ISSN 2306-1561
Automation and Control in Technical Systems (ACTS)
2015, No 3, pp. 94-103.
DOI: 10.12731/2306-1561-2015-3-9

Analytical and Simulation Methods to Evaluate the Rate of Convergence of an Adaptive Algorithm Generating Test Tasks

Gennady Grigorievich Yagudaev

Russian Federation, Doctor of Technical Sciences, branch Manager.

North-Caucasian branch of State Technical University – MADI, 357340, Russian Federation, Stavropol Region, Lermontov, Promyshlennaya Str., 20. Tel.: +7 (7935) 3-78-73, http://skfmadi.ru

gena_yagudaev@mail.ru

Maksim Igorevich Yartzev

Russian Federation, Postgraduate Student, Department «Automated Control Systems».

State Technical University – MADI, 125319, Russian Federation, Moscow, Leningradsky prospekt, 64. Tel.: +7 (499) 151-64-12. http://www.madi.ru

gizmo1654@mail.ru

Andrey Andreevich Ivakhnenko

Russian Federation, Ph. D., Associate Professor, Department «Management».

State Technical University – MADI, 125319, Russian Federation, Moscow, Leningradsky prospekt, 64. Tel.: +7 (499) 151-64-12. http://www.madi.ru

jointlab@mail.ru

Irina Aleksandrovna Evstratova

Russian Federation, Ph. D., Associate Professor, Department «Automated Control Systems».

State Technical University – MADI, 125319, Russian Federation, Moscow, Leningradsky prospekt, 64. Tel.: +7 (499) 151-64-12. http://www.madi.ru

evstratova@inbox.ru

Abstract. In the article the analytical and simulation methods to evaluate the rate of convergence of the adaptive algorithm testing are figured out. The IRT-theory is used to simulate the control test procedures, which aims to find out the analytical expressions that define the likelihood of responses of learners with a given level of knowledge at a given level of difficulty of the test task. The main advantages of this theory are robust scoring of the parameters of the complexity of the test task, which do not affected by the properties of persons to be tested. The Monte Carlo method is used to solve the problems of classification of being tested persons, followed by detection of the maximum likelihood method. A rate of convergence is important to compare the performance of different algorithms of test tasks presentation. Fast convergence allows to achieve the specified accuracy in determining the level of education of the tested person with fewer iterations, and therefore spending less time to conduct testing. The advantages of simulation method to evaluation the rate of convergence against the analytical method are shown.

Keywords: adaptive testing, the rate of convergence of the algorithm for generating a test job, analytical and simulation methods to evaluate the rate of convergence.

ISSN 2306-1561
Автоматизация и управление в технических системах (АУТС)
2015. – № 3. – С. 94-103.
DOI: 10.12731/2306-1561-2015-3-9

УДК 004.8:681.3

Аналитические и имитационные методы оценки скорости сходимости адаптивных алгоритмов формирования тестовых заданий

Ягудаев Геннадий Григорьевич

Российская Федерация, доктор технических наук, директор.

Северо-Кавказский филиал ФГБОУ ВПО «Московский автомобильно-дорожный государственный технический университет (МАДИ)», 357340, Российская Федерация, Ставропольский край, г. Лермонтов, ул. Промышленная, д. 20, Тел.: +7 (7935) 3-78-73, http://skfmadi.ru

gena_yagudaev@mail.ru

Ярцев Максим Игоревич

Российская Федерация, аспирант кафедры «Автоматизированные системы управления».

ФГБОУ ВПО «Московский автомобильно-дорожный государственный технический университет (МАДИ)», 125319, Российская Федерация, г. Москва, Ленинградский проспект, д.64, Тел.: +7 (499) 151-64-12, http://www.madi.ru

gizmo1654@mail.ru

Ивахненко Андрей Андреевич

Российская Федерация, кандидат технических наук, доцент кафедры «Менеджмент».

