# EDULEARN15

7TH INTERNATIONAL CONFERENCE ON EDUCATION AND NEW LEARNING TECHNOLOGIES

BARCELONA (SPAIN)
6TH - 8TH OF JULY, 2015

# CONFERENCE PROCEEDINGS

# EDULEARN15

**7TH INTERNATIONAL CONFERENCE ON EDUCATION AND NEW LEARNING TECHNOLOGIES**

**BARCELONA (SPAIN)**
**6TH - 8TH OF JULY, 2015**

# CONFERENCE PROCEEDINGS

**Published by**
IATED Academy
www.iated.org

**EDULEARN15 Proceedings**
7th International Conference on Education and New Learning Technologies
July 6th-8th, 2015 — Barcelona, Spain

**Edited by**
L. Gómez Chova, A. López Martínez, I. Candel Torres
IATED Academy

**ISBN: 978-84-606-8243-1**
**ISSN: 2340-1117**
**Depósito Legal: V-1653-2015**

Book cover designed by
J.L. Bernat

# THE AGENT-BASED MODELING OF THE ENROLLMENT CAMPAIGN TO RUSSIAN UNIVERSITIES

**Mikhail Nasadkin[1], Eugene Pitukhin[1], Marina Astafyeva[2]**

*[1] Petrozavodsk State University (RUSSIAN FEDERATION)*
*[2] Branch of Dubna International University for Nature, Society and Man in Protvino (RUSSIAN FEDERATION)*

## Abstract

At the present paper we consider the construction of a model of the entrants' behavior for matriculation to institutions of higher education in Russia. All taken into account factors of environment and personal preferences of entrants are indicated. These factors are divisible into 2 groups: deterministic factors of educational environment and stochastic properties of virtual entrants. Utility functions of entry of an applicant to some certain institution on some certain specialty are proposed. All the algorithm of modeling is considered. This paper is of introductory nature: it contains methodology, but not results of modeling.

Keywords: agent-based modeling, modeling of the enrollment campaign, factors of university choice.

## 1 INTRODUCTION

At present time in Russia much attention is paid to career building problems by executive authorities, parents and youth themselves. This is facilitated by the increasing amount of available information on education quality and career building prospects for graduates of Russian universities. From this perspective it is interesting to perform a research on the entrants' reaction to educational environment changing conditions.

Suchlike problems were considered in the following papers: S. Kiselgof "College entrants' choice with quadratic utility functions" [1], I. Prakhov "A Model of College Choice in the Context of the Unified State Examination and the Role of Students' Expectations" [2]. Hierarchical model of applicants choice of university was considered in Ivashko A.A., Konovalchikova E.N., Mazalov V.V. "Game-theory hierarchical choice models" [3].

All investigated papers apply to famous papers of D. Gale, L. S. Shapley "College Admissions and the Stability of Marriaga" [4], M. Balinski, T. Sonmez "A Tale of Two Mechanisms: Student Placement" [5] as a basis.

Besides, we can include following papers as noteworthy examples of sources about factors, which influence on entrants' choice of university: E.A. Morozova "Improving the management of the regional higher education institutions on basis of the students opinion's analysis" [6], I.V. Abankina and others "Multi-stage choice model for forecasting the demand for higher education" [7]. Attraction factors of regions for entrants were investigated in E.A. Pitukhin "Analysis of inter-regional mobility of school-leavers entering to the universities" [8].

This paper describes using of agent-based simulating for solving a problem of modeling of the enrollment campaign to Russian universities. We can emphasize following aspects: methodology of modeling; set of factors which influence on entrants' decision; set of environmental factors; utility functions of entry of an applicant to some certain institution on some certain specialty; whole algorithm of modeling and prospects of development.

## 2 PROCEDURE OF MODELING

Agent-based approach to modeling of enrollment campaign to Russian universities supposes realizations of following steps:

- Generation of applicants population
- Placing created applicants to prepared educational environment
- Calculating utility functions for every applicant


*ISBN: 978-84-606-8243-1*

- Applicants submission to several universities
- Conclusive choice of certain university

Hereafter we shall take a closer look at each of the stages of modeling.

