

УДК 656.7.071:658.386

РАЗРАБОТКА АЛГОРИТМА АДАПТИВНОГО ТЕСТИРОВАНИЯ ДЛЯ АВТОМАТИЧЕСКИХ ОБУЧАЮЩИХ СИСТЕМ ПРОФЕССИОНАЛЬНОЙ ПОДГОТОВКИ ЛЕТНОГО СОСТАВА

д.в. АЙДАРКИН, с.г. КОСАЧЕВСКИЙ

Статья представлена доктором технических наук, профессором Ципенко В.Г.

Разработан алгоритм адаптивного тестирования для использования в автоматизированных обучающих системах профессиональной подготовки летного состава гражданской авиации, обеспечивающий высокую точность измерения уровня знаний обучающихся.

В настоящее время для профессиональной подготовки летного состава все более широкое применение находят автоматизированные обучающие системы (АОС). Такой подход позволяет уйти от модели авторитарного обучения, когда основная роль в учебном процессе отводится преподавателю, а предъявление содержания обучения, управление процессом обучения, его активизация осуществляется преподавателем в вербальной форме, в условиях малопродуктивного с точки зрения эффективности обучения "рассеянного информационного процесса". Управление учебным процессом в этом случае осуществляется преподавателем путем "навязывания" обучающимся своего темпа и траектории обучения. Кроме того, вербальному способу предъявления учебного материала часто придается несвойственная его реальным возможностям функция процесса обучения мыслительного типа, тогда как большинство обучающихся остаются на уровне только информационного восприятия, т.к. они не успевают осмысливать предъявляемую им информацию, понять сущность излагаемого и выполняют только механические конспектирование [1].

Однако "платой" за преимущества обучения с использованием АОС является ряд принципиально новых требований как к организации и методике проведения процесса обучения, так и к самому обучающемуся, который должен обладать более высоким уровнем развития потребностно-мотивационной, познавательной и эмоционально-волевой сфер. С целью компенсации различного уровня развития этих качеств у обучающихся одной из важнейших функций АОС является контрольно-коррекционная. Это связано еще и с тем, что при самостоятельном обучении обучающийся сам контролирует свои познания, что может привести к возникновению так называемой "иллюзии усвоения", когда при первоначальном ознакомлении с материалом у него создается вполне добросовестное заблуждение об усвоении изученного материала и он прекращает предлагаемую ему учебную деятельность либо выполняет ее неполноценно [2].

Поэтому в структурной схеме модели самообучения обязательно должен быть предусмотрен контрольно-коррекционный этап, на котором осуществляется внешний текущий контроль и формирование коррекционной информации для исправления выявленных недостатков. Благодаря внешнему объективному текущему контролю организуется цикл повторного произвольного коррекционного учения. Для повышения эффективности обучения часто предусматривается организация второго контура внешнего текущего самоконтроля и само-коррекции, который, в свою очередь, должен дать толчок для запуска механизма внутреннего саморазвития. В результате получается самообучение с трехконтурным контролем [1].

Для организации внешнего объективного текущего контроля в структуре АОС должна быть предусмотрена система тестового адаптивного контроля (СТАК), которая включает в себя базу контрольных тестов и алгоритмическое обеспечение адаптивного тестирования

обучаемых. СТАК обеспечивает решение задачи адаптивного тестирования, представляющей собой процесс автоматизированного подбора тестовых заданий такого уровня трудности, при котором достигается максимальная точность измерения уровня знаний обучаемого.

При разработке тестовых заданий для контроля качества обучения летного состава необходимо использовать рекомендуемую ИКАО инструктивную модель совершенствования учебного процесса (модель ISD – Instruction Systems Design), в основу которой положено компетентностное обучение и оценивание (competency-based training and assessment) [3].

В модели ISD рекомендуется применение тестов, соотнесенных с критериями, когда результат сравнивается с объективным стандартом, а не с результатами других обучающихся. Если при проведении тестирования обучающийся демонстрирует, что указанный стандарт им достигнут или превзойден, то он получает зачет независимо от сравнения с баллами других обучающихся. При этом результаты тестирования оцениваются по двухбалльной системе "зачет – незачет".

Разработку тестов рекомендуется выполнять до того, как будут подготовлены учебные модули и учебные и учебно-методические пособия. Благодаря этому обеспечивается ориентация на выполнение учебной задачи и предотвращается естественная тенденция составления таких тестов, в которых основное внимание уделяется содержанию учебных материалов, а не эффективности работы.

