

УДК 681.3

## **Об адаптации обучающих и тестирующих систем переподготовки молодых специалистов на предприятиях аэрокосмического комплекса\***

А.В. Наумов, Сай Кхин Аунг Тинт

Для обучающих систем переподготовки молодых специалистов на предприятиях аэрокосмического комплекса предложены средства статистического анализа результатов работы тестирующих программ, на основании которых построены алгоритмы адаптации исходного контента системы в зависимости от уровня подготовки специалистов. В качестве средств статистического анализа в работе использован аппарат проверки статистических гипотез. Приведены результаты апробации предложенной системы для ряда групп специалистов.

**Ключевые слова:** обучающая система; статистический анализ; обратная связь; статистическая гипотеза; виртуальный учитель.

### **1. Введение.**

В условиях коренной модернизации российской системы образования и перехода на новые стандарты обучения, весьма актуальной и значимой на предприятиях становится задача повышения квалификации и переподготовки вновь поступающих молодых кадров. Организация на предприятиях полноценных образовательных комплексов является сложной задачей, требующей привлечения значительных ресурсов. В связи с этим все большую популярность завоевывают системы автоматизации процесса обучения: компьютерные обучающие и тестирующие программы, системы дистанционного обучения и т.д. Развитие подобных систем в

---

\*Работа выполнена при поддержке РФФИ грант № 05-07-0164а

последние годы приобрело значительные масштабы[1-4]. С массовым распространением и

---

\*Работа выполнена при поддержке РФФИ грант № 05-07-0164а

удешевлением ПЭВМ, развитием глобальной сети Интернет существенно расширяется также потенциальная аудитория дистанционно обучаемых и тестируемых.

На предприятия приходят специалисты из различных образовательных учреждений, осуществляющих подготовку кадров для аэрокосмической отрасли, имеющих различные образовательные программы. Тем более, что в течение ряда лет (до 2016 года) выпуск специалистов будет осуществляться на основе государственных образовательных стандартов 2-го поколения (ГОС ВПО 2), но одновременно, ряд вузов уже начал подготовку по ФГОС. Таким образом, на предприятия могут придти выпускники, обучавшиеся по разным, хотя и близким стандартам.

Вузы, осуществляющие подготовку для предприятий аэрокосмического комплекса расположены в различных городах России( Москва, Самара, Уфа, Рыбинск, Казань) и их учебные планы и программы изучаемых дисциплин могут иметь отличия, отражающие специфику предприятий, расположенных в данном регионе. В ФГОС, разработанных УМО АРК, допускается значительная доля вариативной части (от 25-30% для специалистов и до 50% для направлений бакалавриата ). Вузы сами определяют профили подготовки для бакалавров и специализации для специалистов. Поэтому по перечню и объему дисциплин и базовых частей они могут отличаться.

Очень важно для правильной организации последующего процесса дообучения или переподготовки молодых специалистов предложить грамотную систему испытаний и тестов, позволяющих объективно оценить уровень знаний и способности контингента. При этом, открытым остается вопрос о точности, полученной в результате испытаний оценки и информативности (значимости) тестового материала. Следовательно, сам испытательный процесс должен обладать максимальной адаптивностью, т.е. быть чувствительным к ответам обучаемых. Трек вопросов (ответов) должен формироваться так, чтобы выявить реальный уровень знаний и навыков обучаемого, а также дать оценку самому тестовому контенту. Необходимо автоматически получать рекомендации по изменению тестовых заданий (усложнению, упрощению или исключению) в зависимости от группового и индивидуального уровня подготовки учащихся.

«Обратная связь» компьютерной обучающей и тестирующей системы является актуальным фактором объективного контроля знаний. Если обучающий или тестирующий материал, предлагаемый слушателям, не соответствует уровню обучаемой группы, то крайне

затруднительно делать вывод о его усвоении. При этом априорное деление слушателей на «сильных» и «слабых» не всегда возможно, а в условиях, например, удаленного тестирования или заочного обучения нереализуемо. Большинство современных электронных средств обучения обладают теми или иными механизмами адаптивности. Однако оперативность внесения изменений и гибкость многих из них, к сожалению, невысокая. Отчасти это объясняется необходимостью анализа значительных объемов оперативных данных о динамике обучения или тестирования, что, в свою очередь, требует дополнительного времени, повышенной концентрации вычислительных ресурсов, централизованного управления и ручного вмешательства экспертов предметной области.

