

ПРИНЦИПЫ ПОСТРОЕНИЯ СЕАНСА ТЕСТИРОВАНИЯ В КОМПЬЮТЕРНОЙ СИСТЕМЕ ТЕСТИРОВАНИЯ ЗНАНИЙ OPENTEST2

Рассматриваются вопросы формирования частично адаптированного сеанса тестирования в компьютерной системе тестирования знаний OpenTEST2. Сеанс тестирования формируется как репрезентативная выборка из структурированной базы тестовых заданий с использованием методов дискретной оптимизации. Даются рекомендации относительно формирования структуры базы данных тестовых заданий системы OpenTEST2.

Введение

В Харьковском национальном университете радиоэлектроники (ХНУРЭ) разработана и в течение нескольких лет находится в эксплуатации компьютерная система тестирования знаний студентов OpenTEST, а с 2007 года – вторая версия этой системы OpenTEST2. В ней реализуется частично адаптированный алгоритм проведения контроля знаний со случайной выборкой заданий из тестовой базы данных на основе модели учебного материала. Адаптивность состоит в том, что тематическая структура набора тестовых заданий для каждого тестируемого, а также их качественные и количественные характеристики зависят от модели учебного материала.

Тестирование в системе OpenTEST2 осуществляется путем организации сеанса тестирования для каждого тестируемого. Сеанс тестирования представляет собой репрезентативную выборку из генеральной совокупности теста на основе его содержательной структуры, максимально покрывающей модель предметной области.

С точки зрения теории педагогических измерений сеанс тестирования – это фрагмент педагогического теста как системы заданий возрастающей трудности, специфической формы, которая позволяет качественно и эффективно измерить уровень и структуру подготовленности испытуемых. Составной единицей педагогического теста является тестовое задание, отвечающее требованиям к заданиям в тестовой форме и статистическим требованиям известной трудности, дифференцирующей способности (вариации баллов), и положительной корреляции с результатами по тесту в целом [1].

Сеанс характеризуется длиной (количеством N тестовых заданий в сеансе), временем, отводимым на сеанс, и количеством попыток прохождения теста в рамках одного сеанса. Хотя задания в сеансе нумеруются в порядке возрастания трудности, тестируемый может отвечать на них в произвольном порядке.

Формой тестовых заданий в системе OpenTEST2 являются вопросы закрытого типа (типа «выбор одного из нескольких», «выбор нескольких из нескольких», «соответствие») и открытого типа (свободный ввод короткого ответа) с произвольным назначением весомозначности тестовых заданий в баллах в рамках одного теста [2].

Трудность и дифференцирующая способность тестовых заданий

В педагогических измерениях для количественной оценки содержания теста разными авторами используются понятия трудности, сложности и весомозначности тестовых заданий.

1. Весомозначность тестового задания обычно характеризуется баллом за правильный ответ и отражает субъективный взгляд автора теста на относительное влияние конкретного тестового задания на суммарную оценку за ответы на все предложенные тестовые задания.

2. Сложность тестового задания характеризуется количеством умственных или вычислительных операций, необходимых для получения правильного ответа. Эта мера широко используется в учебном процессе при нормировании времени, необходимого на усвоение определенного количества учебного материала.

3. **Трудность** тестового задания – это эмпирическая оценка отношения тестируемых к конкретному тестовому заданию путем правильных или неправильных ответов на него. В качестве меры трудности тестового задания принято использовать частоту (вероятность) получения правильного ответа на i -е тестовое задание (p_i) :

$$p_i = \frac{R_i}{NN_i},$$

где R_i – количество правильных ответов на i -е тестовое задание; NN_i – количество участников i -го тестового задания в испытаниях.

Следует отметить, что p_i на самом деле характеризует легкость тестовых заданий, поэтому между истинной трудностью тестового задания и величиной p_i существует обратная пропорциональная зависимость, т.е. чем меньше значение p_i , тем труднее тестовое задание.

При достаточно большой выборке испытуемых p_i можно считать вероятностью получения правильного ответа, p_i характеризуется положительным числом в интервале от 0 до 1. В системах компьютерного тестирования знаний [3] при анализе трудности тестовых заданий принято выделять три уровня трудности:

- легкие ($0,6 < p_i \leq 0,8$);
- средние ($0,4 < p_i \leq 0,6$);
- трудные ($0,2 < p_i \leq 0,4$).

