Proceedings of the International Multiconference on Computer Science and Information Technology pp. 151–160
ISSN 1896-7094

Implementation and evaluation of the fuzzy system for the assessment of cadastre operators' work

Tomasz Karczyński and Bogdan Trawiński,

Wrocław University of Technology, Institute of Applied Informatics,
Wybrzeże S. Wyspiańskiego 27, 50-370 Wrocław, Poland
bogdan.trawinski@pwr.wroc.pl, tomkarczynski@googlemail.com

Abstract. The Mamdani fuzzy system for the multi-criteria assessment of cadastral system operators' work is presented in the paper. The system comprises five input criteria: productivity, complexity, quality, availability and report production and was implemented using JAVA. Moreover two alternative methods of assessment were incorporated into the system: first based on self-organizing maps and second employing linear function of input criteria. All three methods of operator assessment were evaluated using actual data taken from one cadastre information centre for the period of 70 weeks.

1 Introduction

The methods of computational intelligence are often applied to solve multi-criteria non-linear problems [1–4, 7]. The Mamdani fuzzy model for the multi-criteria assessment of cadastral system operators' work is presented in the paper. Cadastre systems belong to state resources and are designed for the registration of parcels, buildings and apartments as well as their owners and users. Those systems have complex data structures and sophisticated procedures of data processing. There are above 400 information centres located in district local self-governments as well as in the municipalities of bigger towns in Poland. One of critical tasks of cadastre system maintaining is the input of changes into its database. The changes are allowed only on the basis of legal documents such as notarial deeds, extracts from perpetual books, results of surveyors' works, administrative decisions and others. Managers of information centres often complain they have no adequate tool for the assessment of work of cadastre system operators. The first fuzzy models for the assessment of information system operators' work comprising three criteria: productivity, complexity and quality were proposed and evaluated in [5, 6]. The extension of the model of two more criteria i.e. operators' availability and report production as well as the implementation of the assessment system is presented in the paper.

2 Fuzzy Model for the Assessment of Operators' Work

Fuzzy model for the assessment of operators' work is a multi-criteria one that comprises five input variables which values are automatically calculated using operators' activity data recorded in the cadastral database. The input variables are: productivity (further denoted by P) expressed by the number of changes input into the cadastre database by an operator within a given period, complexity of changes (C) specified as the mean number of objects which were modified in the database within a given period, quality of work (Q) defined as the percentage of changes made without any corrections, report production (R) determined by the number of reports generated by an operator within a given period and the availability (A) qualified as the percentage of work time spent on fulfilling tasks using cadastral system.

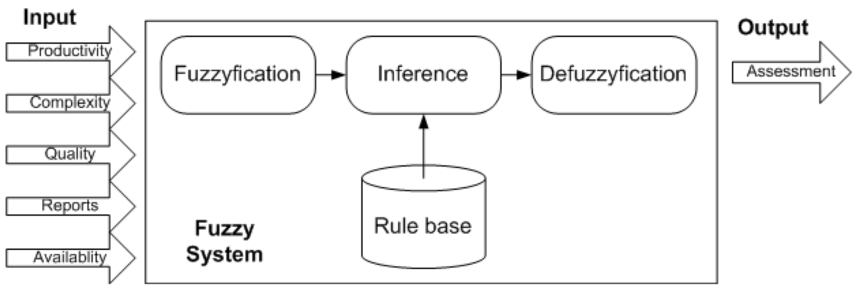

Fig. 1. Classic model of a fuzzy system

The classic model of a fuzzy system (FS) is given in Figure 1. The crisp values of input variables are fuzzyfied, the inference process is performed using rule base and next the fuzzy result of aggregation is defuzzyfied making final crisp assessment value. However in the case of five input criteria no expert is able to create a reliable rule base, so that before implementation the fuzzy model had to be restructured. Having rule bases of selected couples of input variables expressed in the form of two dimensional matrices, eventually seven FSs were constructed and put together in the structure of a tree (see Figure 2). Input of the system is directed to four FSs at level 0 and their output crisp values constitute the inputs to next two FSs at level 1, and in the end the results of the latter are used as the inputs of the terminal FS at level 2. Final assessments depend on all five input variables. The input variables were coupled by an expert who took into account some logical relations between them. For example P and C are related because the more changes completed the more database object modified. However some operators, those more experienced, are able to perform more complex changes whereas others confine to simpler changes. In turn the more complex changes and the bigger number of database updates the higher probability of errors. However some operators are more careful and others less.