Таким образом, актуальным становится разработка алгоритмов статистического анализа результатов тестирования, рассчитанных на максимальную автономность работы (может успешно применяться в распределенных системах) и высокую скорость обработки результатов (“на лету”) в рамках концепции обратной адаптивной связи через виртуального учителя.

## **2. Проверка соответствия сложности тестового контента уровню знаний группы специалистов**

Предлагается что имеется “эталонная” группа из  $n$  специалистов, для которой оцениваются вероятности сдачи теста целиком, а также ряды распределения количества баллов, набранных эталонным слушателем за выполнение каждого задания. Данная эталонная группа формируется экспертами предприятия из числа молодых специалистов в полной мере отвечающих предъявляемым требованиям.

Оценка вероятности успешной сдачи теста задается формулой

$$P_0 = m/n,$$

где  $m$  – число эталонных специалистов, успешно сдавших тест.

Рассмотрим статистику  $T_n$  адаптивного критерия. Пусть случайная величина  $X_n$  – число студентов в оперативно исследуемой группе из  $n$  человек, сдавших тест, тогда

случайная величина  $T_n = \frac{X_n - nP_0}{\sqrt{nP_0(1-P_0)}}$  сходится по распределению[5] к стандартной гауссовской случайной величине  $U \sim N(0,1)$ .

Управляющее воздействие на систему генерации тестового контента формируется следующим образом. Для заданного уровня вероятности  $\alpha$  рассмотрим квантили  $X_{\alpha/2}$  и  $X_{1-\alpha/2}$  стандартной гауссовской случайной величины.

Если  $T_n \in (X_{\alpha/2}, X_{1-\alpha/2})$  то считается, что уровень теста соответствует уровню слушателей в группе и может служить оценкой их знаний. При этом никакие управляющие воздействия на обучающую систему не требуются.

Если  $T_n < X_{\alpha/2}$ , то уровень теста необходимо понизить, а если  $T_n > X_{1-\alpha/2}$ , то уровень теста нужно повышать.

Данный алгоритм имеет несложную программную реализацию, выполняется достаточно быстро, не требует мощных вычислительных ресурсов.

### 3. Проверка соответствия уровня сложности задания тестируемой группе.

Часто бывает, что тестовый контент используется для оценки уровня знаний различных групп слушателей, обучаемых по разным программам, или разными преподавателями. В этом случае очень важно оперативно принимать решение о соответствии отдельных тестовых заданий уровню материала, прочитанного слушателям. Для этого используется аппарат проверки гипотезы о виде закона распределения с помощью критерия Пирсона.

Пусть на основе работы с эталонной группой для каждого тестового задания получены распределения количества баллов набираемых эталонными слушателями за это задание:

$Z_i$	$Z_1$	$Z_2$	...	$Z_{m-1}$	$Z_m$
$p_i$	$p_1$	$p_2$	...	$p_{m-1}$	$p_m$

где  $Z_i, i=1, \dots, m$  – количество баллов, набранных тестируемым за задание (например, от 0 до 1 с шагом 0.2, или 0.25),  $p_i = P(Z=Z_i), i=1, \dots, m$ .

Рассмотрим статистику  $G_n = \sum_{k=1}^m \frac{(n_k - np_k)^2}{np_k}$ , где  $n_k$  - число тестируемых, набравших за это

заданию  $Z_k$  баллов, в группе из  $n$  человек. Величина  $G_n$  сходится при  $n \rightarrow \infty$  по распределению к случайной величине, имеющей распределение  $\chi_{m-1}^2$  (хи-квадрат с  $m-1$  степенью свободы) [5].

Тогда, управляющее воздействие на систему формируется следующим образом. Если реализация статистики  $G_n$  принадлежит доверительной области  $(0, \chi_{\alpha}^2(m-1))$ , где  $\chi_{\alpha}^2(m-1)$ -квантиль уровня  $\alpha$  распределения хи-квадрат с  $(m-1)$  степенью свободы, то уровень сложности задания считается соответствующим уровню группы. В противном случае задание требуется заменить.