Задания, у которых $p_i < 0,2$ (очень сложные) и $p_i > 0,8$ (очень легкие), не рекомендуется использовать в качестве тестовых, если они не несут какой-либо специальной нагрузки. Необходимо иметь в виду, что трудность, как эмпирическая мера количественной оценки тестовых заданий, требует предварительной апробации на достаточно представительной типичной выборке обучаемых. И только после этого тестовые задания могут быть отнесены к определенному уровню трудности.

В качестве меры дифференцирующей способности i -го тестового задания (вариации тестовых баллов) при дихотомической оценке правильных и неправильных ответов принято использовать его дисперсию $D_i = p_i * q_i$, где $q_i = 1 - p_i$. Таким образом, наилучшей дифференцирующей способностью обладают тестовые задания среднего уровня трудности, а максимальная дисперсия будет при $p_i = 0,5$.

Модель тестовой базы данных

В компьютерной системе тестирования знаний OpenTEST2 принята четырехуровневая иерархическая модель предметной области и соответствующая иерархическая структуризация тестовых заданий (вопросов):

«КАТЕГОРИЯ ТЕСТА» -> «ИМЯ ТЕСТА» -> «ИМЯ ТЕМЫ» -> «ВОПРОСЫ ТЕМЫ» -> «КАТЕГОРИЯ ТЕСТА»

– это организационный уровень иерархии, который формируется для каждой конкретной инсталляции OpenTEST2 администратором системы. Система OpenTEST2 адаптирована для использования в учебных учреждениях со сложной структурой, поэтому организационное понятие «категория» связано с организацией учебного процесса в вузе. Под «категорией» обычно понимают кафедру или факультет (деканат). Такая категория позволяет пользователям значительно сократить время поиска тестов в базе данных системы OpenTEST2.

«ИМЯ ТЕСТА» обычно соответствует названию учебной дисциплины или ее части. Доступ к тесту (при организации сеанса тестирования) осуществляется по категории теста и имени теста.

«ИМЯ ТЕМЫ» обычно соответствует структурной единице учебной дисциплины, определенной в ее рабочей программе, и используется при формировании сеанса тестирования из тестовой базы данных.

«ВОПРОСЫ ТЕМЫ» – это тестовые задания (вопросы) разных форм, содержащие текстовые, графические и мультимедийные компоненты.

Три нижних уровня иерархии модели предметной области образуют содержательную структуру теста.

В соответствии с моделью предметной области и принятой системой градаций трудностей тестовых заданий, в системе OpenTEST2 формируется тестовая база данных, каждый тест в которой состоит из L тем, а каждая тема содержит K подразделов (по количеству уровней трудности тестовых заданий). Каждый подраздел содержит тестовые задания (вопросы), трудность которых находится в пределах, заданных для этого уровня трудности. На рис. 1 показана структура базы данных для теста, содержащего 10 тем по три уровня трудности в каждой теме. На данном рисунке $t_{i,j}$ обозначает количество тестовых заданий соответствующего уровня трудности в каждой теме. Исходя из рекомендаций по общей длине теста целесообразно, чтобы соблюдалось условие $t_{i,j} > 10$.

Тестовые задания	ТЕМЫ									
Уровни трудности	1	2	3	4	5	6	7	8	9	10
1 (легкие)	$t_{1,1}$	$t_{1,2}$	$t_{1,3}$	$t_{1,4}$	$t_{1,5}$	$t_{1,6}$	$t_{1,7}$	$t_{1,8}$	$t_{1,9}$	$t_{1,10}$
2 (средние)	$t_{2,1}$	$t_{2,2}$	$t_{2,3}$	$t_{2,4}$	$t_{2,5}$	$t_{2,6}$	$t_{2,7}$	$t_{2,8}$	$t_{2,9}$	$t_{2,10}$
3 (трудные)	$t_{3,1}$	$t_{3,2}$	$t_{3,3}$	$t_{3,4}$	$t_{3,5}$	$t_{3,6}$	$t_{3,7}$	$t_{3,8}$	$t_{3,9}$	$t_{3,10}$

