



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ «Информатика и системы управления»

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
К ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЕ
НА ТЕМУ:

**«Метод прогнозирования итогов приёма в вузы России на
основе агентного моделирования»**

Студент группы **ИУ7-85Б**

(Подпись, дата)

Д.Р. Жигалкин

(И.О. Фамилия)

Руководитель

(Подпись, дата)

О.В. Кузнецова

(И.О. Фамилия)

Консультант

(Подпись, дата)

Д.А. Кузнецов

(И.О. Фамилия)

Нормоконтролер

(Подпись, дата)

Ю.В. Строганов

(И.О. Фамилия)

2022 г.

РЕФЕРАТ

Расчетно-пояснительная записка 46 с., 12 рис., 3 табл., 19 ист., 1 прил.

Ключевые слова: приемная кампания, абитуриенты, механизмы зачисления, ЕГЭ, агентное моделирование.

Объектом разработки является метод прогнозирования итогов приёма в ВУЗы России.

Цель работы – разработать и реализовать метод прогнозирования итогов приёма в ВУЗы России на основе агентного моделирования.

Для достижения поставленной цели необходимо решить следующие задачи:

- проанализировать правила организации приёма в ВУЗы России;
- получить информацию и статистику о ВУЗах России;
- получить информацию по УГСН и связанных с ними предметам ЕГЭ;
- проанализировать полученную информацию, выделить наиболее важные факторы для разработки модели поведения абитуриентов;
- проанализировать существующие методы и механизмы зачисления абитуриентов;
- разработать имитационную модель поведения абитуриентов;
- разработать метод прогнозирования итогов приёма в ВУЗы, используя модель поведения и статистические данные;
- программно реализовать разработанный метод.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	9
1 Аналитическая часть	10
1.1 Организация приема в вузы России	10
1.2 Статистические данные	13
1.3 Построение устойчивых паросочетаний	15
1.4 Бостонский механизм построения неустойчивых паросочетаний	17
1.5 Агентно-ориентированное моделирование	18
1.5.1 Структура агентно-ориентированной модели	18
1.5.2 Агент как элемент структуры модели	19
1.5.3 Взаимодействие агентов	20
1.6 Формализация задачи	21
2 Конструкторская часть	23
2.1 Разработка алгоритма поиска общего названия ВУЗа	23
2.2 Разработка алгоритма приведения данных к официальному перечню направлений подготовки высшего образования	23
2.3 Генерация популяции агентов	25
2.4 Модель поведения и характеристики агента	27
2.5 Разработка метод прогнозирования итогов приема в ВУЗы	27
2.6 Особенности и допущения предлагаемого метода	31
2.7 Используемые структуры данных	32
3 Технологическая часть	33
3.1 Выбор и обоснование языка программирования и среды разработки	33
3.2 Организация клиент-серверного взаимодействия	34
3.3 Структура разработанного ПО	34
3.4 Минимальные требования к вычислительной системе	36
3.5 Формат входных и выходных данных	36
3.6 Тестирование ПО	37
4 Исследовательская часть	40
4.1 Различное количество бюджетных мест в ВУЗе	40
4.2 Различное максимальное число ВУЗов для подачи заявлений	40

4.3	Различные диапазоны результатов сдачи ЕГЭ	40
4.4	Различное соотношение готовых к переезду абитуриентов	40
ЗАКЛЮЧЕНИЕ		41
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ		42
ПРИЛОЖЕНИЕ А Учёт индивидуальных достижений		46

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

ВУЗ – Высшее Учебное Заведение.

ЕГЭ – Единый Государственный Экзамен.

УГСН – Укрупнённые группы специальностей и направлений подготовки.

Проходной балл – минимальный балл, с которым были зачислены абитуриенты на определенное направление подготовки или специальность. Это сумма баллов ЕГЭ или вступительных испытаний и индивидуальных достижений.

Агентное моделирование – это метод имитационного моделирования, который основан на взаимодействии независимых агентов, обладающих собственными характеристиками.

ВВЕДЕНИЕ

В настоящее время изучение выбора абитуриентом высшего учебного заведения остается актуальным. Выбирая место для получения образования, абитуриент должен решить какие специальности для него предпочтительны, какие характеристики вуза играют ключевую роль.

Неразрывно с этим выбором связана и политика вуза, направленная на привлечение лучших абитуриентов. Для достижения этой цели применяются различные методы, одним из которых является моделирование приёмной кампании, позволяющее выявить слабые стороны процесса приёма, а также изучить реакцию абитуриентов на закрытие, либо открытие новой специальности, изменение экономического показателя региона и спрогнозировать спрос на высшее образование.

1 Аналитическая часть

1.1 Организация приема в вузы России

В настоящее время, набор на программы подготовки специалистов и бакалавров в российские ВУЗы проводится по результатам Единого Государственного Экзамена[1], который должны сдавать все школьники по окончании 11 классов общеобразовательного учреждения, причем для получения аттестата необходимо сдать как минимум 2 обязательных предмета: русский язык и математику. До 2021 года прием в вузы проводился по результатам трех (иногда четырех) ЕГЭ, определенных для каждой специальности. Все абитуриенты российских университетов должны были предоставить в приемную комиссию результаты ЕГЭ по русскому языку и профильному ЕГЭ по выбранной специальности, третий предмет ВУЗ выбирал самостоятельно из перечня экзаменов, определенных для данного направления. С 2022 года учебные заведения вправе установить третий ЕГЭ на выбор абитуриента из нескольких предложенных вариантов. Для того, чтобы подать документы в выбранный ВУЗ, необходимо набрать на ЕГЭ по каждому предмету количество баллов, равное или превышающее минимальный балл. ВУЗ сам определяет минимальный балл для каждой специальности и направления, но не может установить его ниже уровня, утвержденного Рособрнадзором[2].

Каждый выпускник имеет право одновременно подать документы в 5 ВУЗов, в каждом можно выбрать 1-10 специальностей по всем формам обучения. Количество специальностей с 2021 года определяется ВУЗом. Схематично это представлено на рисунке 1.

Списки абитуриентов, подавших документы, ранжируются по количеству баллов, то есть более высокие позиции занимают абитуриенты, у которых совокупное количество баллов за ЕГЭ, дополнительные вступительные испытания и индивидуальные достижения выше. Рассматривается сумма баллов без учета индивидуальных достижений, затем профильный предмет и далее в порядке

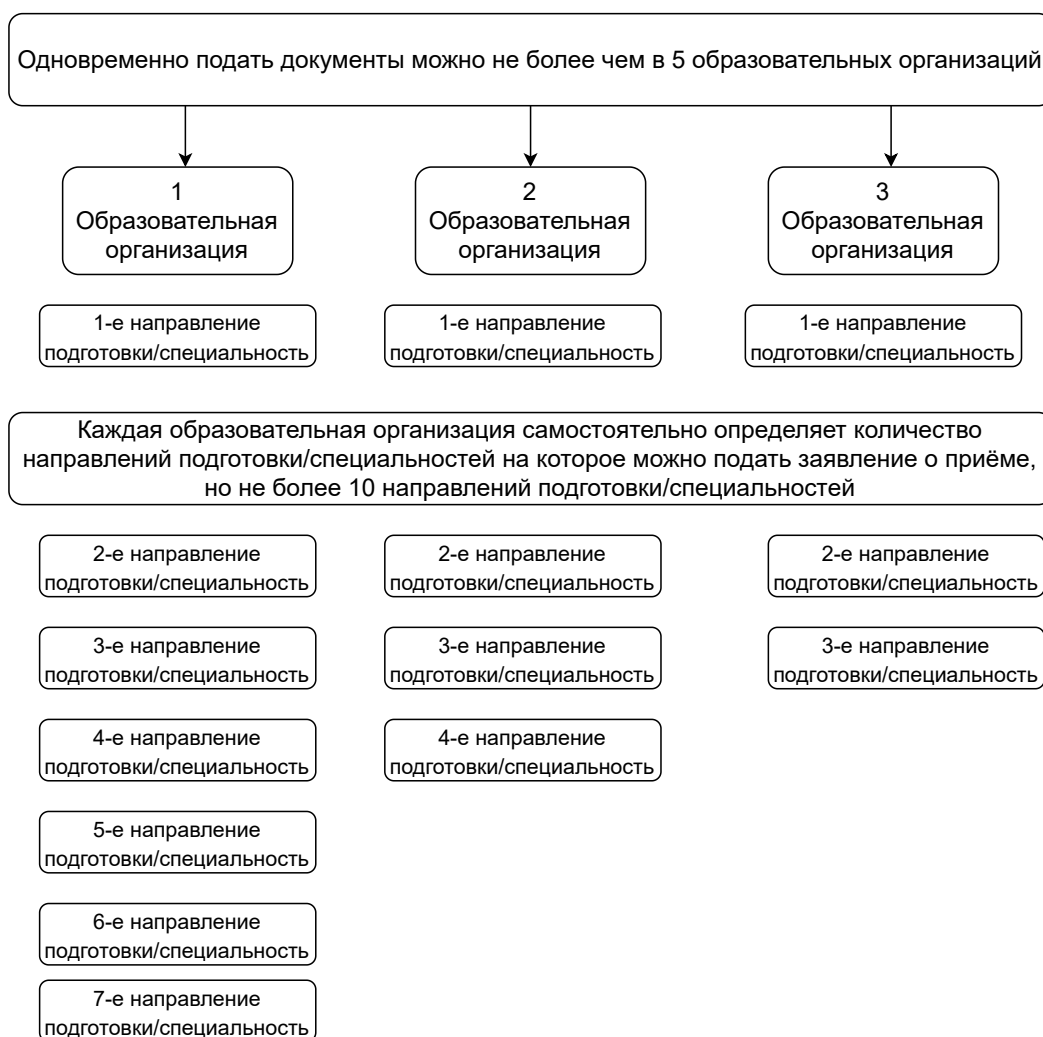


Рисунок 1 – Подача заявления возможна не более чем в 5 образовательных учреждениях

убывания приоритетности. Если у двух абитуриентов совпадает весь перечень, то приоритет отдается тому, у кого есть преимущественное право. При равенстве по всем критериям список ранжируется по индивидуальным достижениям, учитываемым при равенстве поступающих по иным критериям ранжирования. В таблице 1 представлен простой ранжированный список поступающих.