ФГБОУ ВПО «Московский автомобильно-дорожный государственный технический университет (МАДИ)», 125319, Российская Федерация, г. Москва, Ленинградский проспект, д.64, Тел.: +7 (499) 151-64-12, http://www.madi.ru

jointlab@mail.ru

Евстратова Ирина Александровна

Российская Федерация, кандидат технических наук, доцент кафедры «Автоматизированные системы управления».

ФГБОУ ВПО «Московский автомобильно-дорожный государственный технический университет (МАДИ)», 125319, Российская Федерация, г. Москва, Ленинградский проспект, д.64, Тел.: +7 (499) 151-64-12, http://www.madi.ru

evstratova@inbox.ru

Аннотация. В статье сформулированы аналитические и имитационные методы оценки скорости сходимости адаптивного алгоритма тестирования. Для моделирования процедур тестового контроля использована IRT-теория, которая направлена на поиск аналитических выражений, задающих вероятности ответов, обучаемых с заданным уровнем знаний на ТЗ заданного уровня сложности. Основными преимуществами данной теории являются устойчивые статистические оценки параметров сложности ТЗ, на которые не влияют показатели выборки тестируемых. При решении задач классификации тестируемых, используется метод Монте-Карло с последующим распознаванием по методу максимального правдоподобия. Оценка скорости сходимости важна для сравнения эффективности различных алгоритмов предъявления. Быстрая сходимость позволяет добиться заданной точности определения уровня обученности тестируемого за меньшее число итераций, а, следовательно, затратив меньшее время на проведение тестирования. Показаны преимущества имитационного метода оценки скорости сходимости перед аналитическим методом.

Ключевые слова: адаптивное тестирование, скорость сходимости алгоритма формирования тестового задания, аналитические и имитационные методы оценки скорости сходимости.

1. Введение

Принцип работы адаптивного алгоритма предъявления заключается в выборе следующего предъявляемого тестового задания с учетом ответов на предыдущие. В таком случае формирование алгоритма предъявления дает преподавателю больше творческой свободы. Базовая характеристика алгоритма предъявления ТЗ – скорость сходимости алгоритма. Проведя сравнительную оценку скорости сходимости различных алгоритмов предъявления, можно судить об их относительно эффективности. Для количественного определения этой характеристики эффективности различных методов выборки тестовых заданий, следует смоделировать процедуру предъявления и решения тестовых заданий и, в совокупности, педагогических тестов [1 … 15].

Для моделирования процедур тестового контроля наибольший интерес представляет IRT-теория, которая направлена на поиск аналитических выражений, задающих вероятности ответов, обучаемых с заданным уровнем знаний на ТЗ заданного уровня сложности. При решении задач классификации тестируемых можно использовать метод Монте-Карло с последующим распознаванием по методу максимального правдоподобия [3, 6, 7, 8].

IRT-теория служит для оценки латентных свойств тестируемого, базируясь на статистических моделях и логистических кривых задания вероятностей ответов. Основными преимуществами данной теории являются устойчивые статистические оценки параметров сложности ТЗ, на которые не влияют показатели выборки тестируемых [1, 2]. В IRT-теории формируется условная вероятность ответа тестируемым с заданным уровнем знаний Θi при ответе на ТЗ различной сложности β Pi{xij=1|Θi}=f(Θi - β) i=1..N.

Аналогично вводится условная вероятность правильного выполнения j-го задания, сложности βj, различными тестируемыми. Здесь независимой переменной является Θ, а βj - параметр, определяющий трудность j-го задания теста:

Pi{xij=1|βj}=f(Θ- β j) i=1..N, (1)

где xij={0,1}, 1, если ответ i-го испытуемого на j-е задание теста правильный; и 0, если ответ i-го испытуемого на j-е задание, теста неправильный.

N — число испытуемых;

п — количество заданий в тесте.

В теории IRT функции f(β) и φ(Θ) получили название "Item response functions" (IRF). Специальное название имеют и их графики: график функции Pj - это характеристическая кривая j-го задания (ICC), а график функции Pi - индивидуальная кривая i-го испытуемого (PCC). При выборе вида функций Pj и Pi учитываются обстоятельства как эмпирического, так и, математического характера.