And sometimes the operators are not assigned their main tasks but they are asked to generate various kinds of system reports. But both input of changes and report production are good indicators of operators' availability.

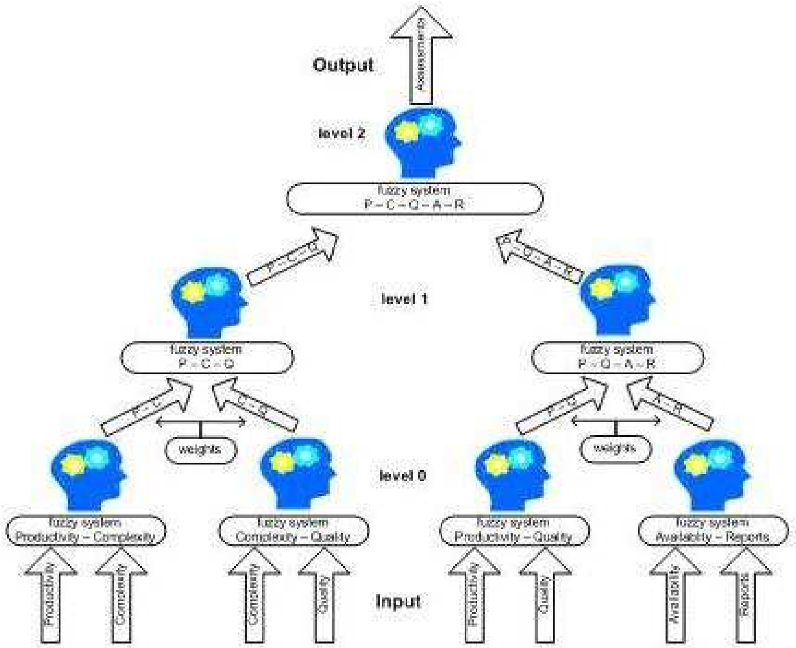

Fig. 2. Tree structure of fuzzy systems.

Mamdani models with fuzzy sets expressed by means of triangular and trapezoidal functions were used to construct fuzzy systems at respective levels. Three different models of input and output variables were used, they were named 7x9, 9x11 and 11x11. In the 7x9 model 7 fuzzy sets determine each input criteria and the output is defined by 9 fuzzy sets. In the 9x11 and 11x11 models there are 9 and 11 fuzzy sets for each input as well as 11 for each output respectively. Fuzzy membership functions used to define P, C, Q, A and R input variables and an output are very similar to those described in [5, 6]. The statistics subsystem provides initial parameters of the model and values of input criteria. The idea of obtaining the final assessment consists in calculating the average value of P, C, Q, A and R criteria taking into account the activity records in a cadastre database for all operators and for long period of time, e.g. a year or a half of year. These average values are used as the middle values of the universe of disclosure for fuzzy sets describing individual input variables. Standard deviations for each criterion are calculated for a long period of time too and they are used for the determination of the width of the basements of triangular and trapezoidal fuzzy sets. The domain for output is an arbitrary assessment scale from 0 to 200, with 0 being the lowest and 200 the highest mark. The initial reference average values and the widths of triangle or trapezoid basements are treated as customization parameters and can be modified by managers during exploitation. The rule base for each model was proposed by an expert and contains simple IF-THEN rules where the condition consists of only two input variables combined by AND operator and the conclusion is built by one variable. An example of a rule is as follows: ***IF Complexity is low AND Productivity is medium THEN Assessment is very low.*** Rule bases for FSs at successive levels are shown in tables 1, 2, and 3, where UL, EL, VL, L, BM, M, AM, H, VH, EH, UH denote Ultra Low, Extremely Low, Very Low, Low, Below Medium, Medium, Above Medium, High, Very High, Extremely High and Ultra High respectively. There is following relationship between the meaning of linguistic values UL $\prec$ EL $\prec$ VL $\prec$ L $\prec$ AL $\prec$ BM $\prec$ M $\prec$ AM $\prec$ BH $\prec$ H $\prec$ VH $\prec$ EH $\prec$ UH. In order to be able to determine the importance of a given pair of input criteria it is possible to assign a weight to a rule base at level 1.

