


Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Комсомольский-на-Амуре государственный университет»

 На правах рукописи

Фролов Дмитрий Олегович

Проектирование и разработка адаптивной системы тестирования

Направление подготовки
09.04.01 «Информатика и вычислительная техника»

**АВТОРЕФЕРАТ
МАГИСТЕРСКОЙ ДИССЕРТАЦИИ**

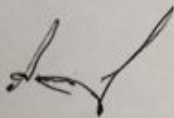
Работа выполнена в ФГБОУ ВО «Комсомольский-на-Амуре
государственный университет»

Научный руководитель:



Кандидат технических наук, доцент,
зав. кафедрой ПУРИС ФГБОУ ВО
«Комсомольский-на-Амуре государ-
ственный университет» Петрова Анна
Николаевна

Рецензент:



Кандидат физико-математических
наук, доцент кафедры Информацион-
ной безопасности, информационных
систем и физики ФГБОУ ВО «Амур-
ский гуманитарно-педагогический
государственный университет» Ани-
симов Антон Николаевич

Защита состоится « 29 » июня 2023 года в 14 часов 00 мин. на заседании
государственной экзаменационной комиссии по направлению 09.04.01 «Ин-
форматика и вычислительная техника» в Комсомольском-на-Амуре государ-
ственном техническом университете по адресу: 681013, г. Комсомольск-на-
Амуре, пр. Ленина, 27, ауд. 321/3

Автореферат разослан 26 июня 2023 г.

Секретарь ГЭК



Е.В. Абрамсон

Работа выполнена в ФГБОУ ВО «Комсомольский-на-Амуре
государственный университет»

Научный руководитель: кандидат технических наук, доцент, зав. кафедрой ПУРИС ФГБОУ ВО «Комсомольский-на-Амуре государственный университет» Петрова Анна Николаевна

Рецензент: кандидат физико-математических наук, доцент кафедры ИБИСиФ ФГБОУ ВО «Амурский гуманитарно-педагогический государственный университет» Анисимов Антон Николаевич

Защита состоится «29» июня 2023 года в 14 часов 00 мин. на заседании государственной экзаменационной комиссии по направлению 09.04.01 «Информатика и вычислительная техника» в Комсомольском-на-Амуре государственном техническом университете по адресу: 681013, г. Комсомольск-на-Амуре, пр. Ленина, 27, ауд. 313/5

Автореферат разослан 23 июня 2023 г.

Секретарь ГЭК

Е.В. Абрамсон

ОБЩАЯ ХАРАКТЕРИСТИКА ДИССЕРТАЦИОННОЙ РАБОТЫ

Актуальность темы исследования. В настоящее время становится ясно, что для успешного обучения студентов необходимо применять системы контроля знаний, которые позволяют определить недостаток подготовки по определенным темам, чтобы корректировать учебную нагрузку и повышать эффективность учебного процесса.

Одним из основных методов контроля знаний является прямое тестирование. При прямом тестировании структура теста не зависит от фактических ответов студента, а оценка его знаний происходит после полного завершения теста. Большинство существующих систем оценивания уровня знаний основано на прямом тестировании. Несмотря на свои преимущества, прямое тестирование имеет некоторые недостатки, такие как фиксированное количество вопросов и низкая вовлеченность студента (необходимость отвечать на вопросы, не соответствующие его уровню подготовки).

Для преодоления проблем, связанных с прямым тестированием, был разработан метод адаптивного тестирования. Адаптивное тестирование - это технология тестирования, при которой каждый следующий вопрос выбирается автоматически на основе ответов, данных на предыдущие вопросы. Такой подход позволяет адаптироваться к уровню знаний каждого студента и избегать слишком сложных или слишком легких вопросов.

Однако на данный момент системы с адаптивным тестированием все еще находятся на стадии развития, и эффективность моделей, используемых в таких системах, еще не изучена.

Цель исследования:

- сокращение временных затрат, которые требуются учебно-преподавательскому составу для разработки тестов;
- сокращение времени, затрачиваемого тестируемым на прохождение теста;
- повышение вовлеченности тестируемых и снижение влияния дополнительных факторов, таких как потеря интереса, отвлечение, утомление и беспокойство;
- увеличение точности результатов тестирования.

Для достижения поставленных целей необходимо решить следующие задачи исследования:

- проанализировать существующие системы адаптивного тестирования и выявить перспективные направления для их развития;
- исследовать применение математических моделей, основанных на теории IRT (Item Response Theory), в системах адаптивного тестирования;

- разработать концепцию и создать систему адаптивного тестирования на основе наиболее эффективной математической модели.

Объектом исследования являются методы и модели контроля знаний.

Предметом исследования являются технологии адаптивного тестирования.

Методологические основы и методы исследования: для решения поставленных задач использовались методы математического моделирования, системного анализа, статистического анализа и численные методы.

Хронологические рамки исследования определяются периодом с 2021 года по 2023 год.

Информационную базу исследования составили материалы опубликованных исследований отечественных авторов, ориентированных на описание математических моделей для систем адаптивного тестирования.

