Using Fuzzy Logic Methods for Predicting Difficulties of the Hybrid Learning System Users

I.f. Astachova, E.I. Kiseleva

Abstract— The article offers a structural model of a hybrid learning system. The system's versatility lies in its independence from the specific subject content, which allows a teacher to upload the required course, and a student to choose an individual learning path. A hybrid system is a system that combines two or more various computer technologies. An algorithm for predicting difficulties of the students learning on the basis of fuzzy logic methods for predicting fuzzy time series is considered.

Keywords: structural model of a hybrid educational system, prediction of students' difficulties, prediction of fuzzy time series.

I. INTRODUCTION

Nowadays, various information and communication technology-based learning systems have gained acceptance. They differ in many parameters: by the degree of the control function distribution between the user and the system (in some of them, the user can independently choose his or her path within the system, while in others this function is partially or fully delegated to the computer); by the degree of combination of the theoretical and practical components; by presence or absence of a control function. The creators of various systems have been choosing a combination of parameters that correspond to the system's purpose. However, all information technology learning systems in use have a potential to accumulate statistical information about the students' path throughout the complex, about their errors in the working process, about their success when passing the control tasks, etc.

II. PROBLEM STATEMENT

The work aimed at designing a hybrid learning and controlling system that enables the user to employ the accumulated data for optimization of the target and content-related component. With this aim in mind, we set the following tasks: creating a mathematical model of a learning system; selecting the means for realization and optimization of the obtained model. With the purpose of developing an automated learning system we chose a hybrid system since it combines the advantages of various technologies that allow solving each problem optimally. A hybrid system is one that combines two or more various computer technologies [1].

III. PROBLEM SOLVING METHODS

A model of a hybrid learning system is shown in figure 1

Each component is presented in several sections of the learning system. The target component manifests itself in the construction of the course goal tree. Prediction using fuzzy logic allows identifying the content sections that present the greatest challenges for the students, which in turn will allow correcting the content of goals. The control component is implemented using Kohonen network that allows classifying the students by three parameters: those who failed the diagnostics, those who demonstrated knowledge at the level of the educational standards, and those who demonstrated knowledge level higher than the educational standards. The statistical component allows saving the information about the students' activity within the system and using this information for optimization.

For more efficient goal-setting, diagnostics and correction, we used the fuzzy logic methods to determine the course sections that present the greatest challenges for the students.

Prediction of the students' difficulties using the fuzzy logic methods

The test result data for the 1st year students of the Faculty of Psychology and Pedagogics of Voronezh State Pedagogical University obtained in 2011–2018 do not allow determining which content sections present the greatest difficulties for the students. It is possible to solve this problem using fuzzy logic methods that allow making predictions based on relatively small data quantities. Prediction consists in estimation of the future states of an object or a system using science-based methods. Using the fuzzy logic methods will allow predicting the number of correct answers at the tests of each course subject. The prediction accuracy is sufficient to single out the subjects with the least number of correct answers that present the greatest difficulties for students.

The following prediction method is used in accordance with the stated problem.

1. Calculation of the variations of the correct answer percentage for each subject as the difference between the correct answer percentage in the current and previous years; determining the universal set *U*.

$$U = [V_{min} - D, V_{max} + D],$$

where V_{min} – the least variation of the students' correct answers, V_{max} – the greatest variation, D – the parameter selected so that to reduce the interval boundaries.

Revised Manuscript Received on April 12, 2019.