#### 4. Примеры использования критериев адаптации.

Пусть следующая таблица отражает исходные данные, полученные по эталонной группе специалистов для некоторого задания.

$X_m$	0	0.25	0.5	0.75	1
$P_m$	0.05	0.1	0.3	0.3	0.25

Здесь  $X_m$  - число баллов, начисленное за это задание, а  $P_m$  - оценка вероятности начисления такого количества баллов, найденная на основе анализа работы эталонной группы специалистов.

В следующей таблице отражены результаты работы тестируемой группы специалистов.

$X_m$	0	0.25	0.5	0.75	1
$n_m$	3	7	12	18	10

Здесь  $n_m$  - число слушателей в тестируемой группе из 50 человек, набравших  $X_m$  баллов за исследуемое задание.

Вычислим значение статистики Пирсона.

$$G_n = \sum_{k=1}^5 \frac{(n_k - np_k)^2}{np_k} = 2.6$$

Выберем уровень значимости  $p = 0,05$ . Тогда доверительная область имеет вид  $(0, 9.49)$ , где  $\chi_{1-p}^2(5-1) = 9,49$  - квантиль уровня  $\alpha = 1 - p = 0.95$  распределения хи-квадрат с 4 степенями свободы.

Проверяем гипотезу

$H_0$ : Уровень сложности исследуемого задания соответствует уровню знаний группы специалистов.

Значение статистики Пирсона принадлежит доверительной области  $G_n = 2.6 < \chi_{1-p}^2(5-1) = 9,49$ , следовательно, проверяемая гипотеза принимается.

Пусть теперь по-прежнему эталонная группа состоит из 50 студентов. Уровень значимости при проверке гипотезы о соответствии сложности контента уровню группы примем равным 0.05.

Оценка вероятности сдачи теста среднестатистическим участником эталонной группы равна  $P_0 = m/n = \frac{40}{50} = 0.8$ .

Пусть  $P$  - вероятность сдачи теста среднестатистическим участником тестируемой группы. Тогда для проверки соответствия уровня сложности контента уровню группы в целом можно проверить гипотезу

$$H_0 = P = P_0$$

Предположим, что в тестируемой группе тест сдали 48 студентов. Вычислим значение статистики  $T_n = \frac{X_n - nP_0}{\sqrt{nP_0(1-P_0)}}$ .

$$T_n = \frac{48 - 50 * 0.8}{\sqrt{50 * 0.8 * 0.2}} = 2.828$$

Доверительная область для проверки гипотезы  $H_0$  при уровне значимости критерия 0.05 имеет вид  $(X_{\alpha/2}, X_{1-\alpha/2})$ , где  $X_{\alpha/2} = -1.92$ , а  $X_{1-\alpha/2} = 1.92$ .

Таким образом,  $T_n = 2.828 > 1.92$ . Следовательно, управляющим воздействием на обучающую систему будет повышение уровня сложности контента.

## 5. Результаты апробации системы обратной связи в обучающих системах переподготовки молодых специалистов на предприятиях аэрокосмического комплекса.

Разработанные критерии адаптации программным образом реализованы в контуре обратной связи (виртуального учителя) обучающей и тестирующей системы, используемой на ряде предприятий аэрокосмического комплекса (ГОС НИИ «Аэронавигация» ФГУП ЦАГИ, ГСКБ ОАО «Алмаз-Антей», ФГУП «Московский институт теплотехники») в процессе дообучения молодых специалистов. Имеется опыт тестирования более двадцати групп специалистов, проходивших повышение квалификации в интерактивном режиме. Результаты тестирования десяти групп по профильной дисциплине предприятия отражены на следующей диаграмме.

	N группы	Количество студентов	объем контента	Число задачи, признанных не соответствующими уровню тестируемых	$T_n$	воверит область уровня 0.95 (-1.96, 1.96)	Изменя контеста
▶	1	20	20	3	-1.05409255338...	$T_n \in (-1.96, 1.96) \dots$	нормальный
	2	20	30	4	1.603567451474...	$T_n \in (-1.96, 1.96) \dots$	нормальный
	3	20	40	10	2.422120283277...	$T_n > 1.96$ - пони...	понижать
	4	20	50	4	0.816496580927...	$T_n \in (-1.96, 1.96) \dots$	нормальный
	5	20	25	2	-2.08893187146...	$T_n < -1.96$ - пов...	повышать
	6	20	30	4	0.768706114785...	$T_n \in (-1.96, 1.96) \dots$	нормальный
	7	20	40	8	2.449489742783...	$T_n > 1.96$ - пони...	понижать
	8	20	30	4	1.673320053068...	$T_n \in (-1.96, 1.96) \dots$	нормальный
	9	20	30	6	1.460593486680...	$T_n \in (-1.96, 1.96) \dots$	нормальный
	10	20	40	3	1.054092553389...	$T_n \in (-1.96, 1.96) \dots$	нормальный