Научная новизна предлагаемых исследований заключается:

1 в поиске наиболее эффективной математической модели IRT (Item Response Theory) для анализа и подбора вопросов в системах адаптивного тестирования. Это означает, что исследование стремится определить и разработать модель, которая обеспечивает максимальную точность и адаптивность в процессе тестирования, учитывая ответы студентов на предыдущие вопросы;

2 в разработке авторского алгоритма для прохождения адаптивного тестирования. Этот алгоритм представляет собой уникальный подход или методологию, которая оптимизирует процесс тестирования, учитывая особенности выбранной математической модели. Такой алгоритм включает в себя специфические правила для выбора следующего вопроса, адаптацию сложности вопросов и другие инновационные компоненты, которые улучшают эффективность и результаты адаптивного тестирования.

Таким образом, научная новизна исследований заключается в предложении новой и эффективной математической модели IRT для систем адаптивного тестирования, а также в разработке авторского алгоритма, который оптимизирует процесс прохождения таких тестов.

Достоверность результатов исследования подтверждается следующими факторами:

1 практическая реализация и успешное применение системы адаптивного тестирования, разработанной на основе предложенного подхода. Это означает, что исследование было применено на практике и показало положительные результаты в реальной среде. Реализация и использование системы адаптивного тестирования учащихся подтверждает эффективность разработанного подхода и его практическую применимость;

2 вычислительные эксперименты, проведенные для проверки эффективности и преимуществ численных методов идентификации вероятностных моделей. Эти эксперименты позволяют оценить и подтвердить правильность и точность методов, используемых в исследовании. Результаты экспериментов подтверждают, что предложенные численные методы являются эффективными и превосходят другие методы идентификации вероятностных моделей.

Таким образом, достоверность результатов исследования подтверждается как практической реализацией и успешным применением системы адаптивного тестирования, так и вычислительными экспериментами, подтверждающими эффективность созданных численных методов идентификации вероятностных моделей.

Практическая значимость и ценность полученных результатов исследования заключаются в следующем: разработанное научно-методическое сопровождение и программно-инструментальная среда адаптивного тестирования имеют широкие возможности для применения в учебном процессе и контроля знаний на всех уровнях образования. Это означает, что результаты исследования могут быть непосредственно применены в практике образовательных учреждений, включая школы, колледжи, университеты и другие образовательные организации.

Основные положения, выносимые на защиту:

- 1 авторский алгоритм прохождения адаптивного тестирования;
- 2 программный комплекс системы адаптивного тестирования учащихся.

Апробация результатов. Результаты работы докладывалась и обсуждались на V, VI, VII Всероссийской национальной научной конференции молодых учёных «Молодежь и наука: актуальные проблемы фундаментальных и прикладных исследований».

Объём и структура диссертационной работы: диссертация включает в себя введение, три основные главы, заключение, список используемой литературы и одно приложение. Объём диссертации составляет 94 страницы. Текст работы содержит 1 таблицу и 35 рисунков. Для написания работы использовалось 30 источников литературы.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении дано обоснование актуальности и характеристика работы, сформулированы цель и задачи, указаны методы исследования, представлены основные положения, показана научная новизна и практическая значимость, описана структура работы.

В первой главе рассмотрена концепция, алгоритмы и основные модели систем адаптивного тестирования. Приведены формы осуществления компьютерного тестирования на рисунке 1.

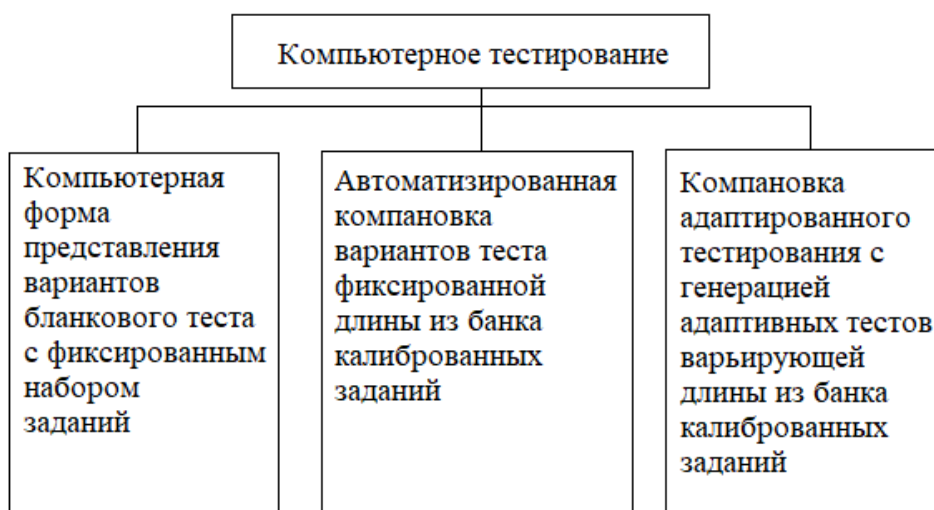


Рисунок 1 – Формы компьютерного тестирования

Описаны достоинства и недостатки компьютерного тестирования.

