

ТЕСТОВАЯ ОБОЛОЧКА, РЕАЛИЗУЮЩАЯ СОВРЕМЕННЫЕ МЕТОДЫ ТЕСТИРОВАНИЯ В СИСТЕМЕ MOODLE

А. И. Безруков, В. Н. Гусятников, Н. С. Вагарина

Саратовский государственный технический университет им. Ю. А. Гагарина, Россия
E-mail: bezr_Alex@mail.ru, victorgsar@rambler.ru, victorgsar@rambler.ru

Предлагается концепция тестовой оболочки, позволяющей использовать современные методы компьютерного тестирования, использовать тесты, накопленные в системе MOODLE, а также создавать новые тесты. Проводится анализ тестовой подсистемы MOODLE. Делается вывод о целесообразности создания отдельной тестовой оболочки. Формулируются требования к оболочке. Предлагается модульная структура тестовой оболочки.

TEST SHELL IMPLEMENTING MODERN TESTING METHODS IN THE MOODLE SYSTEM

A. I. Bezrukov, V. N. Gusyatnikov, N. S. Vagarina

The concept of a test shell is proposed, which allows using modern methods of computer testing, using tests accumulated in the MOODLE system, and also creating new tests. The test subsystem MOODLE is analyzed. The conclusion is made about the expediency of creating a separate test shell. The requirements for the shell are formulated. A modular structure of the test shell is proposed.

Компьютерное тестирование является наиболее удобным объективным методом оценки знаний учащихся. Считается, что каждый испытуемый обладает некой латентной характеристикой «Уровень подготовленности» θ , а каждое задание имеет определенную трудность δ . Вероятность выполнения тестового задания зависит от уровня подготовленности испытуемого и уровня трудности задания. Модель Раша предлагает явный вид параметрического распределения этой вероятности.

$$P(\theta, \delta) = \frac{\exp(\alpha \cdot (\theta - \delta))}{1 + \exp(\alpha \cdot (\theta - \delta))} \quad (1)$$

На базе этой модели построена теория отклика (Item response theory, IRT) [1], адекватно описывающая процесс компьютерного тестирования.

В настоящее время в основном используются два принципиально различных метода организации тестирования: классический и адаптивный. При классическом тестировании требуется с максимальной точностью оценить уровень подготовленности каждого испытуемого. Такое требование возникает, например, при конкурсном отборе, когда необходимо отобрать лучших из претендентов с близкими уровнями подготовленности. Критерий качества такого тестирования – минимум рисков неправильных решений (например, отсеивания перспективного претендента или включение в группу малоперспективного).

При адаптивном тестировании требуется каждого учащегося отнести к одному из заранее заданных классов (например, «неуд», «уд», «хор» или «отл»). При этом желательно использовать как можно меньше заданий. Процесс тестирования заключается в последовательном уточнении вероятностей принадлежности испытуемого к каждому из указанных классов. Каждое следующее задание выбирается с учетом результатов ранее выполненных заданий, так, чтобы результат его выполнения давал максимальную информацию о том, какому классу принадлежит испытуемый. Критерий качества тестирования – минимальная вероятность неверного отнесения. Такой метод удобен, например, при текущем контроле знаний.

Мнение преподавателя (автора теста) о том, какие результаты должен показать испытуемый, чтобы его можно было отнести к одному из заданных классов, формулируется в виде рейтинговой шкалы. При этом тест и рейтинговая шкала должны соответствовать друг другу. Например, отсутствие в тесте заданий требуемой трудности не позволяет достаточно достоверно определить, к какому классу следует отнести испытуемого по результатам тестирования. Отметим, что при использовании классического тестирования для конкурсного отбора, рейтинговая шкала вообще не нужна. А если классическое тестирование используется для текущего контроля знаний (что обычно происходит при использовании существующих систем компьютерного тестирования), мы совершаем нелогичные действия: сначала с максимальной точностью определяем θ каждого студента, а затем закругляем его до интервалов рейтинговой шкалы. В статье [2] проведено сравнение реально проведенного классического тестирования и имитации на этих же результатах адаптивного тестирования. Показано, что достоверность адаптивного теста более чем в 2 раза выше, чем при классическом тестировании, при этом, использовано лишь 65% заданий.

В современной системе тестирования должны быть реализованы различные методы, что позволит организаторам тестирования выбрать и правильно использовать метод, наиболее соответствующий решаемой задаче.

Хорошая система тестирования должна обучаться на результатах ее использования. Например, при создании теста уровень трудности заданий задает разработчик. Но фактический (с точки зрения испытуемых) уровень трудности может сильно отличаться от заданного разработчиком. Поэтому, система должна периодически «пересматривать» трудности заданий в соответствии с тем, насколько часто и кем они были успешно выполнены. Удобные алгоритмы пересчета трудностей заданий предлагает IRT. В работе [3] предложен интерактивный алгоритм пересчета трудностей заданий с учетом уровней подготовленности испытуемых, выполняющих эти задания. Показано, что применение этого алгоритма позволяет более чем вдвое повысить достоверность результатов тестирования.