Table 1. Representation of rule base in matrix form for level 0 (7x9 model)

	VL	**L**	**BM**	**M**	**AM**	**H**	**VH**
VL	EL	EL	VL	L	L	BM	M
L	EL	VL	L	L	BM	M	AM
BM	VL	L	L	BM	M	AM	H
M	L	L	BM	M	AM	H	H
AM	L	BM	M	AM	H	H	VH
H	BM	M	AM	H	H	VH	EH
VH	M	AM	H	H	VH	EH	EH

Table 2. Representation of rule base in matrix form for level 1 (9x11 model)

	EL	**VL**	**L**	**BM**	**M**	**AM**	**H**	**VH**	**EH**
EL	UL	EL	VL	VL	L	L	L	BM	BM
VL	EL	EL	VL	L	L	BM	BM	M	M
L	VL	VL	L	L	BM	BM	M	AM	AM
BM	VL	L	L	BM	BM	M	AM	H	H
M	L	L	BM	BM	M	AM	AM	H	H
AM	L	BM	BM	M	AM	AM	H	H	VH
H	L	BM	M	AM	AM	H	H	VH	VH
VH	BM	M	AM	AM	H	H	VH	VH	EH
EH	BM	M	AM	H	H	VH	VH	EH	UH

Table 3. Representation of rule base in matrix form for level 2 (11x11 model)

	UL	**EL**	**VL**	**L**	**BM**	**M**	**AM**	**H**	**VH**	**EH**	**UH**
UL	UL	UL	EL	VL	VL	L	L	AL	AL	BM	BM
EL	UL	EL	VL	VL	L	L	AL	AL	BM	M	M
VL	EL	VL	VL	L	L	AL	BM	BM	M	AM	AM
L	VL	VL	L	L	AL	BM	BM	M	AM	AM	BH
BM	VL	L	L	AL	MB	BM	M	AM	AM	BH	BH
M	L	L	AL	BM	MB	M	AM	AM	BH	BH	H
AM	L	AL	AL	BM	M	AM	AM	BH	BH	H	H
H	AL	AL	BM	M	AM	AM	BH	BH	H	VH	VH
VH	AL	MB	M	AM	AM	BH	BH	H	VH	VH	EH
EH	BM	M	AM	BH	BH	BH	H	VH	VH	EH	UH
UH	BM	M	AM	BH	BH	H	H	VH	EH	UH	UH

3 Alternatives to the Fuzzy Model

Two alternatives to the fuzzy model were implemented in the system for operators' work assessment: a mechanism based on self-organizing map (SOM) and one employing a linear function of input criteria. The SOM is a type of unsupervised artificial neural network, as only one neuron fires during learning it is called a winner-take-all-network [8]. The architecture of SOM implemented is shown in Fig. 3 where all neurons are arranged in a circular way. The lateral feedback function with the profile of a Mexican hat was employed. The network was trained using the activity data for 10 operators for the period of 70 weeks with 100 neurons and 3000 train iterations.