## 2.1 Generation of entrants population

Every applicant is characterized by such groups of properties like: scores of United State Examination (USE); family financial situation; the applicant's desire to study for free; interest to various education specialties; significance of generally accessible and published indexes of universities and regions; place of residence; attitude of applicant to move to other region or town.

A detailed list of the factors affecting the entrant's choice of university is presented below:

- Set of exams (USE), which entrant takes
- Scores of the exams which form index $E_{i,j}$– summary exams score for entrance on education specialty $j$
- $P_1^i$ – amount of money which applicant's $i$ family is in position to spend on entrant`s study per year
- $P_2^i$ – desire of applicant $i$ to study for free
- $P_3^{i,j}$ – interest of applicant $i$ to educational specialties $j$
- $P_4^{i,edu}$ – significance of index of university educational activity for applicant i $i$
- $P_4^{i,sci}$ – significance of index of university science activity for applicant $i$
- $P_4^{i,inf}$ – significance of index of university infrastructure for applicant $i$
- $P_4^{i,int}$ – significance of index of university international activity for applicant $i$
- $P_4^{i,fin}$ – significance of index of university financial activity for applicant $i$
- $P_4^{i,br}$ – significance of index of university "brand" for applicant $i$
- $P_4^{i,emp}$– significance of index of university graduates employment by profession for applicant $i$
- $P_5^{i,sal}$ – significance of index of region average salary for applicant $i$
- $P_5^{i,unemp}$ – significance of index of region unemployment for applicant $i$
- $P_5^{i,need}$ – significance of index of graduates' being in demand on region's labor market for applicant $i$
- $P_5^{i,cult}$ – significance of index of region cultural level for applicant $i$
- $P_5^{i,crim}$ – significance of index of region criminality level for applicant $i$
- $P_6^{i,k}$ – property describing attitude of applicant $i$ to move to region $s(k)$, where university $k$ is situated

Certain values of each generated applicant properties are calculated through instrumentality of realization of stochastic variable. These stochastic variables are defined with priori fitted distribution functions $F(x, p_0)$, where $p_0$ – vector of distribution parameters, depending on certain distribution type. Afterwards, distribution parameters are chosen with mechanism of verification of model adequacy on retrospective data.

## 2.2 Placing created applicants to prepared educational environment

Created applicants population is placed to educational environment in accordance with realistic distribution of schools graduates by the country. Each applicant is given with actual location. External educational environment is given by set of deterministic properties, which are listed below:

- $B_{j,k}(year-1)$ – "pass" sum of USE scores last year (relative to modeling year) to university $k$ on specialty $j$
- Set of exams (USE) necessary for matriculation on every specialty $j$
- $V_1^{j,k}$ – cost of educating on specialty $j$ in university $k$ per year
- $V_2^{k,edu}$ – index of educational activity of university $k$
- $V_2^{k,sci}$ – index of science activity of university $k$
- $V_2^{k,inf}$ – index of infrastructure of university $k$
- $V_2^{k,int}$ – index of international activity of university $k$
- $V_2^{k,fin}$ – index of financial activity of university $k$
- $V_2^{k,j,emp}$ – index of graduates employment of university $k$ by specialty $j$
- $V_2^{k,br}$ – index of "brand" of university $k$
- $R_1^{s(k),sal}$ – average salary in region $s(k)$ where university $k$ is situated
- $R_1^{s(k),unemp}$ – index of unemployment in region $s(k)$
- $R_1^{s(k),cult}$ – index of cultural level in region $s(k)$
- $R_1^{s(k),crim}$ – index of criminality level in region $s(k)$
- $R_1^{s(k),j,need}$ – index of graduates' being in demand on labor market by specialty $j$ of region $s(k)$

## 2.3 Calculating utility functions of matriculation

Next stage of modeling process is that created agents-applicants (having set of properties and placed to external educational environment) must come arrive to a decision on pairs "university-specialty" which fit them to the fullest extent. We will determine utility of university matriculation consisting of 4 components:

