

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра информатики и программирования

**РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ
«РАСПИСАНИЕ ЗАНЯТИЙ ДЛЯ ВУЗОВ»
НА ОСНОВЕ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ
АВТОРЕФЕРАТ МАГИСТРСКОЙ РАБОТЫ**

студента 2 курса 273 группы

направления 02.03.03 Математическое обеспечение и администрирование
информационных систем

факультета компьютерных наук и информационных технологий

Шевченко Юлии Андреевны

Научный руководитель:

доцент

Кудрина Е. В.

подпись, дата

Зав. кафедрой:

к.ф-м.н , доцент

Огнева М. В.

подпись, дата

Саратов 2020

ВВЕДЕНИЕ

Актуальность темы. Проблема усовершенствования учебного процесса ВУЗа за счет планирования расписания занятий является известной, её решением занимались практики по созданию автоматизированных систем, но в настоящее время эта проблема остается открытой. Актуальность задачи определяется ростом требований к планированию учебного процесса, качеству обучения, размещению студентов в условиях дефицита аудиторного фонда, а также занятости преподавателей.

Достаточно часто расписание в вузах составляется вручную, что занимает очень много времени и сил, потому что необходимо учесть множество критериев, которые могут быть поставлены нечетко или даже противоречить друг друга. Поэтому очень важно автоматизировать этот процесс.

Задача составления расписания относится к задачам, сложность решения которых возрастает с увеличением числа возможных параметров, а также для неё характерно наличие большой объема различной по составу исходной информации и трудно анализируемых требований. Вышеперечисленные факторы делают решение данной задачи полным перебором очень длительным, поэтому для решения могут быть применены нейронные сети или генетические алгоритмы. Более предпочтителен выбор именно генетических алгоритмов для решения данной задачи, так как невозможно определить единственно верное решение для составления расписания, что необходимо для обучения нейронной сети.

Цель магистерской работы - разработка информационной системы для составления расписания занятий для высших учебных заведений на основе генетических алгоритмов.

Поставленная цель определила **следующие задачи:**

1. Рассмотреть теоретические основы генетических алгоритмов.
2. Проанализировать существующие способы составления расписания занятий с использованием генетических алгоритмов.

3. Изучить инструментальные средства и технологии, необходимые для разработки информационной системы.

4. Провести серию экспериментов по выбору метода мутации и скрещивания для эффективного составления расписания, проанализировать полученные данные, сделать выводы.

5. Провести серию экспериментов для оценки производительности приложения, проанализировать полученные данные, предложить и реализовать оптимизационные решения.

6. Результаты, полученные в ходе экспериментов, использовать при реализации информационной системы для составления расписания занятий для вузов.

Методологические основы разработки информационной системы «Расписание занятий для вузов» на основе генетических алгоритмов представлены в работах Т. Панченко [1], Д. Рутковской [2], Л. Гладкова [3], В. Аверченкова [4], В. Кобак [8], И. Астаховой [9] и Ю. Кабальнова [10].

Теоретическая значимость магистерской работы заключается в исследовании и анализе влияния различных методов скрещивания и мутации на эффективность составления расписания на основе генетических алгоритмов.

Практическая значимость магистерской работы заключается в разработке Desktop-приложения «ExtraTimeTable» на языке программирования Kotlin для составления расписания занятий на основе генетических алгоритмов.

Структура и объём работы. Магистерская работа состоит из введения, двух разделов, заключения, списка использованных источников и одного приложения. Общий объём работы – 52 страниц, из них 48 страниц – основное содержание, включая 16 рисунков и 6 таблиц, цифровой носитель в качестве приложения, список использованных источников информации – 22 наименования.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Теоретическая часть» посвящен теоретическим основам генетических алгоритмов, обзору существующих решений составления расписания на основе генетических алгоритмов, а также обзор инструментальных средств для разработки данной информационной системы.

«Введение в теорию генетических алгоритмов» – этот подраздел включает в себя краткий ввод в теоретические основы генетических алгоритмов. Рассмотрены такие понятия, как: генетический алгоритм, популяция, особь, хромосома, ген, фенотип, аллель, локус и другие. В данном подразделе рассмотрен цикл генетического алгоритма, который включает в себя следующие шаги:

1. Создание стартовой популяции.
2. Отбор.
3. Скрещивание.
4. Мутация.