Таблица 1 – Ранжированный список поступающих.

Номер в сводке	Фамилия	Балл	Наличие аттестата	Наличие заявления
1	Гусев	299	Нет	Нет
2	Слепухина	282	Нет	Нет
3	Стрельцов	265	Нет	Да
4	Абасов	248	Да	Да
5	Петрова	221	Нет	Нет
6	Калинин	201	Да	Нет
7	Иванов	194	Да	Да

Согласно правилам приема[3], ВУЗы к результатам ЕГЭ могут добавить своим абитуриентам до 10 баллов за индивидуальные достижения. Каждый университет сам устанавливает количество дополнительных баллов за индивидуальные достижения. Как правило, больше всего баллов университеты добавляют за аттестат с отличием и результаты олимпиад, не использованные для получения особых прав.

После этого начинается зачисление. До 2021 года оно проходило в несколько этапов:

- этап приоритетного зачисления — зачисляют абитуриентов, которые поступают без экзаменов, в рамках особой или целевой квоты[4]. Эти абитуриенты должны подать в ВУЗ оригинал документа о предшествующем образовании и заявление о согласии на зачисление;

- I этап зачисления — на этом этапе ВУЗ может заполнить до 80% бюджетных мест, оставшихся свободными после приоритетного зачисления, по каждой специальности или направлению. Абитуриентов зачисляют в соответствии с позицией, которую они занимают в списке поступающих, — первыми зачисляют тех, кто занимает более высокую позицию. На этом этапе нужно подать оригинал документа о предшествующем образовании и заявление о согласии на зачисление;
- II этап зачисления — вуз заполняет оставшиеся 20% бюджетных мест.

С 2021 года основная волна зачисления только одна[5]. На досрочном этапе ВУЗы принимают олимпиадников, поступающих вне конкурса, абитуриентов, имеющих льготы, и целевиков. Если после основной волны останутся свободные места, ВУЗ вправе принять на них абитуриентов из конкурсного списка.

По итогам конкурса определяется проходной балл — наименьшее количество баллов, которого оказалось достаточно для зачисления. Таким образом, проходной балл меняется каждый год и определяется только после зачисления. Абитуриенты, которые поступают по квотам, имеют право принимать участие и в общем конкурсе. Для этого они также должны набрать количество баллов, равное или превышающее минимальное значение, установленное ВУЗом.

1.2 Статистические данные

Согласно статистике на 2019 год, в России ведёт образовательную деятельность 741 государственный ВУЗ [7]. Публичную информацию о них можно получить на следующих ресурсах:

- мониторинг качества приема в ВУЗы[8];
- информационно-аналитические материалы по результатам проведения мониторинга эффективности деятельности образовательных организаций высшего образования[10].

В таблице 2 приведен пример получаемой информации с первого ресурса по направлениям подготовки.

Из информационно-аналитических материалов второго ресурса можно

Таблица 2 – Пример информации по направлениям подготовки.

Название ВУЗа	Направление подготовки	Год	Средний балл ЕГЭ поступивших на бюджетные места	Число студентов, поступивших на бюджетные места
Орловский гос. ун-т	Биология	2020	59.9	46

получить обширную статистику по многим показателям вуза, выделим основные из них:

- название ВУЗа;
- регион расположения учреждения;
- год статистики;
- наличие общежития;
- средний балл ЕГЭ по очной форме на бюджетные места;
- средний балл ЕГЭ по очной форме на бюджетные места по общему конкурсу;
- усредненный по реализуемым направлениям минимальный балл ЕГЭ по очной форме на программы бакалавриата и специалитета;
- численность студентов, победителей и призеров заключительного этапа всероссийской олимпиады школьников, членов сборных команд Российской Федерации принятых на очную форму обучения на первый курс по программам бакалавриата и специалитета без вступительных испытаний;
- численность студентов, победителей и призеров олимпиад школьников, принятых на очную форму обучения на первый курс по программам бакалавриата и специалитета;
- численность студентов, принятых по результатам целевого приема на первый курс на очную форму обучения по программам бакалавриата

и специалитета;

- удельный вес численности студентов, принятых по результатам целевого приема на первый курс на очную форму обучения по программам бакалавриата и специалитета в общей численности студентов, принятых на первый курс по программам бакалавриата и специалитета на очную форму обучения;
- удельный вес численности иностранных студентов (кроме стран Содружества Независимых Государств), обучающихся по программам бакалавриата, специалитета, магистратуры, в общей численности студентов (приведенный контингент);
- удельный вес численности иностранных студентов из СНГ, обучающихся по программам бакалавриата, специалитета, магистратуры, в общей численности студентов (приведенный контингент);
- средний балл ЕГЭ студентов, принятых на обучение по программам бакалавриата и специалитета, по всем формам обучения;
- доля обучающихся по программам бакалавриата, специалитета, магистратуры в очной форме;
- список реализуемых УГСН.

Статистические данные о распределении абитуриентов по регионам, количестве 100-балльников в каждом из них России можно получить на портале открытых данных [9].

Перечень вступительных испытаний при приеме на обучение по образовательным программам высшего образования - программам бакалавриата и программам специалитета можно получить из приказа Минобрнауки России [6].

1.3 Построение устойчивых паросочетаний

Исследование механизмов построения паросочетаний в ситуациях, когда участники имеют предпочтения и свободу принятия решений, было начато в классической работе[24]. Авторы рассматривали модели построения паросоче-

таний один-к-одному (когда каждый участник получает не более одного партнера) и один-ко-многим (когда участник одной стороны рынка – студенты получают ровно одного партнера, а участники другой стороны рынка – ВУЗы получают более одного партнера). Авторами был предложен алгоритм, позволяющий за конечное число шагов получить паросочетание при известных предпочтениях участников, являющихся линейными порядками на множестве участников противоположной стороны. В литературе этот алгоритм называется алгоритмом отложенного принятия (deferred acceptance algorithm).

Существует две версии алгоритма в зависимости от того, кто является предлагающей стороной, абитуриенты или ВУЗы. Рассмотрим вариант, в котором предлагающей стороной являются студенты. До начала работы алгоритма подаются предпочтения участников – линейные порядки на множестве агентов противоположной стороны. На первом шаге каждый абитуриент обращается в наиболее предпочтительный для себя ВУЗ. ВУЗ рассматривает все полученные заявления и временно принимает q лучших абитуриентов, где q соответствует числу мест. На следующем шаге те абитуриенты, которым было отказано, обращаются в следующие по предпочтительности ВУЗы, они снова рассматривают все полученные заявления (оставленные с первого шага и вновь поступившие). Если ВУЗ получает новое заявление от абитуриента, которого он предпочитает временно зачисленному, то последнему ВУЗ отказывает и на следующем шаге такой абитуриент делает предложение следующему ВУЗу. Эта процедура продолжается до тех пор, пока все абитуриенты не будут зачислены в ВУЗы либо не получат отказ от всех ВУЗов, указанных в предпочтениях. Важным достоинством механизма является то, что он позволяет получить устойчивое паросочетание, то есть такое, что

- никакой участник не хочет отказаться от предложенного партнера и остаться один;
- никакие два участника не желают сговориться и заключить соглашение друг с другом, избежав предписанного распределения.

Кроме того, при использовании варианта механизма, в котором предлагающей стороной являются студенты, им оказывается выгодно представлять свои истинные предпочтения. Работа[24] получила широкое развитие и в теоретических исследованиях, и в прикладном дизайне централизованных механизмов. Механизм отложенного принятия (с незначительными вариациями) был внедрен на программе распределения медицинских интернов в США, системе зачисления в ВУЗы Турции по итогам единого тестирования, а также в департаментах образования Бостона, Нью-Йорка и других городов США для зачисления в муниципальные школы.

1.4 Бостонский механизм построения неустойчивых паросочетаний

Описанный механизм часто именуется "Бостонским механизмом" так как именно этот механизм использовался для распределения детей Бостона по муниципальным школам.

Перед началом работы механизма участники (школы и поступающие) сообщают свои предпочтения в виде линейных порядков (списков) друг относительно друга. На первом шаге каждый абитуриент подает заявку в наилучшую для себя школу. Если школа получила больше заявок, чем имеется мест, то она отбирает лучших абитуриентов. На втором шаге абитуриенты, отвергнутые на первом шаге, подают заявки во вторые по предпочтительности школы. Школы не пересматривают результаты зачисления на первом шаге, даже если вновь обратившиеся абитуриенты для них предпочтительнее уже зачисленных. Каждая школа зачисляет лучших из вновь обратившихся абитуриентов на оставшиеся места. Так продолжается до тех пор, пока все абитуриенты не будут зачислены (либо отвергнуты всеми школами). Такой механизм ведет к созданию неустойчивого паросочетания. Действительно, рассмотрим абитуриента, поставившего на первое и второе места школы X и Y , соответственно. Пусть на первом шаге школа X не приняла абитуриента, так как получила достаточное количество заявлений от более предпочтительных абитуриентов; школа Y также полностью заполнила свои места на первом шаге. Тогда на втором шаге абитуриент подает

заявление в Y . Но даже если Y считает этого абитуриента лучшим среди всех, она не сможет его принять, так как все места заполнены на первом шаге.