Цель адаптивного тестового контроля – повышение эффективности процедуры тестирования в плане сокращения времени и повышения точности идентификации уровня подготовленности тестируемого. В данной работе задача синтеза процедур адаптивного тестового контроля реализована по подобию алгоритмов стохастической аппроксимации, где процедура Роббинса–Монро заключается в нахождении корня уравнения регрессии. Задача построения алгоритмов оценивания уровня обученности сводится к механизмам последовательного выбора сложности предъявляемых тестовых заданий на основе полученных результатов ответов на предыдущие ТЗ [6].

При формировании механизмов предъявления ТЗ вариация реализуется на основе подбора сложности заданий, наиболее близких к уровню обученности текущего тестируемого

β(n+1) = F(n)(β(1),…, β(n)) + ξ (n)(β(1),…, β(n)), (2)

где F(n) –преобразование, которое и определяет механизм предъявления; β(n) - сложность ТЗ; ξ(n)(β(1),…, β(n)) – случайный ответ на ТЗ.

2. Аналитические оценки скорости сходимости

Аналитическая оценка предложенной рекуррентной схемы возможна лишь в частных случаях, когда изначально определяется исходная величина сложности ТЗ β0, а на основании моделирования ответа (да, нет) следующий уровень сложности определяется как:

$$\beta^{(n+1)} = \beta^{(n)} + \frac{1}{n}\left(\xi^{(n)} - \alpha\right), \qquad (3)$$

где ξ(n) – разыгранное значение (1-верно, 0- неверно) случайной величины по закону Бернулли, которое моделирует ответ на ТЗ текущей сложности; β(n) - сложность

ТЗ на заданном шаге рекуррентной схемы; α - порог вероятности ответа на логистической кривой. Т.е. для величины сложности β величина $\xi(n)$ равна

$$P\{\xi(\beta)=0\}=1-L(\beta),\ P\{\xi(\beta)=1\}=L(\beta). \tag{4}$$

Среднее значение величины $\xi(\beta)$ равно $M\{\xi(\beta)\}=R(\beta)$.

Если величина уровня обученности и сложность ТЗ оцениваются в одной количественной шкале, то существует некоторый уровень Θ, такой что $M(\Theta)=\alpha$. Предположим также, что функция $M(\Theta)$ дифференцируема по Θ, а математическое ожидание $M'(\Theta)>0$.

Далее будем полагать, что для сгенерированной последовательности ответов на ТЗ $\{\xi(n)\}$ справедливо свойство независимости, а само распределение последнего ответа зависит лишь от значения сложности последнего ТЗ

$$\begin{aligned}&P\{\xi(n)=1|\beta(1),\ldots,\beta(n),\xi(1),\ldots,\xi(n)\}=R(\beta(n));\\&P\{\xi(n)=0|\beta(1),\ldots,\beta(n),\xi(1),\ldots,\xi(n)\}=1-R(\beta(n)).\end{aligned} \tag{5}$$

В данном случае

$$\lim_{n\to\infty} \boldsymbol{M}\left(\beta^{(n)}-\Theta\right)^2=0, \tag{6}$$

и последовательность $\{\beta(n)\}$ будет сходится к уровню обученности Θ как по вероятности, так и в среднеквадратическом отклонении. Действительно последовательность $x^{(n)}=\boldsymbol{M}\left(\beta^{(n)}-\Theta\right)^2$ сходится к нулю. Так, из выражения (3) можно получить

$$\beta^{(n+1)}-\Theta=\beta^{(n)}-\Theta+\frac{1}{n}\left(\xi^{(n)}-\alpha\right). \tag{7}$$

