

В. М. Московкин

Имитационная экспертная система выбора университетов для обучения

Проведен краткий обзор зарубежных исследований в области моделирования процессов принятия решения о выборе колледжей и университетов для дальнейшего обучения; на концептуальном уровне построена соответствующая имитационная экспертная система.

Ключевые слова: имитационная экспертная система, поисковая машина “Scholar Google”, имитационные модели выбора университетов для обучения, интегральный показатель привлекательности обучения.

В нашей стране в советское время исследования по выбору школьниками и абитуриентами высших учебных заведений для дальнейшего обучения не проводились, так как в этом не было потребности из-за бесплатного и доступного высшего образования, а в постсоветское время традиции проведения таких исследований еще не сложились.

На Западе ситуация совершенно иная. Как показывают наши эксперименты с поисковой машиной “Scholar Google”, проведенные с помощью расширенного поиска с точной фразой, в этой области исследований сложилась устоявшаяся терминология (табл. 1).

Таблица 1

Термины в области исследований по выбору колледжей и университетов для обучения и отклики на них при запросах, сделанных с помощью поисковой машины “Scholar Google”

Названия терминов	Количество откликов	
	08.04.2009	08.07.2009
College choice	7010	7250
College access	3190	3460
University choice	862	982
University access	2390	2580
University decision-making process	75	85
College decision-making process	измерения не проводились	150
University selection process	74	93
College selection process	измерения не проводились	656

Из табл. 1 видно, что за три месяца количество откликов на запросы ключевых терминов по предмету нашего исследования увеличилось приблизительно на 5–10%. Отмечается также закономерность в значительно большем количестве откликов на термины, в которые входит слово “college”, по сравнению с терминами со словом “university”.

Еще 18 лет назад Лари Литтген писал, что процесс принятия решения о выборе колледжа или университета является очень важным и трудным для многих подростков и их родителей. В среднем каждый из них тратит в год свыше 50 часов на сбор и анализ информации, без учета посещения кампуса этих учебных заведений, на что уходит около 1500 долл. США [1].

Наиболее серьезный обзор по данной проблеме был опубликован Л. Хенриксоном в 2002 г. [2].

Он отмечает, что в течение 40 лет по тематике “College choice” и “College access” вышло в свет около 1900 работ, и далее обстоятельно анализирует три категории моделей, сложившиеся в этой области исследований:

- процессные модели (process models),
- теоретические модели (theoretical models),
- методологические модели (methodological models).

В качестве **процессных моделей** рассматриваются трех-пятистадийные модели развития, в частности, классическая трехстадийная модель [3], в которой:

на первой стадии учащийся обдумывает свое желание по поводу того, поступать ему в колледж или нет (на 8–12-й ступени обучения в “high school”);

вторая стадия включает изучение информации о потенциальных колледжах; учащийся определяет множество институтов согласно критериям его отбора, причем последние могут основываться не только на ограниченных характеристиках этих институтов — расположение, размер, программы обучения, студенческое восприятие колледжа и стоимость обучения;

третья стадия включает в себя реальный процесс принятия (или непринятия) студента университетом на учебу (процесс зачисления).

В качестве **теоретических моделей** рассматриваются модели “высшего образования” (“higher education”), социологические и эконометрические модели.

Далее Л. Хенриксон отмечает, что около 60% работ по моделированию процессов выбора учащимися высших учебных заведений для дальнейшего обучения основано на эмпирических данных и их дальнейшей обработке с помощью статистических методов. Этот класс моделей он относит к **методологическим**. 50% работ в нем используют логарифмические вероятностные функции, комбинируемые с максимальными вероятностными оценками, еще 12% работ используют факторный анализ [2].

Перспективы развития в данной области исследований Л. Хенриксон связывает с использованием методов нелинейной динамики и теории сложности (что позволит включить в рассмотрение обратные связи и гетерогенность), а также сложных вычислительных алгоритмов. Некоторые из новых вычислительных методов включают новые

методы хранения данных (например, ГИС) или новые пути компьютерного программирования и моделирования (например, компьютерное моделирование с использованием генетических алгоритмов, нейронных сетей, “agent-based”-моделей или сложных адаптивных систем) [2]. В дальнейшем он подробно останавливается на “agent-based”-моделировании, отмечая, что это совершенно новая вычислительная техника в социальных науках (обнаружено только две статьи), основанная на C** объектно-ориентированном программировании. В качестве агентов в этой модели выступают студент и университет, каждый из которых действует относительно независимо в принятии своих решений, подчиняясь простым правилам взаимоотношений, но вытекающие отсюда макромоделю по принятию абитуриентов в университет (mass enrollment patterns) являются очень сложными.