I.F. ASTACHOVA, Voronezh State University, Voronezh, Russia

E.I. KISELEVA, Voronezh State Pedagogical University, Voronezh, Russia

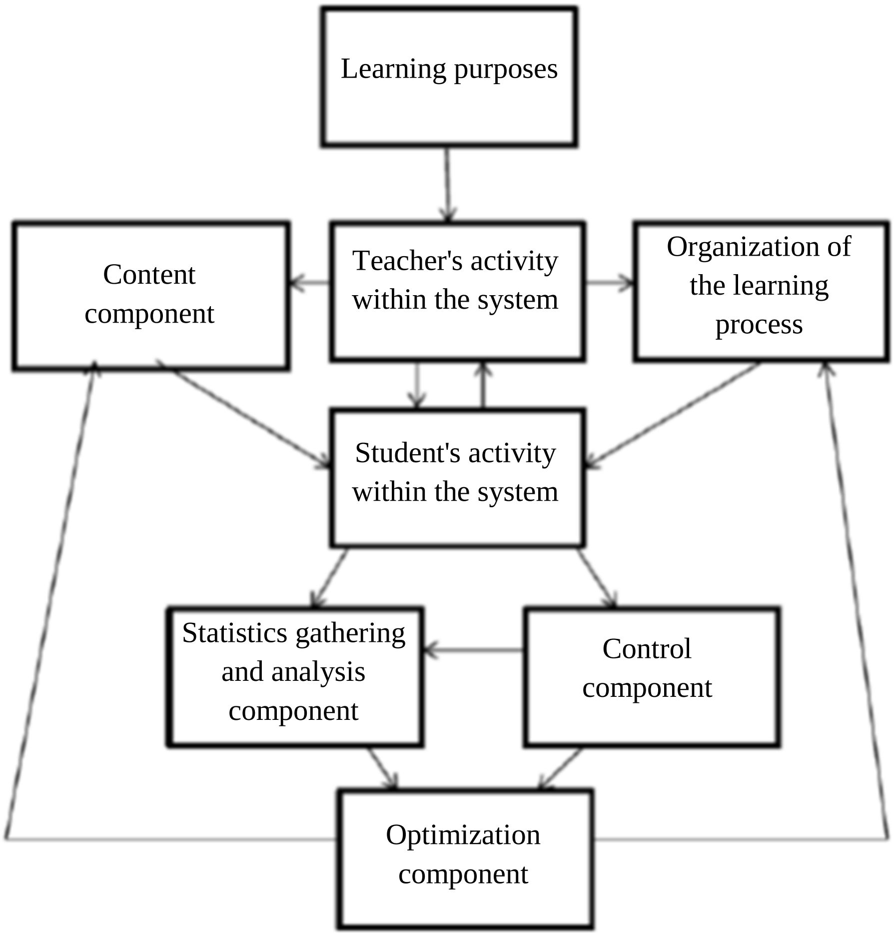

Figure 1 – Structural model of a hybrid learning system & Results

Table 1 – Variations of the students' correct answer percentage for subject "Correspondences between two sets"

Years	Percentage of correct answers	Variation	Fuzzification of variations
2011	10.53		
2012	30.00	19.47	A^{12}={0.07/u_1,0.09/u_2,0.12/u_2,0.17/u_4,0.26/u_5,0.41/u_6,0.67/u_7, 0.96/u_8}
2013	38.10	8.1	A^{13}={0.13/u_1 0.19u_2,0.29/u_2,0.47/u_4,0.76/u_5,1 /u_6,0.84/u_7, 0.53/u_8}
2014	43.18	5.08	A^{14}={0.16/u_1, 0.24u_2,0.39/u_2,0.64/u_4,0.94/u_5, 0.94/u_6,0.64/u_7, 0.39/u_8}
2015	30.35	-12.83	A^{15}={0.82/u_1, 1/u_2,0.78/u_2,0.48/u_4,0.3/u_5, 0.19/u_6,0.13/u_7, 0.1/u_8}
2016	46.22	8.13	A^{16}={0.13/u_1, 0.19/u_2,0.29/u_2,0.47/u_4,0.76/u_5, 1/u_6,0.84/u_7, 0.53/u_8}
2017	44.67	-1.55	A^{17}={0.28/u_1, 0.45/u_2,0.74/u_2,0.99/u_4,0.86/u_5, 0.55/u_6,0.34/u_7, 0.22/u_8}
2018	38.17	-6.5	A^{18}={0.45/u_1, 0.74/u_2,0.99/u_2,0.86/u_4,0.55/u_5, 0.34/u_6,0.22/u_7, 0.15/u_8}