Этот алгоритм был заменен алгоритмом отложенного принятия, что позволило улучшить получаемое распределение для существенной доли абитуриентов и ликвидировать необходимость манипулирования при представлении предпочтений.

1.5 Агентно-ориентированное моделирование

Агентное моделирование – это мощный метод имитационного моделирования, который за последние несколько лет нашел применение в ряде прикладных задач, таких как потребительские и финансовые рынки, организационное поведение, транспортные системы, логистика.

Агентно-ориентированное моделирование – это система которая моделируется как совокупность автономных субъектов принятия решений, называемых агентами. Каждый агент индивидуально оценивает свою ситуацию и принимает решения на основе набора правил. Агенты могут демонстрировать различное поведение, подходящее для системы, которую они представляют, например, производство, потребление или продажа. На простейшем уровне агентно-ориентированная модель состоит из системы агентов и отношений между ними. Даже простая агентно-ориентированная модель может демонстрировать сложные модели поведения и предоставлять ценную информацию о динамике реальной системы, которую она имитирует. Агент – некоторая сущность обладающая активностью, некоторой активностью поведения. Может принимать решение в соответствии с некоторым набором правил взаимодействия с окружением, а также самостоятельно изменяться. Кроме того, агенты могут развиваться, позволяя проявиться непредвиденному поведению. Сложная система иногда включает нейронные сети, эволюционные алгоритмы или другие методы обучения, чтобы обеспечить реалистичное обучение и адаптацию.

1.5.1 Структура агентно-ориентированной модели

Типичная агентная модель состоит из трех элементов:

- набор агентов, их атрибуты и поведение;
- набор агентских отношений и способов взаимодействия;
- среда агентов, где они взаимодействуют между собой.

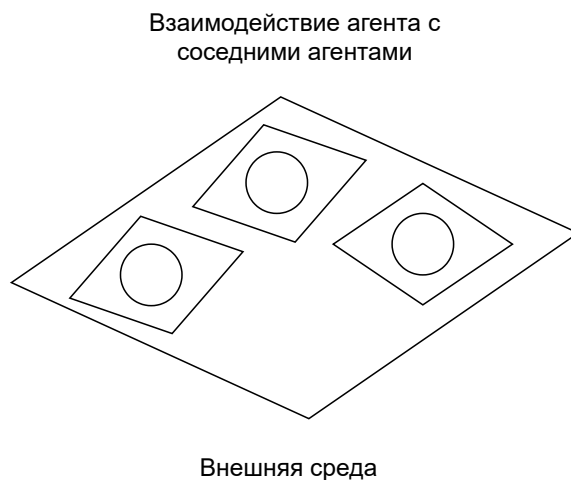


Рисунок 2 – Структура типичной агентно-ориентированной модели

Для создания модели на основе агентов необходимо проанализировать, смоделировать и запрограммировать эти элементы.

Структура типичной модели показана на рисунке 2. Каждый элемент данной структуры будет разобран далее.

1.5.2 Агент как элемент структуры модели

Наиболее важная определяющая характеристика агента – это его способность действовать автономно, то есть действовать самостоятельно в ответ на ситуации, с которыми он сталкивается. Агенты наделены поведением, которое позволяет им принимать независимые решения, обычно они активны, иницируют свои действия для достижения внутренних целей, а не просто пассивны, реагируют на других агентов и окружающую среду.

С точки зрения практического моделирования, агенты обладают некоторыми существенными характеристиками:

- автономность и однозначная идентифицируемость;
- самостоятельность, может независимо функционировать в своей среде и взаимодействовать с другими агентами, по крайней мере, в огра-

ниченном диапазоне ситуаций, представляющих интерес для модели;

- агент имеет состояние, которое представляет основные переменные, связанные с его текущей ситуацией;
- динамичное взаимодействие с другими агентами, влияющими на его поведение. У агентов есть протоколы для взаимодействия с другими агентами, способность реагировать на окружающую среду.

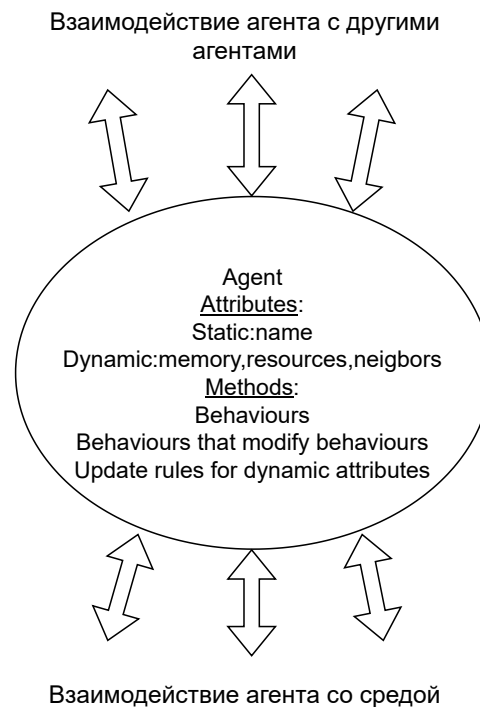


Рисунок 3 – Структура типичного агента

Структура типичного агента показана на рисунке 3. Все связанное с агентом является атрибутом, либо методом, который работает с агентом. Атрибуты могут быть статическими, не изменяемыми во время моделирования, и динамическими. Например, статический атрибут – это имя агента, динамический атрибут - это память агента о прошлых взаимодействиях.

1.5.3 Взаимодействие агентов

Двумя основными проблемами моделирования взаимодействий агентов являются определение того, кто с кем может быть связан, и механизмы динамики взаимодействий. Оба аспекта необходимо учитывать при разработке агентно-ориентированных моделей. Один из принципов сложных систем и агент-

ного моделирования заключается в том, что агенту доступна только локальная информация. Агентные системы – это децентрализованные системы, нет центрального органа, который либо распространяет глобально доступную информацию всем агентам, либо контролирует их поведение в целях оптимизации производительности системы. Агенты взаимодействуют с другими агентами, но не все постоянно взаимодействуют напрямую со остальными, как в реальных системах. Агенты обычно взаимодействуют с подмножеством других агентов, называемых соседями. Локальная информация получается из взаимодействий с соседями агента (не с каким-либо агентом или всеми агентами) и из его локальной среды (а не из какой-либо части всей среды). Как правило, набор соседей агента быстро меняется по мере моделирования. То, как агенты связаны друг с другом, обычно называют топологией или связностью агентно-ориентированной модели. Типичные топологии включают пространственную сетку или сеть узлов (агентов) и связей (отношений).

1.6 Формализация задачи

В соответствии с выбранной темой выпускной квалификационной работы необходимо разработать и реализовать метод прогнозирования итогов приёма в ВУЗы России на основе агентного моделирования с использованием полученных из открытых источников статистических данных о ВУЗах и УГСН, а также распределении абитуриентов по регионам. Необходимо провести обработку полученных статистических данных.

Входные данные для проведения моделирования:

- статистические данные о ВУЗах и их УГСН;
- распределение абитуриентов по регионам России;
- перечень вступительных испытаний и соответствующие им УГСН;
- длительность этапа перераспределения заявлений с копиями аттестатов;
- максимальное количество ВУЗов, в которые можно подать заявление;
- длительность этапа поиска наилучшего места для оригинала аттестата;

та.

— популяция агентов (абитуриентов) с заданными характеристиками.

Результатом проведенного моделирования являются:

- средний балл зачисленных абитуриентов по каждому ВУЗу, участвовавшему в моделировании;
- максимальный балл по каждому ВУЗу;
- минимальный балл по каждому ВУЗу;
- количество зачисленных абитуриентов по каждому ВУЗу;
- средний балл всех зачисленных абитуриентов на конкретный УГСН по каждому ВУЗу;
- максимальный балл по каждому УГСН;
- минимальный балл по каждому УГСН;
- количество зачисленных абитуриентов по каждому УГСН.

Для всех значений баллов в результатах моделирования необходимо показать изменение (в количественном выражении) в сравнении с результатом приема прошлого года.

2 Конструкторская часть

2.1 Разработка алгоритма поиска общего названия ВУЗа

Поскольку количество анализируемых ВУЗов составляет несколько сотен, сбор данных должен производиться в автоматизированном режиме.

В процессе анализа статистических данных было обнаружено, что именованная ВУЗов на аналитических ресурсах различны, что затрудняет их объединение, поскольку название ВУЗа в данном случае является его единственным уникальным идентификатором.

Одним из вариантов решения данной задачи является отправка двух запросов в сеть Интернет, где в качестве параметра поиска передавать названия ВУЗа с информационных источников. Таким образом, будут получены соответствия для двух различных названий ВУЗов и одного унифицированного, с помощью которого можно производить объединение данных.

В таблице 3 представлен пример такой интерпретации.