Каждый тест следует апробировать и убедиться, что он отражает условия и процесс осуществления операций и предусмотренные стандарты. В процессе апробации его необходимо применить к выборочной группе квалифицированных и неквалифицированных участников из числа целевого контингента обучающихся. Большой процент квалифицированного персонала должен пройти данное тестирование, а больший процент неквалифицированного персонала не должен его пройти. Если профессиональные исполнители не получили достаточно высоких баллов при тестировании, то следует убедиться в том, что цели курса, к которому относится данный тест, отвечают реальным потребностям, т.е. что данная задача действительно является частью профессиональных обязанностей.

С целью реализации контрольно-коррекционной функции в разрабатываемых АОС специалистами УВАУ ГА была разработана СТАК по дисциплине "Аэродинамика и динамика полета", позволяющая оперативно проверять знания курсантов-пилотов. При ее разработке использовалась теория Item Response Theory (IRT), основными преимуществами которой являются [4]:

- устойчивость и объективность оценок параметра, характеризующего уровень знаний обучающихся;
- устойчивость и объективность оценок параметра, определяющего трудность заданий теста;
- использование одной шкалы для измерения уровня знаний обучающихся и трудности заданий теста, что позволяет сравнить эти величины.

В пробном тестировании, которое проходило в мае 2007 года, принимали участие 32 курсанта. Каждому испытуемому был предложен в случайном порядке набор из 39 вопросов по всем темам раздела "Аэродинамика и динамика полета легких одномоторных самолетов".

Следует отметить, что тестовые задания составлялись с некоторым избытком, чтобы после проведения тестирования контрольных групп и статистической обработки полученных данных сформировать тест из тех вопросов, которые могли бы хорошо различать курсантов с различным уровнем подготовки по указанной дисциплине. Исходный тест включал задания различного уровня трудности, начиная с самых легких (основные понятия и определения аэrodинамики), на которые должны были ответить правильно почти все испытуемые, и кончая самыми трудными, на которые могут дать верный ответ лишь немногие курсанты.

Каждое тестовое задание оценивалось по двухбалльной системе (1 – дан верный ответ, 0 – дан ошибочный ответ). Исходя из общей суммы набранных баллов определялась итоговая оценка по четырехбалльной шкале.

Предварительный анализ полученных данных показал, что тестовые задания № 11 и № 32 следует убрать из дальнейшего рассмотрения, т.к. ни один из испытуемых не смог дать верный ответ на эти вопросы. Оставшиеся задания были сведены в матрицу тестовых результатов, с помощью которой построена гистограмма относительных частот по данному распределению выборки (рис. 1) и вычислены точечные оценки выборки.

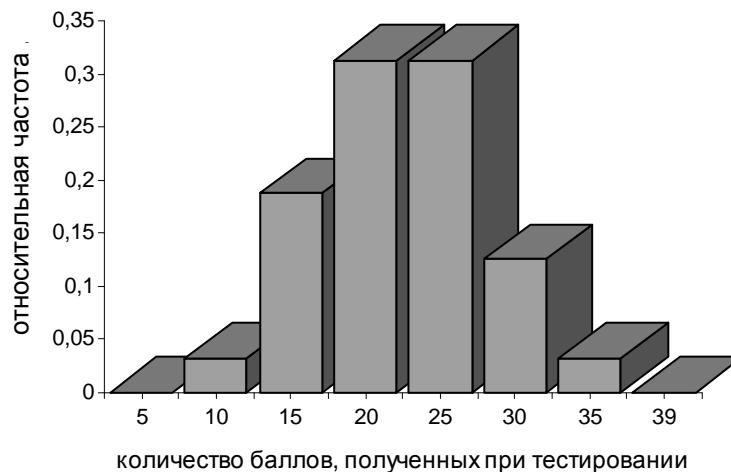


Рис. 1. Гистограмма относительных частот контрольной выборки курсантов

Так как полученное значение выборочного среднего ($\bar{X} \approx 20,2$) примерно равно утроенному значению выборочного стандартного отклонения ($S_x \approx 5,72$), то можно говорить, что полученное распределение близко к нормальному. Полученное значение выборочного стандартного отклонения также указывает на хорошую различающую способность составленных тестовых заданий, т.к. увеличение вариации создает необходимые предпосылки для увеличения надежности теста в целом.