$V_{min} = -12.83$, V_{max} =19.47, $U = [-20, 20]$

2. Dividing the universal set U into several intervals of the same length. In this case, the set was divided into 8 equal intervals $u_1 = [-20, -15]$, $u_2 = [-15, -10]$, $u_3 = [-10, -5]$, $u_4 = [-5,0]$, $u_5 = [0,5]$, $u_6 = [5,10]$, $u_7 = [10,15]$, $u_8 = [15,20]$. Since the average error of prediction of the fuzzy time series method is the smallest, the average points for these intervals are found:

$u^1_{mid} = -17{,}5$, $u^2_{mid} = -12{,}5$, $u^3_{mid} = -7{,}5$, $u^4_{mid} = -2{,}5$, $u^5_{mid} = 2{,}5$, $u^6_{mid} = 7{,}5$, $u^7_{mid} = 12{,}5$, $u^8_{mid} = 17{,}5$.

Linguistic variable "variation of the number of the students' correct answers" is introduced, which takes linguistic values: A_1"sharp decrease in the number of correct answers", A_2 – "significant decrease in the number of correct answers", A_3 "visible decrease in the number of correct answers", A_4 "insignificant decrease in the number of correct answers", A_5 "insignificant increase in the number of correct answers", A_6 "visible increase in the number of correct answers", A_7 "significant increase in the

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

number of correct answers", A_8 "sharp increase in the number of correct answers". A fuzzy variable and a fuzzy set correspond to each linguistic value. Fuzzy sets $A_i, i = 1..m$; are determined by formula [7]

$$\mu_{A_j}(u_i) = \frac{1}{1 + C\left(u_{mid}^j - u_{mid}^i\right)^2},$$

where u_{mid}^j, u_{mid}^i – midpoints of the respective intervals; coefficient C is selected so that to ensure occurrence of numbers $\mu_{A_j}(u_i)$ within interval [0,1]. In this work, *C*=0.01.

3. Fuzzification of the calculated variations implies converting the obtained exact values into fuzzy ones. The membership function reflects qualitative view of the answers in the given group [7]:

$$\mu_{A_t}(u_i) = \frac{1}{1+C(V_t - u_i^{cp})^2} \quad i = 1..n,$$

where *t* – numbers of the years used in the prediction denoted by the last two numbers for convenience, V_t – variation of the year *t*, *C*=const.

4. Fuzzifications of variations A^{12}, A^{13}, A^{14}, A^{15}, A^{16}, A^{17}, A^{18} are shown in table 1.
5. Number of years *w* included into the experimental estimation is selected, and a fuzzy relations matrix is constructed. In this work, *w*=5. Fuzzy relations $A^{13} \to A^{12}$, $A^{14} \to A^{13}$, $A^{14} \to A^{12}$, $A^{15} \to A^{14}$, $A^{15} \to A^{13}$, $A^{15} \to A^{12}$, $A^{16} \to A^{15}$, $A^{16} \to A^{14}$, $A^{16} \to A^{13}$, $A^{16} \to A^{12}$ are considered. The matrix of the relation *R* is a fuzzy relation of the form:

$R = \cup_{i=1}^{10} R_j$, where $R_j = A^s \times A^q, s = 13, \ldots, 16, q = 12, \ldots, 15$, $\cup$ - join operator [7].

$$R = \begin{bmatrix} 0.16 & 0.19 & 0.29 & 0.47 & 0.76 & 0.82 & 0.82 & 0.53 \\ 0.24 & 0.24 & 0.29 & 0.47 & 0.76 & 1 & 0.84 & 0.53 \\ 0.39 & 0.39 & 0.39 & 0.47 & 0.76 & 0.78 & 0.78 & 0.53 \\ 0.64 & 0.64 & 0.64 & 0.48 & 0.64 & 0.64 & 0.64 & 0.53 \\ 0.82 & 0.94 & 0.78 & 0.64 & 0.76 & 0.94 & 0.84 & 0.53 \\ 0.82 & 1 & 0.78 & 0.64 & 0.94 & 1 & 0.84 & 0.53 \\ 0.82 & 0.84 & 0.78 & 0.64 & 0.84 & 0.84 & 0.84 & 0.53 \\ 0.82 & 0.96 & 0.78 & 0.64 & 0.94 & 0.96 & 0.84 & 0.53 \end{bmatrix}$$