2.2 Разработка алгоритма приведения данных к официальному перечню направлений подготовки высшего образования

В данных об укрупнённых группах специальностей и направлений, полученных с мониторинга качества приёма в ВУЗы[10], существует допущение — указаны неофициальные укрупнённые группы, составленные самим мониторингом из официальных групп. Для решения данной задачи необходимо выделить таблицу, в которой будет указано, какую долю имеет официальная группа, в неофициальной. С помощью данной таблицы можно пересчитать количество бюджетных мест по формуле

$$N_{budget} = \sum P_i k_i, \quad (1)$$

где N_{budget} — количество бюджетных мест в официальной группе;
 P_i — количество мест в i -ой неофициальной группе;
 k_i — i -я доля официальной группы в неофициальной.

Средний балл в каждой группе ВУЗа из информационно-аналитических материалов по следующей формуле

$$B_{aver.score} = \frac{\sum P_i k_i B_i}{N_{budget}}, \quad (2)$$

где $B_{aver.score}$ — средний балл в официальной группе;
 B_i — средний балл в i -ой неофициальной группе.

Последним этапом приведения будет масштабирование значений количества бюджетных мест и среднего балла в группе. Его необходимо проводить, если верно следующее:

Таблица 3 – Интерпретация названия ВУЗа поисковой системой Яндекс.

Название ВУЗа на первом ресурсе	Моск. гос. техн. ун-т. им. Н.Э. Баумана
Название ВУЗа на втором ресурсе	федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный технический университет имени Н.Э.Баумана (национальный исследовательский университет)»
Общее название	Московский государственный технический университет им. Н.Э. Баумана

$$N_{official} \neq N_{unofficial}, \quad (3)$$

где $N_{official}$ — сумма бюджетных мест всех официальных групп ВУЗа;
 $N_{unofficial}$ — сумма бюджетных мест всех неофициальных групп ВУЗа.

В данном случае необходимо пересчитать количество бюджетных в каждой группе по описанной выше формуле:

$$N_{budget} = \mu \sum P_i k_i, \quad (4)$$

где $\mu = \frac{N_{official}}{N_{unofficial}}$

Формула для пересчёта среднего балла остается без изменений.

2.3 Генерация популяции агентов

Генерация популяции агентов осуществляется с помощью следующих параметров:

- процентные соотношения численности категорий студентов с различными диапазонами полученных результатов ЕГЭ;
- диапазон результатов ЕГЭ, которые будут заданы агентам каждой категории;
- данные о распределении агентов по регионам;
- данные об связи УГСН и принимаемых предметах ЕГЭ;
- процентные соотношения признака смены региона для каждой категории;
- количество различных УГСН, которые рассматривает агент;
- минимальные баллы ЕГЭ по каждому предмету.

Все агенты разделены на 4 категории: не набравшие минимальный балл, сдавшие ЕГЭ в диапазоне от минимального балла до заданного и 2 категории для конфигурирования агентов, которые сдали ЕГЭ выше результатов двух предыдущих категорий.

В каждой категории, кроме той, где абитуриенты на набрали минимальный балл, агенту проставляются результаты успешно сданных ЕГЭ, а также интересующие его УГСН. Для этого при обработке очередного агента случайным образом определяется его специализация, в зависимости от чего и прописываются результаты ЕГЭ и связанные с ними интересы. Всего таких специализаций 8 и соответствующие им ЕГЭ следующие:

- иностранный язык и литература;
- история и обществознание;
- химия и биология;
- информатика и ИКТ;
- физика;
- физика и информатика и ИКТ;
- биология, химия и информатика и ИКТ;
- история, обществознание, география и литература.

Следует уточнить, что русский язык и математика являются обязательными к сдаче, поэтому они не указывались в списке выше, поскольку уже подразумеваются как базовые.

Первым этапом генерации популяции агентов является создание пустого агента без характеристик и проставление ему домашнего региона и признака готовности к смене домашнего региона в зависимости от параметров конфигурации, полученной от пользователя. Домашний регион определяется из полученного распределения абитуриентов по регионам. Текущий обрабатываемый регион будет для созданного агента домашним.

Вторым этапом является установка каждому агенту результатов сдачи ЕГЭ, в зависимости от специализации, и связанных с ними интересов. Результаты сдачи ЕГЭ проставляются каждому агенту случайным образом из заданного диапазона значений для конкретной категории. Из всех подходящих агенту УГСН выбирается то количество, которое указано в конфигурации, полученной от пользователя.

Полученная популяция агентов сохраняется в базу данных для дальнейшего использования для моделирования.

2.4 Модель поведения и характеристики агента

В данной задаче агентами являются абитуриенты – сущности, обладающие активностью и характеристиками.

Базовыми характеристиками абитуриента являются:

- уникальный идентификатор абитуриента;
- возможность смены региона;
- результаты сдачи ЕГЭ;
- интересующие УГСН, определяющиеся по сданным ЕГЭ.

Каждый абитуриент оценивает свою ситуацию, анализируя свое положение в конкурсном списке (отсортированном по сумме баллов ЕГЭ) и по количеству доступных бюджетных мест на данный УГСН. Если он не попадает в доступные места, то принимается решение о полной смене текущего ВУЗа со всеми заявлениями в нем или о поиске другого ВУЗа, в который можно положить заявление на такой же УГСН.

Оценивание ситуаций производится до тех пор, пока можно подавать новые заявления. Затем абитуриент должен положить оригинал аттестата к одному из своих заявлений. Искомое заявление выбирается исходя из престижности ВУЗа – абитуриенты стремятся попасть в наиболее престижные места. Если выбор производится внутри одного ВУЗа на разных УГСН, то сравниваются величины престижности данных УГСН.

На этапе выбора заявления для подачи оригинала аттестата возможна ситуация, когда абитуриент не имеет ни одного заявления с копией аттестата, соответственно он не может претендовать на возможность обучения в ВУЗе.

2.5 Разработка метод прогнозирования итогов приема в ВУЗы

Ниже представлена IDEF0-диаграмма разрабатываемого метода на рисунке 4.

Разрабатываемый метод прогнозирования состоит из четырех этапов:

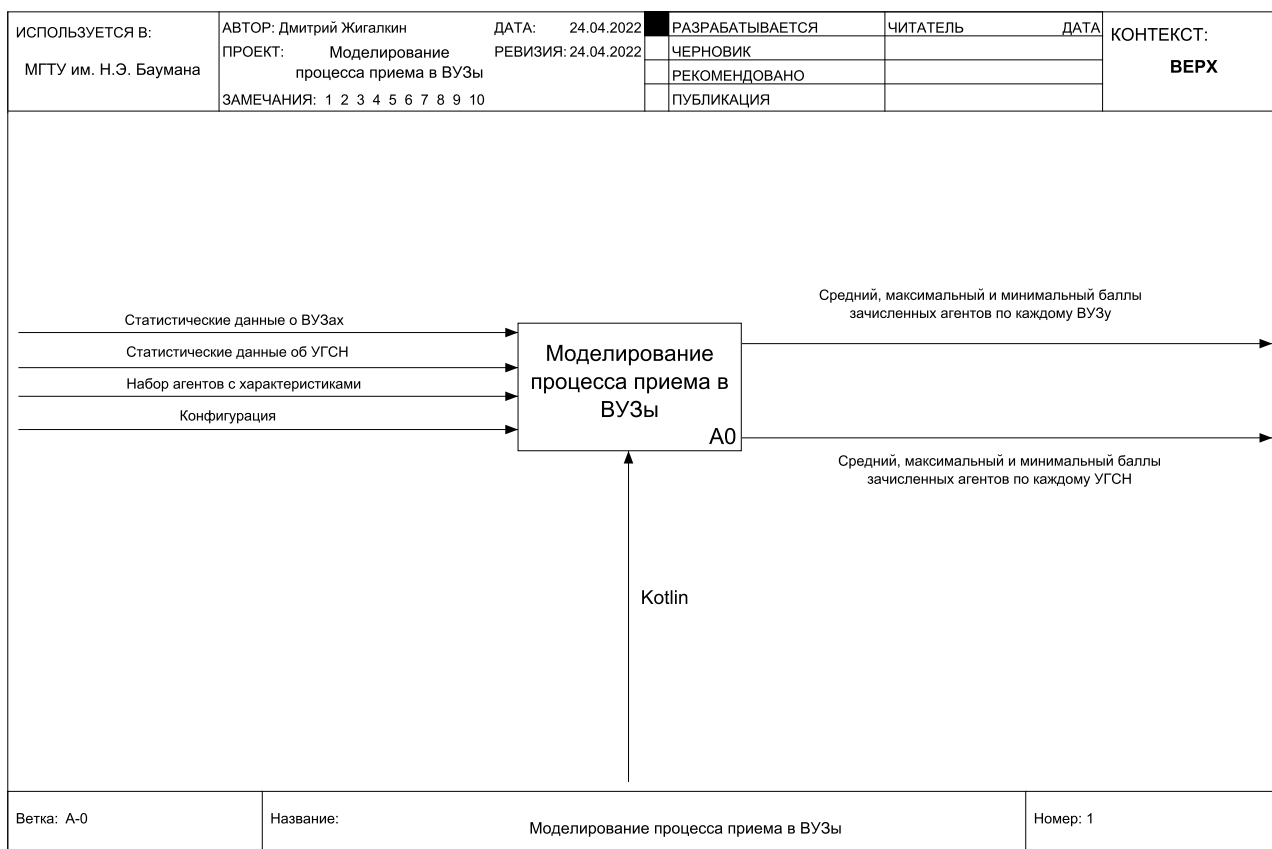


Рисунок 4 – IDEF0 диаграмма ветка A-0

- базовое распределение абитуриентов по ВУЗам на основе проходного балла прошлого года;
- перераспределение агентами поданных заявлений для поиска УГСН, на который имеются шансы пройти по бюджетным местам в текущий момент;
- подача оригинала аттестата к своему заявлению, по результатам сравнения престижности ВУЗов и их УГСН;
- анализ полученных результатов.