Был проведен анализ основных этапов адаптивного тестирования и процесса формирования тестовой базы. Описано, как происходит формирование и структурирование базы вопросов для проведения адаптивного тестирования. Также представлен основной алгоритм для прохождения адаптивного теста, который является циклическим и продолжается до выполнения условия остановки. На рисунке 2 изображена схема, описывающая данный процесс.

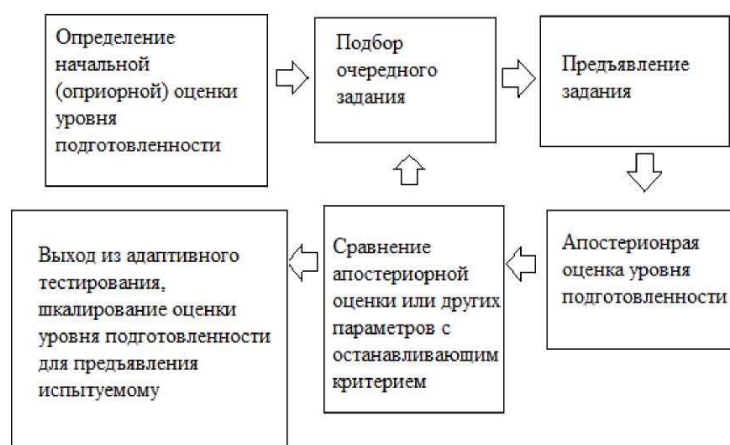


Рисунок 2 - Алгоритм варьирующего многошагового тестирования

На первом этапе определяется априорный или начальный уровень подготовки тестируемого. Это необходимо сделать для выбора оптимальной сложности при начале тестирования. На втором этапе происходит подбор и предъявление задания тестируемому. На третьем этапе происходит обработка ответа от тестируемого и установка апостериорной оценки уровня подготовленности. На четвертом этапе происходит сравнение апостериорной оценки и других критериев с останавливающим критерием (время тестирования, минимальная и максимальная сложность теста). Если условие остановки не выполнено, то процесс тестирования возвращается в третий этап и происходит подбор очередного вопроса. Если условие остановки выполнено, то процесс тестирования переходит в пятый этап – шкалирования оценки уровня знаний и предоставление результатов теста.

Каждый этап алгоритма может быть реализован уникальным образом в зависимости от вида стратегии и имеющихся технологических возможностей. Существующие системы, реализующие данный алгоритм, являются закрытыми и это не позволяет их проанализировать.

На рисунке 3 изображены траектории адаптивного тестирования трех учащихся, начинавших свой вход в адаптивные тестирования без выполнения протест теста.

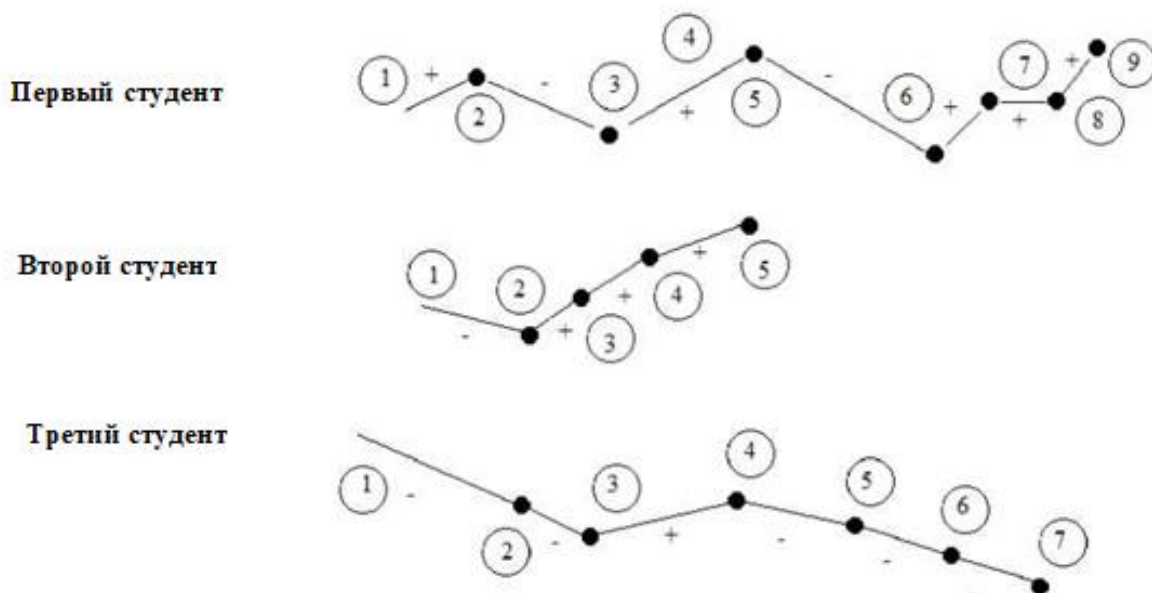


Рисунок 3 - Визуализация индивидуальных траекторий испытуемых

Сложность предоставляемых заданий на рисунке 3 определена высотой ломаной. Продолжительность тестирования и количество вопросов определено длиной ломаной. Для визуального удобства траектории тестирования на

графике отображаются как непересекающиеся ломаные линии. При правильном ответе на тестовый вопрос ломаная линия на графике возрастает, при неправильном ответе - убывает.

