

РАЗРАБОТКА СИСТЕМЫ АДАПТИВНОГО КОНТРОЛЯ

В статье предлагается своеобразный подход к созданию инструментальной системы адаптивного компьютерного тестирования.

Внедрение государственного образовательного стандарта в общеобразовательных школах, включающего основную идею концепции планирования обязательных результатов обучения, ориентация на безусловное достижение всеми учащимися обязательного уровня знаний, в свою очередь предполагает усиления всей контролирующей деятельности в учебном процессе [5]. Кроме этого, показала свою эффективность проблема организации контроля со строгим учетом элементов знаний, на основе которой более эффективно проверяются знания и умения учащихся [2;3]. Но практика деятельности школ доказывает, что устоявшиеся традиционные формы контроля явно недостаточны в решении данных вопросов.

Наиболее актуальным в решении данных проблем нам представляется развитие тестовой методики результатов обучения, в частности, разработка адаптивного компьютерного теста.

На сегодняшний день компьютерные системы контроля знаний приобретают все большую популярность благодаря их объективности, доступности и экономической эффективности. Вопрос лишь состоит в том, что необходимо разработать универсальную систему для создания и проведения различных тестов, а также для обработки результатов тестирования.

Основными требованиями к таким системам являются: интеллектуальность, открытость, гибкость, адаптивность при организации процесса тестирования, наличие дружественного интерфейса. Особое внимание здесь следует уделить адаптивности процесса тестирования. Целесообразность именно этого вида контроля вытекает из необходимости рационализации традиционного тестирования.

При создании компьютерных систем тестирования особую роль играет алгоритм предъявления тестовых заданий. По алгоритму предъявления тестовых заданий в классической теории тестов различают следующие стратегии:

- строгая последовательность, когда автор самостоятельно расставляет задание по тесту;
- случайная выборка тестовые задания выбирается из достаточно большой тестовой базы, что позволяет сформировать равноценные по сложности варианты тестов, то есть тестовая база разбивается на несколько групп сложности, в начале предлагаются простые задания, в случае правильности их выполнения переходят к группе более сложных заданий;

- адаптивный алгоритм в системах тестирования, где используется адаптивный алгоритм, каждое последующее задание выбирается исходя из ответов на предшествующие, что позволяет более качественно проверять знания, используются минимальное количество тестовых заданий.

Создание адаптивных тестов является более сложной задачей, чем создание тестов других типов. Адаптивные тесты komponуются в процессе тестирования, требуют большей базы тестовых заданий и точных расчетов.

Учитывая вышесказанное, в разрабатываемой системе необходимо организовать адаптивный алгоритм предъявления заданий. Адаптивность при этом будет выражаться, в изменении относительных пропорций в предъявлении легких заданий, заданий средней трудности и трудных в зависимости от числа правильных ответов, зарегистрированных в ходе сеанса тестирования.

Исходя из анализа существующих систем контроля знаний, была разработана концептуальная схема автоматизированной адаптивной тестирующей системы, программная часть которой реализована в виде трех отдельных модулей:

- модуля тестирования;

- модуля создания и редактирования тестов;
- модуля статистики и анализа результатов;

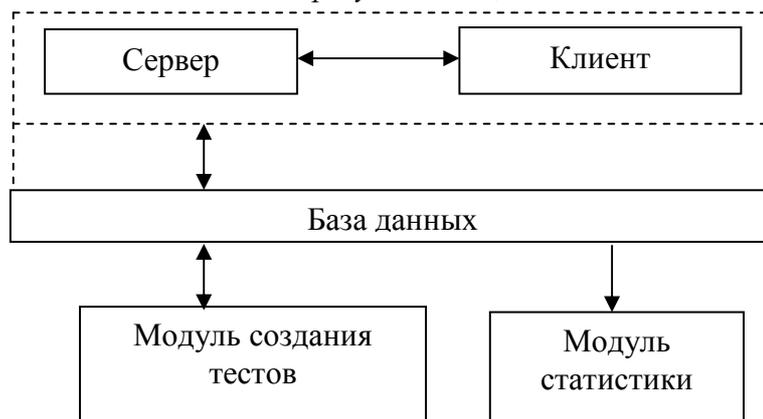


Рис.1 Структурная блок - схема системы адаптивного тестирования.

При этом информационная база данных разрабатывается в рамках концепции «клиент-сервер». База данных должна хранить банк тестовых заданий, параметра надстройки теста и процесса тестирования, информацию для аутентификации пользователей, результаты тестирования, логику трудности и другую информацию по обработке данных.

Принципы построения информационной базы:

- весь информационный массив документов накапливается и хранится в центральной базе данных с удаленным доступом и архитектурой «клиент-сервер»;
- таблицы должны быть спроектированы таким образом, чтобы минимизировать избыточность хранимой информации;
- уменьшение избыточного объема данных не должно достигаться за счет значительного усложнения доступа и скорости обработки информации;
- база данных, сформированная для данной системы, носит максимально обобщенный характер, т. е. может использоваться для создания тестов практически по любой дисциплине.

