

# БАЙЕСОВСКИЙ ПОДХОД К ПОСТРОЕНИЮ ИНДИВИДУАЛЬНОЙ ТРАЕКТОРИИ ПОЛЬЗОВАТЕЛЯ В СИСТЕМЕ ДИСТАНЦИОННОГО ОБУЧЕНИЯ\*

А. В. Босов<sup>1</sup>, Я. Г. Мартюшова<sup>2</sup>, А. В. Наумов<sup>3</sup>, А. П. Сапунова<sup>4</sup>

**Аннотация:** Рассмотрена задача формирования индивидуальной траектории пользователя системы дистанционного обучения (СДО) при смешанной форме ведения образовательной деятельности с организацией самостоятельной работы обучаемых с помощью СДО. По окончании каждого раздела курса обучения происходит разделение пользователей СДО на категории, определяемые решением задачи байесовской классификации. Для каждой категории предлагается индивидуальное задание разного уровня сложности на следующий раздел курса, и таким образом формируется индивидуальная траектория обучаемого. Настройка байесовского классификатора проводится на основе статистики работы пользователей СДО. Приводятся экспериментальные результаты решения задачи на одном из этапов обучения.

**Ключевые слова:** система дистанционного обучения; байесовский классификатор; адаптивные системы; индивидуальная траектория обучения

**DOI:** 10.14357/19922264200313

## 1 Введение

Современный тренд в обучающих системах, в том числе в СДО, — разработка методов построения индивидуальных траекторий пользователей, учитывающих статистику их работы в системе. Современная теория тестирования (item response theory, IRT) часто использует вероятностные модели для адаптации процесса тестирования под уровень знаний пользователей, учитывая специфику обучаемых [1–8]. В качестве случайных параметров в таких моделях выступают обычно уровень знаний пользователей и уровень сложности заданий [1, 4] или время ответа пользователя на задание [7, 8]. Оценивая эти параметры, система проводит самонастройку контента, адаптируя его под контингент пользователей.

Формирование индивидуальных траекторий тестирования или обучения в СДО [3, 5, 6] подразумевает решение в определенные моменты функционирования системы задачи классификации пользователей — отнесение их к различным категориям (классам) на основе статистики работы в системе. По результатам проведенной классификации разным категориям пользователей предлагаются разные траектории дальнейшего тестиро-

вания или обучения. Решение задачи классификации осуществляется, например, с использованием нейросетевых технологий [9–11] или байесовского подхода [12, 13].

В данной работе в модели, учитывающей смешанную форму обучения с проведением нескольких объективных очных рубежных контролей в течение курса, решается задача построения индивидуальной траектории пользователя в используемой для организации его самостоятельной работы СДО. Индивидуальная траектория корректируется после прохождения пользователем каждого раздела предложенного в СДО курса с помощью построения байесовского классификатора. Для определения соответствующих вероятностей используются классические частотные оценки, построенные на основе статистики работы пользователей СДО.

## 2 Задача построения индивидуальной траектории как задача байесовской классификации

В рамках рассматриваемой модели предполагается, что дистанционное обучение проводится

\* Исследование выполнено при частичной финансовой поддержке РФФИ в рамках научного проекта 18-07-00617.

<sup>1</sup> Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; Московский авиационный институт, avbosov@ipiran.ru

<sup>2</sup> Московский авиационный институт, ma1554@mail.ru

<sup>3</sup> Московский авиационный институт, naumovav@mail.ru

<sup>4</sup> Московский авиационный институт, sap2603@mail.ru

совместно с очной формой обучения в виде организации самостоятельной работы студентов посредством использования электронного учебника для выполнения домашнего задания и организации тестирования.

Предлагаемый пользователю курс дистанционного обучения разбит на несколько разделов, которые он должен изучить для освоения данной дисциплины. После каждого раздела пользователю нужно решить тест определенного уровня сложности. Сложность заданий теста определяется экспертом или с помощью специальных алгоритмов, основанных на модели Раша [1]. Для формирования теста из существующих заданий необходимо решить задачу математического программирования, описанную в [14], параметром которой служит суммарный уровень сложности теста. Необходимо определить, какого уровня сложности тест выдавать пользователю в конце каждого раздела, чтобы он соответствовал уровню его подготовки и формировал индивидуальную траекторию обучения.