Надежность в таком тестировании будет зависеть от совокупности факторов. К ним относятся: высокая степень вариативности тестовых заданий, как по уровню сложности, так и по типу; наличие рецензентов для проверки качества заданий; наличие фокус группы для правильного распределения заданий по уровню сложности. На надежность также влияет качество начального контроля (наличие или отсутствие проктор теста) и выбранная математическая модель адаптивного тестирования.

Во второй главе представлена теория IRT, включающая математические модели, основанные на этой теории. Математические модели, представленные в главе, сравниваются по критерию уровня получаемых знаний. Наиболее эффективная математическая модель интегрируется в базовый алгоритм адаптивного тестирования.

На сегодняшний день, при разработке систем адаптивного тестирования, наибольшее распространение получила теория тестовых заданий, также известная как IRT. Она описывает связь между уровнем знаний испытуемых и результатами выполнения тестов, что позволяет определить уровень знаний независимо от сложности заданий.

Основная цель IRT заключается в обеспечении эффективного прогнозирования результатов тестирования для заданий различной сложности. Согласно IRT, вероятность правильного ответа на задание должна быть выше у испытуемого с высоким уровнем знаний по сравнению с испытуемым с низким уровнем знаний. Чем выше уровень знаний, тем выше вероятность правильного ответа на задание определенной сложности. Особенностью применения IRT является то, что ответы множества испытуемых на множество заданий теста прогнозируются на основе математических моделей, используя эмпирически полученную матрицу исходных тестовых баллов. Эмпирический способ получения матрицы исходных тестовых баллов предполагает для распределения вопросов по уровням сложности на основе выполнения их тестовой группой. Традиционно за правильный ответ ставится 1 балл, за неправильный – 0 баллов. Теория IRT является основополагающей в рамках концепции адаптивного тестирования. На основе IRT строятся большинство моделей адаптивного тестирования.

Сравнительный анализ моделей по количеству получаемой информации за ответ приведен в таблице 1.

Таблица 1 – Сравнение информационной ценности ответов

	Количество информации				
	Модель Раша	Двухпараметрическая модель ($a = 0,3$)	Partial Credit Model		
			двух-балльное задание	трёх-балльное задание	четырёх-балльное задание
Максимальное значение	0,25	0,0225	0,67	1,25	2
Незначительное расхождение уровня подготовленности испытуемого и уровней трудности заданий (0,1 лолита)	0,249	0,0225	0,644	1,2	1,901
Значительное расхождение уровня подготовленности испытуемого и уровней трудности заданий (1 лолит)	0,197	0,0220	0,424	0,788	0,837

После проведения анализа моделей адаптивного тестирования были сделаны следующие выводы:

1 тип математической модели оказывает значительное влияние на эффективность адаптивного тестирования.

2 в условиях адаптивного тестирования наиболее эффективной является двухпараметрическая модель, при наличии достаточного количества заданий с высокой способностью дифференциации, равномерно распределенных по всему диапазону уровней подготовки испытуемых.

3 модель РСМ и модель Раша предпочтительны, если вышеперечисленное условие не выполняется.

Для реализации системы адаптивного тестирования будет использоваться модель Раша, поскольку она обладает высокой эффективностью. Еще одним преимуществом модели Раша является ее низкая зависимость от способности заданий к дифференциации, что позволяет достичь максимальной эффективности даже в условиях недостаточной наполненности тестовой базы.

Давайте теперь рассмотрим предлагаемый алгоритм тестирования (рисунок 4), который предлагается испытуемым. Введем обозначения: X –

количество выданных заданий; Y – количество правильных ответов, данных испытуемым; Θ – текущий уровень подготовки испытуемого; i – номер текущего шага; R – оценка испытуемого (неудовлетворительно, удовлетворительно, хорошо или отлично); G – группа сложности, g – вопрос теста. Также для описания алгоритма необходимо ввести понятие критического количества вопросов $X_{\text{крит}}$ – количество вопросов для группы сложности, на которое испытуемый должен ответить (без изменения группы сложности), чтобы выставить ему оценку R . Значение $X_{\text{крит}}$ зависит от общего числа вопросов в базе и в каждой группе сложности и от времени тестирования; на начальном этапе определяется преподавателем. Во избежание постоянных смен уровней сложности при чередовании верных и ошибочных ответах тестируемого введём переменную η – коэффициент верных ответов для сохранения уровня сложности (изменяется от 0 до 1), также определяется преподавателем. Рекомендуется на начальном этапе использовать значение $\eta = 1/3$.