- Correspondence of USE scores of applicant $i$ to "passing" scores in previous years by specialty $j$ in university $k$ will be denoted with $F_{i,j,k}^1$
- Interest of applicant $i$ to different specialties $j$ will be denoted with $F_{i,j}^2$
- Significance of university $k$ factors for applicant $i$ studying on specialty $j$ will be denoted with $F_{i,j,k}^3$
- Significance of region $s(k)$ (where university $k$ is situated) factors for applicant $i$ studying on specialty $j$ will be denoted with $F_{i,j,k}^4$

Then, applicant choice can be made with finding such pairs $j,k$, which make some function (linear combination of components or multiplicative function)

$$f(i,j,k) = f\left(F_{i,j,k}^1, F_{i,j}^2, F_{i,j,k}^3, F_{i,j,k}^4\right)$$

take its maximal value. It bears reminding that applicants choose several pairs university-specialty on this stage. Conclusive choice will be carried out on later stage of modeling, when applicants are familiar with information about contest (number of applicants claiming one state-funded place). This information helps applicants to estimate probabilities of success.

### *2.3.1 Component $F_{i,j,k}^1$*

General property influencing on applicants decision is difference between USE scores (sum of all exams scores divided by number of exams) $E_{i,j}$ and "pass" exams scores of previous year

$B_{j,k}(year-1)$. Each pair $<j,k>$ is available only if applicant has passed all exams necessary for matriculation on specialty $j$.

Besides, assume that applicant tries to matriculate to university if "pass" scores of previous year are lower than his scores of USE

$$E_{i,j} - B_{j,k}(year-1) > 0$$

or amount of money, which applicant's family is is position to spend on study per year is higher than cost of education in chosen university:

$$P_1^i - V_1^{j,k} > 0.$$

Let $J_i$ is a set of educational specialties available for applicant $i$ due to set of exams which applicant passed. Then we can write that applicant can choose only such pairs $<j,k>$, which satisfy following condition:

$$\begin{cases} j \in J_i \\ I\{E_{i,j} - B_{j,k}(year-1) > 0\} + [1 - I\{E_{i,j} - B_{j,k}(year-1) > 0\}] \cdot I\left\{P_1^i - V_1^{j,k} > 0\right\} > 0 \end{cases}$$

Here $I\{E_{i,j} - B_{j,k}(year-1) > 0\}$ – is an indicator function, characterizing that applicant can matriculate and study for free, $I\left\{P_1^i - V_1^{j,k} > 0\right\}$ – indicator function characterizing that financial standing of family allows cover expenses of fee-paying education.

Besides, we should take into account, that applicant may have a possibility to pay for education, but ultimately doesn't want to make it. We can imagine backhanded situation: applicant can matriculate for free in native town or region, but moves to metropolitan city to study in prestige university for money. Mathematically we shall describe such behavior with applicants' property $P_2^i$. Assume $P_2^i = 1$ means that applicant unambiguously wants to study for free, $P_2^i = 0$ means that it is all one to applicant. All other values between 0 and 1 describe attitude to study for free.

Finally, we will calculate $\mathrm{F}^1_{\mathrm{i,j,k}}$ as following:

$$\mathrm{F}^1_{\mathrm{i,j,k}} = \begin{cases} 0, & j \notin J_i \\ I\{E_{i,j} - B_{j,k}(year-1) > 0\} + [1 - I\{E_{i,j} - B_{j,k}(year-1) > 0\}] \cdot I\{P_1^i - V_1^{j,k} > 0\} \cdot P_2^i, & j \in J_i \end{cases}$$

### *2.3.2 Component $F^2_{i,j}$*

Property $F^2_{i,j}$ covers the interest of applicant $i$ to matriculation on specialty $j$, thus it is written easily with provision for possibility of entry only on specialties $j \in J_i$

$$F^2_{i,j} = \begin{cases} 0, & j \notin J_i \\ P_3^{i,j}, & j \in J_i \end{cases},$$

where $P_3^{i,j}$ – stochastic variables, which satisfy following condition: $\sum_{j \in J_i} P_3^{i,j} = 1$.

