

ТЕХНОЛОГИЯ ПРИМЕНЕНИЯ АДАПТИВНОГО ТЕСТИРОВАНИЯ ИТ-СПЕЦИАЛИСТОВ КОММЕРЧЕСКОЙ ФИРМЫ

Колдаев В.Д.¹, Волков А.С.², Павлов Р.И.³

¹Колдаев Виктор Дмитриевич - доктор технических наук;

²Волков Александр Сергеевич – студент;

³Павлов Роман Игоревич - магистрант,

кафедра информатики и программного обеспечения вычислительных систем,
Федеральное государственное автономное образовательное учреждение высшего образования
Национальный исследовательский университет
Московский институт электронной техники,
г. Москва

Аннотация: в статье рассмотрены некоторые наиболее актуальные теоретические вопросы применения тестовых технологий, методика их использования для адаптивного тестирования ИТ-специалистов коммерческой фирмы. Проведен анализ тестовых технологий и указаны пути повышения эффективности.

Ключевые слова: тестовые технологии, адаптивное тестирование, уровни усвоения, конечный автомат, надежность диагностического теста.

Увеличение масштабов тестирования порождает спектр проблем, условий и требований, которые нацелены на совершенствование систем контроля и оценки качества подготовки специалистов коммерческой фирмы. Адаптивное тестирование возможно, если заранее определена трудность тестовых заданий, что предполагает их предварительную апробацию и обработку результатов с использованием методов дисперсионного анализа и теории тестирования для определения надежности, валидности, дифференцирующей способности и трудности. Эффективность контрольно-оценочных процедур повышается при использовании многоступенчатой стратегии отбора и предъявления заданий, основанной на алгоритмах с полной контекстной зависимостью, в которых очередной шаг совершается только после оценки результатов выполнения предыдущего шага. Алгоритм отбора и предъявления заданий строится по принципу обратной (дуплексной) связи, когда при правильном ответе испытуемого очередное задание выбирается более трудным, а неверный ответ влечет за собой предъявление последующего более легкого задания, чем то, на которое испытуемым был дан неверный ответ.

Тест включает три структурных компонента: систему тестовых заданий с возрастающей степенью сложности; набор вариантов ответов, содержащих наряду с правильными и неправильные ответы (дистракторы); систему проверки и анализа результатов. Длительность тестирования ограничивается исходя из соображений: удобства процедуры тестирования; темпа накопления усталости у испытуемых; достижения оптимального уровня дифференциации сотрудников. На надежность тестирования сильно влияет степень гомогенности групп испытуемых, уровень их подготовленности, а также другие факторы, связанные не столько с тестом, сколько с условиями сеанса тестирования [1].

В настоящее время при адаптивном контроле особое внимание уделяется *пирамидальному тестированию*, при котором всем испытуемым выдаются задания средней трудности, а затем, в зависимости от ответов либо более трудное, либо более простое, путем деления шкалы заданий пополам. Для математической формализации задачи адаптивного контроля используются теоретические основы модели Раша и параметризации Item Response Theory (IRT). При статистической обработке ответов адаптивного тестирования первая модель использует аппарат нечетких множеств, а вторая для моделирования вероятностей правильных ответов логистическую кривую. Смысл адаптивности в алгоритме тестирования в том, что тест адаптируется на доступный сотруднику уровень сложности и тем самым быстрее локализует его истинный тестовый балл на шкале измеряемого свойства [1, 2].

Ядром системы тестирования является конечный автомат, реализующий модель поведения ИТ-специалиста и так называемую функцию «уровня знаний». Содержательно база вопросов разделяется на три уровня: мировоззренческий, базовый, углубленный: $A = \{Q, \Sigma, \Delta, I, F\}$, где $Q = \{3...5, True, False\}$ – множество входных данных (3...5 – вопросы соответствующего уровня; $True/False$ – правильность ответа); $\Sigma = \{s_0...s_{12}\}$ – множество состояний автомата; $I = \{s_0\}$ – множество начальных состояний автомата; $F = \{s_4, s_7, s_{10}, s_{12}\}$ – множество конечных состояний автомата; Δ – функция переходов. Необходимый уровень *валидности результатов* достигается за счет нескольких рекурсивных проходов автомата (рис. 1).

