

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

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

Кафедра дискретной математики и информационных технологий

**СПЕКТРАЛЬНЫЙ АНАЛИЗ ГРАФОВОЙ МОДЕЛИ
ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ ФАКУЛЬТЕТА КНИИТ
СГУ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 421 группы
направления 09.03.01 — Информатика и вычислительная техника
факультета КНиИТ
Вишнякова Никиты Алексеевича

Научный руководитель

преподаватель, к. ф.-м. н.

В. В. Кирьяшкин

Заведующий кафедрой

доцент, к. ф.-м. н.

Л. Б. Тяпаев

Саратов 2025

ВВЕДЕНИЕ

Данная бакалаврская работа посвящена спектральному анализу графовой модели образовательных программ факультета компьютерных наук и информационных технологий (КНиИТ). Актуальность темы обусловлена необходимостью применения объективных методов для анализа и оптимизации все более усложняющихся учебных планов в сфере IT-образования. Понимание структуры междисциплинарных связей и выявление ключевых особенностей программ является важной задачей для повышения качества подготовки специалистов. Применение методов спектральной теории графов позволяет получить новые количественные и качественные данные о структуре образовательных программ факультета КНиИТ.

Целью работы является исследование структурных свойств графовой модели образовательных программ факультета КНиИТ с применением методов спектральной теории графов. Для достижения поставленной цели были сформулированы следующие задачи: изучить теоретические основы спектрального анализа графов; разработать программные средства для сбора и систематизации данных учебных планов в единой базе данных; построить графовую модель зависимостей между дисциплинами; реализовать вычисление спектров матриц смежности и лапласианов для построенного графа; провести анализ и интерпретацию полученных спектральных характеристик; сформулировать выводы о структурной организации образовательных программ факультета КНиИТ.

Материалами исследования послужили официальные учебные планы различных направлений подготовки и специальностей факультета КНиИТ, представленные в виде PDF-документов и преобразованные в табличный формат для последующей обработки. На основе этих данных была сформирована единая база данных, содержащая информацию о 357 уникальных дисциплинах и их распределении по семестрам. Для анализа пререквизитных зависимостей также использовались Рабочие программы дисциплин. Программная реализация выполнена на языке Python с использованием библиотек pandas, NetworkX, NumPy, Pyvis и СУБД SQLite.

Данная бакалаврская работа состоит из введения, трех глав, заключения, списка использованных источников и приложений. В первой главе «Теоретические основы и методы графового анализа образовательных программ»

рассматриваются теоретические аспекты исследования. Вторая глава «Разработка и реализация графовой модели учебных планов» посвящена методологии сбора данных, созданию базы данных и построению графовой модели. В третьей главе «Анализ графовой модели учебных планов методами спектральной теории» представлены результаты проведенного спектрального и структурного анализа.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел посвящен теоретическому фундаменту исследования. В нем представлен обзор существующих подходов к использованию графовых моделей в сфере образования. Особое внимание уделено моделированию пререквизитных зависимостей между учебными дисциплинами, так как именно этот аспект является центральным для построения графа в данной работе. Анализ таких графов позволяет визуализировать структуру учебных планов, выявлять критические пути обучения и идентифицировать фундаментальные дисциплины.

Далее излагаются основные понятия и определения из общей теории графов, необходимые для понимания последующих разделов. Вводятся такие термины, как граф, вершина, ребро, степень вершины, путь, цикл, связность и связные компоненты. Различаются ориентированные и неориентированные графы, при этом отмечается, что исходная модель пререквизитов является ориентированной, а для спектрального анализа используется ее неориентированный аналог.

Ключевая часть первого раздела посвящена спектральной теории графов. Подробно рассматриваются три основные матрицы, ассоциированные с неориентированным графом $G = (V, E)$ ($n = |V|$ вершин, $m = |E|$ ребер):