С помощью показателей связи между результатами тестирования по конкретному заданию с суммарным баллом по всему тесту (использовались р-бисериальный коэффициент корреляции и классический коэффициент корреляции Пирсона) был рассмотрен вопрос о валидности каждого задания теста. Полученные расчетные данные для первых десяти тестовых заданий представлены в табл. 1.

В случае значительной положительной корреляции ($r > 0,3$) есть основания считать тестовое задание валидным, т.е. оно отличает испытуемых с высоким уровнем знаний от слабо подготовленных курсантов. Из рассмотренных в табл. 1 первых десяти заданий теста следует удалить задания № 2 и № 10, т.к. они обладают почти нулевой различающей способностью ($r_2 = 0,07$; $r_{10} = 0,04$), а валидность заданий № 3 и № 8 поставить под сомнение.

Затем был рассмотрен вопрос о корреляции заданий не только с суммарным баллом, но и между собой. Полученные результаты корреляционного анализа были сведены в матрицу R (в табл. 2 представлена корреляционная матрица R для первых десяти заданий теста), а затем произведена окончательная "чистка" теста.

Анализ корреляционной матрицы показал, что задания № 2, № 3, № 8 и № 10 необходимо изъять из теста, так как соответствующие им столбцы и строки матрицы R содержат отрицательные значения.

Таблица 1

Номер задания	1	2	3	4	5	6	7	8	9	10
Количество правильных ответов на задание	30	13	29	1	3	16	28	27	13	17
Среднее арифметическое по всему тесту у курсантов, успешно ответивших на данное задание, M_b	20,7	20,7	20,6	30	29	23	21,2	20,7	24,1	20,4
Среднее арифметическое по всему тесту у курсантов, неверно ответивших на данное задание, M_n	12,5	19,9	16,3	19,9	19,3	17,4	13,5	17,6	17,6	20
r -бисериальный коэффициент корреляции, r_{pb}	0,35	0,07	0,22	0,31	0,49	0,49	0,44	0,2	0,56	0,04
Выборочная дисперсия по заданию, SS_{xj}	0,06	0,25	0,09	0,03	0,09	0,26	0,11	0,14	0,25	0,26
Выборочная ковариация (корреляционный момент), SP_{xjy}	0,5	0,2	0,38	0,32	0,85	1,44	0,87	0,42	1,62	0,11
Коэффициент корреляции Пирсона, r_{xy}	0,35	0,07	0,22	0,31	0,5	0,49	0,45	0,2	0,57	0,04

Таблица 2

Номера заданий	1	2	3	4	5	6	7	8	9	10
1	1	0,21	-0,08	0,05	0,08	0	0,68	-0,11	0,21	-0,24
2	0,21	1	-0,17	-0,15	-0,27	0,06	0,12	-0,17	0,09	0,14
3	-0,08	-0,17	1	0,06	0,1	-0,11	-0,12	-0,14	0,05	0,13
4	0,05	-0,15	0,06	1	0,56	0,18	0,07	0,08	0,22	-0,19
5	0,08	-0,27	0,1	0,56	1	0,11	0,12	0,14	0,17	-0,13
6	0	0,06	-0,11	0,18	0,11	1	0	0,09	0,19	-0,06
7	0,68	0,12	-0,12	0,07	0,12	0	1	-0,16	0,31	-0,17
8	-0,11	-0,17	-0,14	0,08	0,14	0,09	-0,16	1	-0,17	0,11
9	0,21	0,09	0,05	0,22	0,17	0,19	0,31	-0,17	1	0,14
10	-0,24	0,14	0,13	-0,19	-0,13	-0,06	-0,17	0,11	0,14	1

Для шкалирования оставшихся тестовых заданий по уровню их трудности (β_j) была использована двухпараметрическая модель Г. Раша [5], согласно которой вероятность правильного ответа (P_{ij}) для i -го испытуемого на задание j определяется выражением:

$$P_{ij} = \frac{e^{a_j(\theta_i - \beta_j)}}{1 + e^{a_j(\theta_i - \beta_j)}},$$

где: a_j – параметр, определяющий дифференцирующую способность j -го задания; θ_i – параметр, характеризующий уровень знаний i -го испытуемого; β_j – параметр, определяющий уровень трудности j -го задания.

Соответствующий график зависимости вероятности правильного ответа на шестое задание от уровня знаний испытуемого приведен на рис. 2.