Решение о том, какой должна быть сложность выдаваемого в конце каждого раздела теста, предлагается принимать экспертно на основе того, к какой объективной категории принадлежит пользователь, например: «отличник», «хорошист» и т. д. Таким образом, на каждом этапе обучения предлагается решать задачу классификации пользователя: отнесение его к одной из категорий. Определяются категории пользователя дискретно в процессе обучения по окончании каждого раздела курса дистанционного обучения. Обозначим моменты времени, когда значение категории может меняться, через  $t = 0, \dots, T$ .

В начальный момент времени  $t = 0$  объективной категории пользователя присваивается некоторая априорная оценка, определяемая на основе предыдущей траектории обучения. Пусть  $t = \{t_1, t_2, t_3\}$  — моменты времени, в которые проводятся очные контрольные мероприятия, на основании которых происходит объективная классификация пользователей СДО с последующей адаптацией их индивидуальных траекторий обучения.

Пусть  $X$  — множество, содержащее  $m$  допустимых категорий пользователя, например  $X = \{2, 3, 4, 5\}$ , где  $m = 4$  — число категорий.

В моменты времени  $t_1, t_2$  и  $t_3$  объективная категория пользователя  $y_{t_i}$  определяется на основе результата очной контрольной работы с помощью традиционной оценки  $k_{t_i} \in Y, i = 1, 2, 3$ , т. е.  $y_{t_i} = k_{t_i}, i = 1, 2, 3$ , где  $k_{t_i} \in Y = \{0, 2, 3, 4, 5\}$ . Случай  $k_{t_i} = 0$  означает, что пользователь не писал контрольную работу в момент времени  $t_i$ . В другие моменты времени объективная категория пользо-

вателя не определена. Однако согласно формуле текущего рейтинга пользователя, предложенной в работе [15] и модифицированной под условия рассматриваемой задачи, можно определить текущую категорию рейтинга  $x_t \in X$  пользователя согласно следующей методике.

Пусть  $B_t$  — текущий (субъективный) рейтинг пользователя в момент времени  $t = 1, \dots, T$ , вычисляемый по формуле:

$$B_t = \frac{\omega_k}{5} \sum_{t_i \leq t} k_{t_i} + \omega_E \frac{\sum_{\tau=1}^t E_{\tau}^{+}}{\sum_{\tau=1}^t E_{\tau}} - \omega_{\delta} \left(1 - \frac{\delta_t}{\Delta_t}\right),$$

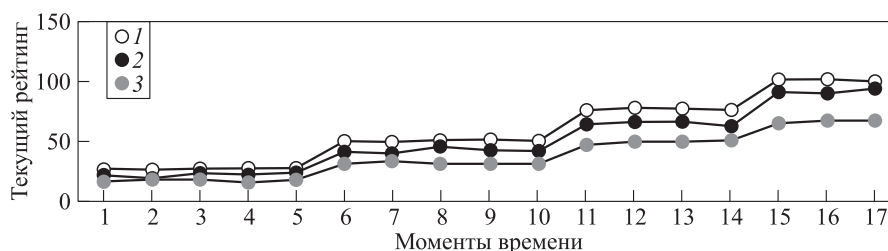
где  $E_{\tau}$  — суммарная сложность теста, выполненного в момент  $\tau$ ;  $E_{\tau}^{+}$  — суммарная сложность правильно решенных заданий теста, выполненного в момент  $\tau$ ;  $\Delta_t$  — общее число очных занятий, прошедших к моменту времени  $t$ ;  $\delta_t$  — число посещений пользователем очных занятий, прошедших к моменту времени  $t$ .

Величины  $\omega_k, \omega_E$  и  $\omega_{\delta}$  — это весовые коэффициенты, которые выбраны так, чтобы максимальное значение текущего рейтинга пользователя в момент времени  $T$  было равно 100. Эти величины отражают важность каждого слагаемого в формуле рейтинга. Например, в [15] выбраны следующие значения:  $\omega_k = \omega_E = 25$  и  $\omega_{\delta} = 15$ . Если при вычислении текущего рейтинга требуется учитывать лишь успехи пользователя в СДО и посещение очных занятий, можно пересчитать коэффициенты, положив  $\omega_k = 0$ .

Пусть  $B_t^{\max}$  — максимальный рейтинг, который может быть набран пользователем к моменту времени  $t$ . Например, для  $t_1 \leq t < t_2$  справедливо  $B_t^{\max} = 50$ . Траектория этой верхней границы, а также несколько примеров траекторий текущего рейтинга  $B_t$  для случая  $T = 17, t_1 = 6, t_2 = 11$  и  $t_3 = 15$  приведены на рис. 1.