Настройки конфигурации включают в себя поля:

- год проведения моделирования;
- количество абитуриентов, участвующих в моделировании;
- длительность этапа перераспределения заявлений с копиями аттестатов;
- максимальное количество ВУЗов, в которое может подать заявления

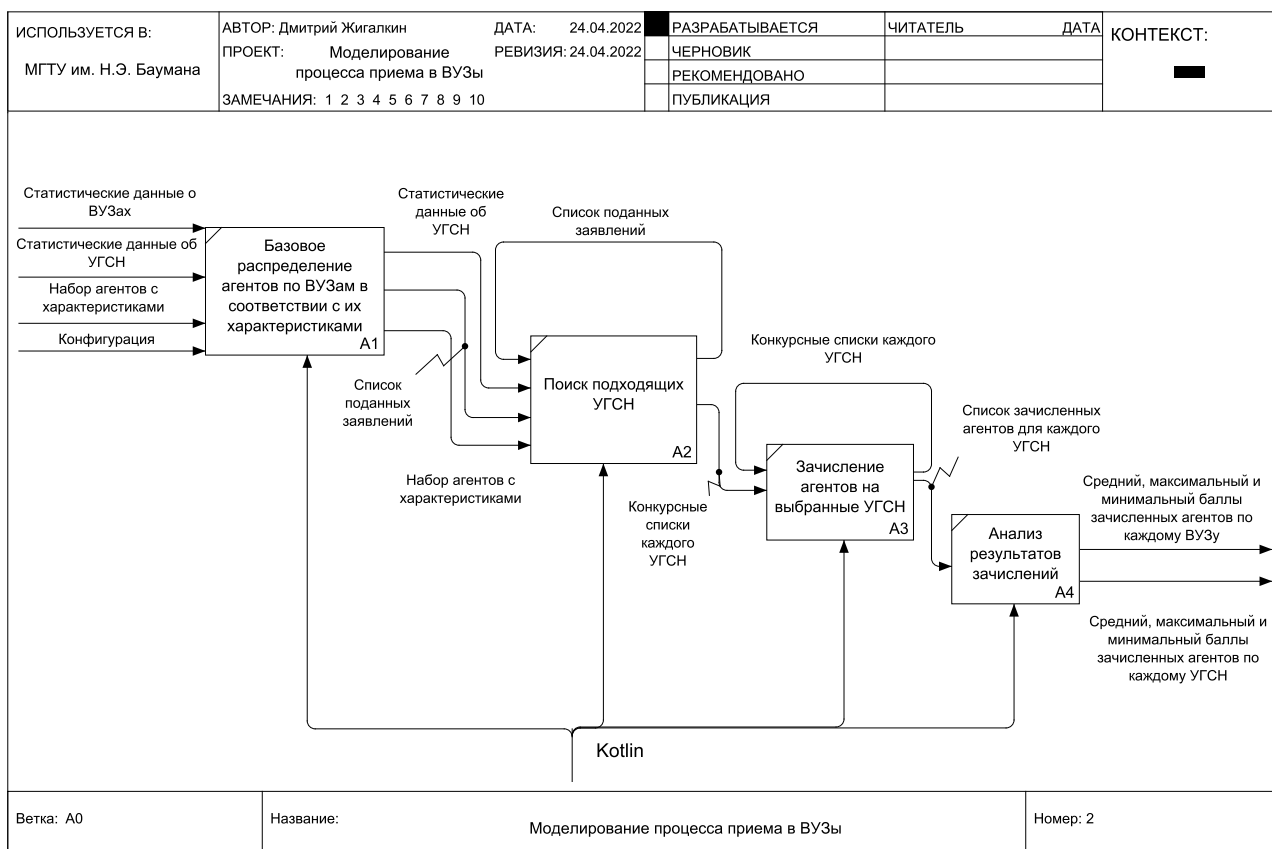


Рисунок 5 – IDEF0 диаграмма ветка A0

агент;

— длительность этапа поиска наилучшего места для оригинала аттестата.

Первым этапом моделирования является базовое распределение абитуриентов по УГСН в ВУЗах на основе среднего балла по УГСН в прошлом году. Абитуриенты просматривают ВУЗы, начиная с самого престижного, проверяют наличие интересующего УГСН, какие результаты ЕГЭ необходимы для подачи заявления, сравнивают со своими результатами (средним баллом по предметам, необходимым для поступления на данный УГСН) и кладут копию аттестата к заявлению.

Второй этап моделирования представлен в виде схемы алгоритма поиска подходящих УГСН на рисунке 6

Третий этап – подача оригинала аттестата к своему заявлению и зачисления. Студенты, не имеющие заявлений ни в каком ВУЗе не рассматриваются. Ана-

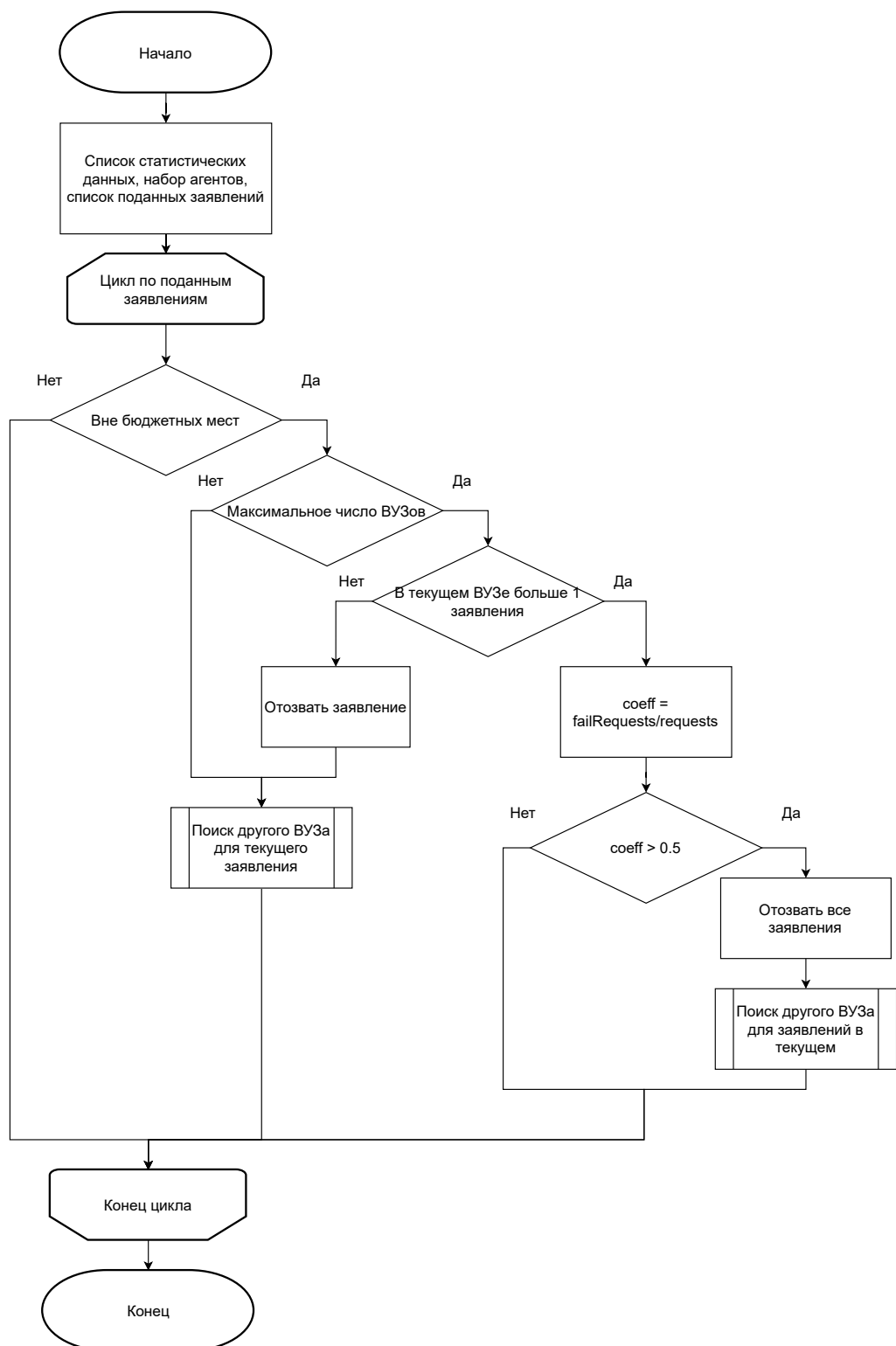


Рисунок 6 – Схема алгоритма поиска подходящего УГЧН

лизируем каждое заявление абитуриента, проходит ли он в текущий момент по бюджетным местам среди других абитуриентов с оригиналами аттестатов. Отмечаем данную заявку как лучшую, если абитуриент проходит на нее и она лучше предыдущего выбора. Сравнение текущего и предыдущего выбора производится на основе престижности ВУЗа, если сравниваются УГСН в разных ВУЗах и престижности УГСН, если сравнение производится внутри одного ВУЗа. Зачисление на доступные бюджетные места абитуриентов с наиболее высокими баллами в конкурсном списке.

Завершающий этап - анализ сложившейся ситуации в конкурсном списке каждого УГСН, вычисление минимального, максимального, среднего балла по данному УГСН и ВУЗу в целом и сравнение результата моделирования с итогами приема абитуриентов предыдущего года.