При этом для среднего значения справедливо и

$$\boldsymbol{M}\left(\beta^{(n+1)}-\Theta\right)^2=\boldsymbol{M}\left(\beta^{(n)}-\Theta\right)^2-\frac{2}{n}\boldsymbol{M}\left(\beta^{(n)}-\Theta\right)\left(\xi^{(n)}-\alpha\right)+\frac{1}{n^2}\boldsymbol{M}\left(\xi^{(n)}-\alpha\right)^2. \tag{8}$$

А если предположить, что $d_n=M[(\beta(n)-\Theta)(\xi(n)-\alpha)]$ и $e_n=M(\xi(n)-\alpha)^2$, то на основании введенных соотношений можно показать, что

$$x^{(n+1)}-x^{(n)}=-2\sum_{i=1}^{n}\frac{d_i}{i}+\sum_{i=1}^{n}\frac{e_i}{i^2}. \tag{9}$$

А так как справедливо неравенство $0\le e_n=M(\xi(n)-\alpha)^2\le 1$, то последовательность $\sum_{i=1}^{n}\frac{e_i}{i^2}$ возрастает. Кроме того, она ограничена, а сам ряд $\sum_{i=1}^{\infty}\frac{e_i}{i^2}$ является сходящимся. Также справедливо и соотношение $\forall\beta\ (\beta-\theta)(R(\beta)-\alpha)\ge 0$, которое приводит к следующему равенству

$$\begin{aligned}&d_n=M[(\beta(n)-\Theta)(\xi(n)-\alpha)]=\\&=M\{M[(\beta(n)-\Theta)(\xi(n)-\alpha)|\beta(1),\ldots,\beta(n),\xi(1),\ldots,\xi(n)]\}=\\&=M[(\beta(n)-\Theta)(R(\xi(n))-\alpha)]\ge 0.\end{aligned} \tag{10}$$

Это в свою очередь приводит к справедливости

$$2\sum_{i=1}^{n}\frac{d_i}{i} = x^{(1)} - x^{(n+1)} + \sum_{i=1}^{n}\frac{e_i}{i^2} \le x^{(1)} + \sum_{i=1}^{n}\frac{e_i}{i^2}. \tag{11}$$

В результате, ряд $\sum_{i=1}^{n}\frac{d_i}{i}$ ограничен и не убывает. В связи с этим ряд $\sum_{i=1}^{\infty}\frac{d_i}{i}$ сходится. Следовательно, справедлив предел $\lim_{n\to\infty} x^{(n+1)} = x^{(1)} - 2\sum_{i=1}^{n}\frac{d_i}{i} + \sum_{i=1}^{n}\frac{e_i}{i^2}$, который доказывает утверждение о сходимости процедуры адаптивного тестового контроля.

3. Имитационная модель оценки скорости сходимости

Однако, аналитические соотношения для оценки скорости сходимости адаптивных алгоритмов и рекуррентных последовательностей дают достаточно грубые оценки. Поэтому на практике более целесообразно использовать имитационные процедуры, которые обладают большей адекватность и точностью, тем более, если при проектировании тестов не требуется оперативность решений по выбору сложностей заданий.

В плане оценки скорости сходимости адаптивных алгоритмов разработана вероятностная рекуррентная схема (рисунок 1) поиска корня уравнения регрессии в условиях шума, построенная по принципу процедуры стохастической аппроксимации Роббинса-Монро [4, 9]. Представленная модель генерирует выборочную траекторию случайной последовательности сложностей предъявляемых тестовых заданий. Для генерации траектории задается начальная сложность ТЗ и уровень знаний [5]. Параметр α является корнем уравнения регрессии, т.е. подбирается такое значение β при котором для заданного уровня знаний Θ вероятность правильного ответа на ТЗ равно 0.5.