- Матрица смежности A , элементы которой $A(u, v)$ равны 1, если вершины u и v смежны, и 0 в противном случае. Обсуждается ее симметричность и свойства ее спектра $\text{Spec}(A) = \{\mu_0, \dots, \mu_{n-1}\}$, включая связь наибольшего собственного значения со степенями вершин, соотношение $\sum \mu_i^2 = 2m$ и симметрию спектра для двудольных графов.
- Комбинаторный лапласиан $L = D - A$, где D – диагональная матрица степеней вершин $\deg(v)$. Отмечается, что L является симметричной и положительно полуопределенной. Рассматриваются свойства его спектра $\text{Spec}(L) = \{0 = \nu_0 \leq \nu_1 \leq \dots \leq \nu_{n-1}\}$: кратность $\nu_0 = 0$ равна числу связных компонент, а ν_1 (или первое ненулевое ν_k) определяет алгебраическую связность (число Фидлера). Также приводится свойство $\sum \nu_i = 2m$.
- Нормализованный лапласиан $\mathcal{L} = D^{-1/2}LD^{-1/2} = I - D^{-1/2}AD^{-1/2}$. Его собственные значения $0 = \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{n-1} \leq 2$. Аналогично комбинаторному лапласиану, кратность $\lambda_0 = 0$ равна числу связных

компонент, а λ_1 (или λ_k) является мерой связности. Важным свойством является то, что $\lambda_{n-1} = 2$ указывает на двудольность одной из компонент графа, а сумма $\sum \lambda_i$ равна числу неизолированных вершин.

Теоретические положения, изложенные в данном разделе, служат основой для выбора методов анализа и интерпретации результатов, полученных в практической части работы.

Второй раздел описывает методологию и практические шаги по созданию графовой модели образовательных программ факультета КНиИТ. Вначале приводится обзор использованных технологий и программных инструментов. В качестве основного языка программирования выбран Python (версия 3.x) благодаря его гибкости и наличию специализированных библиотек. Для работы с табличными данными использовалась библиотека `pandas`, для взаимодействия с базой данных – встроенный модуль `sqlite3`, для построения и анализа графов – `NetworkX`, для численных расчетов – `NumPy`, а для интерактивной визуализации – `Pyvis`. В качестве среды разработки применялась `PyCharm`.

Далее следует описание исходных данных. Основой послужили официальные учебные планы различных направлений подготовки факультета КНиИТ, изначально представленные в формате PDF. Описан процесс их преобразования в структурированный табличный формат (эквивалентный Excel) для последующей автоматизированной обработки. Отмечаются сложности, связанные с разнородностью форматов исходных документов и необходимостью унификации данных.

Центральным элементом подготовки данных стало формирование единой реляционной базы данных `all_direction.db` на СУБД `SQLite`. Подробно описывается структура этой базы, включающая таблицы для хранения информации о направлениях (`direction`), курсах (`course`), семестрах (`semester`), уникальных дисциплинах (`subject`), их характеристиках в конкретных семестрах (`subject_semester`), а также о выявленных пререквизитных зависимостях (`subject_subject`). Представлена схема базы данных, иллюстрирующая таблицы и связи между ними.

Затем детально излагаются алгоритмы обработки данных. Описывается процесс извлечения информации из подготовленных Excel-файлов учебных планов, включая нормализацию наименований дисциплин и форм кон-

троля, и загрузку этих данных в промежуточные индивидуальные базы данных SQLite для каждого направления. После этого подробно рассматривается этап консолидации данных из промежуточных баз в единую базу `all_direction.db`, с решением задач по обеспечению уникальности сущностей, генерации глобальных идентификаторов и обновлению агрегированной информации о дисциплинах. Ключевой частью данного раздела является описание методики выявления пререквизитных зависимостей между дисциплинами. Этот процесс состоял из двух этапов: ручного анализа Рабочих программ дисциплин (РПД) и последовательности изучения курсов с формированием Excel-таблицы зависимостей, и последующей автоматизированной загрузки этих данных в таблицу `subject_subject` с помощью Python-скрипта. Приведены примеры логики скрипта, включая нормализацию наименований дисциплин и их идентификацию в базе данных для корректного установления связей. Всего было выявлено и загружено 1251 пререквизитная связь.

Завершается раздел описанием непосредственного построения графовой модели. На основе данных из таблиц `subject` и `subject_subject` с использованием библиотеки NetworkX был создан ориентированный граф G_{orig} , где вершины – это дисциплины, а ребра – пререквизиты. Каждой вершине и ребру были присвоены атрибуты, используемые для анализа и визуализации (например, курс дисциплины для атрибута `level` вершин). Для целей спектрального анализа граф G_{orig} преобразовывался в неориентированный G_{undir} . Описана процедура интерактивной визуализации графа с помощью библиотеки Pyvis, обеспечивающей иерархическое отображение дисциплин по курсам. Приведен пример общего вида полученной визуализации.