2.6 Особенности и допущения предлагаемого метода

Данный метод не предусматривает многопоточную реализацию процесса моделирования. Для оптимизации работы с базой данных необходимо использовать массовую вставку записей в одной транзакции, вместо множества запросов в транзакциях.

При выборе из нескольких УГСН, которые подходят по результатам ЕГЭ не учитываются индивидуальные предпочтения отдельно взятого агента, предпочтение отдается наиболее престижному УГСН, на который абитуриент может поступить.

При моделировании учитываются только бюджетные места для обучения, поскольку сведений об их количестве для каждого УГСН нет в открытых источниках.

Агенты кладут оригинал аттестата только на те УГСН, на которые гарантированно смогут поступить, рискованные и нечестные стратегии не предусмотрены.

2.7 Используемые структуры данных

Для хранения списка агентов необходимо использовать стандартную реализацию массива фиксированной длины, поскольку количество студентов не меняется в процесса моделирования.

Для эффективного доступа к статистической информации о ВУЗе, применяется хэш-таблица, которая позволяет хранить пары вида ключ-значение и выполнять три операции: операцию добавления новой пары, операцию поиска и операцию удаления пары по ключу.

Кеширование значений предметов ЕГЭ для каждого УГСН производить с помощью хэш-таблицы, значением которой является множество – коллекция неупорядоченных элементов.

Конкурсные списки с информацией на какой УГСН подал заявление агент хранить в потокобезопасной реализации хэш-таблицы для избежания ошибок при модификации одной коллекции разными итераторами.

3 Технологическая часть

3.1 Выбор и обоснование языка программирования и среды разработки

Для реализации разработанного метода используется язык программирования Kotlin – статически типизированный, объектно-ориентированный язык программирования[11], поскольку имеется опыт ведения разработки на данном языке, что сократит время кодирования программы.

Ключевой особенностью языка Kotlin является то, что его код сначала транслируется в специальный байт-код, независимый от платформы. А затем этот байт-код выполняется виртуальной машиной JVM (Java Virtual Machine). В этом плане Java отличается от стандартных интерпретируемых языков, код которых сразу же выполняется интерпретатором. В то же время Kotlin не является и чисто компилируемым языком.

Подобная архитектура обеспечивает кроссплатформенность и аппаратную переносимость программ, благодаря чему подобные программы без recompilации могут выполняться на различных платформах - Windows, Linux, Mac OS и т.д. Для каждой из платформ может быть своя реализация виртуальной машины JVM, но каждая из них может выполнять один и тот же код.

В качестве среды разработки была выбрана одна из самых популярных – IntelliJ IDEA от компании JetBrains[26].

Среди преимуществ данной программы можно выделить следующие:

- наличие студенческой лицензии для ведения некоммерческой разработки;
- простые и удобные средства отладки программного кода;
- интеграция с системой контроля версий;
- работа с базой данных, получение выгрузок и резервных копий;
- встроенные средства автоматизации сборки проекта.

Для реализации фронтенд части проекта был использован JavaScript –

один из самых популярных языков программирования в мире. Он используется для создания интерактивных веб-страниц.

3.2 Организация клиент-серверного взаимодействия

Поскольку для взаимодействия с конечным пользователем было решено использовать веб-интерфейс, который предусматривает дальнейшее расширение и масштабирование приложения, необходимо организовать передачу сообщений от сервера к клиенту в реальном времени без перезагрузки страницы.

Для решения данной задачи была выбрана технология Server-Send Events (SSE)[25]. Основная концепция заключается в том, что клиент подписывается на события, присылаемые сервером, в свою очередь сервер шлет их с помощью специального объекта, который именуется как SSE Emitter. На рисунке 7 показан пример такого взаимодействия.

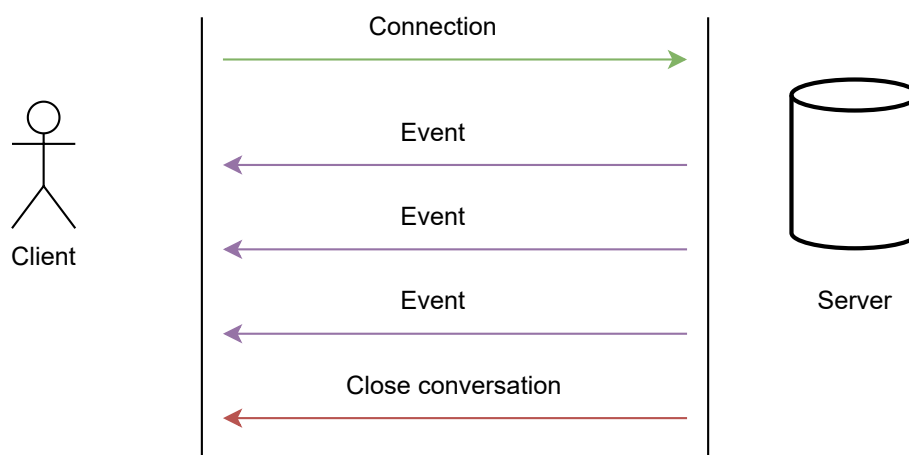


Рисунок 7 – Организация взаимодействия по технологии SSE

3.3 Структура разработанного ПО

Структура разработанного программного обеспечения представлена на рисунке 8.

Каждый модуль, изображенный на диаграмме, содержит в себе сгруппированные по функциональному значению соответствующие классы. Модуль взаимодействия с пользователем отвечает за пользовательский интерфейс. Модуль управления связывает модули разработанного метода и модуль взаимо-

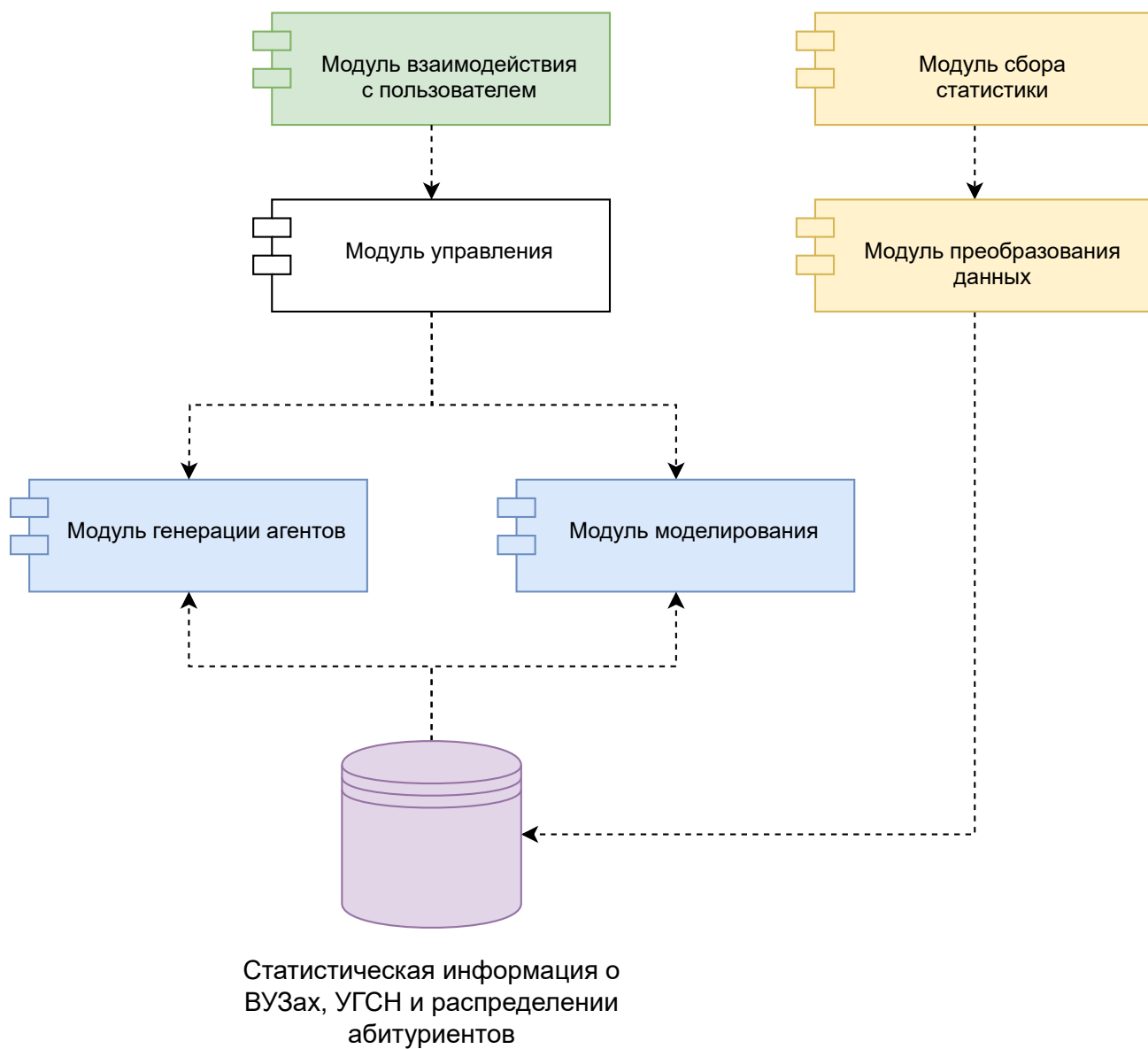


Рисунок 8 – Структура программного обеспечения

действия с пользователем.