Модуль тестирования, которая строится с использованием технологии «клиент-сервер». На рабочей станции в качестве клиента находится интерфейсная часть, функции которой преимущественно заключаются в получении вопроса от сервера и отправке ответа. Тестируемые учащиеся при работе с данными модулем могут просмотреть свою статическую статистику по предыдущим, ранее пройденным тестам.

Выполнения же всего основного объема исчислений, обработка данных, интерпретации параметров теста, взаимодействие с базой данных и управление процессом адаптивного тестирования предполагается на сервере.

Используя базы данных, размещенную на сервере, модуль тестирования проводит аутентификацию зарегистрированных или регистрацию новых пользователей, выбор теста, осуществляет тестирование испытуемого посредством вывода на экран вопроса и последующего получения ответа, обрабатывает поступившие данные и записывает полученные результаты тестирования в базу данных на сервер для возможности дальнейшего анализа и использования преподавателем. По окончании выполнения теста (либо по истечении времени, если оно ограничено в данном тесте) для тестируемого выводится результат тестирования и краткий комментарий.

Обработка результатов тестирования достаточно сложная задача и требует специального математического аппарата, который должен решать задачу реализации интеллектуальных методов оценки знаний студента. Здесь необходима разработка математических методов и моделей поведения обучаемого, разработка критериев оценки его знаний, способностей и умений.

Исторически выделяются два основных подхода к созданию тестов и к интерпретации результатов его выполнения. В первом подходе, получившем развитие в

рамках классической теории тестов, уровень знаний тестируемых оценивается с помощью их индивидуальных баллов. Балл тестируемого вычисляется как алгебраическая сумма оценок выполнения каждого задания теста.

Второй подход, основанный на Item Response Theory (IRT), который и предлагается использовать в разрабатываемой системе тестирования, нацелен на оценивание латентных качеств личности и параметров заданий теста на основе математических моделей измерения.

К преимуществам данного подхода можно отнести [3, с.114]:

- устойчивые, объективные оценки параметра, характеризующего уровень знаний тестируемых;
- устойчивые, объективные оценки параметра трудности заданий, не зависящие от свойств выборки тестируемых;
- измерение знаний параметров тестируемых и заданий теста в одной и той же шкале, имеющее свойства интервальной шкалы;
- возможность предсказать вероятность правильного выполнения заданий теста любым испытуемым в выборке.

В отличие от классической теории, где индивидуальный балл тестируемого учащегося рассматривается как постоянное число, в IRT латентный параметр трактуется как некоторая переменная.

В рамках IRT можно рассматривать различные математические модели, в данном случае выбор был сделан в пользу однопараметрической модели G. Rasch [2]. Данная модель основывается на взаимодействии двух множеств латентных параметров, значения которых порождают результаты выполнения теста. Элементы этого множества определяет уровень знаний i -го испытуемого θ_i где $i = 1, 2, \dots, N$, элементы второго множества – трудность j -го задания теста β_j где $j = 1, 2, \dots, n$. На практике всегда ставится задача: по ответам тестируемых оценить значения латентных параметров θ и β .

Значения параметра θ_i можно рассматривать как положение i -го тестируемого, а значения β_j – как положение j -го задания по одной и той же оси. В этом случае абсолютная величина разности $|\theta_i - \beta_j|$ – это расстояние, на котором находится тестируемый с уровнем знаний θ_i от заданий трудностью β_j . Если эта разность велика по модулю и отрицательна, то задание бесполезно для измерения уровня знаний i -го учащегося: учащийся наверняка не сможет выполнить задание правильно. С другой стороны, большое положительное значение этой разности тоже не представляет интереса i -го для процесса контроля j -го учащегося: задание данной сложности им освоено.

Существует некоторая математическая модель взаимосвязи между результатами тестирования и значениями латентных переменных θ и β . Можно рассматривать условную вероятность правильного выполнения i -м учащимся с уровнем знаний θ_i различных по трудности заданий теста, считая параметром θ_i i -го студента, а β – независимой переменной. В этом случае условная вероятность будет функцией латентной переменной β :

$$P_i \{x_{ij} = 1 | \theta_i\} = f(\theta_i - \beta), i = 1, 2, \dots, N. \quad (1)$$

Аналогично вводится условная вероятность правильного выполнения j -го задания трудностью β_j различными учащимися. Здесь независимой переменной является θ , а β_j – параметр, определяющий трудность j -го задания:

$$P_j \{x_{ij} = 1 | \theta_i\} = f(\theta - \beta_j), j = 1, 2, \dots, n, \quad (2)$$

где $x_{ij} = \begin{cases} 1, & \text{если } i\text{-го тестируемого на } j\text{-е задание теста правильный} \\ 0, & \text{если } i\text{-го тестируемого на } j\text{-е задание теста неправильный} \end{cases}$

Аналитическое задание однопараметрической модели G. Rasch может представлено

в виде формул следующим образом:

$$P_j(\theta) = \{1 + \exp[-1,7(\theta - \beta_j)]\}^{-1};$$

$$P_j(\beta) = \{1 + \exp[-1,7(\theta_i - \beta)]\}^{-1}.$$

В первом случае вероятность правильного выполнения j -го задания теста P_j является возрастающей функцией переменной θ . Это означает, что чем выше уровень знаний тестируемого, тем больше вероятность правильного выполнения им j -го задания теста.