### *2.3.3 Component $F^3_{i,j,k}$*

Component $F^3_{i,j,k}$ covers the significance of university factors for applicant. Properties of Russian universities activity efficiency (educational activity, science activity, infrastructure, international activity, financial activity, "brand" of university and graduates' employment) are published [9] by the Ministry of Education and Science of the Russian Federation. For convenience let $M = \{edu, sci, int, inf, fin, br\}$. Then we can write $F^3_{i,j,k}$ a following linear combination

$$F^3_{i,j,k} = \sum_{m \in M} P_4^{i,m} \cdot V_2^{k,m} + P_4^{i,emp} \cdot V_2^{k,j,emp}.$$

In this formula summand, which carries responsibility for graduates' employment, is detached: this property is used by specialties, whereas all other properties are used by universities. Indexes of significance of factors $P_4^{i,m}$ satisfy condition of normalizing

$$\sum_{m \in M} P_4^{i,m} + P_4^{i,emp} = 1.$$

### 2.3.4 Component $F_{i,j,k}^4$

Component $F_{i,j,k}^4$ covers influence of region $s(k)$ factors and is calculated in the same manner as previous component. Assume $N = \{sal, unemp, cult, crim\}$ for convenience, we write formula for calculating component $F_{i,j,k}^4$. Only difference here is in the fact that every applicant in model has different attitude to move to other region. Mathematically it is expressed with existence of property $P_6^{i,k}$. This property takes value equal 1, if the region $s(k)$ is most preferable for applicant

$$F_{i,j,k}^4 = \left( \sum_{n \in N} P_5^{i,n} \cdot R_1^{s(k),n} + P_5^{i,need} \cdot R_1^{s(k),j,need} \right) \cdot P_6^{i,k}.$$

Condition of normalizing of significance properties $P_5^{i,n}$ is written in the same manner

$$\sum_{n \in N} P_5^{i,n} + P_5^{i,need} = 1.$$

## 2.4 Applicants submission to several universities and conclusive choice

On the basis of calculating utility functions every applicant draws up sorted lists of pairs "university-specialty" $< j, k >$. Agents submit applications to some random (less than 15) pairs of university and specialty, which have best utility function values.

Universities publish all information about applicants who submitted documents: that are lists of applicants sorted by exams scores and number of state-funded places and places for fee. Learning this information and utility functions for all attractive pairs "university-specialty" (which they have chosen of previous stage) every applicant in model can estimate probabilities of successful matriculation. Table 1 shows an example of information which should be analyzed by every agent-applicant in model.

Table 1

| University | Specialty | Utility function value | Estimate of successful matriculation probability |
|---|---|---|---|
| U1 | Spec1 | 0,954 | 0,93 |
| U2 | Spec2 | 0,932 | 0,87 |
| U1 | Spec3 | 0,911 | 0,95 |
| … | … | … | … |

Most confident applicants (having large probabilities of success) submit originals of applications to universities, which are best for them. Thereby these applicants make their conclusive choice of university and don't take part in further modeling. All other applicants (not satisfying criteria of confidence in matriculation) get refreshed information about free state-funded places in universities and recalculate probabilities of successful matriculation.

Such iteration procedure continuous till all state-funded places and places for fee become occupied by applicants.

## 3 VERIFICATION OF MODEL ADEQUACY AND PARAMETERS IDENTIFICATION

During the modeling a population of applicants is generated. Each of this created agents-applicants has own stochastic properties, but every of these properties is a realization of some stochastic variable with distribution function with parameters. We shall identify these parameters as initial and denote them as $p_1^0, p_2^0, \dots, p_l^0$.

We should identify something like a result of modeling: let it be contest (number of applicants claiming for one state-funded place) in each university $k$ on each specialty $j$. Let us denote contest as $\widehat{C_{j,k}}(p_1^0, p_2^0, \dots, p_l^0)$, because it depends on parameters we used to generate a population of applicants.

To verify adequacy of a constructed model it is enough to carry out modeling on retrospective data and compare real contest (which was fixed on retrospective) $C_{j,k}$ with modeled contest $\widehat{C_{j,k}}(p_1^0, p_2^0, \dots, p_l^0)$.