Отвечающие за сбор и обработку информации модули напрямую взаимодействуют с энергонезависимым хранилищем информации для сохранения и дальнейшего переиспользования другими модулями.

3.4 Минимальные требования к вычислительной системе

Для того, чтобы воспользоваться программным продуктом, достаточно только ЭВМ, на которой будет установлена JRE (Java Runtime Environment) — минимальная реализация виртуальной машины, необходимая для исполнения Kotlin-приложений, без компилятора и других средств разработки.

Минимальные требования к вычислительной системе:

- операционная система Windows 7 или младше, Ubuntu 16.04, MacOS Monterey 12.0.1;
- оперативная память – 8 Гб.

При указанных выше параметрах программный продукт будет штатно функционировать.

3.5 Формат входных и выходных данных

Входными данными для генерации агентов являются:

- процентные соотношения численности категорий студентов с различными диапазонами полученных результатов ЕГЭ;
- диапазон результатов ЕГЭ, которые будут заданы агентам каждой категории;
- процентные соотношения признака смены региона для каждой категории;
- количество различных УГСН, которые рассматривает агент;
- минимальные баллы ЕГЭ по каждому предмету.

Входными данными для процесса моделирования являются:

- год проведения моделирования;
- количество абитуриентов, участвующих в моделировании;
- длительность этапа перераспределения заявлений с копиями аттеста-

тов;

- максимальное количество ВУЗов, в которое может подать заявления агент;
- длительность этапа поиска наилучшего места для оригинала аттестата.

Выходными данными является список строк формата: именование университета, полученный в ходе моделирования средний балл по ВУЗу, средний балл прошлого года, изменение среднего балла по сравнению с прошлым годом, минимальный балл, максимальный балл, количество поступивших абитуриентов. А также связанные с данным ВУЗом результаты каждого УГСН в формате: именование УГСН, полученный в ходе моделирования средний балл по УГСН, средний балл прошлого года, изменение среднего балла по сравнению с прошлым годом, минимальный балл, максимальный балл, количество поступивших абитуриентов.

3.6 Тестирование ПО

Адекватность реализованного метода проверялась путем проведения моделирования на статистических данных ВУЗов и УГСН за 2020 год и сравнения с результатами приемной кампании того же года. На рисунке 11 представлены результаты тестирования для выборки технических ВУЗов. Результаты работы с разбиением по медицинским и гуманитарным категориям представлены на рисунках ниже.

На рисунках видно боковую динамику, резких изменений среднего балла не наблюдается.

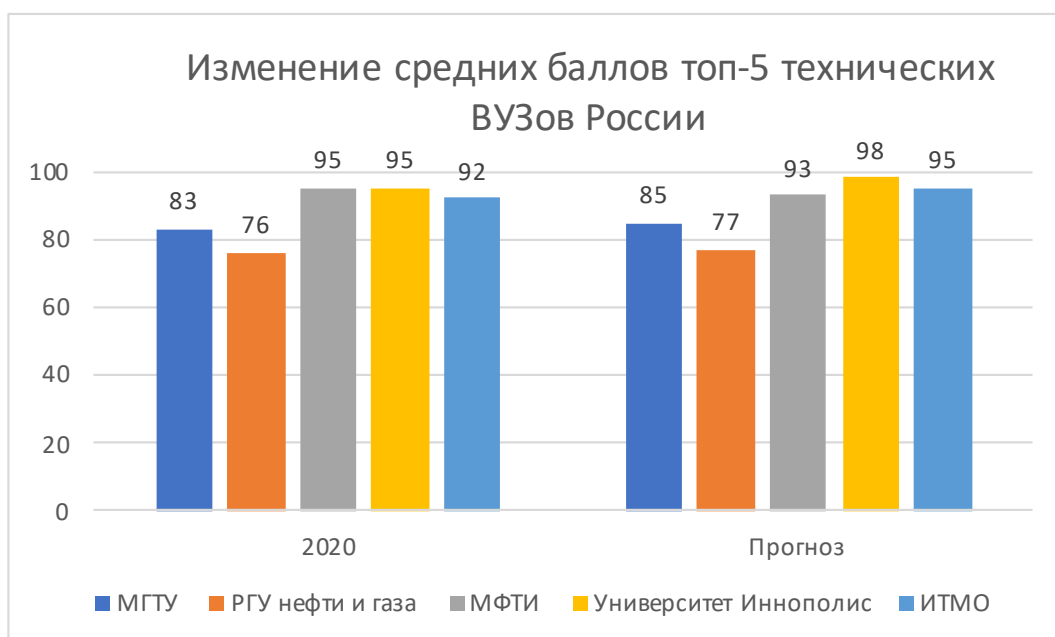


Рисунок 9 – Изменение средних баллов ТОП-5 технических ВУЗов



Рисунок 10 – Изменение средних баллов ТОП-5 медицинских ВУЗов

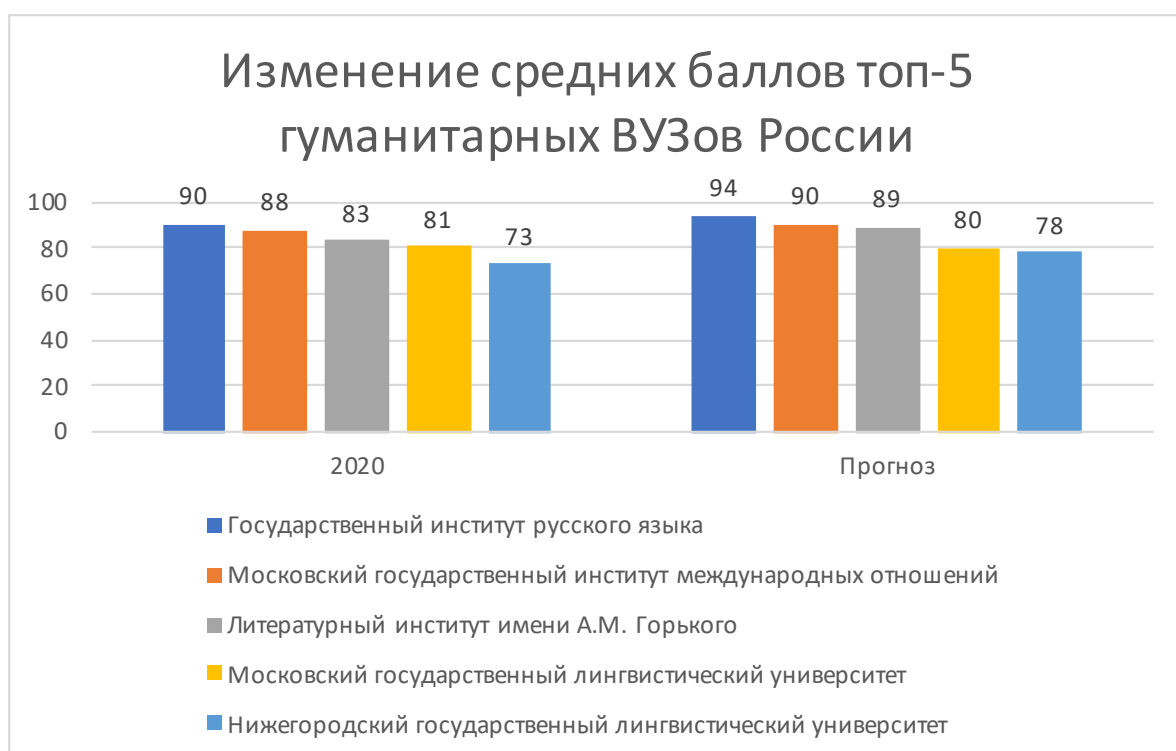


Рисунок 11 – Изменение средних баллов ТОП-5 гуманитарных ВУЗов

4 Исследовательская часть

4.1 Различное количество бюджетных мест в ВУЗе

Перед началом процесса моделирования, необходимо сгенерировать популяцию абитуриентов, согласно распределению по регионам. Всего было сгенерировано 700.000 тыс. агентов с индивидуальными характеристиками.

4.2 Различное максимальное число ВУЗов для подачи заявлений

4.3 Различные диапазоны результатов сдачи ЕГЭ

4.4 Различное соотношение готовых к переезду абитуриентов

ЗАКЛЮЧЕНИЕ

В рамках данной выпускной квалификационной работы был разработан и реализован метод прогнозирования итогов приёма в ВУЗы России на основе агентного моделирования.

В результате проделанной работы были выполнены все поставленные задачи:

- проанализированы правила организации приёма в ВУЗы России;
- получена информация и статистика о ВУЗах России;
- получена информацию по УГСН и связанных с ними предметам ЕГЭ;
- проанализирована полученная информация, выделены наиболее важные факторы для разработки модели поведения абитуриентов;
- проанализированы существующие методы и механизмы зачисления абитуриентов;
- разработана имитационную модель поведения абитуриентов;
- разработан метод прогнозирования итогов приёма в ВУЗы, с использованием разработанной модели поведения и статистических данных;
- программно реализован разработанный метод.