Рис. 1. Структура теста в базе данных в системе OpenTEST2

Покрывание учебного материала тестовыми заданиями

Для численной оценки степени покрытия учебного материала тестовыми заданиями определим $C_{i,j}$ – коэффициент покрытия учебного материала j -й темы содержательной структуры теста, где $i = 1, K$, $j = 1, L$. Он является чисто качественной характеристикой и отражает субъективную точку зрения автора теста на степень покрытия содержательной структуры j -й темы теста заданиями разных уровней трудности. Коэффициент $C_{i,j}$ задается автором теста для каждого уровня трудности темы или для всего теста целиком. В дальнейшем для простоты изложения будем считать $C_{i,j}$ равными для всех j тем теста, таким образом: $C = (c_1, c_2, \dots, c_K)$.

Определим, что предметная область покрывается полностью (на 100%), если в сеанс тестирования из каждой темы попадает хотя бы по одному заданию каждого уровня трудности. На практике это часто недостижимо, потому что длина сеанса является ограниченной и условие $N \geq K * L$ трудновыполнимо. Но, с другой стороны, правильный ответ на задание 3-го уровня трудности предполагает наличие у студентов какого-то минимума знаний учебного материала, покрываемого заданиями 1-го и 2-го уровней трудности. Например, автор теста может принять, что задания первого уровня покрывают 30% учебного материала ($c_1=0,3$) данной темы или теста в целом, второго – 40% ($c_2=0,4$), третьего – 50% ($c_3=0,5$). Любые два уровня в сумме не покрывают 100% учебного материала, а все три уровня в сумме покрывают 100%, но это не арифметическая сумма, а учет взаимного «перекрывания» учебного материала различных уровней трудности.

Определим пересечение («перекрывание») уровней трудности E сумму частных коэффициентов пересечения e_i для каждого уровня трудности:

$$E = (e_1, e_2, \dots, e_K), \quad E = \sum_{i=1}^K e_i, \quad E = \sum_{i=1}^K c_i - 1,$$

где K – количество уровней трудности тестовых заданий.

Значения отдельных e_i назначаются экспертом по аналогии c_i . Значение $E < 0$ говорит о неправильном выборе автором (экспертом) теста значений c_i . Если $E = 0$, то «перекрывание» между заданиями разного уровня трудности отсутствует, а это не совсем точно отражает содержательную структуру учебного материала (от простого к сложному).

Дополнительно определим групповые коэффициенты пересечения нескольких уровней трудности учебного материала для одной темы:

$$c_{i,j} = \sum_{i=1}^{i,j} c_i - \sum_{i=1}^{i,j} e_i.$$

Например, для $c_1=0,3$, $c_2=0,4$, $c_3=0,5$ $E = (0,3 + 0,4 + 0,5) - 1 = 0,2$.

При этом можно определить $e_1 = 0,06$, $e_2 = 0,07$, $e_3 = 0,07$.

Следовательно, групповые коэффициенты покрытия вычисляются:

$c_{1,2} = (0,3 + 0,4) - (0,06 + 0,07) = 0,57$, $c_{1,3} = (0,3 + 0,5) - (0,06 + 0,07) = 0,67$,

$c_{2,3} = (0,4 + 0,5) - (0,07 + 0,07) = 0,76$, $c_{1,2,3} = 1$ (по определению).

Алгоритм формирования сеанса тестирования

Исходя из сказанного выше, в сеанс тестирования из тестовой базы данных должны включаться задания разных уровней трудности с максимальной дифференцирующей способностью и максимальным покрытием всех разделов учебного материала.

Формальная постановка задачи: построить выборку, содержащую N тестовых заданий с максимальными значениями дисперсии D_i и коэффициентов покрытия учебного материала C_j , где $i = \overline{1, N}$, $j = \overline{1, L}$.

Ограничения: количество тем в тесте L , количество уровней сложности в каждой теме K , из каждого уровня сложности любой темы в сеанс тестирования выбирается одно тестовое задание, при условии $K * L \leq N$.

