i\}$ where $M_i, N_i \in U$. Now select one of the dependencies, which includes an attribute with fuzzy parameters, and define it as a $W \to V$, where W and V, in global case, may be sets. Relation R^1 contains one dependence $F' = \{A_1, A_2, A_3 \to A_1, A_2, A_3\}$. Now we can obtain an equivalent set $F' = \{A_1, A_2, A_3 \to A_1; A_1, A_2, A_3 \to A_2; A_1, A_2, A_3 \to A_3\}$ using output axioms.

Let's assume that the parameter corresponds to the fuzzy attribute A_2, then to determine the type of communication we need to get a set $F = F \cup F'$ and study two cases affecting the rules of normalization. 1. $A_2 \in W$ - search for parent relations: if functional relations and $\xi \to \zeta$, and $\omega \to \zeta$ take place - then the dependence of $\omega \to \zeta$ is incomplete. 2. $A_2 \in V$ - search for transitively dependent elements: if the functional relations $\xi \to \omega$ and $\omega \to \zeta$ take place - then an element ζ is transitive dependent. The existence of such relations will allow to make correct decomposition and to establish connection between U and R^1 databases. If $A_2 = W$ or $A_2 = V$, then the process of decomposition leads to the second or third normal form. If equation is not satisfied, then it's impossible to organize support of connected data uniqueness, because the association between relations will be "many-to-many".

Thus, in general to analyze data that's stored in relational databases, it is enough to build a fuzzy relation and establish a connection with the attribute (attributes) whose values must be properly analyzed.

Conclusions

Based on analysis of relational data model's features we described the main problems of building integrated information systems that combine the database as a source of primary data mining tools that enable to extract knowledge from data. This article proposes an efficient method for designing relational data model schema to represent the membership functions of linguistic variables. The proposed approach describes data analysis using fuzzy queries.

REFERENCES

[1] Zadeh L.A. The concept of a linguistic variable and its application to approximate reasoning // Information Sciences. - 1975. - 8. - P. 199 - 249.

[2] Codd E.F., The Relational Model for Database Management – version 2, reading, MA: Addison-Wesley.

Authors: *prof. Dr. Valentin Filatov, Kharkov University of Radioelectronics, Lenina av., 14, 61166, Kharkov, Ukraine, E-mail: filatov_val@ukr.net, Ph.D. Natalia Kasankina, Deputy Head of Technical Science Department Supreme Attestation Commission of Ukraine, Khreshchatyk st., 34 01001, Kiev, Ukraine. E-mail: nkasatkina@ukr.net*