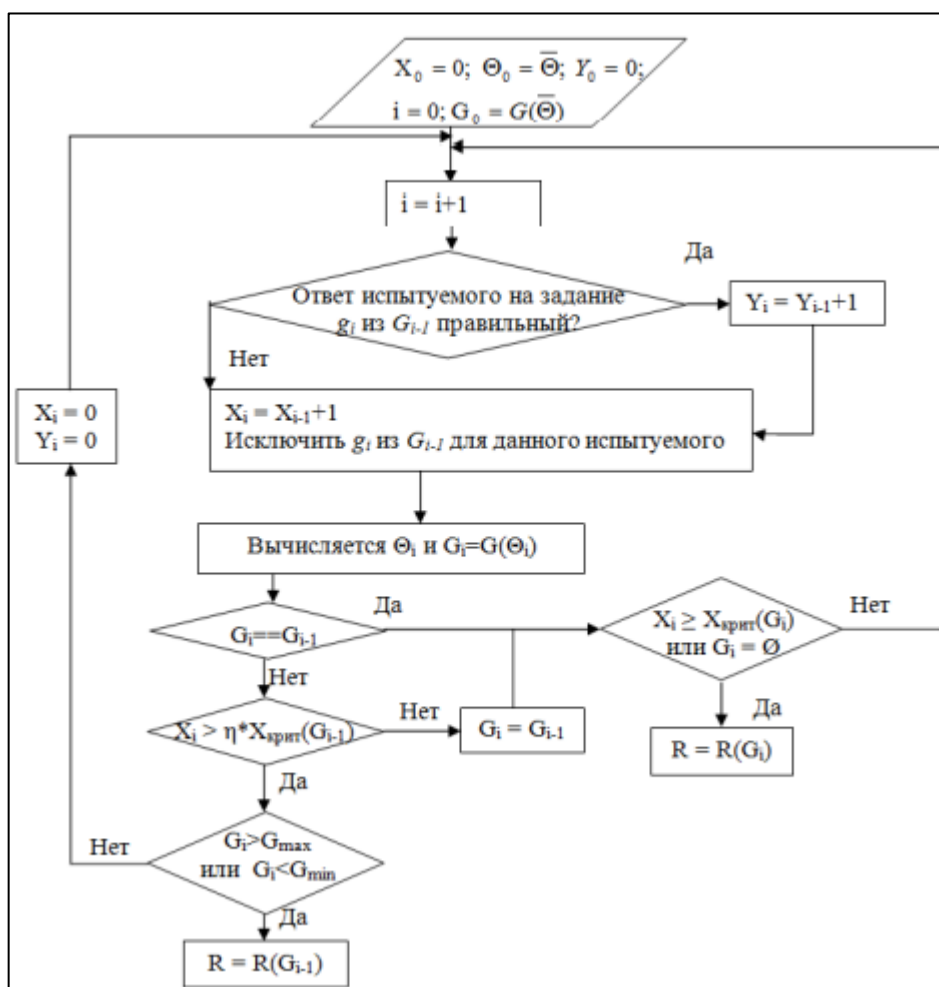


Рисунок 4 – Алгоритм адаптивного тестирования

Процедура оценки уровня подготовки испытуемых (рисунок 4):

1 выбирается задание из текущей группы сложности вопросов $g_i \in G_i$ предлагается для ответа испытуемому;

2 при успешном ответе испытуемого, увеличивается количество правильных ответов;

3 увеличивается количество выданных заданий, выданное задание исключается из набора заданий для данного испытуемого;

4 вычисляется новый уровень подготовки испытуемого Θ_i по модели Раша ранее описанным способом;

5 вычисляется группа сложности в зависимости от значения Θ_i . Принадлежность Θ_i к группе сложности определяется путем сравнения значения Θ_i с максимальным и минимальным значениями трудности заданий для каждой группы сложности. Если $\Theta_i > \beta_{\min}(G_k)$ и $\Theta_i < \beta_{\max}(G_k)$, то $G_i = G_k$ (k – индекс перебора по всем группам сложности, k изменяется от 1 до 4);

6 если группа сложности изменилась, количество выданных заданий превышает η -ую часть от $X_{\text{крит}}$ и новая группа сложности больше максимальной (оценка 5) или меньше минимальной (оценка 2), то выставляется оценка в соответствии с группой сложности на предыдущем шаге (5 или 2);

7 если группа сложности изменилась, количество выданных заданий превышает η -ую часть от $X_{\text{крит}}$ и новая группа сложности не больше максимальной (оценка 5) и не меньше минимальной (оценка 2), то количество выданных вопросов и правильных ответов обнуляется и происходит возврат к шагу 1;

8 если группа сложности изменилась, а количество выданных заданий не превышает η -ую часть от $X_{\text{крит}}$, то группа сложности возвращается к предыдущему значению (изменения не принимаются);

9 если группа сложности не изменилась, не превышено количество критических вопросов и текущая группа сложности содержит еще хотя бы 1 вопрос, то происходит возврат к шагу 1;

10 если группа сложности не изменилась, но при этом превышено количество критических вопросов или все вопросы текущей группы сложности уже были выданы испытуемому ранее, то выставляется оценка в соответствии с текущей группой сложности.

В третьей главе представлен процесс создания константной диаграммы (рисунок 5) с участниками процесса создания теста. Описываются данные для работы в адаптивной системе тестирования.