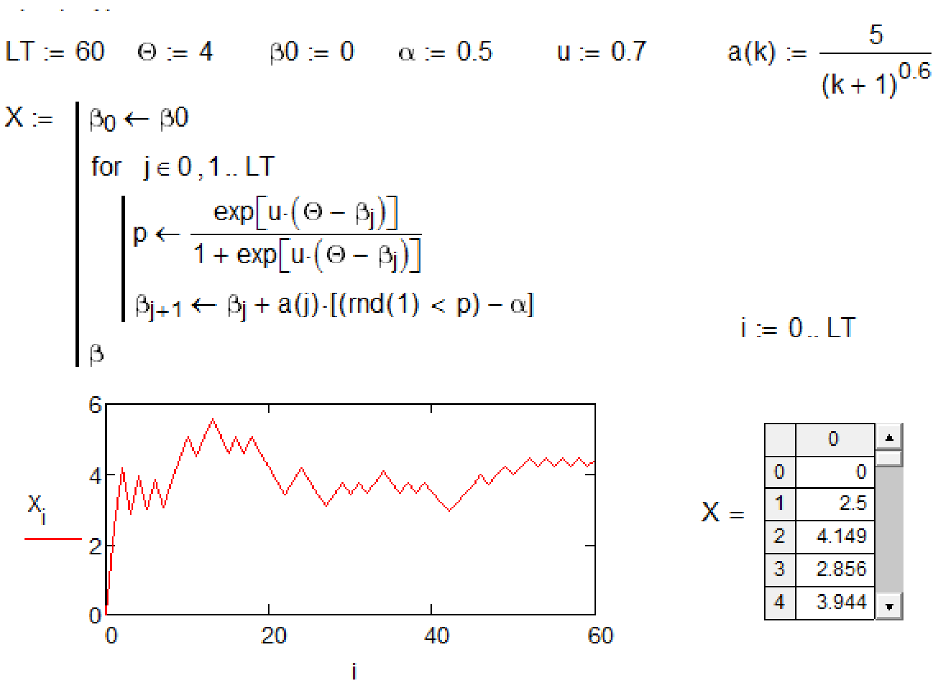
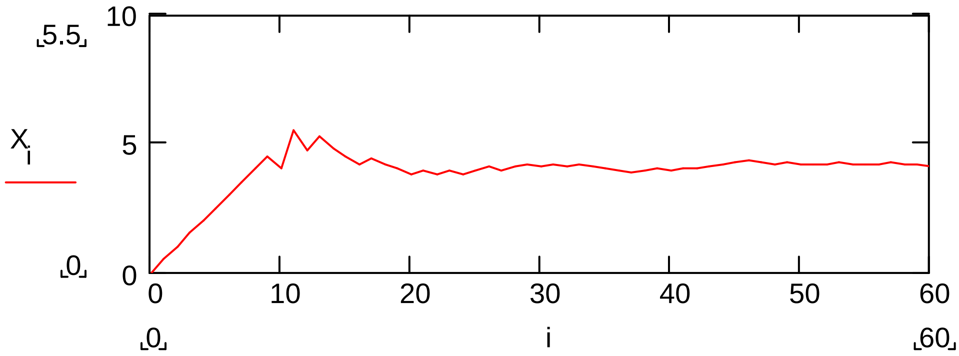


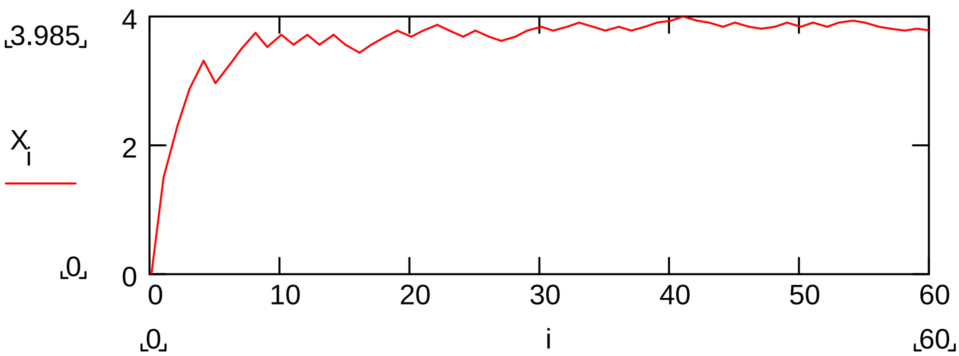
Рисунок 1 – Рекуррентная схема генерации сложности ТЗ

