

УДК 004.94

АГЕНТНОЕ МОДЕЛИРОВАНИЕ ПОВЕДЕНИЯ АБИТУРИЕНТОВ ПРИ ВЫБОРЕ ВУЗА В РОССИИ

¹Насадкин М.Ю., ¹Питухин Е.А., ²Астафьева М.П.

¹*Петрозаводский государственный университет, Петрозаводск,
e-mail: mnasad@petsru.ru, eugene@petsru.ru;*

²*Филиал «Протвино» Международного университета природы, общества и человека «Дубна»,
Протвино, e-mail: marines2000@mail.ru*

В статье рассматривается построение агентной модели поведения абитуриентов при выборе вуза России. Приведены все учитываемые факторы внешней среды и личных предпочтений абитуриентов, которые можно разделить на две группы: детерминированные факторы среды и стохастические показатели виртуальных абитуриентов. Предлагается схема расчета функций полезности от поступления абитуриента в конкретный вуз на определенную специальность. Предложена итерационная процедура моделирования подачи оригиналов документов в выбранный вуз, основывающаяся на критериях уверенности абитуриентов в поступлении. Рассматривается вся схема моделирования. Указаны способы проверки адекватности построенной модели на ретроспективных данных, заключающиеся в рассмотрении полученного при моделировании конкурса в вузы в качестве моделируемого параметра. Статья носит ознакомительный характер: рассматривается методология без приведения результатов моделирования.

Ключевые слова: агентное моделирование, моделирование приемной кампании, факторы выбора вузов

THE AGENT-BASED MODELING OF ENTRANTS' UNIVERSITY CHOICE IN RUSSIA

¹Nasadkin M.Y., ¹Pitukhin E.A., ²Astafeva M.P.

¹*Petrozavodsk State University, Petrozavodsk,
e-mail: mnasad@petsru.ru, eugene@petsru.ru;*

²*Branch of International University for Nature, Society and Man «Dubna»,
Protvino, e-mail: marines2000@mail.ru*

In the present paper we consider the construction of a model of the entrants' behavior for admission to institutions of higher education in Russia. The agent-based modeling is taken as a methodology of computer simulation. Each entrant is presented by virtual agent, which has a set of stochastic properties. All factors can be divided by two groups: deterministic factors of educational environment and stochastic factors of entrants. We suggest the schema of calculating utility functions for each virtual entrant which determine prospective pairs «university-specialty» and an iterate procedure of matriculation. The procedure is based on confidence criteria. Each entrant evaluates probability of successful matriculation using knowledge on sorted lists of all entrants. Full schema of modeling is presented in paper. Adequacy of model may be verified on retrospective data if we use contest (number of entrants claiming one state-funded place) as a result property. The paper's character is introductory: we present methodology but not results of modeling.

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

В настоящее время в России усилилось внимание общества к вопросам построения карьеры, как со стороны исполнительной власти, так и со стороны родителей и самой молодежи. Этому способствует увеличение количества доступной информации о качестве образования в российских вузах и возможных перспективах построения карьеры. С этой точки зрения интересным представляется исследование того, как абитуриенты реагируют на характеристики и изменения внешней образовательной среды.

Схожие задачи по моделированию выбора вузов абитуриентами решались в работах С.А. Кисельгоф «Выбор вуза абитуриентами с квадратичной функцией полезности» [3], И.А. Прахова «Модель вы-

бора вуза в условиях ЕГЭ и роль ожиданий абитуриентов» [5] и др. Во всех изученных работах в качестве методологии в основу брались известные результаты D. Gale, L.S. Shapley «College Admissions and the Stability of Marriage» [6]. Кроме того, укажем и ряд других стоящих внимания работ об источниках информации о влияющих на выбор абитуриента факторах. И.В. Абанкина и другие в работе «Модель многоступенчатого выбора для прогнозирования поведения спроса на высшее образование» [1] рассматривали некоторые из факторов выбора вузов. Факторы образовательной привлекательности регионов для абитуриентов рассматривались в работе Е.А. Питухина «Анализ межрегиональной мобильности

выпускников школ при поступлении в высшие учебные заведения» [4].