Текущая категория рейтинга  $x_t$  для заданного множества  $X = \{2, 3, 4, 5\}$  определяется текущим рейтингом  $B_t$ , например согласно табл. 1, предложенной в [15]. При этом границы изменения рейтинга в каждой из категорий определяются экспертами.

Таким образом, в каждый момент времени  $t$  необходимо по текущей категории рейтинга  $x_t$  принять решение об отнесении пользователя к одной из объективных категорий по результатам следующей за  $t$  контрольной работы, т. е. требуется прогнозировать оценку, которую получит пользователь на следующей контрольной работе, имея текущую категорию рейтинга  $x_t$ . Этот прогноз позволит наилучшим образом подобрать сложность предлагаемого в текущий момент времени теста, чтобы



**Рис. 1** Траектории текущего рейтинга: 1 — максимальный рейтинг; 2 — рейтинг хорошо успевающего студента; 3 — рейтинг студента, успевающего удовлетворительно

**Таблица 1** Соответствие между категорией и рейтингом

$x_t$	$b_t = \frac{B_t}{B_t^{\max}}$
2	[0; 0,6]
3	[0,6; 0,75]
4	[0,75; 0,91]
5	[0,91; 1]

максимально способствовать повышению успеваемости пользователя. Для подбора сложности теста в этой ситуации используются различные педагогические приемы, описание которых выходит за рамки данной статьи. В данной работе для прогнозирования предлагается использовать метод байесовской классификации. Пусть  $P(y_{t_k})$  — вероятность того, что наугад выбранный пользователь в момент времени  $t_k$  проведения контрольной работы, следующей за моментом времени  $t$ , будет иметь объективную категорию рейтинга  $y_{t_k} \in Y$ ;  $P(x_t|y_{t_k})$  — условная вероятность того, что наугад выбранный пользователь в момент времени  $t$  имеет текущую категорию рейтинга, равную  $x_t \in X$ , при условии что в момент времени  $t_k$  проведения контрольной работы, следующей за моментом времени  $t$ , он будет иметь объективную категорию рейтинга  $y_{t_k} \in Y$ .

Рассмотрим алгоритм классификации  $a(\cdot) : X \rightarrow Y$ . Он ставит в соответствие пользователю с текущей категорией рейтинга  $x_t$  объективную категорию  $s$  на момент  $t_k$  следующей контрольной работы. Рассмотрим множество  $A_s = \{x_t \in X | a(x_t) = s, s \in Y\}$ . Вероятность того, что пользователь имеет текущую категорию рейтинга  $x_t \in X$  и алгоритм  $a(\cdot)$  отнесет его к объективной категории  $s$ , равна  $P(y_{t_k})P(A_s|y_{t_k})$ , где  $P(A_s|y_{t_k})$  — условная вероятность событий  $x_t \in A_s$  при условии что в момент времени  $t_k$  пользователь будет иметь объективную категорию рейтинга  $y_{t_k}$ . Каждой паре  $(y_{t_k}, s) \in Y \times Y$  поставим в соответствие величину потери  $\lambda_{y_{t_k}s}$  при отнесении пользователя к объективной категории вместо  $y_{t_k}$ . Тогда функционал

среднего риска — ожидаемая величина потери при классификации алгоритмом  $a(\cdot)$  — будет иметь вид:

$$R(a(\cdot)) = \sum_{s \in Y} \sum_{y_{t_k} \in Y} \lambda_{y_{t_k}s} P(y_{t_k}) P(A_s|y_{t_k}).$$

Согласно теории байесовской классификации, если известны вероятности  $P(y_{t_k})$  и  $P(x_t|y_{t_k})$ , то минимум среднего риска  $R(a(\cdot))$  достигается алгоритмом [13]:

$$a(x_t) = \arg \min_{s \in Y} \sum_{y_{t_k} \in Y} \lambda_{y_{t_k}s} P(y_{t_k}) P(x_t|y_{t_k}). \quad (1)$$