6. The prediction of the number of correct answers is calculated as follows [7]:

$F(2018) = A^{17} \circ R,$

where $F(2018)$ – predicted variation of the number of correct answers in year 2018, $\circ$ – operation of max-min composition.

$$F(2018) = [0.28 \quad 0.45 \quad 0.74 \quad 0.99 \quad 0.86 \quad 0.55 \quad 0.34 \quad 0.22]$$

$$\circ \begin{bmatrix} 0.16 & 0.19 & 0.29 & 0.47 & 0.76 & 0.82 & 0.82 & 0.53 \\ 0.24 & 0.24 & 0.29 & 0.47 & 0.76 & 1 & 0.84 & 0.53 \\ 0.39 & 0.39 & 0.39 & 0.47 & 0.76 & 0.78 & 0.78 & 0.53 \\ 0.64 & 0.64 & 0.64 & 0.48 & 0.64 & 0.64 & 0.64 & 0.53 \\ 0.82 & 0.94 & 0.78 & 0.64 & 0.76 & 0.94 & 0.84 & 0.53 \\ 0.82 & 1 & 0.78 & 0.64 & 0.94 & 1 & 0.84 & 0.53 \\ 0.82 & 0.84 & 0.78 & 0.64 & 0.84 & 0.84 & 0.84 & 0.53 \\ 0.82 & 0.96 & 0.78 & 0.64 & 0.94 & 0.96 & 0.84 & 0.53 \end{bmatrix} =$$

$$= [0.82 \quad 0.86 \quad 0.78 \quad 0.64 \quad 0.76 \quad 0.86 \quad 0.84 \quad 0.53]$$

7. Transition from the obtained fuzzy values to quantitative ones (defuzzification) is performed according to [6], by the formula:

$$V(t) = \frac{\sum_{i=1}^{m} \mu_t(u_i) u_i^{mid}}{\sum_{i=1}^{m} \mu_t(u_i)}$$

where $\mu_t(u_i)$ – calculated membership functions for the specified year, u_i^{mid} – midpoints of the specified intervals.

The variation calculated by this formula *V*(2018)=-0.7266. If the obtained predicted value of the variation of the number of correct answers is added to the number of correct answers in the previous year, we obtain the predicted number of correct answers for the year under consideration:

$$N_{pred}^{2018} = N_{actual}^{2017} + V(2018) = 44.67 - 0.7266 = 43.9474$$

Погрешность метода рассчитывалась по формуле:

$$\delta(t) = \frac{\left|N_{actual}^t - N_{pred}^t\right|}{N_{actual}^t} \cdot 100\%$$

where N_{actual}^t – actual number of correct answers for the subject, N_{pred}^t – predicted number of correct answers.

The average error of the method calculated in this way was 12.83 %

The system interface looked as follows:

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

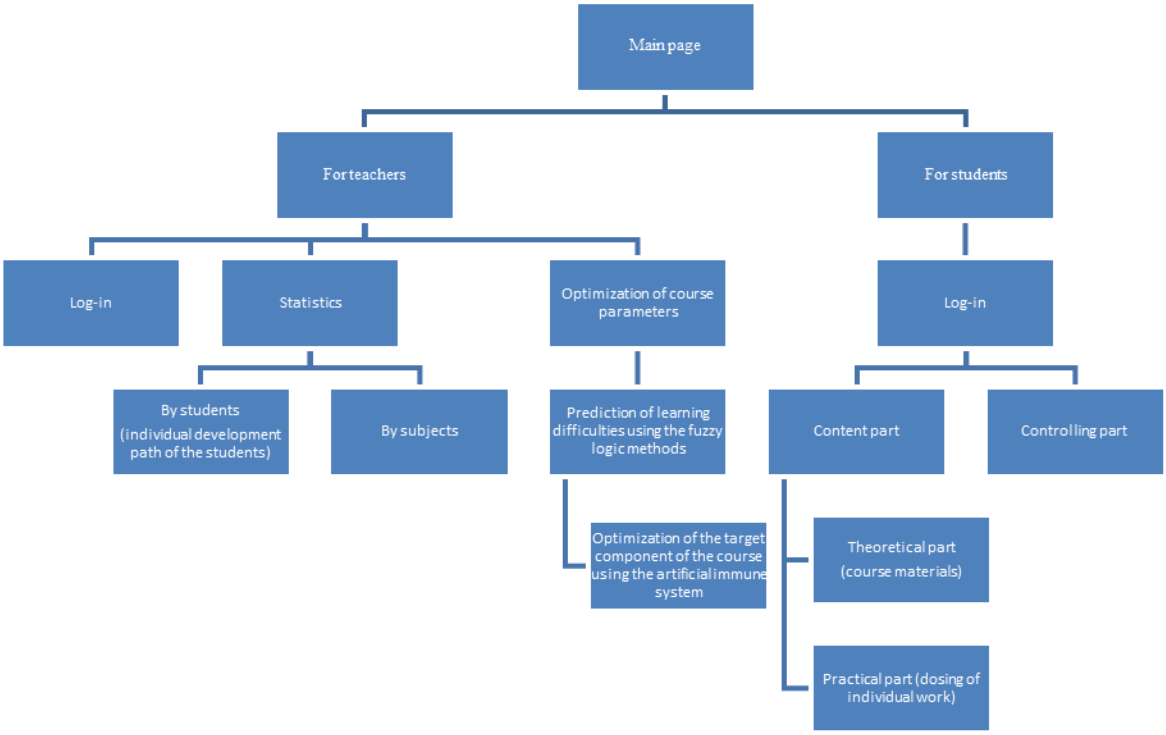

Figure 2 – Interface of a hybrid learning system

The user registered (logged in) as a teacher has access to the accumulated course statistics and can predict the material learning difficulties based on the fuzzy logic methods, and then optimize the course objectives.

IV. CONCLUSION

The present work describes a hybrid system used for teaching students and controlling their knowledge. The article considers a structural model of the learning and controlling hybrid system, which provides an opportunity to optimize the learning process using statistical information obtained in the course of the system's operation.

The system's peculiarity is the use of various intelligent technologies for its development. In order to predict the number of correct students' answers at the tests, fuzzy time series prediction methods have been used that allow drawing conclusions based on a small data quantity. The obtained prediction of the number of correct answers allowed identifying the course sections with the least number of correct answers that present the greatest difficulties for the students, which in turn allowed correcting the system's operation based on the data obtained.

V. REFERENCES

1. Batyrshin I.Z. Fuzzy hybrid systems. Theory and practice / edited by N.G. Iarushkina. – Moscow: FIZMATLIT. – 2007. – 208 pp.
2. Gorban A.N. Personal computer neural networks / A.N. Gorban, D.A. Rossiev. – Novosibirsk: Nauka, 1996. – 276 pp.
3. Borisov A. N. Linguistic variable-based decision models / A.N. Borisov, A.V. Alekseev, O. A. Krumberg, et al. – Riga: Zinatne, 1982. – 256 pp.
4. Iurkov N.K. Intelligent computer learning systems, monograph. Penza, Penza State University Publishing House, 2010.
5. Monakhov V.M. Technological bases for design and construction of educational process. – Volgograd: Peremena, 1995.
6. Zade L.A. The concept of a linguistic variable and its application to approximate solutions. - M.: Mir, 1976.
7. Mamedova M.G., Dzhabrailova Z.G. Application of fuzzy logic in demographic projection // Informatsionnyie Tekhnologii. – 2004. – № 3. – pp. 45-53.