В данной статье для решения задачи выбора вуза применяется агентное моделирование. Рассматривается методология моделирования; набор факторов, влияющих на решения абитуриентов; набор факторов внешней среды; вводятся функции полезности от поступления конкретного абитуриента на специальность конкретного вуза; указывается весь алгоритм проведения моделирования и перспективы развития.

Схема моделирования

Генерация популяции абитуриентов

Каждый из абитуриентов характеризуется такими группами показателей, как баллы Единого Государственного Экзамена (ЕГЭ); финансовое положение семьи; стремление абитуриента обучаться бесплатно; интерес абитуриента к различным образовательным специальностям; значимость общедоступных и публикуемых в России показателей вузов и региона; отношение выпускника к переезду в другой регион. Ниже представлен подробный список факторов, влияющих на выбор вуза абитуриентом: набор экзаменов (ЕГЭ), которые сдает абитуриент; баллы ЕГЭ за каждый экзамен и показатель E_{ij} – суммарный балл ЕГЭ для поступления на специальность j ; P_1^i – сумма, которую семья абитуриента i готова тратить на обучение в год; P_2^i – стремление абитуриента i обучаться бесплатно; $P_3^{i,j}$ – интерес абитуриента i к специальности j ; $P_4^{i,edu}$ – значимость показателя образовательной деятельности вуза для абитуриента i ; $P_4^{i,sci}$ – значимость показателя научно-исследовательской деятельности вуза для абитуриента i ; $P_4^{i,inf}$ – значимость показателя инфраструктуры вуза для абитуриента i ; $P_4^{i,int}$ – значимость показателя международной деятельности вуза для абитуриента i ; $P_4^{i,fin}$ – значимость показателя финансово-экономической деятельности для абитуриента i ; $P_4^{i,br}$ – значимость показателя бренда вуза для абитуриента i ; $P_4^{i,emp}$ – значимость показателя трудоустройства по специальности для абитуриента i ; $P_5^{i,sal}$ – значимость показателя средней заработной платы в регионе для абитуриента i ; $P_5^{i,unemp}$ – значимость показателя безработицы в регионе для абитуриента i ; $P_5^{i,need}$ – значимость востребованности на рынке труда в регионе для абитуриента i ; $P_5^{i,cult}$ – значи-

мость показателя культурного уровня в регионе для абитуриента i ; $P_5^{i,crim}$ – значимость показателя уровня преступности в регионе для абитуриента i ; $P_6^{i,k}$ – показатель, описывающий стремление абитуриента i к переезду в регион $s(k)$, в котором находится вуз k .

Конкретные значения показателей каждого из генерируемых абитуриентов – агентов формируются с помощью реализации случайных величин. Эти случайные величины описываются с помощью заранее подобранных функций распределения $F(x, p_0)$, где p_0 – вектор параметров распределения, зависящий от конкретного вида распределения. Впоследствии параметры распределений подбираются с помощью проверки адекватности модели на ретроспективных данных.

Помещение абитуриентов во внешнюю среду

Созданная популяция абитуриентов помещается во внешнюю среду в соответствии с реальным распределением. Каждому абитуриенту задается его фактическое местоположение. Внешняя среда задается набором детерминированных показателей, которые представлены ниже: $B_{j,k}^{(year-1)}$ – проходной балл ЕГЭ прошлого по отношению к моделируемому году в вуз k по специальности j ; набор экзаменов (ЕГЭ), необходимых для поступления на каждую из специальностей j ; $V_1^{j,k}$ – стоимость обучения по специальности j в вузе k за год; $V_2^{k,edu}$ – показатель образовательной деятельности вуза k ; $V_2^{k,sci}$ – показатель научно-исследовательской деятельности вуза k ; $V_2^{k,inf}$ – показатель инфраструктуры вуза k ; $V_2^{k,int}$ – показатель международной деятельности вуза k ; $V_2^{k,fin}$ – показатель финансово-экономической деятельности вуза k ; $V_2^{k,j,emp}$ – показатель трудоустройства выпускников вуза k по специальности j ; $V_2^{k,br}$ – показатель «бренда» вуза k ; $R_1^{s(k),sal}$ – средняя заработная плата в регионе (городе) $s(k)$; $R_1^{s(k),unemp}$ – уровень безработицы в регионе (городе) $s(k)$; $R_1^{s(k),cult}$ – культурный уровень в регионе (городе) $s(k)$; $R_1^{s(k),crim}$ – уровень преступности в регионе (городе) $s(k)$; $R_1^{s(k),j,need}$ – востребованность на рынке труда региона (города) $s(k)$ выпускников по специальности j .