Решать задачу оптимизации (1) не представляется возможным, поскольку вероятности  $P(y_{t_k})$  и  $P(x_t|y_{t_k})$  неизвестны. Вместо этого для классификации можно использовать эмпирические оценки вероятностей  $\hat{P}(y_{t_k})$  и  $\hat{P}(x_t|y_{t_k})$ , найденные по обучающей выборке. Под обучающей выборкой понимается статистика работы пользователей СДО, обучающихся по аналогичной программе за определенный промежуток времени. Эта информация становится доступной после выполнения СДО первого курса обучения. Таким образом, указанные оценки можно найти как относительные частоты:

$$\hat{P}(y_{t_k}) = \frac{l(y_{t_k})}{l}, \quad i = 2, 3, 4, 5; \quad \hat{P}(x_t|y_{t_k}) = \frac{l_{y_{t_k}}^{x_t}}{l(y_{t_k})},$$

где  $l(y_{t_k})$  — число пользователей, имевших по результатам теста в СДО в момент  $t_k$  объективную категорию рейтинга  $y_{t_k}$ ;  $l$  — общее число пользователей СДО, обучавшихся по данной программе, по результатам обучающей выборки;  $l_{y_{t_k}}^{x_t}$  — число пользователей СДО, имевших к моменту времени текущую категорию  $x_t$ , у которых объективная категория рейтинга на момент времени  $t_k$  была равна  $y_{t_k}$ .

В силу малого числа элементов множеств  $X$  и  $Y$  задача (1) может быть решена в каждый момент времени  $t$  простым перебором. Результат решения сильно зависит от выбора величины потери  $\lambda_{y_{t_k}s}$ . Традиционно в качестве  $\lambda_{y_{t_k}s}$  выбирается модуль разности  $y_{t_k}$  и  $s$  или его квадрат. Этот

выбор приводит к формированию классификатора на основе условного математического ожидания  $\sum_{y_{t_k} \in Y} y_{t_k} P(y_{t_k} | x_t)$ , симметрично усредняющего ошибки как с завышением, так и с занижением объективного рейтинга пользователя. Однако с точки зрения процесса обучения и построения индивидуальной траектории пользователя представляется, что заметно хуже ошибочно занизить категорию пользователя, чем ошибочно ее завысить. Поэтому предлагается использовать разные штрафы для неположительных  $y_{t_k} - s$ , когда рейтинг ошибочно завышается (например, использовать модуль разности или корень), и для положительных, когда рейтинг ошибочно занижается (можно предложить  $\lambda_{y_{t_k} s} = 2^{y_{t_k} - s}$ , т.е. кратно увеличивать штраф за занижение рейтинга).

Представленный в табл. 1 способ формирования текущей категории рейтинга пользователя отражает его текущую успеваемость по классической пятибалльной шкале. Для более тонкой настройки процесса адаптации индивидуальной траектории пользователя с применением описанного классификатора можно расширить множество возможных значений текущего рейтинга пользователя  $x_t$  до  $K$  элементов, дискретизировав соответствующим образом величину  $b_t \in [0, 1]$ . Например, для  $K = 20$  текущей категории рейтинга можно присваивать значение  $k$ , т.е.  $x_t = k$ , если  $b_t$  будет удовлетворять условию  $(k - 1)0,05 < b_t \leq 0,05k$ ,  $k = 1, \dots, 20$ .

В качестве альтернативы предложенному классификатору можно рассматривать вариант

$$a(x_t) = \arg \max_{y_{t_k} \in Y} P(y_{t_k} | x_t), \quad (2)$$

т.е. оценку, обеспечивающую максимум апостериорной вероятности  $P(y_{t_k} | x_t)$  объективной категории относительно текущего рейтинга  $x_t$ . В реализации эта вероятность также может быть оценена

**Таблица 3** Оценки условного относительно  $x_8$  распределения категории  $y_{t_{10}}$

$x_8$	$y_{t_{10}}$				
	0	2	3	4	5
1	0,30	0,42	0,22	0,05	0,01
2	0,18	0,54	0,23	0,04	0,01
3	0,33	0,36	0,27	0,03	0,01
4	0,11	0,34	0,47	0,06	0,02
5	0,05	0,28	0,42	0,17	0,08
6	0,06	0,24	0,41	0,19	0,10
7	0,05	0,09	0,33	0,32	0,21
8	0,04	0,05	0,31	0,41	0,19
9	0,02	0,06	0,17	0,34	0,41
10	0,01	0,02	0,09	0,39	0,49

соответствующей частотой по обучающей выборке. Это метод, так же как и метод условного среднего, имеет недостаток симметричного учета штрафа за ошибки завышения/занижения рейтинга, что и демонстрирует представленный далее численный эксперимент.

