IOP Conf. Series: Materials Science and Engineering **91** (2015) 012069 doi:10.1088/1757-899X/91/1/012069

Methods for control over learning individual trajectory

A A Mitsel [1], N V Cherniaeva [2]

[1] Professor, Doctor of Engineering Sciences,
Yurga Institute of Technology Tomsk Polytechnic University
[2] Postgraduate student, Tomsk Polytechnic University

e-mail: maa@asu.tusur.ru, nina.turalina@yandex.ru

Abstract. The article discusses models, methods and algorithms of determining student's optimal individual educational trajectory. A new method of controlling the learning trajectory has been developed as a dynamic model of learning trajectory control, which uses score assessment to construct a sequence of studied subjects.

1. Introduction

In connection with the transition of the educational system to a competence-oriented approach, the problem of learning outcomes assessment and creating an individual learning trajectory of a student has become relevant. Its solution requires the application of modern information technologies.

The third generation of Federal state educational standards of higher professional education (FSES HPE) defines the requirements for the results of mastering the basic educational programs (BEP). According to FSES HPE up to 50% of subjects have a variable character, i.e. depend on the choice of a student. It significantly influences on the results of developing various competencies.

The object of research is the process of formation of individual student's learning trajectory and the subject is dynamic model of formation of individual student's learning trajectory based on linear dynamic model of portfolio management of securities under constraints.

The problem of the research is development of a model (algorithm) to form individual learning student's trajectory. The aim of research work is to develop dynamic models of formation of individual student's learning trajectory.

Scientific novelty consist in adaptation of the linear dynamic model of portfolio management of securities under constraints to our domain to build individual student's learning trajectory.

The article discusses models, methods and algorithms of determining student's optimal individual educational trajectory. A new method of controlling the learning trajectory has been developed as a dynamic model of learning trajectory control, which uses score assessment to construct a sequence of studied subjects.

[1] A.A. Mitsel, Russia, Kemerovo region, Yurga, Leningradskaya street 26

[2] N.V. Cherniaeva, Russia, Kemerovo region, Yurga, Kirova street 27B, apt. 18

2. Approaches to solve the problem of formation and selection of student's individual learning trajectory

DSSs are often used when building individual learning path, because this task can be considered as a discrete multi-criteria problem, creating a significant burden on the decision maker (DMs). In the paper [1] three ways of representing data in the DSS were analyzed, including tables, heatmaps and parallel coordinate plots. The data in a DSS are represented both by objective and subjective indicators. In heatmaps each column represents the criterion, and each row represents an alternative. The cell color represents the significance of a criterion for a concrete solution. The criterion with low significance is indicated red, the medium one is indicated yellow, and the high one is indicated green. When using parallel coordinate plots alternatives are depicted as a line connecting points on the corresponding axes. All feasible solutions are superimposed on each other. Therefore, multiple criteria can be displayed, and reducing the load on the DMs.

In the experiment, the authors found out that the sequential structure of the spatial information representation facilitates decision-making provided that the amounts of data are large.

The work by G. Setlak [2] is of a significant interest for our research. It considered the Neuro-Fuzzy Method for Knowledge Modeling. The model is based on neural networks and artificial intelligence, which can be used to build fuzzy inference systems, aimed at forming an individual learning path of a student.

The fuzzy neural network, used for classification in this paper, has the following features:

- each neuron represents one fuzzy IF-THEN rule,
- the number of neurons equals to the number of rules in the base of rules,
- weights of the neurons have an interpretation concerning parameters of the membership functions of the corresponding neuro-fuzzy system.

In [3], the author compares various algorithms of curriculum formation: KBS, LS-Plan and IWT. For comparison the following criteria and metrics were used: Overall Effort metrics, Overall Acquired Knowledge metrics and Overall effort metrics. The curriculum was presented as an algorithm or graphs. It was stated that the LS-Plan has the longest distance of educational trajectories and the greatest number of errors, whereas the algorithm IWT generates the shortest path.

We can classify incoming data using the method of hierarchies analysis. An example of this classification is presented in [4] by J. Sivic, B. Russell et al.

Another way to classify data are Data Mining Techniques, successfully applied in [5] for the Student Recruitment Classification.

The following techniques are applied to profiling and classifying the student database:

- K-Means Clustering;
- J-48 Algorithm;
- LMT (Logistic Model Trees) Algorithm;
- Bayesian Network Algorithm.

The Bayes algorithm is often used for processing and analyzing the data. In [6] this method is used to classify text information, reducing uncertainty and ambiguity of natural language.

The proposed ensemble method is based on Bayesian Model Averaging, where both uncertainty and reliability of each single model are taken into account. The authors resolved the issue of classifier selection by proposing a "greedy" approach that evaluates the contribution of each model to the ensemble. Experimental results on "gold standard" datasets show that the proposed approach outperforms both traditional classification and ensemble methods.