Расчет функций полезности поступления

Следующий этап процесса моделирования заключается в том, что агенты-абитуриенты, обладающие наборами случайных величин и помещенные во внешнюю

среду, должны определиться, какие возможные пары «вуз – специальность» им подходят в наибольшей мере. Будем рассматривать полезность поступления в вуз из 4 составляющих:

- Соответствие баллов ЕГЭ абитуриента i проходным баллам прошлых лет по специальности j в вузе k будем обозначать $F_{i,j,k}^1$.

- Интерес абитуриента i к различным специальностям j будем обозначать как $F_{i,j}^2$.

- Значимость для абитуриента i факторов вуза (эффективность деятельности) k при обучении по специальности j будем обозначать как $F_{i,j,k}^3$.

- Значимость для абитуриента i факторов региона $s(k)$, в котором находится вуз k при обучении по специальности j , будем обозначать как $F_{i,j,k}^4$.

Тогда выбор абитуриента осуществляется путем нахождения таких пар j, k , при которых некоторая функция (линейная комбинация или мультипликативная)

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

будет принимать максимальные значения. Напомним, что на данном этапе моделирования абитуриент выбирает несколько вузов, куда он подаст свои заявления. Окончательный выбор будет осуществляться на более позднем этапе, когда абитуриенты будут иметь информацию о конкурсе в вузы, которая позволит им оценить вероятности успешного поступления. Опишем подробнее вид каждой из указанных составляющих.

Основной показатель, влияющий на решение абитуриента, – это разница между его баллом ЕГЭ (суммой, деленной на количество предметов) E_{ij} и проходным баллом прошлого года $B_{j,k}(year - 1)$ в вуз k . Каждая пара $\langle j, k \rangle$ возможна лишь в том случае, если абитуриент сдавал ЕГЭ по предметам, необходимым для поступления на специальность j . Кроме того, будем считать, что абитуриент подает документы только в том случае, если проходной балл в прошлом году был меньше баллов ЕГЭ абитуриента: $E_{ij} - B_{j,k}(year - 1) > 0$ или сумма, которую семья может позволить себе тратить на обучение, превосходит стоимость обучения в вузе $P_1^i - V_1^{j,k} > 0$.

Таким образом, пусть J_i – это множество специальностей, на которые абитуриент может поступить с теми экзаменами, которые он сдавал. Тогда получаем, что абитуриент может выбрать только такие пары $\langle j, k \rangle$, для которых выполняется следующее условие (используются обозначения в индикаторах):

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

Кроме того, учтем и то, что абитуриент может иметь возможность обучаться платно, но принципиально не хочет этого делать. Или наоборот, человек может поступить на бесплатное обучение в родном городе, но уезжает учиться в мегаполис. Математически за это будет отвечать показатель абитуриента P_2^i . Будем считать, что если $P_2^i = 0$, то абитуриент однозначно не хочет учиться платно, если $P_2^i = 1$, то ему все равно. Таким образом, описывать вид составляющей полезности $F_{i,j,k}^1$ будем следующим образом в случае $j \in J_i$ (иначе значение будет равно 0):

$$F_{i,j,k}^1 = I\{E_{i,j} - B_{j,k}(year - 1) > 0\} + [1 - I\{E_{i,j} - B_{j,k}(year - 1) > 0\}] \times I\{P_1^i - V_1^{j,k} > 0\} \cdot P_2^i.$$

Показатель $F_{i,j}^2$ отражает интерес абитуриента i к поступлению на специальность j , поэтому имеет простой вид $F_{i,j}^2 = P_3^{i,j}$ с учетом возможности поступления только на специальности $j \in J_i$ (иначе 0). $P_3^{i,j}$ – случайные величины, которые подчиняются условию $\sum_{j \in J_i} P_3^{i,j} = 1$.