### 3 Результаты численного эксперимента

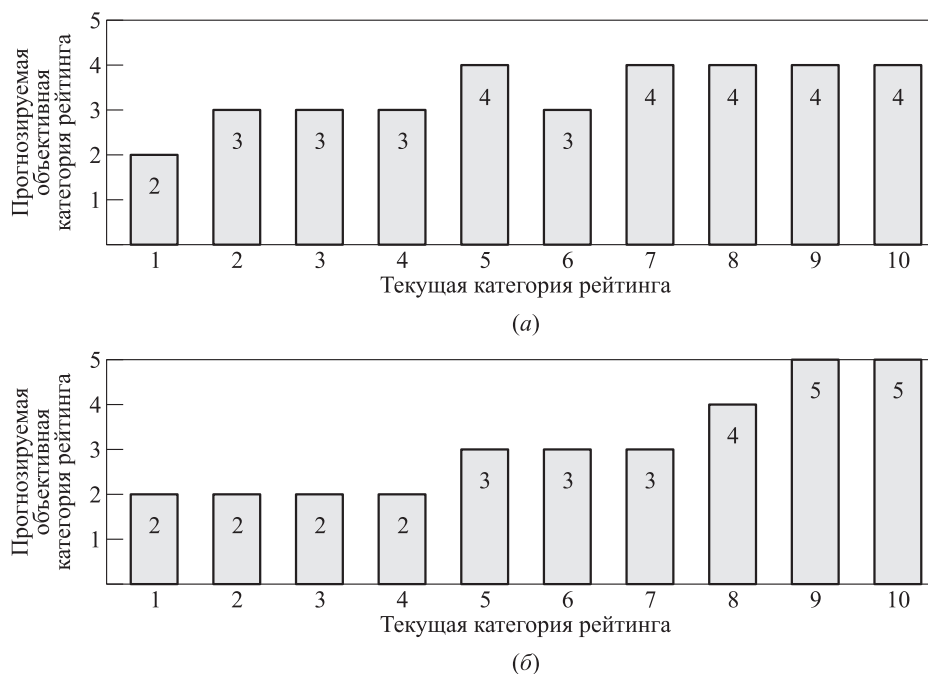
Представленная в статье методика формирования индивидуальной траектории обучения была опробована на примере использования электронного учебника по курсу «Теория вероятностей и математическая статистика» в системе дистанционного обучения CLASS.NET [16] МАИ при очной форме преподавания этой дисциплины студентам инженерных факультетов и наличии трех контрольных работ в процессе обучения. Использовались оба приведенных выше классификатора для числа дискретных значений текущей категории рейтинга  $K = 10$ . Электронный учебник содержит 14 разделов. Очные контрольные работы проводятся после 5-го, 10-го и 14-го разделов. Таким образом,  $T = 17$ . Рассмотрим решение задачи классификации перед 8-м разделом учебника. Соответствующие оценки вероятностей, необходимые для решения задач классификации, отражены в табл. 2–4.

**Таблица 2** Оценки распределения объективной категории  $y_{t_{10}}$

$y_{t_{10}}$	$\hat{P}(y_{t_{10}})$
0	0,09
2	0,11
3	0,38
4	0,28
5	0,14

**Таблица 4** Оценки условного относительно  $y_{t_{10}}$  распределения категории  $x_8$

$y_{t_{10}}$	$x_8$				
	0	2	3	4	5
1	0,11	0,23	0,01	0,01	0
2	0,15	0,09	0,03	0,01	0
3	0,14	0,17	0,02	0,03	0
4	0,09	0,13	0,02	0,02	0,01
5	0,11	0,09	0,10	0,10	0,01
6	0,12	0,05	0,13	0,09	0,01
7	0,09	0,03	0,23	0,14	0,13
8	0,06	0,05	0,24	0,26	0,24
9	0,07	0,09	0,13	0,21	0,31
10	0,06	0,07	0,09	0,13	0,29



**Рис. 2** Результаты несимметричной байесовской классификации (а) и классификации по максимуму апостериорной вероятности (б)

Эти оценки получены по обучающей выборке объемом 1500 элементов, составленной из наблюдений за схожим контингентом пользователей СДО, обучавшихся по аналогичной программе в предыдущем семестре.