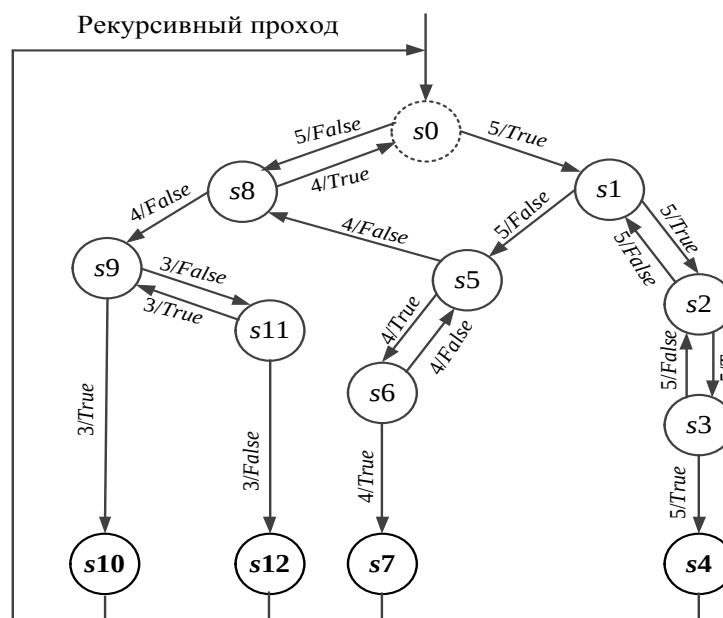
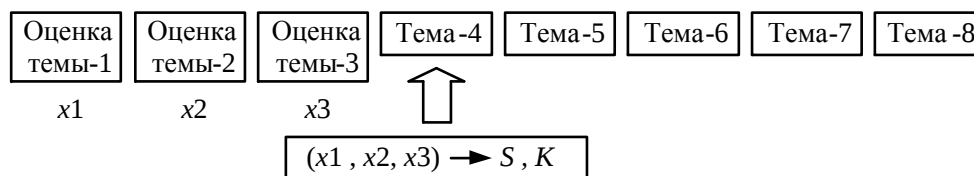


Рис. 1. Графическое изображение конечного автомата

Надежность теста связана с понятием стандартной ошибки: чем выше надежность, тем меньше стандартная ошибка измерений. Существует несколько различных понятий надежности диагностического теста. *Надежность параллельных форм* – характеристика, получаемая с помощью параллельных форм (вариантов) теста, проведенных на одной и той же выборке ИТ-специалистов. *Ретестовую надежность* определяют при повторном использовании того же самого теста спустя некоторый промежуток времени. *Надежность расщепленных частей теста* исследуется путем анализа устойчивости результатов отдельных совокупностей тестовых заданий при однократном тестировании.

Разработанный *алгоритм компьютерного адаптивного тестирования* (АКАТ) содержит восемь тем. Он автоматически выбирает обучающий метод по следующим критериям: по тестовому опросу или предыдущему тестированию. В каждом тематическом разделе предоставляется множество материалов (теоретические сведения или практические навыки) и тестов (самоконтроль или контрольные задания).

В первых разделах генерируются вопросы и задания для проверки знаний по первым трем темам, которые являются вводными – содержащими общие понятия классификации и определения. В предлагаемом алгоритме тестирования при формировании вопросов для очередной темы используются результаты выполнения контрольных вопросов из трех предыдущих тем (рис. 2). Результаты (оценки) первых трех тем прямо пропорциональны сложности заданий четвертой темы и обратно пропорциональны количеству заданий четвертой темы [3].



S = сложность вопросов, K = количество вопросов

Рис. 2. Схема определения сложности и количества вопросов

Принцип генерирования заданий для последующих тем (5, 6, 7, 8) – одинаковый: для формирования заданий по теме 5, необходимо учитывать результаты предыдущих трех тем (2, 3, 4); для генерирования заданий и контрольных вопросов по теме 6 – результаты предыдущих трех тем (3, 4, 5) и т.д. Анализ результатов выполнения тестов и контрольных вопросов по темам (1, 2, 3) осуществляется с использованием метода минимума, предложенного Такаги–Сугено [3, 4]. Затем, по методу среднего центра, определяются количество (K) и сложность (S) контрольных вопросов: $S = S(x_1, x_2, x_3)$ – сложность заданий; $K = K(x_1, x_2, x_3)$ – количество заданий. На рис. 3 показан алгоритм генерирования вариантов вопросов и структуры заданий. Если сотрудник получает оценку «отлично» за предыдущие задания, то количество вопросов $K = 4$, а сложность вопросов – $S_6 = \{1, 4, 6, 7\}$. Если получена оценка «удовлетворительно» за предыдущие задания, то количество вопросов $K = 5$, а сложность вопросов – $S_1 = \{1, 2, 3, 4, 5\}$.

