ON FUZZY SEARCH IN DATA SETS

S. Shumikhin[1], *I. Trymorush*[2]

[1] IMEM, Odessa I. I. Mechnikov National University,
Department of Methods of Mathematical Physics
Dvoryanskaya str., 2, 65082 Odessa, Ukraine

shum@onu.edu.ua,

[2] BERG Faculty, Technical University of Kosice,
B. Nemcovej 3, 04200 Kosice, Slovakia

iryna.trymorush@tuke.sk

In this work we study the problem of performing an efficient search in the unstructured data sets. One of the approaches to analysis and search across big data sets is the clusterization. To provide more flexibility, the fuzzy clusterization can be used. The big amount of the clusters by itself does not resolve the search problem. The set of clusters is not structured, and this results in inefficient search procedure. The possible solution of such problem is to structure the fuzzy data sets. One of the most efficient search data structures is a binary search tree. In this work we propose how to build the fuzzy binary search tree on the unstructured data sets for performing the efficient fuzzy search.

The creation of the binary search tree is based on the order of the sets and the equivalence relations. We have defined the α-level order and equivalence relations, based on the one introduced by [1]. The order relation reflects the order between two fuzzy sets, it is some kind of 'less', 'more' for the ordinary sets; and the equivalence relation reflects how similar two fuzzy sets are. We have also determined several types of clusters that represents the tree's nodes. Having all aforesaid the fuzzy binary search tree can be build and the efficient binary search can be performed.

Naturally, the order and the equivalence relations can be defined in some other way, but it has to satisfy some rules, such as reflexivity, symmetry, transitivity [2] or reflexivity, anti-symmetry, transitivity respectively [3, 4].

A sample application has been written in $C\#$ using the .Net Framework, as the object-oriented programming concepts are most appropriate for the creating entities with necessary characteristics.

References

[1] Kham, L. D. Some extensions in the theory of relational fuzzy databases. St. Petersburg, 2003. Available at `http://www.lib.ua-ru.net/diss/cont/57860.html` (in Russian).

[2] Ibrahimov, B. *Elements of Fuzzy Mathematics*. Science, Baku, 2010. Available at `http://anl.az/el_ru/i/iv_enm.pdf` (in Russian).

[3] Zadeh, L. Fuzzy sets. *Information and Control*, vol. 8, no. 3, 1965, pp. 338–353.

[4] Zadeh, L. *The Concept of Linguistic Variable and Its Application to Approximate Reasoning.* Mir, 1976 (in Russian).