Second alternative implemented to calculate the assessment was the linear function of input criteria expressed by means of the following equation

$$Ass(x) = 100 + (w_p * \frac{P_x - P_n}{P_n} + w_c * \frac{C_x - C_n}{C_n} + w_q * \frac{Q_x - Q_n}{Q_n} + w_a * \frac{A_x - A_n}{A_n} + w_r * \frac{R_x - R_n}{R_n}) * 100$$

where P_x, C_x, Q_x, A_x, R_x are activity data for user X determined for a given week, and P_n, C_n, Q_n, A_n, R_n are average values of productivity, complexity, quality, availability and report production calculated for all users for a long period. Due to the lack of actual assessments it was not possible to determine coefficients by individual variables using multiple regression. So that criterion importance weights W_p, W_c, W_q, W_a, and W_r were used and those weights could be set by centre managers.

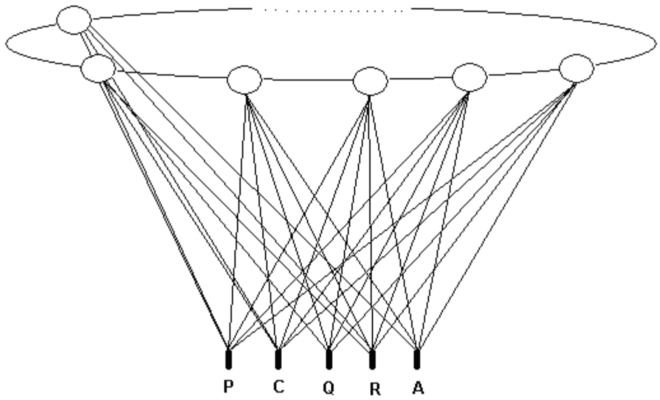

Fig. 3. The architecture of self-organizing map implemented

4 Implementation of the System

The system for the assessment of operators' work was developed for Windows environment using JAVA programming language and SQL Server 2005 database management system. It is composed of following modules: Data gain, Statistics, Assessment, Visualization, Customization and Help subsystems (see Figure. 4).

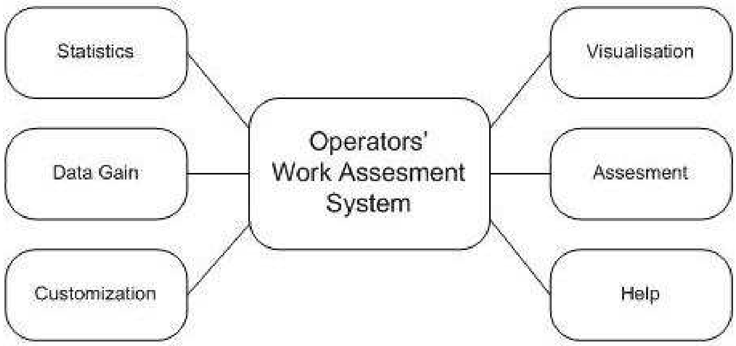

Fig. 4. Subsystems of the system for operators' work assessment.

The task of data gain subsystem (see Figure 5) is to scan periodically, eg. once a day or once a week, the database of the cadastral system, draw operators' activity data for successive days, process them and save into a proprietary database of the subsystem. In order to calculate the input values for FSs a period of day and night was divided into 48 time quanta each 30 minutes long and the data gain subsystem, calculates the frequencies within five dimensions: productivity, complexity, quality, availability and report production and assigns them to corresponding time quanta. Such distribution of activity data assures the assessment system to be flexible and easily parameterized and customized. Especially the manager can determine time scope i.e. what period is an unit of the assessment: a week or a month, what time of a day should be taken into account: either only working hours, or also overtime in the afternoon or early coming to work or night work. There is a possibility to define what days are interesting for monitoring users activity: either only working days or also Saturdays and holidays.