Система OpenTEST2 построена таким образом, что все параметры сеанса тестирования и тестовой базы данных, за исключением количества уровней трудности K , задаются пользователем произвольно. Поэтому за базовый параметр при формировании сеанса тестирования выберем количество уровней трудности тестовых заданий.

Для решения поставленной задачи будем использовать методы дискретной оптимизации. Определим множество вариантов сеансов тестирования как K -мерное пространство $M(m_1, m_2, \dots, m_K)$ с ограничениями $m_1 + m_2 + \dots + m_K \leq N$ и $0 \leq m_i \leq L$, где m_i – количество тестовых заданий соответствующего уровня трудности [4].

Определим количество элементов пространства M первоначально без учета верхнего ограничения на m_i , т.е. только с учетом условия $0 \leq m_i$. Из практики решения комбинаторных задач [5] известно, что количество элементов такого пространства будет равно числу целых неотрицательных решений уравнения $x_1 + x_2 + \dots + x_K = N$.

Если имеются целые неотрицательные числа $x_1, x_2, \dots, x_K$ такие, что $x_1 + x_2 + \dots + x_K = N$, то из них можно составить сочетания из K элементов по N , взяв x_1 элементов первого типа, x_2 элементов второго типа, ..., x_K – K -го типа. Наоборот, имея сочетание из K элементов по N , получим решение уравнения $x_1 + x_2 + \dots + x_K = N$ в целых неотрицательных числах, где x_1 элементов первого типа, x_2 элементов второго типа, ..., x_K – K -го типа. Следовательно, между множеством всех сочетаний из K элементов по N с повторениями и множеством всех целых неотрицательных решений уравнения $x_1 + x_2 + \dots + x_K = N$ устанавливается взаимно-однозначное соответствие. Поэтому число решений равно $f_K^N = C_{N+K-1}^N$ или C_{N+K-1}^{K-1} (по свойству симметрии биномиальных коэффициентов $C_n^k = C_n^{n-k}$).

Рассмотрим проекцию трехмерного пространства $M(m_1, m_2, m_3)$ на плоскость (m_1, m_2) , при этом учитывая, что $m_3 = N - (m_1 + m_2)$. Если на данную проекцию наложить ограничение $m_i \leq L$, то получим число элементов пространства M , равное:

$$Z = C_{N+K-1}^{K-1} - K \cdot C_{N-L+K-2}^{K-1} + K \cdot C_{N-2(L+1)+K-1}^{K-1}. \quad (1)$$

Вывод этой формулы достаточно сложен, а на алгоритм формирования сеанса практически не влияет, поэтому ввиду ограниченности объема статьи может быть опущен.

С точки зрения учета содержательной структуры теста, уровней трудности тестовых заданий и реальной длины сеанса тестирования в качестве примера рассмотрим структуру сеанса при $K=3$, $L=10$, $N=20$. При этом количество элементов пространства $M(m_1, m_2, m_3)$ будет:

$$Z = C_{20+3-1}^{3-1} - 3 \cdot C_{20-10+3-2}^{3-1} + 3 \cdot C_{20-2(10+1)+3-1}^{3-1} = 121 - 3 \cdot C_{11}^2 + 3 \cdot C_1^2 = 121 - 3 \cdot \frac{11 \cdot 10}{2} + 0 = 231 - 165 = 66.$$

Если число тем в тесте $L < \frac{N}{K}$, то сеанс, в котором из каждого уровня трудности в тест попадает только одно задание, не существует. В этом случае из каждой темы в сеанс попадает больше, чем K заданий, но этот вариант в данной статье не рассматривается. Отметим, что для $N = 20$ минимальное число тем в тесте не может быть меньше 7.

На рис. 2 проекция пространства M на плоскость (m_1, m_2) без верхних ограничений на m_i показана серым цветом без штриховки, а с ограничениями $m_i \leq 10$ – заштрихованная область. Справа от этой области показаны координаты трех различных элементов пространства M .