This way we handle identification of model parameters – parameters of distribution functions $p_1^0, p_2^0, \dots, p_l^0$. It is realized by carrying out long series of simulations with different sets of parameters with least square method

$$\sum_{j,k} \left(\widehat{C_{j,k}}(p_1, p_2, \dots, p_l) - C_{j,k}\right)^2 \to min.$$

Model may be denoted as adequate when sum of modules of differences between real and modeling contest satisfy some criteria ($C_{cr} = const$)

$$\sum_{j,k} \left|\widehat{C_{j,k}}(p_1, p_2, \dots, p_l) - C_{j,k}\right| < C_{cr}.$$

## 4 CONCLUSION

This paper covers the modeling of enrollment campaign to Russian universities methodology, which draws upon agent-based modeling. Advantages of such approach to solving considered problem consist of following:

- Maximum extent feasible model - degree of circumstantiation is bounded only by computers resources
- There are no difficult formulas which describe different streams of applicants like in system dynamics. Everything is described with easy rules of applicants thinking processes. Same time it is possible to get not obvious results on aggregate level.
- Modeling of many social or economic processes often is a problem with nonlinear behavior of agents. Such models are filled with rules “if-then-else”. Sometimes such aspects hardly can be described with difference equations with continuous or discrete time. Agent-based modeling allows to perform very versatile adjustment of a system.

Constructed model of enrollment campaign can be used as by each of universities and public authorities in the field of education. It allows to forecast distributions of applicants of oncoming years by universities, reactions of applicants on some changes of educational environment in country like closing or creation of new universities or specialties, changing of social and financial properties of regions and universities. Presented model can be useful in generation of management decisions in field of higher education.

Besides, it would be interesting to use presented methodology of calculating matriculation utility functions for applicants and their families. It can be expressed by possible creating Internet resource, oriented on preparation matriculation recommends adjusted for personal preferences of applicant (all properties presented in this paper).

## REFERENCES


[1] Kiselgof, S. College entrants' choice with quadratic utility functions: Working paper WP7/2011/01 [Text] / S. Kiselgof ; Higher School of Economics. – Moscow : Publishing House of the Higher School of Economics, 2011. – 44 p.

[2] Prakhov, I. A Model of College Choice in the Context of the Unified State Examination and the Role of Students' Expectations : Working paper WP10/2010/06 [Text] / I. Prakhov

[3] Ivashko A.A., Konovalchikova E.N., Mazalov V.V. Game-theory hierarchical choice models. // Transactions of XIII All-Russian Conference on Governance of VSPU-2014, 16-19 June 2014. Moscow: IPU RAN, 2014. P. 8308-8313.

[4] D Gale, L.S. Shapley. College Admissions and the Stability of Marriage. // The American Mathematical Monthly, Vol. 69, No. 1 (Jan., 1962), 9-15.

[5] M. Balinski, T. Sonmez. A Tale of Two Mechanisms: Student Placement. // Journal of Economic Theory 84, 73-94 (1999).

[6] E.A. Morozova. Improving the management of the regional higher education institutions on basis of the students opinions' analysis. [Text] / E.A.Morozova // University management: practice and analysis. – 2014. - № 6. – p. 102-109.

[7] I.V. Abankina. Multi-stage choice model for forecasting the demand for higher education. [Text] / Abankina I.V., Abankina T.V., Alekserov F.T., Derkachev P.V., Egorova L.G., Zinkovsky K.V., Nikolaenko E.A., Ogorodniychuk D.L., Seroshtan E.S., Filatova L.M. // University management: practice and analysis. – 2014. - № 4-5. – p. 84-94.

[8] E.A. Pitukhin. Analysis of inter-regional mobility of school-leavers entering to the universities. [Text] / Pitukhin E.A., Semenov A.A. // University management: practice and analysis. – 2011. - № 3. – p. 82-89.

[9] Informational and analytical materials according to the analysis of performance indicators of educational institutions of higher education in Russia (n.d.). Retrieved from http://miccedu.ru/monitoring/2014/