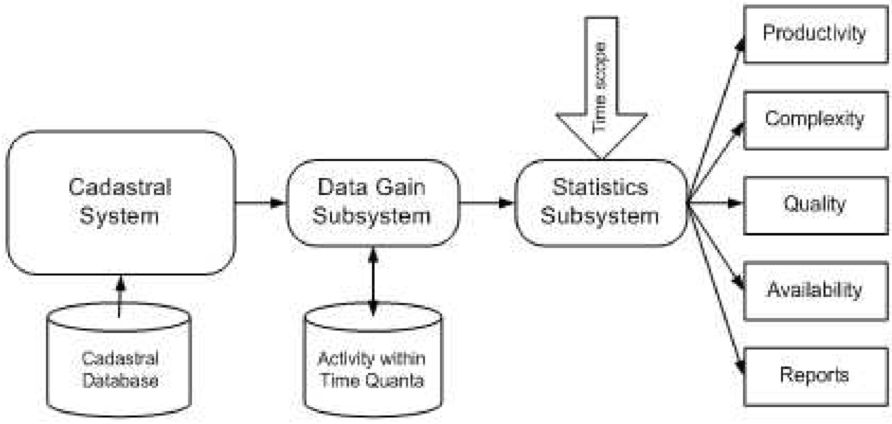

Fig. 5. Preparation of input values for FSs.

The main goal of the Customization subsystem is to provide tools to define work parameters of each operator and whole information centre as well as to set the parameters of the fuzzy models and the SOM neural network. It is possible to define

for each operator his weekly work schedule, leaves, sickness absences etc. As for the centre you can determine working hours, free days, and holidays. Finally the shapes of fuzzy sets, inference and defuzzyfication method, rule importance weights can be set for FSs and the number of neurons, variation range of neurons, number of iterations in the training phase, etc.

5 Evaluation of the System

The evaluation of the system was carried out using cadastre database taken from one information centre and the activity records for 10 operators working with the system during the period of 71 weeks from September 2005 till January 2007 were analyzed. In Fig 7 final assessments produced by fuzzy mechanism and

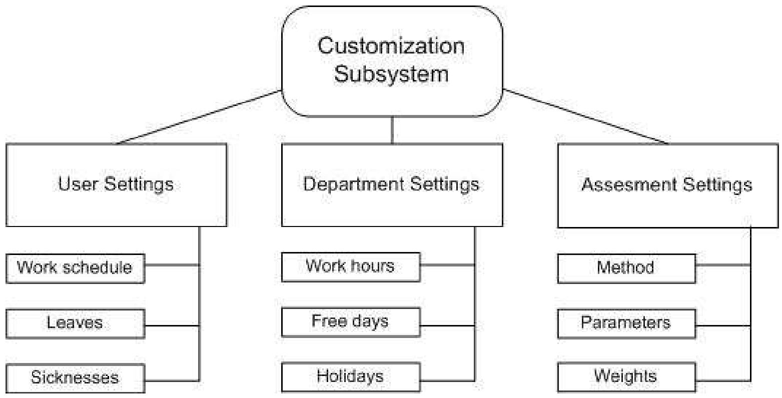

Fig. 6. Preparation of input values for FSs.

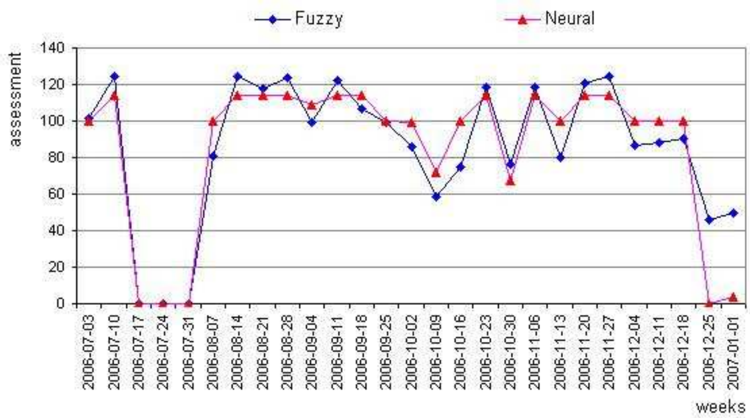

Fig. 7. Comparison of assessments produced by fuzzy set and SOM mechanisms for one user