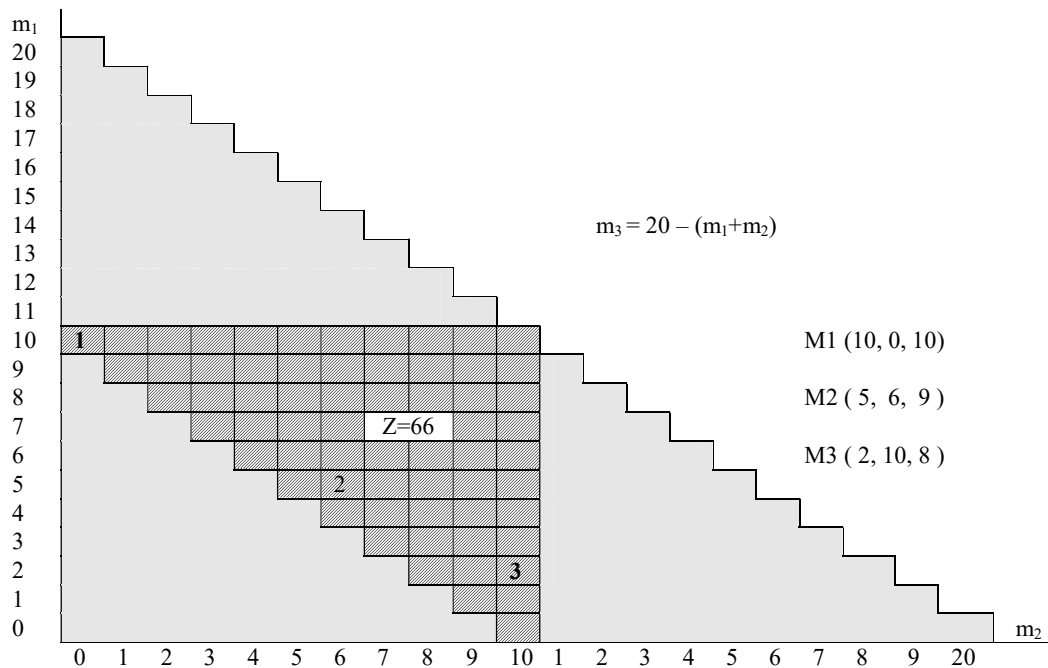


Рис. 2. Модель трехмерного пространства $M(m_1, m_2, m_3)$

Путем несложных расчетов можно рассчитать число элементов Z для любой длины сеанса. Результаты для N от 10 до 20 приведены в таблице.

N	10	11	12	13	14	15	16	17	18	19	20
Z	66	75	82	87	90	91	90	87	82	75	66

Представим табличную модель сеанса тестирования (рис.3) с учетом, что $k_{ij} = \{0, 1\}$. Для упрощения изложения в качестве первого примера будем рассматривать середины диапазонов уровней трудности тестовых заданий, для которых $p_1 = 0,7$, $p_2 = 0,5$, $p_3 = 0,3$, хотя в базе данных присутствуют реальные коэффициенты трудности.

Уровни	Коэффициенты			Темы									
	p	c	e	1	2	3	4	5	6	7	8	9	10
1 (m_1)	0,7	0,3	0,06	$k_{1,1}$	$k_{1,2}$	$k_{1,3}$	$k_{1,4}$	$k_{1,5}$	$k_{1,6}$	$k_{1,7}$	$k_{1,8}$	$k_{1,9}$	$k_{1,10}$
2 (m_2)	0,5	0,4	0,07	$k_{2,1}$	$k_{2,2}$	$k_{2,3}$	$k_{2,4}$	$k_{2,5}$	$k_{2,6}$	$k_{2,7}$	$k_{2,8}$	$k_{2,9}$	$k_{2,10}$
3 (m_3)	0,3	0,5	0,07	$k_{3,1}$	$k_{3,2}$	$k_{3,3}$	$k_{3,4}$	$k_{3,5}$	$k_{3,6}$	$k_{3,7}$	$k_{3,8}$	$k_{3,9}$	$k_{3,10}$

Рис. 3. Модель сеанса тестирования в системе OpenTEST2

Данная модель строится на основе координат элемента пространства $M(m_1, m_2, m_3)$. В i -ю строку, соответствующую уровню сложности i , записывается число единиц ($k_{ij} = 1$),

равное значению координаты m_i . Для остальных $(L - m_i)$ элементов i -й строки $k_{ij} = 0$. На рис. 4 показана модель сеанса тестирования для точки 2 на рис. 2, для которой $M_2(5, 6, 9)$.