Также данный подраздел включает в себя определения различных видов скрещивания, таких как:

- Панмиксия
- Инбридинг
- Аутбридинг
- Ассоциативное скрещивание
- Положительно ассоциативное скрещивание
- Отрицательно ассоциативное скрещивание
- Селективное скрещивание

Рассмотрены различные виды мутации в генетических алгоритмах, такие как:

- Точечная мутация
- Двухточечная мутация

- Инверсия
- Сильная мутация

«Обзор существующих решений составления расписания на основе генетических алгоритмов» – в этом подразделе рассматриваются существующие подходы, которые направлены на решение задачи автоматизированного составления расписания. Рассмотрены работы И. Астаховой и А. Фирас «Составление расписания учебных занятий на основе генетического алгоритма» и Ю. Кабальнова, Л. Шехтман, Г. Низамовой, Н. Земченковой «Композиционный генетический алгоритм составления расписаний учебных занятий».

В данном подразделе выделены основные преимущества и недостатки данных решений. Представлены критерии, которых не хватает в существующих решениях, но они могли бы успешно применяться в реальных условиях, например: расположение корпусов, вместимость аудиторий, специфические особенности аудиторий (количество компьютеров, наличие проектора, меловой доски и т.д.).

«Инструментальные средства для разработки информационной системы» – данный подраздел включает в себя описания выбранных технологий для разработки информационной системы, а именно – язык программирования Kotlin [8] и фреймворк TornadoFX [9], также приведены доводы, почему для разработки данной системы были выбраны именно такие технологии.

Дополнительно в данном подразделе рассмотрены этапы парсинга данных, такие как:

- Поиск данных
- Извлечение информации
- Сохранение данных

Представлен краткий обзор библиотеки для парсинга Excel-файла для JVM-приложения – Apache POI, которая была использована при разработке приложения.

Второй раздел «Практическая часть» посвящен реализации приложения «ExtraTimeTable» для составления расписаний для ВУЗов на основе генетических алгоритмов. Представлена архитектура разработанной информационной системы, рассмотрена реализация генетического алгоритма в данном приложении. Проведены эксперименты выбору по наиболее эффективным методам скрещивания и мутации для эффективного составления расписания, а также эксперименты, направленные на улучшение производительности приложения в целом. Представлены и внедрены оптимизационные решения, которые значительно помогли ускорить генерацию расписания в разработанном приложении.

«Архитектура информационной системы» – данный подраздел включает в себя полное дерево проекта, а также описание основных модулей. Приложение построено на принципах трехслойной архитектуры и написано в реактивном стиле, то есть ориентировано на потоки данных и распространение изменений. Это позволяет приложению выполнять операции асинхронно.

Также в данном подразделе содержатся функциональные требования разрабатываемого приложения, а именно:

- Приложение должно работать с тестовыми данными из базы данных (далее БД), а также реальными данными из заявки на 2019-2020 учебный год на обеспечение учебных курсов факультета КНиИТ в формате Excel.
- Приложение должно уметь составлять расписание на основе имеющихся данных.
- Приложение должно позволять просматривать составленное расписание в удобном виде.
- В приложении должна быть возможность вручную устанавливать такие параметры как оптимальное и максимальное количество пар в день, выбор полугодия, количество особей в популяции, количество повторений эксперимента для замеров времени,

количество циклов генетического алгоритма, а также вероятность мутации.

Данный подраздел включает в себя критерии для оценивания эффективности расписания. Выделенные критерии представлены ниже:

- Ограничение на количество проводимых занятий в день
- Соответствие количества занятий в день оптимальному количеству
- Минимизация окон у студентов
- Вместимость аудиторий
- Соответствие аудитории дисциплине

«Реализация информационной системы» – данный подраздел демонстрирует работы Desktop-приложения «ExtraTimeTable» для составления расписания для ВУЗов на основе генетических алгоритмов. Представлены скриншоты работы приложения. Также представлен фрагмент файла заявки с реальными данными для составления расписания факультета КНИИТ за 2019-2020 год. Рассмотрен механизм парсинга данных из Excel-файла с помощью библиотеки Apache POI. Рассмотрена реализация генетического алгоритма для составления расписания в разработанном приложении. Представлены фрагменты кода, которые описывают хромосому, ген, особь и другие важные для понимания реализации генетического алгоритма классы. Приведены фрагменты кода, которые описывают вычисление фитнес-функции для особи, генерацию стартовой популяции, процесс мутации и т.д.

«Выбор методов скрещивания и мутации для эффективного составления расписания» – в данном подразделе проводятся эксперименты по выбору наиболее эффективного для данной реализации генетического алгоритма методов скрещивания и вероятности мутации. Были рассмотрены следующие виды скрещивания – панмиксия, позитивно ассоциативное скрещивания, а также негативно ассоциативное скрещивание. В ходе экспериментов было выявлено, что метод скрещивания **панмиксия** является

более эффективным для данной реализации. Также были проведены эксперименты по выбору вероятности мутации. Были рассмотрены следующие значения вероятности мутации – 0.01, 0.15, 0.3, 0.4, 0.5, 0.7. Эмпирическим путем было выявлено, что значение равное 0.3 оказалось самым эффективным для данной реализации генетического алгоритма.