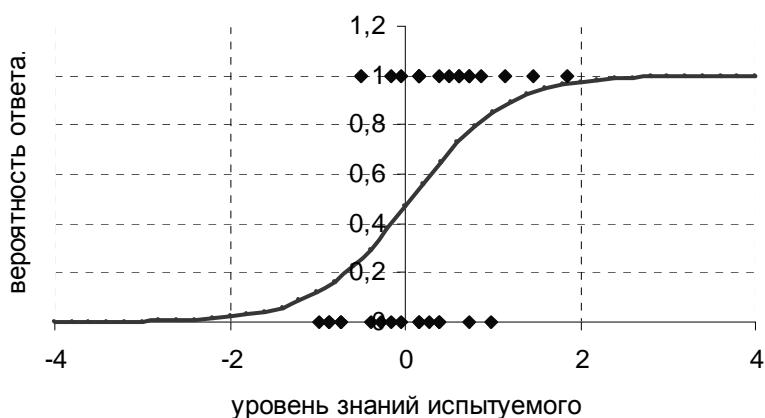


Рис. 2. Зависимость вероятности правильного ответа на задание №6 от уровня знаний испытуемого

С помощью метода наибольшего правдоподобия были вычислены числовые значения параметров a_j и β_j для каждого тестового задания (в табл. 3 представлены данные для оставшихся заданий из первых десяти, расположенных в порядке возрастания трудности задания).

Таблица 3

Номера заданий	1	7	6	9	5	4
Разрешающая способность j -го задания, a_j	2,17	2,43	1,87	2,30	91,36	2,31
Трудность j -го задания, β_j	-1,40	-0,97	0,07	0,29	1,07	1,77

Зная значения уровня трудности β_j для всех заданий теста, сравнительно легко реализовать оптимизационный подход при тестировании. Суть этого подхода заключается в том, что количество предлагаемых заданий в teste резко сокращается, причем каждому испытуемому предлагаются лишь те задания, уровень трудности которых примерно равен уровню знаний испытуемого. В этом случае будут исключены ситуации, когда хорошо подготовленному курсанту будут предложены только легкие задания, или противоположный случай, когда слабо подготовленный курсант практически не может правильно ответить на трудные задания теста.

Если предварительные сведения об уровне подготовленности курсанта отсутствуют, то тестирование можно проводить по следующему алгоритму. В начале тестирования всем испытуемым предлагаются задания средней сложности, а затем сложность заданий увеличивается или уменьшается в зависимости от правильности ответов на предыдущие задания до тех пор, пока не будет достигнут момент стабилизации (т.е. трудность предложенного вопроса приблизительно равна уровню знаний тестируемого).

Результатом выполненной работы является апробированный алгоритм проведения испытаний прототипа теста на репрезентативной выборке обучаемых, селекции неинформативных тестовых заданий, ранжирования оставшихся заданий по уровню трудности β_j или разрешающей способности a_j и формирования эффективного теста с реализацией оптимизационного подхода при тестировании.

ЛИТЕРАТУРА

1. Репьев Ю.Г. Интерактивное самообучение. – М.: Логос, 2004.

-
2. **Беспалько В.П.** Основы теории педагогических систем / Проблемы и методы психолого-педагогического обеспечения технических обучающих систем. – Воронеж: Издательство Воронежского университета, 1977.
3. Training. Procedures for Air Navigation Services. Doc 9868. First Edition. ICAO. 2006.
4. **Аванесов В.С.** Основы научной организации педагогического контроля в высшей школе. – М.: МИСиС, 1989.
5. **Baker F.** The Basics of Item Response Theory. – College Park, MD.: ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, 2001.

ADAPTIVE TEST ALGORITHM DESIGN FOR FLIGHT PERSONNEL COMPUTER BASED TRAINING APPLICATION

Aidarkin D.V., Kosachevskiy S.G.

The adaptive test algorithm for civil aviation flight personal computer based training providing high accuracy measurements of trainees' knowledge level has been devised.

Сведения об авторах

Айдаркин Дмитрий Викторович, 1969 г.р., окончил МГУ (1994), доцент кафедры естественно-научных дисциплин УВАУ ГА, автор 10 научных работ, область научных интересов – разработка компьютерных обучающих систем и тестирующих программ.

Косачевский Сергей Григорьевич, 1951 г.р., окончил КИИ ГА (1975), кандидат технических наук, доцент, проректор по научной работе УВАУ ГА, автор более 70 научных работ, область научных интересов – профессиональная подготовка летного состава гражданской авиации, применение в учебном процессе новых технологий обучения и методов математического моделирования, практическая аэродинамика.