Уровни	Коэффициенты			ТЕМЫ									
	p	c	e	1	2	3	4	5	6	7	8	9	10
1 (m_1)	0,7	0,3	0,06	1	1	1	1	1	0	0	0	0	0
2 (m_2)	0,5	0,4	0,07	1	1	1	1	1	1	0	0	0	0
3 (m_3)	0,3	0,5	0,07	1	1	1	1	1	1	1	1	1	0

Рис. 4. Модель сеанса тестирования для $M_2(5, 6, 9)$

В данной работе весомозначность всех тем в тесте считается одинаковой, а способ распределения $k_{ij} = 1$ по темам принят произвольным (случайным).

Из подраздела темы, соответствующего определенному уровню трудности, задания в сеанс тестирования выбираются по случайному алгоритму. Рекомендуется, чтобы в каждом подразделе было не менее 10 тестовых заданий.

Введем оценку достоверности репрезентативной выборки из тестовой базы данных при формировании сеанса тестирования – коэффициент S . Примем, что среди множества вариантов формирования сеансов тестирования максимальной достоверностью, с точки зрения дифференцирующей способности тестовых заданий и максимального покрытия учебного материала, обладает тот вариант, у которого значение S максимально.

Коэффициент S вычисляется как сумма для всех L тем теста произведений средней дисперсии коэффициентов трудности тестовых заданий всех K уровней трудности на коэффициент покрытия учебного материала для каждого уровня трудности темы :

$$S = \sum_{j=1}^L D_j C_j,$$

$$D_j = \frac{\sum_{i=1}^K k_{i,j} p_i q_i}{\sum_{i=1}^K k_{i,j}} = \frac{k_{1,j} p_1 q_1 + k_{2,j} p_2 q_2 + k_{3,j} p_3 q_3}{k_{1,j} + k_{2,j} + k_{3,j}},$$

$$C_j = k_{1,j}(1 - k_{2,j})(1 - k_{3,j})c_1 + k_{2,j}(1 - k_{1,j})(1 - k_{3,j})c_2 + k_{3,j}(1 - k_{1,j})(1 - k_{2,j})c_3 + \\ + k_{1,j}k_{2,j}(1 - k_{3,j})(c_1 - e_1 + c_2 - e_2) + k_{1,j}k_{3,j}(1 - k_{2,j})(c_1 - e_1 + c_3 - e_3) + \\ + k_{2,j}k_{3,j}(1 - k_{1,j})(c_2 - e_2 + c_3 - e_3) + k_{1,j}k_{2,j}k_{3,j},$$

где L – количество тем в тесте; K – количество уровней трудности тестовых заданий; c_j , e_i – коэффициенты покрытия и пересечения учебного материала для каждого уровня трудности.

Отметим, что $S_{\max} \Rightarrow 0,25 * L \Rightarrow 2,5$, при условии $p_1 = p_2 = p_3 = 0,5$.

Вычисление значений коэффициента S выполняется для всех вариантов сеанса тестирования (элементов пространства M), количество которых приведено в таблице. Так как количество вариантов в общем случае меньше 100, то использовать какие-либо специальные алгоритмы дискретной оптимизации нецелесообразно, достаточно использовать полный перебор вариантов для нахождения максимальных значений коэффициента S .

На рис. 5–8 показаны результаты расчета достоверности 10 сеансов тестирования для разных вариантов p и c с максимальными величинами S для значений $K=3$, $L=10$, $N=20$.

Уровни	Коэффициенты			Структура сеанса (задания по уровням трудности)									
	p	c	e	1	2	3	4	5	6	7	8	9	10
1 (m_1)	0,7	0,3	0,06	0	1	1	2	2	3	2	3	4	3
2 (m_2)	0,5	0,4	0,07	10	9	10	8	10	7	9	10	6	8
3 (m_3)	0,3	0,5	0,07	10	10	9	10	8	10	9	7	10	9
S				1.748	1.726	1.721	1.705	1.695	1.684	1.67	1.669	1.662	1.649