Третий раздел посвящен результатам спектрального и структурного анализа построенной графовой модели G_{undir} . Вначале излагается алгоритм и реализация спектрального анализа, реализованного в виде функции на Python, которая вычисляет спектры трех основных матриц графа (смежности, комбинаторного и нормализованного лапласианов) с использованием функций библиотеки NetworkX, а также проводит анализ ключевых спектральных характеристик и структурный анализ компонент. Приведены фрагменты кода, иллюстрирующие вызов этих функций.

Далее представляется общая характеристика неориентированного графа G_{undir} , состоящего из $n = 357$ вершин (дисциплин) и $m = 1091$ ребер.

Отмечается, что количество ребер в неориентированном графе меньше исходного числа выявленных зависимостей (1251) из-за наличия 160 двусторонних зависимостей в первоначальной модели. Рассчитаны средняя степень вершин $\langle k \rangle \approx 6.11$ и низкая плотность графа $\rho \approx 0.017$, свидетельствующая о его разреженности.

Последовательно излагаются результаты анализа каждого из трех спектров, сопровождаемые гистограммами распределения их собственных значений:

- Спектр матрицы смежности ($\text{Spec}(A)$). Минимальное собственное значение $\mu_0 \approx -7.24$, максимальное $\mu_{n-1} \approx 14.63$. Отмечено отсутствие симметрии спектра относительно нуля, что указывает на то, что граф в целом не является двудольным. Проверено выполнение теоретического соотношения $\sum \mu_i^2 = 2m = 2182$.
- Спектр комбинаторного лапласиана ($\text{Spec}(L)$). Установлено наличие 52 собственных значений ν_i , близких к нулю, что соответствует 52 связным компонентам графа. Алгебраическая связность (первое ненулевое собственное значение после нулевых) $\nu_{52} \approx 0.047$, что указывает на слабую связность как минимум одной из крупных компонент. Максимальное собственное значение $\nu_{n-1} \approx 65.08$. Подтверждено свойство $\sum \nu_i = 2m$.
- Спектр нормализованного лапласиана ($\text{Spec}(\mathcal{L})$). Также подтверждено наличие 52 связных компонент. Алгебраическая связность $\lambda_{52} \approx 0.011$ также имеет малую величину. Максимальное собственное значение $\lambda_{n-1} \approx 2.0$ свидетельствует о наличии двудольных подструктур в графе. Сумма $\sum \lambda_i \approx 307$ совпала с числом неизолированных вершин в графе.

Результаты спектрального анализа были сопоставлены с прямым структурным анализом связных компонент графа, который также выявил 52 компоненты. Установлено, что 50 из них являются изолированными дисциплинами (приведены примеры). Выделены две более крупные компоненты: основная, объединяющая 305 дисциплин, которая не является двудольной, и малая компонента из 2 дисциплин («Русский язык и культура речи» и «Иностранный язык»), являющаяся двудольной.

В завершение раздела обобщаются основные результаты спектрального и структурного анализа, подчеркиваются выявленные особенности графа

учебных планов: его разреженность, высокая фрагментация, наличие большого ядра дисциплин со сложной структурой и слабой внутренней связностью, а также присутствие малых изолированных групп и отдельных дисциплин.

ЗАКЛЮЧЕНИЕ

В ходе выполнения выпускной квалификационной работы была успешно решена поставленная цель – исследование структурных свойств графовой модели образовательных программ факультета КНиИТ с применением методов спектральной теории графов.

На основе официальных учебных планов факультета была разработана методология сбора и систематизации данных, результатом которой стало создание единой реляционной базы данных `all_direction.db`. С использованием языка Python и специализированных библиотек (`pandas`, `NetworkX`, `NumPy`, `Pyvis`, `sqlite3`) были разработаны и реализованы программные средства для построения графовой модели пререквизитных зависимостей между дисциплинами и для последующего вычисления спектров ассоциированных с графом матриц.