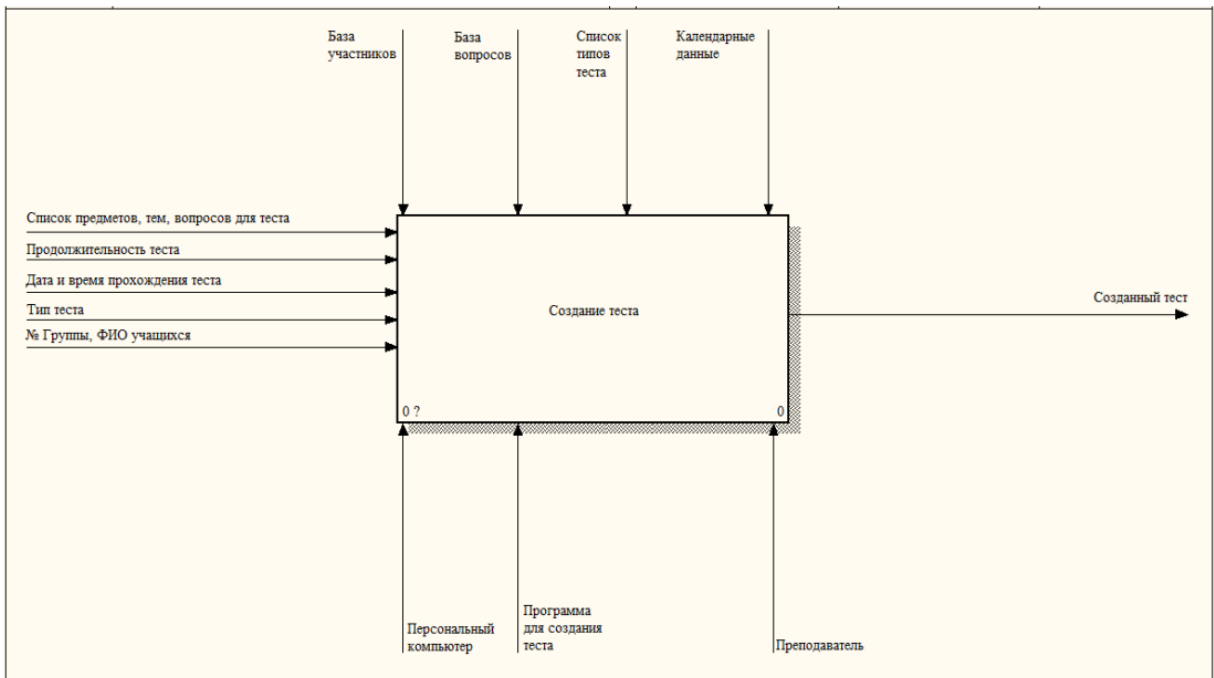


Рисунок 5 – Контекстная диаграмма

Так же созданы диаграммы декомпозиции, результат представлен на рисунке 6.

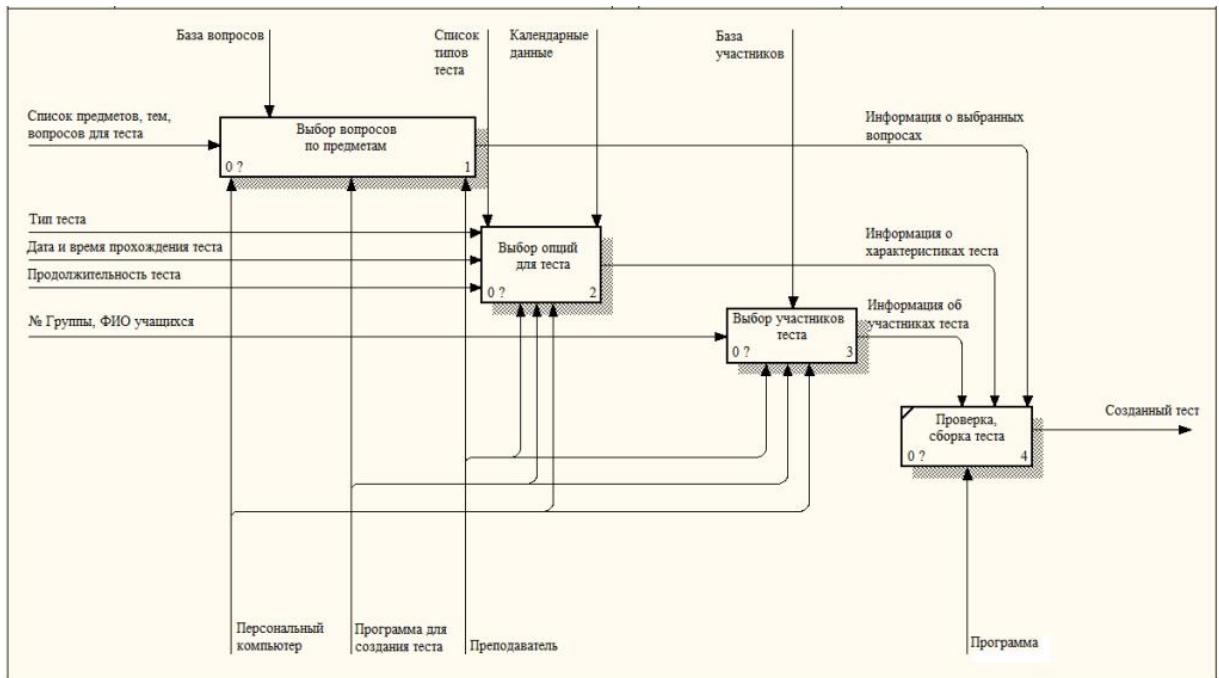


Рисунок 6 – Диаграмма декомпозиции для верхнего узла

В информационной системе адаптивного тестирования основными функциональными подсистемами являются следующие: авторизация,

регистрация, база тестов, база заданий, база учебных групп и проигрыватель теста.