Рис. 5. Структура сеанса тестирования (вариант 1) по уровням трудности

	Коэффициенты			Структура сеанса (задания по уровням трудности)									
Уровни	p	c	e	1	2	3	4	5	6	7	8	9	10
1 (m_1)	0,8	0,3	0,06	0	1	1	2	3	2	4	2	5	3
2 (m_2)	0,5	0,4	0,07	10	10	9	10	10	8	10	9	10	9
3 (m_3)	0,2	0,5	0,07	10	9	10	8	7	10	6	9	5	8
	S			1.746	1.742	1.741	1.738	1.738	1.735	1.733	1.731	1.729	1.728

Рис. 6. Структура сеанса тестирования (вариант 2) по уровням трудности

	Коэффициенты			Структура сеанса (задания по уровням трудности)									
Уровни	p	c	e	1	2	3	4	5	6	7	8	9	10
1 (m_1)	0,7	0,4	0,15	5	4	5	4	3	3	2	2	1	1
2 (m_2)	0,5	0,5	0,15	5	6	10	10	7	10	8	10	9	10
3 (m_3)	0,3	0,6	0,2	10	10	5	6	10	7	10	8	10	9
	S			1.746	1.742	1.741	1.738	1.738	1.735	1.733	1.731	1.729	1.728

Рис. 7. Структура сеанса тестирования (вариант 3) по уровням трудности

	Коэффициенты			Структура сеанса (задания по уровням трудности)									
Уровни	p	c	e	1	2	3	4	5	6	7	8	9	10
1 (m_1)	0,8	0,4	0,15	5	4	3	2	1	0	1	6	2	5
2 (m_2)	0,5	0,5	0,15	10	10	10	10	10	10	9	10	8	9
3 (m_3)	0,2	0,6	0,2	5	6	7	8	9	10	10	4	10	6
	S			1.575	1.567	1.56	1.552	1.545	1.537	1.516	1.506	1.494	1.478

Рис. 8. Структура сеанса тестирования (вариант 4) по уровням трудности

На рис. 9,10 показаны результаты расчетов S для $K=3$, $L=10$, $N=15$.

	Коэффициенты			Структура сеанса (задания по уровням трудности)									
Уровни	p	c	e	1	2	3	4	5	6	7	8	9	10
1 (m_1)	0,7	0,3	0,06	0	1	0	0	2	0	1	1	0	0
2 (m_2)	0,5	0,4	0,07	5	4	10	6	3	9	10	5	7	8
3 (m_3)	0,3	0,5	0,07	10	10	5	9	10	6	4	9	8	7
	S			1.399	1.377	1.374	1.363	1.356	1.348	1.347	1.342	1.328	1.323

Рис. 9. Структура сеанса тестирования (вариант 5) по уровням трудности

	Коэффициенты			Структура сеанса (задания по уровням трудности)									
Уровни	p	c	e	1	2	3	4	5	6	7	8	9	10
1 (m_1)	0,8	0,3	0,06	0	1	2	0	1	3	2	0	0	0
2 (m_2)	0,5	0,4	0,07	10	10	10	9	9	10	9	8	5	6
3 (m_3)	0,2	0,5	0,07	5	4	3	6	5	2	4	7	10	9
	S			1.279	1.257	1.235	1.234	1.213	1.196	1.191	1.19	1.179	1.174

Рис. 10. Структура сеанса тестирования (вариант 6) по уровням трудности

Анализ приведенных результатов позволяет сделать три основных вывода :

- 1) наивысшую достоверность имеют сеансы тестирования, у которых хотя бы по одному из уровней трудности заполнены все темы;
- 2) уменьшение дисперсии трудностей тестовых заданий (увеличение разброса значений p_i) смещает структуру сеанса в область заданий среднего уровня трудности;
- 3) увеличение значений коэффициентов покрытия учебного материала смещает структуру сеанса в область заданий повышенного уровня трудности.

Таким образом, можно сформулировать алгоритм формирования частично адаптивного сеанса тестирования.

1. На основании значений K , L , N находится элемент пространства M (или несколько элементов) с максимальным значением достоверности.

2. На основании координат выбранной точки пространства M формируется модель сеанса тестирования (см. рис. 4).