Многие зарубежные эмпирические исследования, проводимые в рамках построения социологических моделей, основаны на типизации информационных источников, связанных с выбором университетов для дальнейшего обучения, и количественных оценках факторов выбора университетов на основе самостоятельного генерирования учащимися весов учебных заведений [4, 5]. Например, в работе [4] используется семибалльная шкала оценок важности информационных источников (от 1 до 7) и десятибалльная шкала (от 0 до 10) оценок значимости весовых факторов выбора колледжей для дальнейшего обучения.

Данный обзор литературных источников по исследованиям в области выбора учащимися высших учебных заведений приводит нас к пониманию, какие основные факторы влияют на процесс выбора университетов для обучения, когда учащиеся могут сами оценивать значимость этих факторов. Это очень важно при построении имитационных моделей выбора университетов для обучения. Однако мы не обнаружили ни одной работы, которая бы синтезировала возможности имитационного моделирования с онлайн-инструментами с целью построения имитационных экспертных систем поиска и выбора университетов для обучения в глобальном масштабе. Ниже нами строится такая система — на концептуальном уровне.

* * *

Будем полагать, что абитуриенты при выборе университетов мира на качественном уровне обычно принимают во внимание следующие факторы: рейтинг (престижность) университета, стоимость обучения, дороговизну жизни (прожиточный минимум, стоимость потребительской корзины) в городе, где расположен университет, условия проживания (жилищные условия) и условия безопасности в рассматриваемом городе.

На основе этих пяти факторов построим простейшую формализованную процедуру определения привлекательности обучения в i -м университете. Для этого все указанные факторы необходимо представить в виде количественных индикаторов. Первые три индикатора соответствуют измеряемым (расчетным) статистическим данным (hard data), остальные два собираются на основе данных опросов (survey data). Рассмотрим их по порядку.

1. Рейтинг университета определяется с помощью интегрального показателя одного из мировых университетских рейтингов (китайский академический рейтинг, испанский вебометрический

рейтинг, рейтинг Times Higher Education (Великобритания) и др.). Обозначим его для i -го университета через I_i . Если отсутствуют данные по интегральным показателям, то можно использовать сами рейтинги (места в ранжировке университетов). Например, для TOP-500 китайского академического рейтинга положим $I_1=1, I_{500}=500$.

2. Стоимость обучения в i -м университете обозначим через C_i , долл. США.

3. Прожиточный минимум (living wage) в городе, где расположен i -й университет, обозначим через LW_i , долл. США. Здесь следует иметь в виду, что в одном городе могут быть расположены несколько университетов, поэтому значения LW_i для этих университетов будут одинаковыми.

4. Условия проживания (жилищные условия, living conditions) в i -м университете определяются на основе, например, пятибалльных оценок — BLC_i . Так, $BLC_i=5$ соответствует отличным условиям проживания, $BLC_i=1$ — очень плохим условиям проживания (могут использоваться также семи- или десятибалльные оценки).

5. Условия безопасности (security conditions) в городе, где расположен i -й университет, определим также на основе пятибалльных оценок — BSC_i ; $BSC_i=5$ соответствует очень высокому уровню безопасности, $BSC_i=1$ — очень низкому уровню безопасности.

Теперь необходимо использовать процедуру стыковки фактических измеряемых (расчетных) данных по первым трем индикаторам с балльными оценками двух последних индикаторов. При этом следует обратить внимание, что балльные оценки последних двух индикаторов построены в виде “стимуляторов”. Интегральный показатель мирового рейтинга университетов также является “стимулятором” (чем выше интегральный показатель, тем выше привлекательность университета для обучения), его перевод в пятибалльную шкалу можно произвести по формуле:

$$BI_i = \frac{4(I_i - I_{\min})}{(I_{\max} - I_{\min})} + 1. \quad (1)$$

Если I_i означает места в ранжировке, являясь “дестимулятором”, то

$$BI_i = \frac{4(I_i - I_{\max})}{(I_{\min} - I_{\max})} + 1. \quad (2)$$

Например, при китайском академическом ранжировании университетов ($I_1 = I_{\min} = 1, I_{\max}=500$) Гарвардский университет, занимающий первое место, будет оцениваться лучшим баллом ($BI_1=5$), а последний университет в этой ранжировке — худшим баллом ($BI_{500}=1$).