The analysis of scientific works relevant for our research problem provided the foundation for developing our own dynamic limitation model to form student's individual learning trajectory.

3. The dynamic model for control over student's learning individual trajectory

Let us denote the number of disciplines by $N_t,\ t = 1,...,T$, student has to study per semester t . Here, T is the period of study (number of semesters).

As the results of the studying disciplines a student acquires many competencies. The structure of the model of graduate competence will be presented in the form of three levels, as written in FSES -3:

- The first level — competence;
- The second level — general cultural and professional competences;
- The third level — private competence.

The competence of the student can be evaluated on the basis of multiple estimates obtained by the student in the process of learning $N = \sum_{t=1}^{T} N_t$ disciplines of the chosen specialty.

We denote the valuation on subjects as $V_j(t),\ j = 1,...,N_t$, where N_t – the number of courses to be studied in the semester t in according to the curriculum.

Variables V_j can be estimated in points, for example on a 100-point scale. Here traditional assessment offered in the examination sheet, is determined by each University according to it own scale.

Integral score of student $V(t)$ at the time t equal $V(t) = \sum_{j=1}^{N_t} w_j V_j(t),\ t = 1,...,T$, where w_j – the importance weight of discipline.

The dynamics of the student's academic performance in discrete time will be described by the equation:

$$V_j(t+1) = \left(1 + \mu_j(t) + \eta_j(t)\right)\left(V_j(t) + u_j(t)\right),\ j = 1,...,N_t. \quad (1)$$

Here $\mu_j(t)$ is the average complexity of learning discipline j; $\eta_j(t)$ – the random component (deviation) of the learning discipline j complexity with parameters: $M\left(\eta_i(t)\right) = 0,\ M\left(\eta_i(t)\eta_k(t)\right) = \Sigma_{ik}(t),\ i,k = 1,...,n$, where $\Sigma_{ik}(t)$ – the covariance matrix of the development disciplines complexity. Values $\mu_j(t)$ are determined on the basis of semester evaluation's historical data. $u_j(t)$ – Scores in the discipline obtained during the semester ($u_j(t) > 0$), or penalty (minus) scores ($u_j(t) < 0$).

The complexity $\rho_{jk}(t)$ of the learning discipline j by the student k in this semester of the year t we define as $\rho_{jk}(t) = \dfrac{1}{V_{jk}(t)}$, where $V_{jk}(t)$ is the final score in the discipline in the semester t. Then, the average value μ_j and the covariance matrix Σ_{ij} are calculated by the formulas $\mu_j = \dfrac{1}{T_h \cdot N_g(t)} \sum_{k=1}^{N_g(t)} \sum_{t=1}^{T_h} \dfrac{1}{V_{kj}(t)}$, $\Sigma_{ij} = \dfrac{1}{T_h \cdot N_g(t) - 1} \sum_{k=1}^{N_g(t)} \sum_{t=1}^{T_h} \left(\rho_{ik}(t) - \mu_i\right)\left(\rho_{jk}(t) - \mu_j\right)$.

Here, T_h – historical horizon (quantity of years); t – sequence number of year; $N_g(t)$ – the number of students in the group in the year t.

Let us introduce "reference" integrated scoring and write the equation of the reference student, as follows:

$$V^0(t+1) = [1 + \mu_0(t)]V^0(t), \quad (2)$$

Where, $\mu_0(t)$ - the given complexity of the reference student (set by an expert).

The initial condition $V^0(0) = V(0) = 0$, i.e., in the initial moment of time the score of the reference student, as well as the scoring actual student equal to zero.

The task of managing the student's learning trajectory is to select disciplines and tasks on the basis of the outcome's estimates of the curriculum so that the generated learning trajectory is to follow the reference one on the horizon management T, where T – the period of time for which the student develops a program of the specialty.

Let us introduce the vector $y(t) = \left(V_1, ..., V_N\right)^T$ and vector $z(t) = \left(y(t),\ V^0(t)\right)^T$. Then, equations (1), (2) can be rewritten in the form

$$z(t+1) = A(t)z(t) + B(t)u(t), \tag{3}$$

Where, $A(t) = \overline{A}(t) + \widetilde{A}(t)$; $\overline{A}(t),\ \widetilde{A}(t)$ – diagonal matrix of dimension $(N+1)\times(N+1)$.