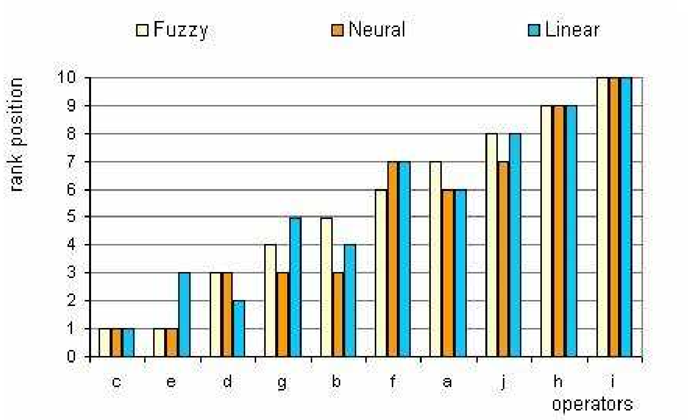

Fig. 8. Comparison of operator rankings made by fuzzy, neural and polynomial methods

self-organizing map for the user a for the period of half year are compared. The results produced by both methods are similar and they describe the variation of operators' work well. The summer leave period in August and holidays in December where no or minor activity was detected can be easily noticed in the chart.

Rank positions of individual operators produced by the system using fuzzy, neural and polynomial methods are presented in Figure. 8. It can be seen that all three methods equally recognized the best and the worst operators. This is similar to the results presented in [5] where the manager was able to distinguish easily only utmost operators and had substantial difficulties to assign different assessment to average ones.

6 Conclusions and Future Works

The model for the multi-criteria assessment of cadastral operators' work has been designed and programmed with the aid of an expert and information centre managers. The evaluation of the model has proved that it is reasonable to use operators' work statistics as the initial parameters and the input values for the model. The model basing on operators' activity records saved in cadastral database produces the assessments for defined periods of time automatically. Two additional assessment methods were implemented in the system, namely one based on self-organizing map and second employing linear function of input variables. The architecture of proposed model enables the centre managers to modify system parameters to customize them to local circumstances. It is planned to carry out further evaluation experiments with the active participation of the centre managers. They will be instructed how the fuzzy system operates and will be got familiar with the statistics of operators' work within a given time. Moreover they will be able to determine the importance weights to successive criteria in order to adjust the system to their preferences. Subjective opinions of the usefulness of the fuzzy system will also be gathered.

References

1. Ajith A.: *Neuro-Fuzzy Systems: Sate-of-the-Art Modelling Techniques,* Proceedings of the 6th International Work-Conference on Artificial and Natural Neural Networks, pp. 269–276, (2001).
2. Bonissone, P. P., Yu-To Chen, Goebel K., Khedkar P. S.: *Hybrid soft computing systems: industrial and commercial applications,* Proceedings of the IEEE,vol. 87, no. 9, pp. 1641–1667 (1999).
3. Cordón O., Gomide F., Herrera F., Hoffmann F., Magdalena L.: *Ten years of genetic fuzzy systems: current framework and new trends,* Fuzzy Sets and Systems 141 5–31 (2004).
4. Herrera F.: *Genetic Fuzzy Systems: Status, Critical Considerations and Future Directions,* International Journal of Computational Intelligence Research 1(1) pp. 59–67 (2005).
5. Król D., Kukla G., Lasota T., Trawiński B.: *Investigation of the fuzzy system for the assessment of cadastre operators' work,* In: Advances in Web intelligence and data minning. M. Last et al. (eds.) Studies in Computational Intelligence, pp. 131–140. Springer, Berlin (2006).
6. Król D., Kukla G., Lasota T., Trawiński B.: *Fuzzy model for the assessment of operator's work in a cadastre information system,* LNAI, vol. 4253, pp. 774–781. Springer, Heidelberg (2006).
7. Rutkowski L.: *Methods and Techniques of Artificial Intelligence,* Computational Intelligence (in Polish: Metody i techniki sztucznej inteligencji. Inteligencja obliczeniowa). PWN Warszawa (2005).
8. Skapura D. M.: *Building Neural Networks,* Addison Wesley (1996).