Результаты классификации с использованием формулы (1) для случая

$$\lambda_{y_{t_k}s} = \begin{cases} s - y_{t_k}, & y_{t_k} - s \leq 0, \\ 2^{y_{t_k}-s}, & y_{t_k} - s > 0, \end{cases}$$

отражены на диаграмме рис. 2, а, а с использованием формулы (2) — на диаграмме рис. 2, б. По горизонтальной оси расположены значения текущей категории рейтинга  $x_8$ , а по вертикальной оси — объективная категория рейтинга по результатам классификации. Для приведенных исходных данных задачи (1) и (2) имеют единственные решения для каждого значения  $x_8$ .

Результаты использования байесовского классификатора (см. рис. 2, а) выглядят более сбалансированно и в среднем завышенно по сравнению с использованием классификатора, максимизирующего условные вероятности (см. рис. 2, б). Использование весовых коэффициентов  $\lambda_{y_{t_k}s}$  добавляет гибкости инструменту байесовской классификации.

Предложенный байесовский классификатор был использован для построения индивидуальных

траекторий обучения и адаптации функционирования СДО в осеннем семестре 2019 г. в педагогическом эксперименте с использованием 5 экспериментальных групп и 12 контрольных.

Успеваемость в экспериментальных группах по итогам семестра оказалась выше, чем в контрольных группах, в среднем на 10%–15%. При этом наблюдался рост успеваемости в экспериментальных группах по сравнению с итогами их обучения в предыдущем семестре.

## 4 Заключение

В работе предложен адаптивный способ формирования индивидуальной траектории пользователя СДО, используемой при очной форме ведения занятий. Индивидуальность траектории достигается последовательным изменением суммарной сложности теста в СДО в дискретные моменты времени на основе решения задачи классификации пользователя — оценки категории, соответствующей уровню знаний. Адаптивность процесса происходит за счет проведения в некоторые из указанных дискретных моментов времени, соответствующих разделам курса дистанционного обучения, очных контрольных работ, которые позволяют оценить объективную категорию пользователя с последующим использованием этой информации в процессе поэтапной классификации. Результаты про-



веденного численного эксперимента на примере использования СДО МАИ CLASS.NET в процессе преподавания физико-математических дисциплин показали эффективность предложенной методики построения индивидуальной траектории обучения.

## Литература

1. *Rasch G.* Probabilistic models for some intelligence and attainment tests. — Chicago, IL, USA: The University of Chicago Press, 1980. 224 p.
2. *Van der Linden W. J., Scrams D. J., Schnipke D. L., et al.* Using response-time constraints to control for differential speededness in computerized adaptive testing // *Appl. Psych. Meas.*, 1999. Vol. 23. Iss. 3. P. 195–210.
3. *Куравский Л. С., Мармалюк П. А., Алхимов В. И., Юрьев Г. А.* Новый подход к построению интеллектуальных и компетентностных тестов // *Моделирование и анализ данных*, 2013. № 1. С. 4–28.
4. *Кибзун А. И., Иноземцев А. О.* Оценивание уровней сложности тестов на основе метода максимального правдоподобия // *Автоматика и телемеханика*, 2014. № 4. С. 20–37.
5. *Куравский Л. С., Мармалюк П. А., Юрьев Г. А., Думин П. Н., Панфилова А. С.* Вероятностное моделирование процесса выполнения тестовых заданий на основе модифицированной функции Раша // *Вопросы психологии*, 2015. № 4. С. 109–118.
6. *Kuravsky L. S., Margolis A. A., Marmalyuk P. A., Panfilova A. S., Yuryev G. A., Dumin P. N.* A probabilistic model of adaptive training // *Applied Mathematical Sciences*, 2016. Vol. 10. Iss. 48. P. 2369–2380.
7. *Наумов А. В., Мхитарян Г. А.* О задаче вероятностной оптимизации для ограниченного по времени тестирования // *Автоматика и телемеханика*, 2016. № 9. С. 124–135.
8. *Наумов А. В., Мхитарян Г. А., Черыгова Е. Е.* Стохастическая постановка задачи формирования теста заданного уровня сложности с минимизацией количества времени выполнения // *Вестник компьютерных и информационных технологий*, 2019. № 2. С. 37–46.
9. *Каллан Р.* Основные концепции нейронных сетей / Пер. с англ. — М.: Вильямс, 2001. 287 с. (*Callan R.* The essence of neural networks. — Prentice Hall Europe, 1999. 232 p.)
10. *Воробьев Е. В., Пучков Е. В.* Классификация текстов с помощью сверточных нейронных сетей // *Молодой исследователь Дона*, 2017. № 6(9). С. 2–7. [http://mid-journal.ru/upload/iblock/8ed/1.-vorobev\\_-puchkov.pdf](http://mid-journal.ru/upload/iblock/8ed/1.-vorobev_-puchkov.pdf).
11. *Микрюков А. А., Бабаиш А. В., Сизов В. А.* Классификация событий в системах обеспечения информационной безопасности на основе нейросетевых технологий // *Открытое образование*, 2019. Т. 23. № 1. С. 57–63.
12. *Дьяконов А. Г.* Методы решения задач классификации с категориальными признаками // *Прикладная математика и информатика. Труды факультета Вычислительной математики и кибернетики МГУ имени М. В. Ломоносова*, 2014. № 46. С. 103–127.
13. *Воронцов К. В.* Лекции по алгоритмическим композициям, 2015. 45 с. <http://www.ccas.ru/voron/download/Composition.pdf>.
14. *Наумов А. В., Иноземцев А. О.* Алгоритм формирования индивидуальных заданий в системах дистанционного обучения // *Вестник компьютерных и информационных технологий*, 2013. № 6. С. 46–51.
15. *Мартыушова Я. Г., Лыкова Н. М.* Организация рефлексивно-оценочной деятельности студентов университетов средствами электронного учебника // *Психолого-педагогические исследования*, 2018. Т. 10. № 2. С. 125–134. doi: 10.17759/psyedu.2018100211.
16. *Наумов А. В., Джумурат А. С., Иноземцев А. О.* Система дистанционного обучения математическим дисциплинам CLASS.NET // *Вестник компьютерных и информационных технологий*, 2014. № 10. С. 36–44.