Показатель $F_{i,j,k}^3$ отражает значимость факторов, обусловленных вузом для абитуриента. Показатели эффективности деятельности вузов в России (образовательная деятельность, научно-исследовательская деятельность, инфраструктура, международная деятельность, финансовая деятельность, трудоустройство выпускников, бренд) публикуются [2] Министерством образования и науки и доступны для абитуриентов. Для удобства записи вида показателя введем обозначение $M = \{edu, sci, int, inf, fin, br\}$. Тогда значение показателя $F_{i,j,k}^3$ запишем в виде линейной комбинации

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

В указанной формуле выделено слагаемое, отвечающее за трудоустройство выпускников: этот показатель используется при расчетах в разрезе специальностей,

тогда как остальные только в разрезе вузов. Для показателей значимости факторов $P_4^{i,m}$ выполняется условие нормировки

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

Аналогичным образом рассчитывается и показатель $F_{i,j,k}^4$, отвечающий за учет влияния факторов региона $s(k)$, в котором находится вуз k . Вводя обозначение $N = \{sal, unemp, cult, crim\}$, укажем формулу для расчета показателя $F_{i,j,k}^4$. Единственное отличие заключается в том, что каждый абитуриент имеет различное отношение к переезду в другой регион, что отражается в существовании показателя $P_6^{i,k}$. Указанный показатель принимает значение, равное 1, если регион $s(k)$ наиболее предпочтителен для абитуриента.

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

Поддача заявлений в вузы и основной выбор абитуриента

На основании расчета функций полезности для каждого абитуриента составляются ранжированные списки по всем возможным парам специальностей и вузов $\langle j, k \rangle$. Абитуриент подает заявления в некоторое случайное число вузов на специальности, наиболее подходящие ему. Вузы публикуют информацию обо всех поданных заявлениях – списки абитуриентов, отсортированные по их баллам ЕГЭ, а также информацию о количестве бюджетных и платных мест. На основании этой информации, абитуриенты могут оценить вероятности поступления в вузы, которые они выбрали на первом этапе (пример представлен в таблице).

Пример информации, анализируемой каждым виртуальным агентом

Вуз	Специальность	Значение функции полезности	Оценка вероятности успешного поступления
Вуз 1	Специальность 1	0,954	0,93
Вуз 2	Специальность 2	0,932	0,87
Вуз 1	Специальность 3	0,911	0,95
...

Зная оценки вероятностей поступления, наиболее уверенные абитуриенты подают оригиналы документов в те вузы, которые им подходят в наибольшей степени. Все остальные абитуриенты (те, кто не удовлетворяют критерию уверенности в поступлении), получают эту информацию и на

основании занятых мест в вузах пересчитывают вероятности своего поступления. Такая итерационная процедура продолжается до тех пор, пока все места на бюджетной и платной основах не будут заняты абитуриентами.

Проверка модели на адекватность и идентификация параметров

В ходе моделирования генерировалась популяция абитуриентов, каждый из которых имел свои собственные случайные показатели. Каждый из этих показателей – реализация случайной величины, имеющей некоторый заранее определенный закон распределения с параметрами, которые мы будем называть начальными и обозначим их как $p_1^0, p_2^0, \dots, p_l^0$.

В качестве результата моделирования можно рассматривать количество поданных заявлений к общему количеству мест по каждой специальности j каждого вуза k (конкурс – количество человек, претендующих на одно место). Будем обозначать этот показатель как $\widehat{C}_{j,k}(p_1^0, p_2^0, \dots, p_l^0)$.

Для того чтобы показать адекватность построенной модели, достаточно провести моделирование на ретроспективных данных и сравнить реальный конкурс $C_{j,k}$ с полученным при моделировании $\widehat{C}_{j,k}(p_1^0, p_2^0, \dots, p_l^0)$. В таком случае идентификация параметров модели – параметров распределений $p_1^0, p_2^0, \dots, p_l^0$ – будет осуществляться путем проведения многих серий экспериментов с различными наборами параметров методом наименьших квадратов:

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

Заключение

В данной статье представлена методология моделирования приемной кампании в российские вузы, которая основана на использовании агентного подхода. Достоинства такого подхода к рассматриваемой задаче заключаются, прежде всего, в следующем:

● Максимально приближенная к реальности модель – степень детализации ограничивается возможностями компьютеров.

● Нет сложных формул, описывающих различные потоки абитуриентов. Все представляется с помощью простых правил, которые описывают мышление абитуриентов. В то же время возможно получение неочевидных результатов на агрегированном уровне.