Рис. 5.1. Экранная форма апробации тестов по профильной дисциплине

Результаты тестирования эталонной группы специалистов по профильной дисциплине представлены на рис. 5.2. Вертикальная линия красного цвета наглядно показывает “пороговое значение” оценки, установленное экспертом. Цифра на стрелке соответствует числу слушателей эталонной группы, успешно завершивших испытание (сдавших тест).



Рис. 5.2. Графическая иллюстрация результатов апробации тестов. Эталонная группа.

На рис. 5.3 ,5.4 представлены графические иллюстрации результатов тестирования двух групп молодых специалистов с той же “пороговой оценкой” успешной сдачи теста, что и в эталонной группе. Здесь цифрой под красной стрелкой показано количество слушателей, сдавших тест и рекомендация “виртуального учителя” о действиях, которые необходимо предпринять с контентом, т.е. не изменять, повысить или понизить сложность предлагаемых в процессе тестирования вопросов.



Рис. 5.3. Результаты тестирования по профильной дисциплине, группа (1)



Рис. 5.4. Результаты тестирования по профильной дисциплине, группа (2)

Результаты апробации показывают, что уровень подготовки молодых специалистов различный и не всегда соответствует требованиям, предъявляемым экспертами предприятия. Система обратной связи, основанная на использовании аппарата проверки статистических гипотез, позволяет адаптировать тестовый контент и обеспечить, таким образом, максимальную объективность в проверке уровня знаний и способности молодых специалистов к дальнейшему дообучению на предприятии.

#### **4. Заключение.**

В работе на основе аппарата проверки статистических гипотез решена актуальная задача построения контура обратной связи в тестирующих программах обучающих систем. Разработанные критерии адаптации позволяют эффективно корректировать в автоматическом режиме исходный тестовый контент и обеспечивать максимальную объективность получаемых оценок уровня знаний и способностей тестируемых. Апробация программной реализации разработанной адаптивной тестирующей системы при организации процесса доучения и переподготовки молодых специалистов предприятий аэрокосмического комплекса показала эффективность контура обратной связи в подобных обучающих и тестирующих системах.

#### **Библиографический список**

1. *Юрков А.В.* Обзор отечественных систем дистанционного обучения. // Компьютерные инструменты в образовании. 2003. № 1. С. 8-14.
2. *Сай Кхин Аунг Тинт* Архитектура и сценарии работы обучающей системы. // Вестник Московского авиационного института. 2009. Т. 16. №7. с. 90-95.
3. *Наумов А.В., Сай Кхин Аунг Тинт* Алгоритм статистического анализа сложности тестов в адаптивной обучающей системе. // Материалы VIII Международной конференции по неравновесным процессам в соплах и струях (NPNJ'2010).- М.:Изд-во МАИ-ПРИНТ. 2010. С. 593-594.
4. *Кибзун А.И., Наумов А.В.* Электронный учебно-методический комплекс по курсу «Теория вероятностей и математическая статистика» для дистанционного обучения. // Вестник компьютерных и информационных технологий. 2008. № 10. С. 43-48.
5. *Панков А.Р., Платонов Е.Н.* Практикум по математической статистике. –М: МАИ. 2006.

#### **Сведения об авторах**

**Наумов Андрей Викторович**, доцент Московского авиационного института (государственного технического университета), к.ф.-м.н.

Тел. (499)158-4113, e-mail [naumovav@mail.ru](mailto:naumovav@mail.ru)

**Сай Кхин Аунг Тинг**, аспирант Московского авиационного института  
(государственного технического университета).

Тел. (499)158-4998, e-mail [mar@k806.mainet.msk.su](mailto:mar@k806.mainet.msk.su)