Для разработки пользовательской части приложения был использован фреймворк Vue.js на языке JavaScript. Для разработки серверной части приложения был выбран фреймворк Laravel на языке PHP. В качестве реляционной базы данных для хранения данных была использована PostgreSQL. Общая архитектура приложения представлена на рисунке 7.

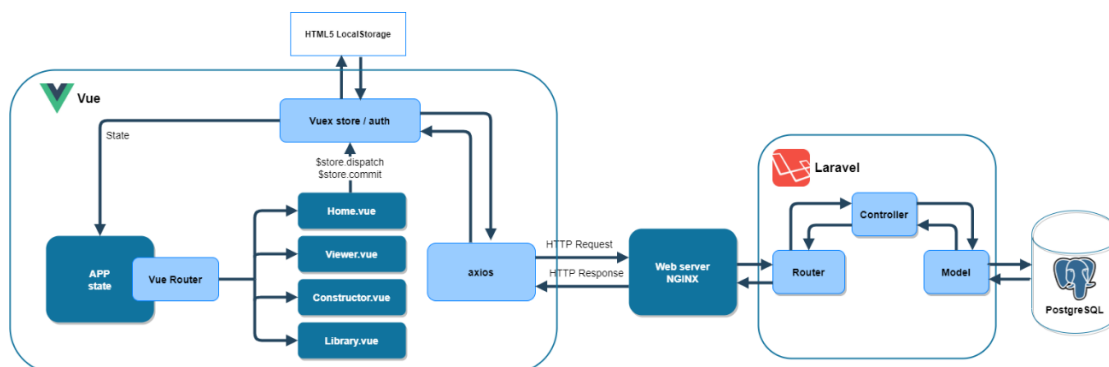


Рисунок 7 – Архитектура ИС адаптивного тестирования

Основной используемой единицей Vue являются компоненты - многократно используемые экземпляры Vue. Мы можем использовать компоненты как пользовательские элементы внутри корневого экземпляра Vue. Компоненты могут быть переиспользованы и вложены в другие компоненты, как на рисунке 8.

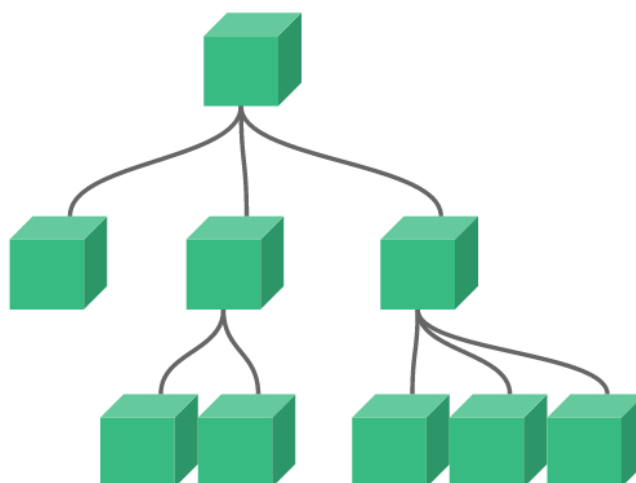


Рисунок 8 – Вложенность компонентов

На основе созданных моделей была разработана компонента программного обеспечения, которая предназначена для адаптивного тестирования учащихся (рисунок 9).