Еще один пример самообучения – выявление некорректных заданий. В предположении, что большинство заданий в тесте корректны, можно считать, что оценка уровня подготовленности, сделанная по всему тесту, более надежна,

чем оценка, сделанная по одному заданию. Таким образом, зависимость вероятности выполнения корректного задания от уровня подготовленности испытуемых (определенного по результатам выполнения всего теста) должна монотонно возрастать. Отклонение от этой монотонности может свидетельствовать о некорректности задания [4].

В соответствии с современными представлениями, целью обучения является формирование у учащихся совокупности компетенций, необходимых ему в будущей деятельности. Компетенции – способность применить полученные знания для решения практических проблем, возникающих в ходе профессиональной деятельности выпускника. Необходимые компетенции формируются на протяжении всего процесса обучения, при изучении каждой дисциплины, как правило, формируются несколько компетенций, поэтому инструменты контроля качества обучения должны обеспечивать оценку уровня сформированности сразу нескольких компетенций.

В классической IRT предполагается, что все задания посвящены одной изучаемой теме, а уровни подготовленности испытуемого θ и трудность задания δ можно выразить числами. Для оценки нескольких компетенций необходимо предположить, что уровни сформированности различных компетенций у испытуемого разные (число θ превращается в набор чисел: $\theta \rightarrow \theta_1, \theta_2, \theta_3 \dots$). Аналогично, на успешность выполнения задания компетенции влияют в различной степени: ($\delta \rightarrow \delta_1, \delta_2, \delta_3 \dots$). В работах [5, 6] предложено многомерное расширение классической IRT, позволяющее создавать тесты, выявляющие уровни сформированности сразу нескольких компетенций за один сеанс тестирования.

Одной из самых популярных систем организации обучения является MOODLE. Модуль тестирования в этой системе называется Quiz (викторина) [7]. Такое название во многом отражает взгляд авторов на задачи и функции компьютерного тестирования. Разработчики Quiz¹ много внимания уделили визуальному представлению заданий различных (иногда весьма экзотических) типов, а также методам организации викторины и подведению ее итогов. Однако реализация в Quiz других методов тестирования вызывает существенные трудности. Так, многочисленные попытки реализовать в MOODLE адаптивное тестирование не привели к созданию общепризнанного инструмента, доступного всем пользователям системы. [8, 9]. По мнению самих разработчиков, внедрение новых методов и инструментов тестирования потребует переосмысления архитектуры всей системы, кардинального изменения базы данных и структуры программных модулей. В настоящее время основные разработчики не готовы к столь кардинальным изменениям и планируют реализацию только небольшой части из них в новых версиях системы.

В результате широкого применения системы MOODLE в России, в ее

¹ Официально MOODLE является системой с открытым кодом и формально любой человек может участвовать в ее разработке. На практике, основные разработчики знают друг друга по имени, постоянно общаются и не очень-то доверяют новым участникам

формате были созданы и апробированы многочисленные тесты, позволяющие контролировать знания учащихся по многим дисциплинам. Разработка этих тестов потребовала огромного количества интеллектуального труда, а их применение позволило накопить бесценный опыт проверки знаний с помощью компьютера.

В данных обстоятельствах для реализации современных методов тестирования целесообразно разработать специальную программную оболочку для проведения тестирования и сделать ее максимально совместимой с MOODLE на информационном, методическом и идеологическом уровнях.

Такая оболочка должна:

1. Реализовывать следующие методы тестирования: классическое, адаптивное и многомерное, а также соответствующие методы обработки и представления результатов тестирования

2. Иметь собственную базу данных, хранящую всю информацию, необходимую для проведения тестирования, обработки, представления и анализа его результатов

3. Иметь интерфейс с системой MOODLE позволяющий:

а. Выгружать из MOODLE тесты и всю информацию необходимую для проведения тестирования и анализа его результатов

б. Формировать из выгруженной информации банк тестовых заданий (БТЗ)

с. Загружать в MOODLE результаты тестирований, проведенных в оболочке

4. Иметь методическую подсистему, позволяющую:

а. Выбрать метод тестирования и сформировать из заданий, хранящихся в БТЗ тест, соответствующий поставленной задаче,

б. Оценить соответствие теста и используемой рейтинговой шкалы

5. Иметь подсистему ведения банка тестовых заданий, позволяющую:

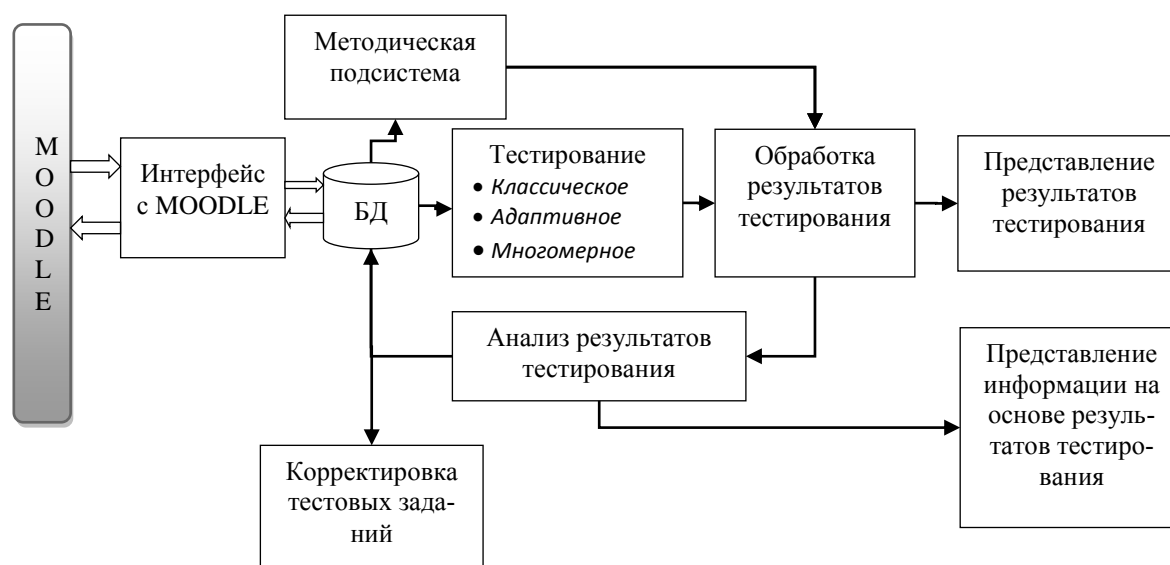
а. Периодически пересчитывать фактические трудности заданий и формировать рекомендации по пополнению БТЗ заданиями требуемой тематики и трудности

б. Выявлять и редактировать некорректные задания

с. Своевременно выявлять факты «взлома» тестов и обеспечивать эффективное противодействие этому.

Структура тестовой оболочки представлена на рисунке.

Работа выполнена при финансовой поддержке гранта РФФИ (проект № 20-013-00783.).



Структура тестовой оболочки

СПИСОК ЛИТЕРАТУРЫ

1. *Partchev I.* A visual guide to item response theory – Jena : Friedrich-Schiller-Universitat, 2004. 61 p. [Electronic resource]. URL: <https://docplayer.net/20748000-A-visual-guide-to-item-response-theory.html> (date of application: 06.09.2023).
2. *Bezrukov A. I., Akimova S. A.* Моделирование адаптивного теста по результатам классического тестирования. Современные информационные технологии и ИТ-образование. 2023. Vol. 19. No. 2. [Электронный ресурс]. URL: <http://sitito.cs.msu.ru/index.php/SITITO/article/view/928> (дата обращения: 06.09.2023).
3. *Безруков А. И., Акимова С. А., Погужильская Г. Г.* Имитационная модель для оценки достоверности и точности результатов тестирования // Сб. математическое моделирование и информационные технологии в исследованиях по физике и педагогике. 2017. С. 8-15.
4. *Безруков А. И., Погужильская Г. Г.* Алгоритм визуализации статистики результатов выполнения тестовых заданий, накопленных в системе АСТ // Известия Российского экономического университета им. Г. В. Плеханова. 2015. № 4 (22). С. 178-188.
5. *Соколова Т. Н., Гусятников В. Н., Безруков А. И., Каюкова И. В.* Методика оценки набора компетенций на основе результатов тестирования // Фундаментальные исследования. 2020. № 12. С. 209-215.
6. *Гусятников В. Н., Соколова Т. Н., Безруков А. И., Каюкова И. В.* Адаптивная модель тестирования нескольких компетенций на основе алгоритма Байеса // Современные наукоемкие технологии. 2022. № 1. С. 40-46.
7. База данных Moodle LMS 3.9. [Электронный ресурс]. URL: https://moodleschema.zoola.io/tables/quiz_attempts.html (дата обращения: 12.09.2023).
8. Moodle plugins directory: Adaptive Quiz: CAT implementation for Moodle. [Electronic resource]. URL: https://moodle.org/plugins/mod_adaptivequiz (date of application: 12.09.2023).
9. Исследовательский проект “Разработка плагина для адаптивного тестирования на платформе Moodle” OODLEion for чисел наших работости сразу нескольких компетенций за один сеанс тестирования. [Электронный ресурс]. URL: sowa-ru.com/item-work/2019-194/ (дата обращения: 12.09.2023).