В качестве направлений дальнейшей работы над методом можно выделить следующие:

- использование многопоточности для оптимизации генерации агентов;
- разработка нечестных и рискованных стратегий поведения абитуриента;
- масштабирование разработанного метода до полноценного интернет сервиса.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Федеральный институт педагогических измерений – (дата обращения 2.12.2021). // [Электронный ресурс]. – Режим доступа: <https://fipi.ru/ege>
2. Приказ Министерства науки и высшего образования РФ от 13 августа 2021 г. N 753 – (дата обращения 2.12.2021). // [Электронный ресурс]. – Режим доступа: <https://www.garant.ru/products/ipo/prime/doc/402694290/>
3. Порядок приема на обучение по программам высшего образования – (дата обращения 2.12.2021). // [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/420216602>
4. Постановление Правительства РФ от 13 октября 2020 г. № 1681 “О целевом обучении по образовательным программам среднего профессионального и высшего образования” – (дата обращения 3.12.2021). // [Электронный ресурс]. – Режим доступа: <https://www.garant.ru/products/ipo/prime/doc/74665624/>
5. Приказ Министерства науки и высшего образования РФ об утверждении Порядка приёма на обучение по образовательным программам высшего образования от 14 сентября 2020 г. – (дата обращения 3.12.2021). // [Электронный ресурс]. – Режим доступа: <https://cdnimg.rg.ru/pril/article/195/62/75/25411494948547594417.pdf>
6. Приказ Минобрнауки России от 30.08.2019 N 666 ”Об утверждении перечня вступительных испытаний при приеме на обучение по образовательным программам высшего образования - программам бакалавриата и программам специалитета”. – (дата обращения 3.12.2021). // [Электронный ресурс]. – Режим доступа: <http://pushgu.ru/wp-content/uploads/2020/09/prikaz-minobrnauki-rossii-ot-30.08.2019-n-666.pdf>
7. Российский Статистический Ежегодник 2019 – (дата обраще-

- ния 10.12.2021). // [Электронный ресурс]. – Режим доступа: https://gks.ru/bgd/regl/b19_13/Main.htm
8. Мониторинг качества приёма в вузы – (дата обращения 1.11.2021). // [Электронный ресурс]. – Режим доступа: <https://ege.hse.ru>
 9. Сведения о сдаче единого государственного экзамена – (дата обращения 1.11.2021). // [Электронный ресурс]. – Режим доступа: <https://data.gov.ru/opendata/7710539135-ege>
 10. Информационно-аналитические материалы по результатам анализа показателей эффективности образовательных организаций высшего образования – (дата обращения 1.11.2021). // [Электронный ресурс]. – Режим доступа: <https://monitoring.miccedu.ru>
 11. Язык программирования Kotlin – (дата обращения 5.12.2021). // [Электронный ресурс]. – Режим доступа: <https://kotlinlang.org>
 12. Система управления базами данных PostgreSQL – (дата обращения 5.12.2021). // [Электронный ресурс]. – Режим доступа: <https://www.postgresql.org>
 13. Насадкин М.Ю., Агентное моделирование поведения абитуриентов при выборе вуза в России / Насадкин М.Ю., Питухин Е.А., Астафьева М.П. // Фундаментальные исследования. – 2015. – № 8-2. – С. 307-311.
 14. Абанкина И.В. Модель многоступенчатого выбора для прогнозирования спроса на высшее образование / И.В. Абанкина и др. // Университетское управление: практика и анализ. – 2014. – № 4–5. – С. 84-94.
 15. Кисельгоф С.А. Выбор вузов абитуриентами с квадратичной функцией полезности: препринт WP7/2011/01; Высшая школа экономики. – М.: Изд. дом Высшей школы экономики, 2011. – 44 с.

16. Питухин Е.А. Анализ межрегиональной мобильности выпускников школ при поступлении в высшие учебные заведения / Е.А. Питухин, А.А. Семенов // Университетское управление: практика и анализ. – 2011. – № 3. – С. 82-89.
17. Прахов И.А. Модель выбора вуза в условиях ЕГЭ и роль ожиданий абитуриентов: препринт WP10/2010/06; Гос. ун-т – Высшая школа экономики. – М.: Изд. дом Гос. ун-та – Высшей школы экономики, 2010. – 56 с.
18. Евдонин Г. А. Математическое моделирование и управление социально-экономическими и политическими процессами : учеб. пособие / Г. А. Евдонин. СПб. : Издательство СЗИУ РАНХиГС, 2012. 322 с.
19. Рыченков М. В. Исследование факторов, оказывающих влияние на выбор вуза абитуриентами, на различных этапах процесса поступления / Рыченков М. В., Рыченкова И. В., Киреев В. С. // Современные проблемы науки и образования. М., 2013. № 6.
20. Тимохович А. Н. Российский абитуриент вуза в условиях неопределенности // Вестник Университета (Государственный университет управления). М., 2012. № 1. С. 181–185.
21. Теплов Е. В., Филинова И. М. Факторы выбора абитуриентом образовательного учреждения // Среднее профессиональное образование. М., 2013. № 10. С. 43–44.
22. Берёза А.Н. Поддержка принятия решения при планировании набора абитуриентов в вузе на основе нечетких моделей/ Берёза А.Н., Ершова Е.А. // Известия ЮФУ. - Технические науки. - 2011. - №7. – С. 131-136.
23. Новосадова Н.О. Моделирование приемной кампании вузов с различным качеством и реализация модели в программной среде matlab // Фундаментальные и прикладные исследования в современной науке: сборник статей

II Международной научно-практической конференции (14 апреля 2018 г., г. Самара). - Самара: ЦНИК, 2018. – С. 3-6.

24. Gale D., Shapley L.S. College Admission and the Stability of Marriage. The American Mathematical Monthly. – Vol. 69. – № 1 (Jan., 1962). – P. 9-15.
25. Server push technology – (дата обращения 11.05.2022). // [Электронный ресурс]. – Режим доступа: <https://docs.nestjs.com/techniques/server-sent-events>
26. IntelliJ IDEA – (дата обращения 11.05.2022). // [Электронный ресурс]. – Режим доступа: <https://www.jetbrains.com/ru-ru/idea/>

ПРИЛОЖЕНИЕ А Учёт индивидуальных достижений

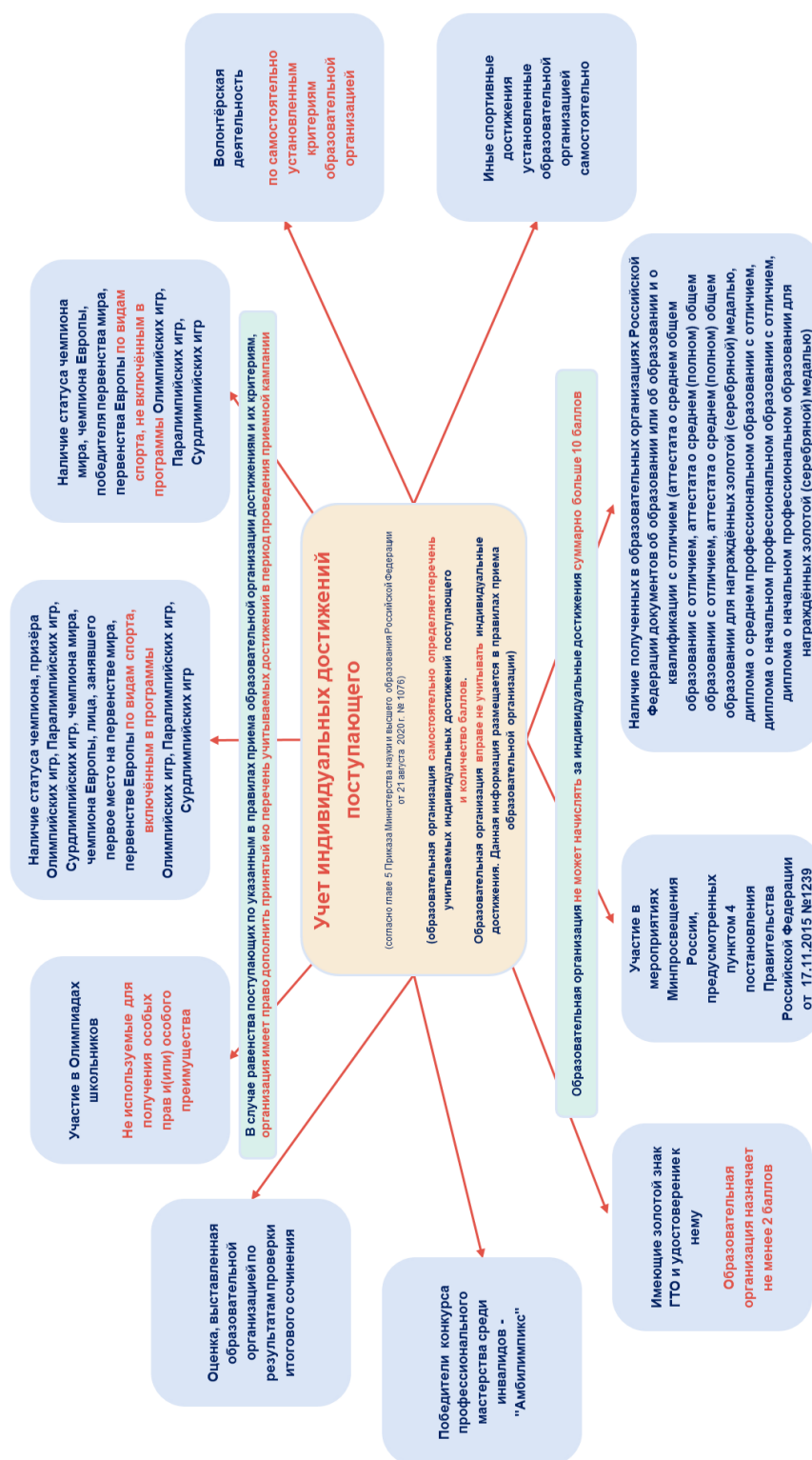


Рисунок 12 – Учет индивидуальных достижений