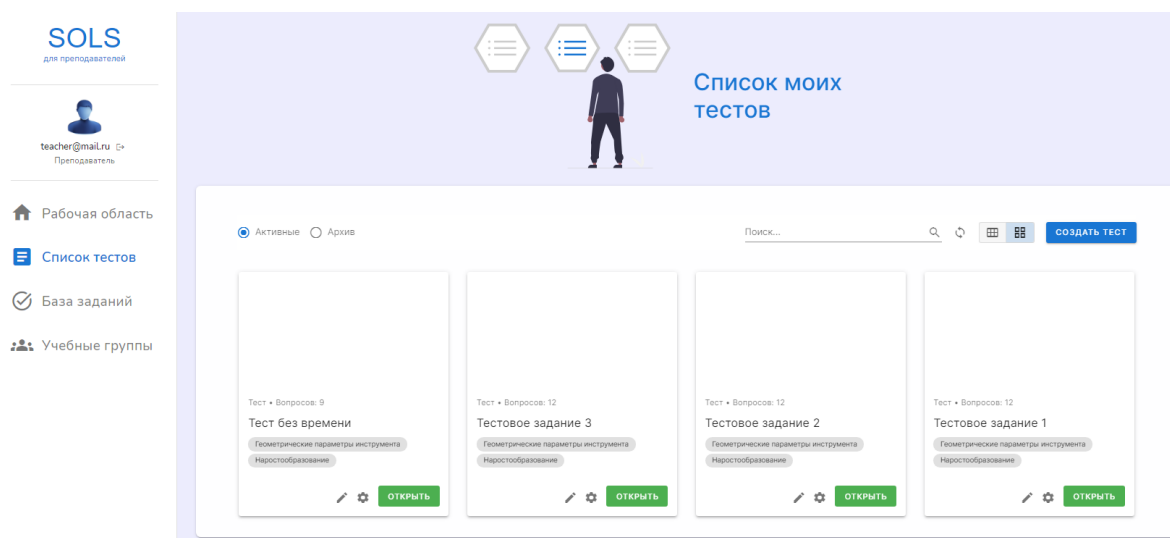


Рисунок 9 – Программная компонента

Так же был разработан функционал для изменения существующего теста. Данная возможность позволит преподавателю изменять параметры для созданного теста (рисунок 10).

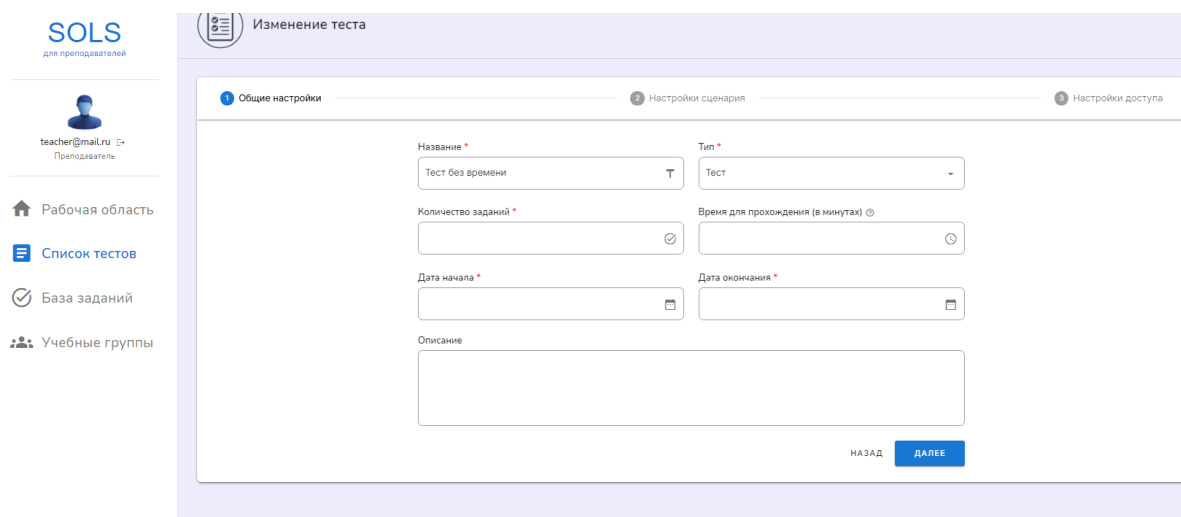


Рисунок 10 – Изменение параметров созданного теста

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

Анализ уже существующих моделей адаптивного тестирования позволил выявить перспективные направления их развития. Был обоснован выбор объекта и предмета исследования. Проведение исследования эффективности работы существующих моделей адаптивного тестирования показало, что основным недостатком большинства моделей является сильная зависимость количества информации от разности уровня подготовленности испытуемых и уровня сложности заданий. В результате, модель Раша оказалась наиболее устойчивой к изменению этого показателя.

На основе авторского алгоритма и выбранной модели (Раша) был разработан программный комплекс для адаптивного тестирования с индивидуальной траекторией прохождения для каждого испытуемого в области комплексов программ. Полученные результаты исследования имеют практическую значимость и ценность, поскольку разработанное научно-методическое сопровождение и программно-инструментальная среда для реализации адаптивного тестирования могут быть применены в качестве системы контроля знаний в различных учебных заведениях и использоваться в процессе повышения квалификации профессорско-преподавательского состава.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ИССЛЕДОВАНИЯ

Фролов Д.О. Адаптивная система тестирования / Д.О. Фролов, А.Н. Петрова // молодежь и наука: актуальные проблемы фундаментальных и прикладных исследований материалы V Всероссийской. нац. науч. конф. студентов, аспирантов и молодых ученых, Комсомольск-на-Амуре, (Комсомольск-на-Амуре, апрель 2022 г.).

Фролов Д.О. Разработка адаптивной системы тестирования / Д.О. Фролов, А.Н. Петрова // молодежь и наука: актуальные проблемы фундаментальных и прикладных исследований материалы VI Всероссийской. нац. науч. конф. студентов, аспирантов и молодых ученых, Комсомольск-на-Амуре, (Комсомольск-на-Амуре, декабрь 2022 г.).

Фролов Д.О. Алгоритм прохождения теста в адаптивной системе тестирования / Д.О. Фролов, А.Н. Петрова // молодежь и наука: актуальные проблемы фундаментальных и прикладных исследований материалы VII Всероссийской. нац. науч. конф. студентов, аспирантов и молодых ученых, Комсомольск-на-Амуре, (Комсомольск-на-Амуре, апрель 2023 г.).