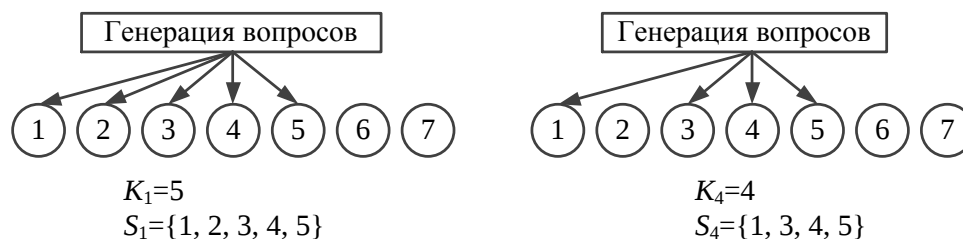


Рис. 3. Варианты генерации тестов и контрольных вопросов

В алгоритме адаптивного тестирования различают пять уровней усвоения производственного материала: нулевой уровень – *понимание* новой информации; первый уровень – *опознание* (выделение) изучаемого объекта; второй уровень – *воспроизведение* информации по памяти, по образцу; третий уровень – *применение* алгоритма для решения нетиповых задач; четвертый уровень – *творческая деятельность* при разработке алгоритма решения задачи [3, 4].

Исследование возможности повышения эффективности в адаптации тестов связано с уровнем подготовленности специалистов и их психологическим типом: сотрудники-синтетики (для них трудны тесты с возможностью множественного выбора вариантов ответа; тесты с открытыми вопросами дают возможность проявить свои творческие способности); сотрудники-аналитики (выигрывают от тех видов деятельности, которые позволяют им углубиться в детали, проанализировать, сделать вывод; тесты на решение определенных задач позволяют применять свои способности к анализу); импульсивные сотрудники (хорошо справляются с традиционными тестами с ограниченным сроком выполнения); рефлексивные сотрудники (потребуют тесты, в которых не задан лимит времени); индуктивные сотрудники (утомляются объяснениями руководителя, не имея возможности попробовать сделать это самим); дедуктивные сотрудники (живут правилами и любят слушать объяснения руководителя) [4, 5].

Методика адаптивного тестирования предполагает пошаговую переоценку уровня подготовленности специалиста, через обучающие тесты градуированного (уровневого) характера. По результатам исследования (50 сотрудников) преимущества адаптивного тестирования знаний одобряют большинство – 84%; 16% выразили отрицательное отношение. Применение рейтинговой системы стимулирует работу (45%); позволяет более объективно оценивать знания (20%); создает заинтересованность в самостоятельной работе (11%). В результате разработанной технологии получена динамика значительного роста профессионального и личностного развития IT-специалистов фирмы.

Список литературы

1. Аванесов В.С. Композиция тестовых заданий. М.: Учебная книга, 2002. 240 с.
2. Лебедева И.П. Математические модели как средство обучения. М.: Педагогика, 2004. № 2. С.11-19.
3. Колдаев В.Д. Технология адаптивного тестирования структурно-содержательной модели индивидуального образовательного маршрута студента вуза // «Избранные вопросы современной науки». Монография / Под общей ред. д.п.н., проф. С.П. Акутиной. М.: Из-во «Перо», 2011. С. 5-37.
4. Нейман Ю.М., Хлебников В.А. Введение в теорию моделирования и параметризации педагогических тестов. М.: Прометей, 2000. 168 с.
5. Колдаев В.Д. Модерация дуплексной системы управления процессом обучения на основе графовых моделей // Инновационная деятельность в системе образования. Монография / Под общей ред. д.п.н., проф. Г.Ф. Гребенщикова. М.: Издательство Перо, 2011. С.182-211.