ITM Web of Conferences 6, 02006 (2016)
DOI: 10.1051/itmconf/20160602006

Support of decision-making in the conditions of uncertainty of different types

Ha D. Doan[1,a], Alexander N. Pylkin[1], Alexander V. Kroshilin[1], Svetlana V. Kroshilina[1] and Valeriya V. Tishkina[1]

[1]*Ryazan State Radio Engineering University, 390005 Ryazan, Russia*

Abstract. At the present stage the solution to the problem of decision-making support in the conditions of uncertainty is seen in the application of decision-making support systems (DMSS), which are based on the theory of fuzzy sets and fuzzy logic. One of the most important stages of decision-making support in the conditions of uncertainty is the analysis of the accumulated statistical data. At present, the fuzzy algorithms, including the algorithms of fuzzy clustering of data, are successfully applied to effective research of such data. In recent years, for the effective solution to problems of fuzzy clustering, the device of fuzzy relations that is based on building fuzzy relations of data objects and their properties is widely and successfully used.

1 Introduction

The process of decision-making consists in the generation of possible alternative decisions, their assessment and the selection of the better variant. In the conditions of society informatization, support of decision-making has become the most topical. At the present stage the solution to this problem is seen in the application of decision-making support systems (DMSS).

Modern computer systems allow accumulating, storing information and managing huge volumes of data for resolving problems of processing and analysis. Unfortunately, the presence of data in itself is still not enough, as the decision-maker often needs so-called hidden information, in connection with this, it is necessary to transform available data into useful information for decision-making.

Uncertainty is an integral part of decision-making processes, as the choice of some variant of the decision demands the accounting of a large number of uncertain and inconsistent factors. In this connection, uncertainty can be classified as follows:

- the uncertainty is connected with the incompleteness of knowledge about the problem, to which the decision must be made;
- the uncertainty is connected with the impossibility of the full accounting of environment reaction to the made decisions;
- the uncertainty is connected with the wrong understanding of the purposes by the person, who makes the decision.

Inconsistency appears because of ambiguity of the assessment of situations, mistakes in the choice of priorities that, finally, strongly complicates decision-making. Researches show that the decision-makers without additional analytical support, as a rule, use simplified, sometimes inconsistent rules of decision selection.

The device of fuzzy relations that is based on the theory of fuzzy sets and fuzzy logic is widely and successfully applied to resolve above problem now. Thus the fuzziness of actions in decision-making process is presented in the form of fuzzy algorithms, on the basis of which DMSS realize formalization and processing of fuzzy information.

2 The problems of decision-making support systems

The main functions of DMSS are accumulation, storage, analysis and data processing to provide the decision-maker with useful information in the subject area, and also the ability to choose the best variant of possible decisions. Therefore, for the effective functioning of DMSS, it is necessary to provide the solution to 3 main tasks: data input; data storage and data analysis [1].

In solving the problem of data input, transaction data processing is carried out with using common database management systems (DBMS).

For data storage, modern DBMS and data warehousing concept are used.

Data analysis can be performed by the following methods:

- Information retrieval analysis of data on the basis of relational DBMS and static queries with using SQL (Structured Query Language);
- Operational data analysis with applying the technology of operative analytical data processing OLAP

[a] Corresponding author: d2h.rus@mail.ru

(Online analytical processing), which uses the concept of multi-dimensional data representation;

- Methods and algorithms of Data Mining.

DMSS do not generate the correct decision but only provide the decision-maker with tools and data for the analysis and assessment of a concrete situation in the subject area, thus for the effective use of DMSS, the decision-maker must have both the corresponding qualification and ability to use data, which are presented in the machine form.

The analysis of accumulated statistical data, which volume continuously increases, plays an important role in activity of DMSS in the conditions of uncertainty. In researching of the accumulated statistical data, there is a high probability to skip the potential helpful information, which is implicit and hidden. To effectively research the accumulated statistical information, it is necessary to apply the special methods of intellectual analysis, in which the methods of data clustering take an important place [2].

3 Fuzzy clustering of data in the decision-making support systems in the conditions of uncertainty

One of the most important stages of decision-making support in the conditions of uncertainty is the analysis of accumulated statistical data. Today for effective research of them the fuzzy algorithms that are based on the theories of fuzzy sets and fuzzy logic, including the algorithm of fuzzy clustering of data, are successfully applied [3].

Data clustering is the partitioning a set of data objects into groups of similar objects (clusters). And objects in various clusters are considerably different from each other. The main purpose of data clustering is to search the cluster structure for data understanding and information post-processing. The main feature of data clustering consists in that the number of clusters isn't known beforehand.