«Оптимизация производительности информационной системы» – данный подраздел включает в себя основные принципы оптимизации программного кода [10], а также анализ разработанного приложения. Были выявлены «узкие» места, которые было принято оптимизировать.

Предложенные оптимизационные решения успешно внедрены в приложение, с помощью них повысилась производительность составления расписания. Проанализировано время выполнения составления расписания на различных входных данных до оптимизации и после, сделаны выводы, что оптимизация помогла ускорить время выполнения генерации стартовой популяции примерно на 45.6%, а время выполнения генетического алгоритма – в среднем на 13,3%.

ЗАКЛЮЧЕНИЕ

В рамках выпускной квалификационной работы были выполнены все поставленные задачи, а именно: изучены теоретические основы генетических алгоритмов, проанализированы существующие решения для составления расписаний занятий с использованием генетических алгоритмов, изучены инструментальные средства для разработки информационной системы, проведены и проанализированы эксперименты по выбору метода мутации и скрещивания для эффективного составления расписания. Это позволило разработать приложение «ExtraTimeTable» на языке программирования Kotlin для составления расписания учебных занятий для ВУЗов на основе генетических алгоритмов. Также были проведены эксперименты производительности приложения, предложены и реализованы оптимизационные решения для оптимизации работы приложения.

В приложении «ExtraTimeTable» генетический алгоритм реализован универсально, чтобы он был масштабируемый, а также мог быть использован в другом приложении или в другом контексте.

В дальнейшем разработку можно использовать на практике для генерации расписания учебных занятий в СГУ. Также в будущем можно развивать информационную систему, добавляя в неё новые параметры для оценки расписания.

Основные источники информации:

1. Панченко Т.В. Генетические алгоритмы: учебно-методическое пособие – Астрахань: Издательский дом «Астраханский университет», 2007 – 87 с.
2. Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Рутковская Д., Пилиньский М., Рутковский Л. – М.: Горячая линия – Телеком, 2006 – 383 с.
3. Гладков. Л.А. Генетические алгоритмы: учебное пособие / Гладков Л.А., Курейчик В.В., Курейчик В.М. – М.: ФИЗМАТЛИТ, 2006 – 320 с.

4. Аверченков В.И. Эволюционное моделирование и его применение / Аверченков В.И., Казаков П.В. – М.: ФЛИНТА, 2016. – 200 с.
5. Кобак В.Г. Исследование эффективности генетической модели при использовании различных видов мутации / Кобак В.Г., Чижов А.Ю., Плешаков Д.В. – М.: Известия ВУЗов, 2012 – С. 16-19 [Электронный ресурс] – Режим доступа: <https://cyberleninka.ru/article/n/issledovanie-effektivnosti-geneticheskoy-modeli-pri-ispolzovanii-razlichnyh-vidov-mutatsii> -
Дата обращения: 14.05.2020.
6. Астахова И.Ф. Составление расписания учебных занятий на основе генетического алгоритма / Астахова И.Ф., Фирас А.М. // Вестник ВГУ, Серия: Системный анализ и информационные технологии. – 2013. – С.93-99 [Электронный ресурс] – Режим доступа: <http://www.vestnik.vsu.ru/pdf/analiz/2013/02/2013-02-17.pdf> - Дата обращения: 14.05.2020.
7. Кабальнов Ю.С. Композиционный генетический алгоритм составления расписаний учебных занятий / Кабальнов Ю.С., Шехтман Л.И., Низамова Г.Ф., Земченкова Н.А // Вестник УГАТУ. Информационные технологии. – 2006. – С. 99-107 [Электронный ресурс] – Режим доступа: <https://cyberleninka.ru/article/n/kompozitsionnyy-geneticheskiiy-algoritm-sostavleniya-raspisaniya-uchebnyh-zanyatiy> – Дата обращения: 14.05.2020.
8. Жемеров Д. Kotlin в действии / Жемеров Д., Исакова С. – М.: ДМК Пресс, 2018 – 402 с.
9. TornadoFX [Электронный ресурс] – Режим доступа: <https://tornadofx.io> – Дата обращения: 14.05.2020.
10. Макконнелл С. Совершенный код. / Макконнелл С. — М.: Русская редакция, 2005. — 896 с.