Поступила в редакцию 01.06.20

## BAYESIAN APPROACH TO THE CONSTRUCTION OF AN INDIVIDUAL USER TRAJECTORY IN THE SYSTEM OF DISTANCE LEARNING

A. V. Bosov<sup>1,2</sup>, Ya. G. Martyushova<sup>2</sup>, A. V. Naumov<sup>2</sup>, and A. P. Sapunova<sup>2</sup>

<sup>1</sup>Institute of Informatics Problems, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation

<sup>2</sup>Moscow State Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125933, Russian Federation

**Abstract:** The paper considers the task of forming an individual user path for a distance learning system (LMS) with a mixed form of conducting educational activities with organization of independent work of students using LMS. At the end of each section of the training course, the LMS users are divided into categories determined by

the solution of the Bayesian classification problem. For each category, an individual task of a different level of complexity is proposed for the next section of the course, thus forming the individual trajectory of the student. The Bayesian classifier is set up based on statistics of work of the users of LMS. The experimental results of solving the problem at one of the stages of training are presented.

**Keywords:** distance learning system; Bayesian classifier; adaptive systems; individual learning path

**DOI:** 10.14357/19922264200313

## Acknowledgments

The reported study was partially funded by RFBR, project No. 18-07-00617.

## References

1. Rasch, G. 1980. *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: University of Chicago Press. 224 p.
2. Van der Linden, W. J., D. J. Scrams, and D. L. Schnipke. 1999. Using response-time constraints to control for differential speededness in computerized adaptive testing. *Appl. Psych. Meas.* 23(3):195–210.
3. Kuravsky, L. S., P. A. Marmalyuk, V. I. Alkhimov, and G. A. Yuryev. 2013. Novyy podkhod k postroeniyu intellektual'nykh i kompetentnostnykh testov [A new approach to constructing intellectual and competence-based tests]. *Modelling Data Analysis* 1:4–28.
4. Kibzun, A. I., and A. O. Inozemtsev. 2014. Using the maximum likelihood method to estimate test complexity levels. *Automat. Rem. Contr.* 75(4):607–621.
5. Kuravsky, L. S., P. A. Marmalyuk, G. A. Yuryev, P. N. Dumin, and A. S. Panfilova. 2015. Veroyatnostnoe modelirovanie protsessa vypolneniya testovykh zadaniy na osnove modifitsirovannoy funktsii Rasha [Probabilistic modeling of the test tasks based on the modified Rasch function]. *Voprosy psikhologii* [Psychology Issues] 4:109–118.
6. Kuravsky, L. S., A. A. Margolis, P. A. Marmalyuk, A. S. Panfilova, G. A. Yuryev, and P. N. Dumin. 2016. A probabilistic model of adaptive training. *Applied Mathematical Sciences* 10(48):2369–2380.
7. Naumov, A. V., and G. A. Mkhitarian. 2016. On the problem of probabilistic optimization of time-limited testing. *Automat. Rem. Contr.* 77(9):1612–1621.
8. Naumov, A. V., G. A. Mkhitarian, and E. E. Cherygova. 2019. Stokhasticheskaya postanovka zadachi formirovaniya testa zadannogo urovnya slozhnosti s minimizatsiei kvantili vremeni vypolneniya [Stochastic statement of the problem of generating tests with defined complexity with the minimization of quantile of test passing time]. *Vestnik komp'yuternykh i informatsionnykh tekhnologiy* [Herald of Computer and Information Technologies] 2:37–46.