На каждом шаге рекуррентной схемы для текущего значения сложности в данной последовательности на основании логистической кривой вычисляется вероятность ответа на ТЗ и затем путем генерации случайной величины с вычисленным параметром вероятности по закону Бернулли генерируется вариант ответа (да -1, нет – 0). В зависимости от значения разыгранной величины следующее значение сложности либо увеличивается, либо уменьшается на величину приращения сложности u(k) в зависимости от номера шага последовательности k.

Проведена серия экспериментов и показаны различные виды сходимости последовательностей для различных исходных параметров рекуррентной схемы (рисунок 2).



а) β=0, s=0,6



б) β=0, s=0,7

Рисунок 2 - Примеры сходимости последовательностей

Графики показывают, что, начиная с 10-го – 20-го шага, предъявляемая сложность колеблется вокруг истинного значения уровня обученности θ=4. Значение величины шага приращения определяется на основании условий процедур стохастической аппроксимации, которые обеспечивают сходимость.

Последовательность значений приращения шага сложности должна удовлетворять соотношениям

$$\sum_{1<k<\infty} a_k = \infty, \qquad \sum_{1<k<\infty} a_k^2 < \infty \tag{12}$$

Сама процедура Роббинса-Монро предполагает использование рекуррентной схемы определения значений уровня обученности с использованием уравнения баланса.

В ряде работ доказывается нормальность процедуры Роббинса-Монро, когда величина шага приращения сложности равна

$$a_k = \frac{a_0}{k^c + l}, \tag{13}$$

где k - номер шага процедуры; a0, l, c - положительные постоянные;

Анализ предложенной рекуррентной последовательности основывался на проведении повторных экспериментов по методу Монте-Карло с последующим усреднением всех значений сложности тестовых заданий на всех шагах генерируемых последовательностей.

При различных значениях параметров начального значения сложности и величины приращения шага была проведена серия экспериментов, которая показала очень высокую скорость сходимости (рисунок 3).

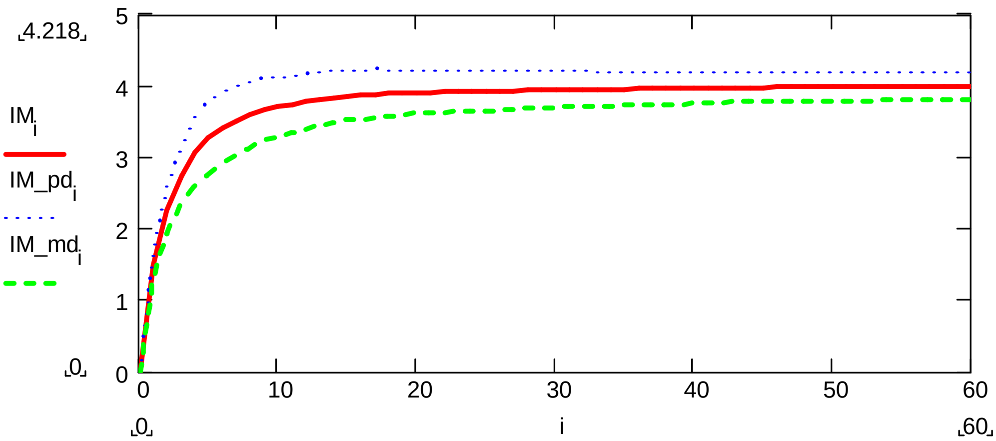


Рисунок 3 – Доверительные интервалы рекуррентной схемы

Из рисунка видно, что после предъявления 15-20 ТЗ значение сложности дает практически несмещенную оценку истинного уровня знаний.

4. Заключение

Оценка скорости сходимости важна для сравнения эффективности различных алгоритмов предъявления. Быстрая сходимость позволяет добиться заданной точности определения уровня обученности тестируемого за меньшее число итераций, а, следовательно, затратив меньшее время на проведение тестирования. Показаны преимущества имитационного метода оценки скорости сходимости перед аналитическим методом.