3. На основании указанной модели из тестовой базы данных выбираются задания из соответствующих тем и уровней трудности.

Отметим, что в каждой теме могут быть свои соотношения числовых значений p_i и c_i , но это не влияет на общий алгоритм формирования сеанса тестирования.

В системе OpenTEST2 преподаватель может выбрать упрощенный неадаптивный алгоритм формирования сеанса тестирования (если количественные оценки уровней трудности всех тем фиксированы, и автор теста не задает коэффициентов покрытия учебного материала).

1. В сеанс тестирования из всех тем выбираются задания средней трудности с максимальной дисперсией.

2. Если после этого сеанс не заполнен, то в него из всех тем выбираются трудные задания.

3. Если и после этого сеанс не заполнен, то он дополняется легкими заданиями.

Данный алгоритм формирования сеанса также рекомендуется применять для тестов неизвестного происхождения, которые импортируются в систему OpenTEST2 из внешних источников.

В заключение отметим, что весомость отдельных тем в тесте и способ распределения $k_{ij} = 1$ по темам существенно влияют на структуру сеанса тестирования. Но это предмет дальнейших исследований.

Выводы

Анализ расчетов достоверности для сеансов тестирования с разными параметрами показал, что при формировании сеанса многое зависит от соотношения параметров длины сеанса (N) и количества тем в тесте (L) при фиксированном количестве уровней трудности тестовых заданий.

При $N > KL$ сеанс, формируемого по принципу «из каждого подраздела темы в сеанс попадает по одному заданию» не существует, а преподавателю, проводящему тестирование, следует задуматься о соотношении длины сеанса и тематической разбивки учебного материала.

При $N < L$ длина сеанса такова, что задания из некоторых тем вообще не попадают в сеанс. Если эта ситуация преподавателя устраивает, то все в порядке, а если нет, то следует увеличить длину сеанса хотя бы до числа тем в тесте.

При $L \leq N \leq KL$ длина сеанса считается оптимальной. Кроме того, данное соотношение может помочь автору теста в тематической разбивке учебного материала при заданных временных и количественных ограничениях на длину сеанса тестирования.

Список литературы: 1. *Аванесов В.С.* Композиция тестовых заданий. Учебная книга, 3 изд. доп. М.: Центр тестирования, 2002. 240 с. 2. *Шкиль А.С., Чумаченко С.В., Напрасник С.В., Гаркуша Е.В.* Построение тестовых заданий в системе компьютерного тестирования знаний в OpenTEST2 // АСУ и приборы автоматики. 2006. Вып. 137. С.21-32. 3. *Карпенко Д.С., Карпенко О.М., Шлихунова Е.Н.* Система автоматического повышения качества тестовых заданий и мониторинг процесса усвоения знаний // Телекоммуникации и информатизация образования. 2001. № 1. С. 42-57. 4. *Сергиенко И.В., Каспицкая М.Ф.* Модели и методы решения на ЭВМ комбинаторных задач оптимизации. Киев: Наукова думка, 1981. 288 с. 5. *Ежов И.И., Скороход А.В., Ядренко М.И.* Элементы комбинаторики. М.: Наука, 1977. 80 с.

Поступила в редколлегию 07.12.2006

Шкиль Александр Сергеевич, канд. техн. наук, доцент кафедры АПВТ ХНУРЭ. Научные интересы: проектирование и диагностика цифровых устройств, создание электронных учебников, технологии дистанционного образования. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-326.

Чумаченко Светлана Викторовна, канд. физ.-мат. наук, доцент кафедры АПВТ ХНУРЭ. Научные интересы: технологии тестирования знаний. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-326. E-mail: ri@kture.kharkov.ua.

Напрасник Сергей Владимирович, инженер тестового центра ХНУРЭ. Научные интересы: системное администрирование сетей Windows и UNIX, проектирование и разработка Internet/Intranet приложений. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-326.

Бабич Анна Витальевна, канд. техн. наук, старший преподаватель кафедры АПВТ ХНУРЭ. Научные интересы: сетевые технологии, технологии дистанционного образования. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-326.

Гаркуша Елена Владимировна, преподаватель кафедры автоматических систем управления войсками факультета «Телекоммуникации» Полтавского военного института связи.