Построенная графовая модель, включающая 357 уникальных дисциплин и 1091 пререквизитную связь (в неориентированном представлении), была подвергнута комплексному анализу. Исследование спектров матрицы смежности, комбинаторного лапласиана и нормализованного лапласиана, а также прямой структурный анализ связанных компонент, позволили сделать следующие ключевые выводы о структуре учебных планов факультета КНиИТ:

- Граф учебных планов характеризуется низкой плотностью связей ($\rho \approx 0.017$), что свидетельствует о разреженной структуре пререквизитных зависимостей.
- Выявлена значительная степень фрагментации образовательного пространства: граф состоит из 52 независимых связанных компонент. Данный результат был согласованно получен как спектральными методами (на основе кратности нулевых собственных значений лапласианов), так и путем прямого структурного анализа.
- Большинство компонент (50 из 52) представляют собой изолированные дисциплины. Их наличие может быть обусловлено спецификой курсов (факультативы, дисциплины по выбору, общеобразовательные блоки) или особенностями формализации пререквизитов в учебных документах.
- Обнаружены две основные крупные компоненты: первая, объединяю-

щая 305 дисциплин, образует ядро учебных программ и имеет сложную недвудольную структуру (подтверждено асимметрией спектра матрицы смежности, $\mu_0 \approx -7.24$, $\mu_{n-1} \approx 14.63$). Вторая крупная компонента состоит всего из двух дисциплин и является двудольной, что согласуется с максимальным собственным значением нормализованного лапласиана $\lambda_{n-1} \approx 2.0$.

- Низкие значения алгебраической связности, полученные из спектров лапласианов ($\nu_{52} \approx 0.047$ для комбинаторного и $\lambda_{52} \approx 0.011$ для нормализованного), указывают на относительно слабую внутреннюю связность основных блоков дисциплин и наличие потенциальных "узких мест" в образовательных траекториях.

Все теоретические соотношения для спектров матриц были подтверждены вычислениями, что свидетельствует о корректности использованных методов и надежности полученных данных.

Таким образом, применение методов спектральной теории графов позволило провести глубокий и объективный анализ структуры образовательных программ факультета КНИИТ. Выявленные характеристики, такие как фрагментарность, наличие крупного ядра и специфические особенности связности, предоставляют ценную информацию для возможной дальнейшей оптимизации учебных планов и повышения эффективности образовательного процесса. Разработанный в ходе работы программный инструментарий и предложенная методика могут быть адаптированы для анализа учебных планов других факультетов или образовательных учреждений.

Основные источники информации:

- 1 Chung, F. R. K. Spectral Graph Theory / F. R. K. Chung. – Providence, RI: American Mathematical Society, 1997. - 207 p. - (CBMS Regional Conference Series in Mathematics; Vol. 92).
- 2 Теория графов. Термины и определения в картинках [Электронный ресурс] URL: <https://habr.com/ru/companies/otus/articles/568026/> (дата обращения: 01.06.2025)
- 3 Граф (математика) [Электронный ресурс] URL: [https://ru.wikipedia.org/wiki/Граф_\(математика\)](https://ru.wikipedia.org/wiki/Граф_(математика)) (дата обращения: 01.06.2025)
- 4 Типы графов [Электронный ресурс] URL: <https://education.yandex.ru/handbook/math/article/tipi-grafov> (дата обращения: 01.06.2025)

- 5 Теория Графов. Часть 1 Введение и классификация графов [Электронный ресурс] URL: <https://habr.com/ru/articles/564594/> (дата обращения: 01.06.2025)
- 6 Спектральная теория графов [Электронный ресурс] URL: https://ru.wikipedia.org/wiki/Спектральная_теория_графов (дата обращения: 01.06.2025)
- 7 Ma, Y. Deep Learning on Graphs / Y. Ma, J. Tang. – Cambridge : Cambridge University Press, 2021. – XVI, 274 p.
- 8 An introduction to the normalized Laplacian [Электронный ресурс] URL: https://www.d.umn.edu/dfroncek/MCCCC_2018/abstracts/Butler_slides.pdf (дата обращения: 01.06.2025)
- 9 Laplacian matrix [Электронный ресурс] URL: https://en.wikipedia.org/wiki/Laplacian_matrix (дата обращения: 01.06.2025)