9. Callan, R. 1999. *The essence of neural networks*. Prentice Hall Europe. 232 p.
10. Vorob'ev, E. V., and E. V. Puchkov. 2017. Klassifikatsiya tekstov s pomoshch'yu svertochnykh neyronnykh setey [Classification of texts using convolutional neural networks]. *Molodoy issledovatel' Dona* [Young Researcher of the Don] 6(9):2–7. Available at: [http://mid-journal.ru/upload/iblock/8ed/1.-vorobev\\_-puchkov.pdf](http://mid-journal.ru/upload/iblock/8ed/1.-vorobev_-puchkov.pdf) (accessed July 23, 2020).
11. Mikryukov, A. A., A. V. Babash, and V. A. Sizov. 2019. Klassifikatsiya sobytii v sistemakh obespecheniya informatsionnoy bezopasnosti na osnove neyrosetevykh tekhnologiy [Classification of events in information security systems based on neural networks]. *Otkrytoe obrazovanie* [Open Education] 23(1):57–63.
12. D'yakonov, A. G. 2015. Solution methods for classification problems with categorical attributes. *Computational Mathematics Modeling* 26(3):408–428.
13. Vorontsov, K. V. 2015. Lektsii po algoritmicheskim kompozitsiyam [Lectures on algorithmic compositions]. Available at: <http://www.ccas.ru/voron/download/Composition.pdf> (accessed July 23, 2020).
14. Naumov, A. V., and A. O. Inozemtsev. 2013. Algoritmi formirovaniya individual'nykh zadaniy v sistemakh distantsionnogo obucheniya [The algorithm for generating the individual tasks in distance learning]. *Vestnik komp'yuternykh i informatsionnykh tekhnologiy* [Herald of Computer and Information Technologies] 6:46–51.
15. Martyushova, Ya. G., and N. M. Lykova. 2018. Organizatsiya refleksivno-otsenочноy deyatel'nosti studentov universitetov sredstvami elektronnoy uchebnika [Organization of reflexive-evaluative activity of university students by using the learning management system]. *Psikhologo-pedagogicheskie issledovaniya* [Psychological-Educational Studies] 10(2):125–134. doi: 10.17759/psyedu.2018100211.
16. Naumov, A. V., A. S. Dzhumurat, and A. O. Inozemtsev. 2014. Sistema distantsionnogo obucheniya matematicheskimi distsiplinami CLASS.NET [Distance learning system for mathematical disciplines CLASS.NET]. *Vestnik komp'yuternykh i informatsionnykh tekhnologiy* [Herald of Computer and Information Technologies] 10:36–44.

Received June 1, 2020

## Contributors

**Bosov Alexey V.** (b. 1969) — Doctor of Science in technology, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov

Str., Moscow 119333, Russian Federation; Moscow State Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125933, Russian Federation; AVBosov@ipiran.ru

**Martyushova Yanina G.** (b. 1970) — Candidate of Science in pedagogy, senior lecturer, Department of Probability Theory and Computer Simulations, Moscow Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125933, Russian Federation; ma1554@mail.ru

**Naumov Andrey V.** (b. 1966) — Doctor of Science in physics and mathematics, professor, Moscow Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125933, Russian Federation; naumovav@mail.ru

**Sapunova Anastasia P.** (b. 1998) — Master student, Moscow Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125933, Russian Federation; sap2603@mail.ru