Matrix $B(t)$ with dimension $(N+1)\times N$ has the structure:

$$B(t) = \begin{pmatrix} A_{11}(t) & 0 & ... & 0 \\ 0 & A_{22}(t) & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & A_{NN}(t) \\ 0 & 0 & ... & 0 \end{pmatrix}.$$

As the objective function a linear functional is chosen

$$J = M\left\{\sum_{t=1}^{T-1}\left(V(t) - V^0(t)\right) + \sum_{t=0}^{T-1} b^T(t)\cdot u(t) + \left(V(T) - V^0(T)\right)\right\} \to \min_{u(t)},$$

Where, $b(t) = \left(\mu_1(t)d_{1t}, ..., \mu_N d_{Nt}(t)\right)^T$.

Using $z(t)$, let's rewrite $\left(V(t) - V^0(t)\right)$ in the form $\left(V(t) - V^0(t)\right) = Cz(t)$, where $C = \left(1,1,...,1,-1\right) \in R^{N+1}$. The quality criterion J takes the form

$$J = M\left\{\sum_{t=1}^{T-1} Cz(t) + \sum_{t=0}^{T-1} b^T(t)\cdot u(t) + Cz(T)\right\} \to \min_{u(t)}. \tag{4}$$

So, we have the optimal control task, in which the equation of state is described by a multistep process (3), and the functional of quality by expression (4). The control vector is given by $u(t)$. The problem is solved with the constraint $V(t) \geq V^0(t)$ or $C\cdot z(t) \geq 0$.

4. The restrictions of the task

The restriction related to the prohibition of penalty (minus) points, looks $u(t) \geq 0,\ t = 0, ..., T-1$.

Let us introduce restrictions on scoring disciplines

$$c_j^{\min} \leq y_j(t) + u_j(t) \leq c_j^{\max},\ j = 1, ..., N. \tag{5}$$

Here $c_j^{\min}$, $c_j^{\max}$ - the minimum and maximum number of points that a student can get to receive a grade in the gradebook ($c_j^{\min} = 55,\ c_j^{\max} = 100$).

In terms of $z(t)$ restrictions (5) take the form of a $c^{\min} \leq z(t) + Y(t)u(t) \leq c^{\max},\ t = 0, ..., T-1,$

where Y – matrix with dimension $(N+1)\times N$ with elements $Y=\begin{pmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \\ 0 & 0 & ... & 0 \end{pmatrix}$.

The restriction of the semester discipline complexity can be viewed as

$$\frac{M_{\min}^{g}(t)}{N_t} \le \frac{1}{100N_t}\sum_{j=1}^{N} k_j d_{jt} y_j(t) \le \frac{M_{\max}^{g}(t)}{N_t},\ t=1,...,T, \tag{6}$$

Where, k_j – the number of credits in the discipline j;

$N_t=\sum_{j=1}^{N} d_{jt}$ – the number of disciplines studied in semester t.

$M_{\min}^{g}(t),\ M_{\max}^{g}(t)$ – the minimum and maximum values of credits on subjects defined by the curriculum. For example, for a bachelor degree in technical specialties humanities, social and economic cycles are allocated ($g=1$), mathematical and natural-science cycle ($g=2$) and professional cycle ($g=3$). In each of these cycles, as well as in FSES -3 the minimum and maximum number of credits are determined.

Let us introduce the vector $D(t)=\frac{1}{100N_t}\left(d_{1,t}k_1, d_{2,t}k_2,...,d_{N,t}k_N;0\right)$.

Then the restriction (6) takes the form $\frac{M_{\min}^{g}(t)}{N_t} \le D(t)z(t) \le \frac{M_{\max}^{g}(t)}{N_t},\ t=1,...,T$.

For accounting courses-prerequisites we will enter the coefficients r_{ij}. These coefficients are called the coefficients of the narrowness of interdisciplinary communication. The coefficients r_{ij} will be equal to 0 or 1. The coefficients $r_{ij}=1$, if for studying of discipline j you must learn discipline with number i and $r_{ij}=0$ — vice versa. For accounting disciplines-co requisites we introduce a coefficient f_{ij}, which is equal to 1 if discipline i will be used further to study discipline j, and $f_{ij}=0$ — vice versa. Using these coefficients it is determined by the order of the studies, which specifies the individual educational plan. If $r_{ij}=0$ and $f_{ij}=0$, then there may be a parallel study of subjects.