In the tasks of clear clustering, a row of assumptions is laid: clusters have the certain form; each data object always belongs to only one cluster; the limits among clusters are well defined and etc. In general, the form of clusters can be random, the limits among clusters can be indefinite, and some objects belong to different clusters. In practice, the obtained data has fuzzy character, which is caused by the inaccuracy, insufficiency, unreliability of information resources. They lead to the difficulty in obtaining adaptive decision.

The above-stated shortages were overcome by the ideas of fuzzy clustering: clusters are the fuzzy subsets of initial sets of data, and each data object belongs to different clusters with various belonging degrees that correspond with various values of belonging function, which can be found in the segment [0, 1].

The basic algorithm of fuzzy clustering of data Fuzzy C-Means (FCM) carries out partitioning a set *X* of *n* data объектов into *m* clusters by specifying the belonging matrix *U* of dimensions $n \times m$. In this situation, each data object *x* that is contained in *X* belongs to any cluster with a belonging degree, which is an element of belonging matrix *U* in the interval [0, 1]. The belonging degrees of data objects are calculated on the basis of relationships among data objects and “cluster centroids”. This process is iteratively carried out until the objective function achieves the minimization value.

Besides the algorithm FCM, there is a considerable quantity of other known methods of fuzzy clustering, which are its variety, for example, the Gyustafson-Kessel's method, the K-medium methods, etc. In general, they are simple in realization and work quickly enough. However, when these methods are applied for initial information, which has a weak structure, there are some shortages: all clusters have the certain form, which is defined by the algorithm; the cluster centroids always exist, but in practice they can be missing.

In recent years, for the effective solution to problems of fuzzy clustering, the device of fuzzy relations is widely and successfully used. This approach is based on building the fuzzy relations of data objects and their properties [1].

Introduction of the concept of "the normal measure of proximity" that is based on distances between two any attributes *x*, *y* of the fuzzy set *X* of initial data:

$$\mu_x(y) = 1 - \frac{d(x,y)}{\max\limits_{z\in X} d(x,z)}, \qquad (1)$$

where: $x, y, z \in X$; $d(x,y)$ – the distance between *x* and *y*; $max\, d(x,z)$ – the maximum distance between *x* and *z*. The distance $d(x,y)$ can be calculated by the known measures of proximity: the Euclidean distance, the Hamming distance, etc. At that time, *x* and *y* are really different from each other when $\mu_x(y) = 0$, and are absolutely close when $\mu_x(y) = 1$.

The relative measure of proximity $\mu_x(y,z)$ of two attributes *y* and *z* to the third *x* is defined:

$$\xi_x(y,z) = 1 - |\mu_x(y) - \mu_x(z)| \qquad (2)$$

On the whole set *X* the proximity measure of all attributes is built:

$$\xi(y,z) = T\left[\xi_{x_1}(y,z), \xi_{x_2}(y,z), \ldots, \xi_{x_{|X|}}(y,z)\right] \qquad (3)$$

here: $\xi_{x_i}(y,z)$ – the relative measure of proximity, $x_i \in X$, $i = \overline{1,|X|}$, T – operation *t-norm*.

T – operation *t-norm*, which answers the *min-norm* minimization according to Zadeh that is used in this case:

$$T(y,z) = \min(y,z) \qquad (4)$$

It follows from (3) and (4) that the proximity measure of two attributes on *X* is presented in the following type:

$$\xi(y,z) = \min\left[\xi_{x_1}(y,z), \xi_{x_2}(y,z), \ldots, \xi_{x_{|X|}}(y,z)\right] \qquad (5)$$

The relationship that receives from the expression (5) is called the fuzzy soft relationship, which transitive closure in cycle must be calculated by the following formula:

$$R_\xi^t = R_\xi^{t-1} \circ R \qquad (6)$$

where: $R_\xi = \xi(y,z)$, $t = \overline{1,|X|}$. The received result $R_\xi^{|X|}$ is the relationship of fuzzy equivalence.

Thus, the relationship gradation of fuzzy equivalence generates the family of equivalence relations in classical sense that partitions an initial set of data, which is intended for research, into classes of equivalence. The more level of relationship, the more a set X of data is partitioned in detail.

References

1. A.V. Kroshilin, A.N. Pylkin and S.V. Kroshilina, *2014 International conference on computer technologies in physical and engineering applications (ICCTPEA)*, pp. 89-90 (2014)
2. D.H. Doan, A.V. Kroshilin and S.V. Kroshilina, *The VIII International Academic Congress "Development of Countries in Asia, Africa and Europe: Past, Present and Future"* **II**, 256 (Seoul, 2015).
3. L. Demidova, E. Nikulchev and Yu. Sokolova, International Journal of Advanced Computer Science and Applications, **7** (2016)