Список информационных источников

[1] Проблемы создания автоматизированных обучающих и тестирующих систем: Сборник науч. трудов / Ред. колл. Иванченко А.И. и др. – Новочеркасск. – 2001. – 199 с.

[2] Аванесов В.С. Применение тестовых форм в e-learning с проведением дистракторного анализа / Образовательные технологии. – 2013. – № 3. – С. 125-135.

[3] Звонников В.И. Современные средства оценивания результатов обучения: учеб. пособие для студентов высш. учебных заведений / М.: ИЦ «Академия». – 2007. – 118 с.

[4] Челышкова М. Б. Адаптивное тестирование в образовании (теория, методология, технология) / М.: ИЦПКПС. – 2001. – 165 с.

[5] Шамова Т.И., Белова С.Н., Ильина И.В., Подчалимова Г.Н., Худин А.Н. Современные средства оценивания результатов обучения в школе: учебное пособие. – М.: Педагогическое общество России. – 2007. – 78 с.

[6] Никитин М.М, Строганов В.Ю., Карташев М.И. Адаптивный тестовый контроль в системах дистанционного образования / В мире научных открытий. – 2011. – № 9 (21). – С. 118-126.

[7] Современное образование: новые методы и технологии в организации образовательного процесса: материалы междунар. науч.-метод. конф., 31 января - 1 февраля 2013 г., Россия, Томск. / Томск: Томск. Гос. Ун-т систем упр. и радиоэлектроники. – 2013. – 305 с.

[8] Ярцев М. И., Милов Л. Т. Опыт использования центра дистанционного обучения МАДИ для проведения аудиторных занятий на примере дисциплины «Основы теории управления» / Молодой ученый. – 2011. – №4. – Т.3. – С. 29-32.

[9] Николаев А.Б., Ярцев М.И. Способы комбинирования аудиторного и дистанционного компьютерного обучения и тестирования / Электронное обучение и дистанционные образовательные технологии. – 2013. – № 1.

[10] Ярцев М.И., Остроух А.В., Николаев А.Б., Строганов В.Ю. Автоматизация процедуры аттестации персонала дорожно-строительных управлений и предприятий // Промышленные АСУ и контроллеры. – 2015. – №4. – С. 9-16.

[11] Ostroukh A.V., Popov D.I., Demidov D.G., Surkova N.E. Development of the rules base for an expert system choice adaptive learning strategy // ARPN Journal of Engineering and Applied Sciences. 2015. Vol. 10. No 10. pp. 4430-4435.

[12] Ostroukh A.V., Yartsev M.I., Surkova N.E. Automation of Personnel Certification Roadbuilding Departments and Enterprises // International Journal of Advanced Studies (iJAS). 2015. Vol. 5. No 2. pp. 4-11. DOI: 10.12731/ 2227-930X-2015-2-1.

[13] Остроух А.В. Электронные образовательные ресурсы в профессиональном образовании / А.В. Остроух, Н.Е. Суркова. – Saarbrucken, Germany: LAP LAMBERT Academic Publishing, 2011. – 184 p. – ISBN 978-3-8433-2216-4.

[14] Остроух А.В. Корпоративное обучение. Автоматизация процессов управления подготовкой персонала промышленных предприятий / А.В. Остроух, П.А. Петриков, Н.Е. Суркова. – Saarbrucken, Germany: LAP LAMBERT Academic Publishing, 2012. – 147 p. – ISBN 978-3-659-16272-5.

[15] Остроух А.В. Корпоративное обучение. Подготовка персонала предприятий на основе виртуальной модели профессионального сообщества и грид-технологий / А.В. Остроух, М.И. Исмоилов, А.М. Меркулов. – Saarbrucken, Germany: LAP LAMBERT Academic Publishing, 2012. – 129 p. – ISBN 978-3-659-23865-9.