Restrictions related to the procedure of study subjects in each semester t, we write in the form

$$\begin{aligned} N_t c_k^{\min} &\le \sum_{i=1}^{N} r_{ik} d_{it} y_i(t) \le c_k^{\max} N_t,\ k=1,...,N_t; \\ N_t c_i^{\min} &\le \sum_{k=1}^{N} f_{ik} d_{kt} y_k(t) \le c_i^{\max} N_t,\ i=1,...,N_t. \end{aligned} \tag{7}$$

Here $c_k^{\min}, c_k^{\max}$ – the minimum number of scores allowed in the discipline. If you restrict your studies during the entire period of study T is not attributed to a specific semester, then instead of (7) we will have

$$Nc_k^{\min} \le \sum_{i=1}^{N}\sum_{t=1}^{T} r_{ik} d_{it} y_i(t) \le c_k^{\max} N,\ k = 1,...,N;$$
$$Nc_i^{\min} \le \sum_{k=1}^{N}\sum_{t=1}^{T} f_{ik} d_{kt} y_k(t) \le c_i^{\max} N,\ i = 1,...,N. \quad (8)$$

We introduce the vectors: $R^k(t) = \left(r_{1,k}d_{1,t}, r_{2,k}d_{2,t},...,r_{N,k}d_{N,t},0\right)$, $F^i(t) = \left(f_{i,1}d_{1,t}, f_{i,2}d_{2,t}..., f_{i,N}d_{N,t},0\right)$ and we can rewrite the constraints (7) in terms of $z(t)$. Get: $N_t c_k^{\min} \le R^k(t)z(t) \le c_k^{\max} N_t,\ k = 1,...,N$, $N_t c_i^{\min} \le F^i(t)z(t) \le c_i^{\max} N_t,\ i = 1,...,N$.

Restrictions (8) take the form of a: $Nc_k^{\min} \le \sum_{t=1}^{T} R^k(t)z(t) \le c_k^{\max} N,\ k = 1,...,N$, $Nc_i^{\min} \le \sum_{t=1}^{T} F^i(t)z(t) \le c_i^{\max} N_t,\ i = 1,...,N.$

For solving the problem of tracking, we must specify the initial state of the system $z(0) = \begin{pmatrix} y(0) \\ V^0(0) \end{pmatrix}$. The initial scores of real and reference student are considered equal to zero $V(0) = V^0(0) = 0$.

So, finally we will state the problem for control of an individual training trajectory:

$$J = M\left\{\sum_{t=1}^{T-1} Cz(t) + \sum_{t=0}^{T-1} b^T(t)\cdot u(t) + Cz(T)\right\} \to \min_{u(t)}, \quad (9)$$

$$z(t+1) = A(t)z(t) + B(t)u(t), \quad (10)$$

$$\begin{aligned}
& C\cdot z(t) \ge 0,\\
& c^{\min} \le z(t) + Yu(t) \le c^{\max},\ t = 0,...,T-1;\\
& N_t c_k^{\min} \le R^k(t)z(t) \le c_k^{\max} N_t,\ k = 1,...,N;\\
& N_t c_i^{\min} \le F^i(t)z(t) \le c_i^{\max} N_t,\ i = 1,...,N;\\
& \frac{M^g_{\min}(t)}{N_t} \le D(t)z(t) \le \frac{M^g_{\max}(t)}{N_t},\ t = 1,...,T;\\
& u(t) \ge 0,\ t = 0,1,...,T-1
\end{aligned} \quad (11)$$

5. Solution of the problem

Task (9) - (11) is a linear dynamic programming task. It can be solved by Bellman method or converted to an equivalent model of linear programming, for which a reliable solution methods are developed. To solve large-scale problems, you can use the control method with the predictable model.

6. Conclusion

Today, many of the issues in the domain have been solved, but still there are contradictions in the application of the proposed models in practice, i.e. how formed individual learning trajectory meets the interests and need of a student as well as the requirements of the Federal State education standards.

The problem of forming student's learning trajectory is analyzed in general and the choice of an individual direction was studied in details. Various methods, models and algorithms of the student's individual learning trajectory formation were described.

A new model of management learning trajectory based on dynamic models for tracking the reference trajectory was developed.

References

[1] Gettinger J Kiesling E Stummer C Vetschera R A comparison of representations for discrete multi-criteria decision problems *Decision Support Systems* 54 (2013) pp 976–985
[2] Setlak J G The fuzzy-neuro classifier for decision support *Information Theories & Applications* 15 (2008) pp 21-26
[3] Limongelli C Sciarrone F Temperini M Vaste G Comparing Curriculum Sequencing Algorithms for Intelligent Adaptive (e)-Learning *Conference ICL2010* September 15 -17 2010 Hasselt Belgium pp 1013 – 1020
[4] Sivic J Russell B Zisserman A Freeman W and Efros A Unsupervised discovery of visual object class hierarchies *IEEE Conference on Computer Vision and Pattern Recognition* pp 1–8 2008
[5] Kumar A Singhal S Dhankher P Gulia A Student Recruitment Classification using *Data Mining Techniques IJCSMS (International Journal of Computer Science &Management Studies)* Vol 14 Issue 08 (August 2014) pp 24-31
[6] Fersini E Messina E Pozzi F A Sentiment analysis: Bayesian Ensemble Learning *Decision Support Systems* 68 (2014) pp 26–38