● Моделирование многих социальных процессов зачастую является задачей, в которой поведение индивидов нелинейно, наполнено правилами «если-то». Такие особенности едва ли можно удовлетворительно описывать дифференциальными уравнениями, рекуррентными соотношениями и т.д.

Построенная модель может использоваться как вузами, так и органами государственной власти в области образования. Она позволяет проводить прогнозирование того, как абитуриенты будущего года будут распределяться по вузам, изучать реакции абитуриентов на закрытие вузов, закрытие или создание новых специальностей, изменения социально-экономических показателей региона и др. Кроме того, интересным представляется и использование приведенного аппарата расчета полезности поступления в вузы России и для самих абитуриентов и их родителей. Это может выразиться в возможном создании Интернет-ресурса, нацеленного на выработку рекомендаций по поступлению с учетом личных предпочтений абитуриента.

Список литературы

1. Абанкина И.В. Модель многоступенчатого выбора для прогнозирования спроса на высшее образование / И.В. Абанкина и др. // Университетское управление: практика и анализ. – 2014. – № 4–5. – С. 84–94.
2. Информационно-аналитические материалы по результатам анализа показателей эффективности образовательных организаций высшего образования [Электронный ресурс]. – Режим доступа: <http://miccedu.ru/monitoring/2014> (дата обращения: 01.07.2015).
3. Кисельгоф С.А. Выбор вузов абитуриентами с квадратичной функцией полезности: препринт WP7/2011/01;

Высшая школа экономики. – М.: Изд. дом Высшей школы экономики, 2011. – 44 с.

4. Питухин Е.А. Анализ межрегиональной мобильности выпускников школ при поступлении в высшие учебные заведения / Е.А. Питухин, А.А. Семенов // Университетское управление: практика и анализ. – 2011. – № 3. – С. 82–89.

5. Прахов И.А. Модель выбора вуза в условиях ЕГЭ и роль ожиданий абитуриентов: препринт WP10/2010/06; Гос. ун-т – Высшая школа экономики. – М.: Изд. дом Гос. ун-та – Высшей школы экономики, 2010. – 56 с.

6. Gale D., Shapley L.S. College Admissions and the Stability of Marriage. The American Mathematical Monthly. – Vol. 69. – № 1 (Jan., 1962). – P. 9–15.

References

1. Abankina I.V. Model mnogostupenchatogo vzbora dlja prognozirovanija sprosa na vysshee obrazovanie / I.V. Abankina i dr. // Universitetskoe upravlenie: praktika i analiz. 2014. no. 4–5. pp. 84–94.
2. Informacionno-analiticheskie materialy po rezultatam analiza pokazatelej jeffektivnosti obrazovatelnyh organizacij vysshego obrazovanija [Elektronnyj resurs]. Rezhim dostupa: <http://miccedu.ru/monitoring/2014> (data obrashhenija: 01.07.2015).
3. Kisel'gof S.A. Vyb'or vuzov abiturientami s kvadraticnoj funkciej poleznosti: preprint WP7/2011/01; Vysshaja shkola jekonomiki. M.: Izd. dom Vysshej shkoly jekonomiki, 2011. 44 p.
4. Pituhin E.A. Analiz mezhregionalnoj mobilnosti vypusknikov shkol pri postuplenii v vysshie uchebnye zavedenija / Pituhin E.A., Semenov A.A. // Universitetskoe upravlenie: praktika i analiz. 2011. no. 3. pp. 82–89.
5. Prahov I.A. Model vzbora vuzav v uslovijah EGJe i rol ozhidaniy abiturientov: preprint WP10/2010/06; Gos. un-t Vysshaja shkola jekonomiki. M.: Izd. dom Gos. un-ta Vysshej shkoly jekonomiki, 2010. 56 p.
6. Gale D., Shapley L.S. College Admissions and the Stability of Marriage. The American Mathematical Monthly. Vol. 69. no. 1 (Jan., 1962). pp. 9–15.

Рецензенты:

Соколов А.А., д.ф.-м.н., ведущий научный сотрудник отдела экспериментальной физики, Национальный исследовательский центр «Курчатовский институт» ФГБУ ГНЦ РФ «Институт физики высоких энергий», г. Москва;

Клименко К.Г., д.ф.-м.н., профессор кафедры «Математика и естественные науки» филиала «Протвино», ГБОУ ВО Московской области университета «Дубна», г. Протвино.