Второй и третий индикаторы являются “дестимуляторами”, и для них будут справедливы следующие формулы для перевода в баллы:

$$BC_i = \frac{4(C_i - C_{\max})}{(C_{\min} - C_{\max})} + 1, \quad (3)$$

$$BLW_i = \frac{4(LW_i - LW_{\max})}{(LW_{\min} - LW_{\max})} + 1. \quad (4)$$

В результате расчетов по формулам (1)–(4) округление до целых балльных оценок не производится, иначе существенно разные значения первых трех индикаторов будут приводить к одним и тем же балльным оценкам.

Отметим, что формула типа (1) используется для перевода фактических статистических данных в семибалльные оценки в методологии расчета перспективного индекса конкурентоспособности стран Всемирного экономического форума.

Теперь на аддитивной основе можно построить интегральный показатель привлекательности обучения в i -м университете с учетом весов частных индикаторов (university attractive indicator, UAI):

$$UAI_i =$$

$$\gamma_1 \cdot BI_i + \gamma_2 \cdot BC_i + \gamma_3 \cdot BLW_i + \gamma_4 \cdot BLC_i + \gamma_5 \cdot BSC_i, \quad (5)$$

где $\sum_{i=1}^5 \gamma_i = 1$.

Можно показать, что

$$\min\{UAI_i\} = \sum_{i=1}^5 \gamma_i = 1,$$

$$\max\{UAI_i\} = 5 \sum_{i=1}^5 \gamma_i = 5,$$

т. е. интегральный показатель привлекательности обучения изменяется в том же диапазоне, что и его частные индикаторы. Поэтому на основе равномерной шкалы для этого показателя можно построить следующую пятиуровневую классификацию университетов по степени их привлекательности (табл. 2).

Таблица 2

Классификация университетов по степени привлекательности для обучения

Диапазон изменения интегрального показателя UAI_i	Степень привлекательности университетов для обучения
0-1	очень низкая
1-2	низкая
2-3	средняя
3-4	высокая
4-5	очень высокая

Для обобщенных оценок привлекательности обучения в университетах мира весовые коэффициенты могут быть раз и навсегда определены методами попарных сравнений и экспертных оценок, но для индивидуальных оценок в этом нет никакой необходимости.

* * *

Абитуриент может сам подбирать для себя приемлемые веса. Например, состоятельный абитуриент на первые места поставит престижность университета и безопасность проживания, а все остальные индикаторы будут для него менее значимы. Варьируя веса и испытывая различные системы ранжирования университетов мира, абитуриент может проигрывать различные сценарии выбора университета, т. е. мы приходим к имитационной экспертной системе выбора университетов для обучения.

Сам выбор университетов должен исходить из предпочтений абитуриента: с помощью ограниченных, задаваемых на частные индикаторы. Например: из первых двухсот университетов, входящих в TOP-500 по одной из систем ранжирования университетов мира, абитуриент хочет выбрать университеты, удовлетворяющие следующим ограничениям: $C_i \leq 4$ тыс. долл. США в год; $LW_i \leq 50$ долл. США/сут.; $BLC_i = 3, 4$; $BSC_i = 4, 5$. На этом множестве ограничений на частные индикаторы следует искать максимум интегрального показателя UAI_i , т. е. мы приходим к задаче линейного программирования.

СПИСОК ЛИТЕРАТУРЫ

1. Litten L. H. Ivy Bound: High Ability Students and College Choice.— New York: College Board, 1991.
2. Henrickson L. Old Wine in a New Wineskin. College Choice, College Access Using Agent-Based Modelling // Social Science Computer Review.— 2002.— Vol. 20, № 4.— P. 400-419.
3. Litten L. H. Different strokes in the applicant pool: Some refinements in a model of student college choice // Journal of Higher Education.— 1982.— Vol. 55, № 4.— P. 383-402.
4. Galotti K. M., Mark M. C. How Do High School Students Structure an Important Life Decision? A Short-Term Longitudinal Study of the College Decision-Making Process // Research in Higher Education.— 1994.— Vol. 35, № 5.— P. 589-607.
5. Velontsou C., Lewis J. W., Paton R. A. University Selection: Information Requirements and Importance // International Journal of Educational Management.— 2004.— Vol. 18, № 3.— P. 160-171.

Материал поступил в редакцию 09.07.09.