Вероятность правильного выполнения i -м тестируемым различных по трудности заданий P_i является убывающей функцией переменной β . Это означает, что с ростом трудности заданий значения вероятности будут уменьшаться.

Если $\theta = \beta_j$ для формулы (1) и $\beta = \theta_i$ для формулы (2), то соответствующие вероятности P_j и P_i будут равны 0,5. Таким образом, тестируемый с уровнем знаний, равным трудности j -го задания теста, ответит на него правильно с вероятностью 0,5.

Начальная оценка уровней знаний i -го тестируемого в логитах находится по формуле: $\theta_i^0 = \ln(P_i/q_i)$, $i = 1, \dots, N$, где

P_i - доля правильных ответов всех i -го тестируемого на все задания теста;

q_i - соответственно доля неправильных ответов, причем $P_i = 1 - q_i$,

N - число тестируемых.

Аналогично начальное значение параметра

$$\beta_i^0 = \ln(q_i/P_j), i = 1, \dots, n, \text{ где}$$

P_j - доля правильных ответов всех тестируемых групп на j -е задание теста;

q_j - соответственно доля неправильных ответов, причем $P_j = 1 - q_j$,

n - число заданий.

Уничтожение эффекта влияния трудности задания на оценки тестируемых студентов:

$$\theta_i = \beta + X\theta_i^0, i = 1, \dots, N, X = (1 + W^2/2,89)^{\sqrt{2}},$$

где β - среднее значение логитов трудности заданий теста,

N - число тестируемых;

W - стандартное отклонение распределения начальных значений параметра β .

Для нахождения β_j вводится совершенно аналогичная формула с соответствующими параметрами.

Модуль создания и редактирования тестов обращается к базе данных, в которых хранятся тестовые задания и ответы к каждому тесту, а также его параметры и позволяет учителю создать новый тест, изменить настройки имеющегося теста, редактировать вопросы и ответы, использовать в одном тесте вопросы других, т.е. обеспечивать слияние тестов.

Модуль статистики и анализа результатов тестирования должен представлять учителю возможность просмотреть тестирование отдельного учащегося или целого класса по одному или нескольким тестам. Таким образом, вся полученная информация хранится в базе данных в любой момент и может быть использована.

Основной и самой сложной частью данной системы является серверная часть модуля тестирования, на которую ложится вся нагрузка по обработке данных и обеспечению функционирования алгоритмов, основанных на различных интеллектуальных методах.

Подводя итог, можно сказать, что разрабатываемая система соответствует современным требованиям, предъявляемым классу подобных систем, как в области педагогического тестирования, так и в области информационных технологий. Необходимо отметить универсальность данной системы, ее модульность и, следовательно, открытость,

что позволяет говорить о дальнейшей модернизации и оптимизации работы всех функциональных модулей системы.

При процессе разработки инструментальной системы адаптивного компьютерного тестирования выделено следующее:

- предъявляемому испытуемому задания зависят от результатов его ответов на предыдущие задания; вследствие этого учащимся предъявляется гораздо меньше заданий с сохранением диагностической способности целого объемного теста;
- возможность более точно измерить уровень обученности;
- значительное сокращение времени тестирования;
- возможность выявления тем, который тестируемый знает недостаточно хорошо, и подбор соответствующих заданий во время тестирования;
- лучшую защищенность теста, так как каждый учащийся получает свою траекторию набора заданий во время тестирования.

Литература:

1. Аванесов В.С. Методологические и теоретические основы тестового педагогического контроля. Дисс... д-ра пед. наук:13.00.01. –М., 1994.-339 с.
2. Ашыров Э.Т. Разработка заданий для оценки качества знаний на различных уровнях усвоения/ Э.Т. Ашыров// Высшее образование КР. -2011.-Выпуск 2/12.
3. Калдыбаев С.К., Ажыбаев Д.М., Бекежанов М.М. Компьютерная диагностика результатов обучения в общеобразовательной школе: Практико-ориентированная монография.- Бишкек, 2003. – 352 с.
4. Чельшкова М.Б. Разработка педагогических тестов на основе современных математических моделей: Учебное пособие / М.Б.Чельшкова. -М.: Исследовательский центр проблем качества подготовки специалистов, 1995.-32 с.
5. Временный государственный образовательный стандарт. Общее среднее образование. Математика. –Б., 2008.

Примечание: Исследование частично выполнено Грантом №11-06-90907, Российского фонда Фундаментальных научных исследований, по проекту «Мобильность Молодых ученых».