

Министерство образования и науки Российской Федерации  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра информатики и программирования

**ИСПОЛЬЗОВАНИЕ ГРАФОВЫХ БАЗ ДАННЫХ В СОЦИАЛЬНЫХ  
СЕТЯХ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 441 группы

направления 02.03.03 Математическое обеспечение и администрирование  
информационных систем

факультета компьютерных наук и информационных технологий

Антонова Сергея Александровича

Научный руководитель:

Старший преподаватель

М.С. Портенко

---

(подпись, дата)

Зав. кафедрой:

К.ф.-м.н.

М.В. Огнева

---

подпись, дата

Саратов 2017

## **ВВЕДЕНИЕ**

**Актуальность темы.** С 1980-х годов реляционные системы управления базами данных стали занимать доминирующее положение среди средств хранения данных. Несмотря на их многофункциональность и универсальность, существуют классы задач, которые требуют специфического представления и обработки объектов предметной области. В последнее время все чаще используются хранилища, реализующие альтернативные модели данных. Основная их цель – расширить возможности баз данных в тех областях, где реляционная модель недостаточно гибка. Одним из наиболее актуальных видов нереляционных хранилищ являются графовые базы данных, в связи с тем, что графовая модель представления данных оказалась естественной и востребованной в современном мире различных социальных связей.

**Цель бакалаврской работы** – сравнение эффективности использования реляционной и графовой баз данных для обработки информации о пользователях социальной сети.

Поставленная цель определила **следующие задачи**:

1. Проанализировать структуру данных пользователей социальных сетей.
2. Изучить основные принципы работы с графовыми базами данных.
3. Реализовать хранение, добавление, изменение, удаление информации и выполнение запросов в реляционной и графовой базах данных.
4. Сравнить графовый и реляционный подход по основным действиям работы с базами данных.

**Практическая значимость бакалаврской работы.** В ходе выполнения работы были решены все поставленные задачи, что позволило достигнуть заявленной цели – сравнить эффективность использования реляционной и графовой баз данных для обработки информации о пользователях социальной сети. Результаты исследования позволяют получить представление о достоинствах и недостатках современных графовых баз данных в вопросе хранения и обработки структуры систем, модель которых представляет собой

множество узлов с разнородными данными и большим количеством различных связей между ними.

**Структура и объём работы.** Бакалаврская работа состоит из введения, четырёх разделов, заключения, списка использованных источников и приложения. Общий объём работы – 61 страница, из них 57 страниц – основное содержание, включая 18 рисунков и 8 таблиц, цифровой носитель в качестве приложения, содержащий код программной реализации веб-приложения и генератора данных, список использованных источников информации – 22 наименования.

## **КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ**

**Первый раздел «КЛАССИФИКАЦИЯ БАЗ ДАННЫХ ПО МОДЕЛИ ДАННЫХ»** посвящен описанию моделей данных, используемых различными системами управления базами данных. Раздел содержит несколько подразделов.

Подраздел «ИЕРАРХИЧЕСКАЯ МОДЕЛЬ» содержит раскрытие основных понятий и принципов иерархической модели данных, включая определения элемента, агрегата, группы, группового отношения. [1]

В подразделе «РЕЛЯЦИОННАЯ МОДЕЛЬ» описываются теоретические основы реляционных баз данных, раскрываются понятия отношения, атрибута, кортежа, первичного и внешнего ключей. [2-3]

Подраздел «ОБЪЕКТНАЯ МОДЕЛЬ» содержит основные понятия объектной модели, используемой в объектно-ориентированных базах данных. [4]

Подраздел «ДОКУМЕНТНАЯ МОДЕЛЬ» описывает основные принципы работы и назначение документно-ориентированных баз данных.

Подраздел «ГРАФОВАЯ МОДЕЛЬ» включает в себя описание графовой модели данных и её основных составляющих: узлов и отношений, описываются принципы перехода от реляционной модели к графовой. [5-6]

**Второй раздел «ГРАФОВЫЕ БАЗЫ ДАННЫХ»** содержит информацию об областях применения графовых баз данных, обзор существующих графовых систем управления базами данных, подробное описание графовой СУБД Neo4j. Раздел содержит несколько подразделов.

Подраздел «ОБЛАСТИ ИСПОЛЬЗОВАНИЯ» содержит информацию о том, как графовые базы данных могут быть использованы в различных областях, таких как обнаружение мошенничества, геоинформационные системы и логистика, авторизация и контроль доступа, компьютерные сети, социальные сети, системы рекомендаций. [7]

Подраздел «ОБЗОР СУЩЕСТВУЮЩИХ СУБД» включает в себя обзор существующих графовых СУБД (Neo4j, OrientDB, ArangoDB, AllegroGraph), а также обоснование выбора системы Neo4j для использования в работе. [8-12]

Подраздел «NEO4J. ЯЗЫК ЗАПРОСОВ И СРЕДСТВА ВИЗУАЛИЗАЦИИ» содержит описание основных конструкций языка запросов Cypher графовой СУБД Neo4j, а также содержит информацию о встроенных в Neo4j средствах визуализации результатов запросов. [13-15]

**Третий раздел «ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА ВЕБ-ПРИЛОЖЕНИЯ»** посвящен разработке веб-приложения, реализующего наиболее часто используемые элементы и функции социальных сетей. Раздел состоит из нескольких подразделов.

Подраздел «АРХИТЕКТУРА И ФУНКЦИОНАЛЬНОСТЬ ПРИЛОЖЕНИЯ» включает в себя описание трехслойной архитектуры приложения, структуры данных пользователей социальной сети и функциональности веб-приложения. [16]

Подраздел «ASP.NET MVC. ОБЩИЕ СВЕДЕНИЯ» содержит общую информацию о фреймворке для создания веб-приложений ASP.NET MVC, реализующем шаблон Model-View-Controller (Модель-Представление-Контроллер). [17]

Подраздел «АРХИТЕКТУРА MVC» включает в себя описание основных принципов архитектуры Model-View-Controller, поясняет смысл и предназначение компонентов данной архитектуры: моделей, представлений, контроллеров. [17]

Подраздел «РАСШИРЯЕМОСТЬ» содержит информацию о том, как разработчик может расширять функциональность фреймворка ASP.NET MVC, благодаря тому, что фреймворк построен как ряд независимых компонентов. [17]

Подраздел «МАРШРУТИЗАЦИЯ» описывает систему маршрутизации в фреймворке ASP.NET MVC, которая позволяет разработчику использовать в своем приложении URL-адреса, не сопоставляемые с определенными файлами на веб-узле. [18]

Подраздел «ASP.NET WEB API» содержит описание одной из частей платформы ASP.NET, позволяющей создавать веб-службы, которые

предоставляют API для широкого диапазона HTTP-клиентов, включая браузеры и мобильные устройства. [19]

Подраздел «ДВИЖОК ПРЕДСТАВЛЕНИЙ И RAZOR» описывает возможности фреймворка ASP.NET MVC для разметки веб-страниц. [20]

**Четвёртый раздел «СРАВНЕНИЕ ИСПОЛЬЗОВАНИЯ РЕЛЯЦИОННОЙ И ГРАФОВОЙ БАЗ ДАННЫХ»** посвящен сравнению эффективности использования реляционной и графовой баз данных. Раздел состоит из нескольких подразделов.

Подраздел «ГЕНЕРАЦИЯ ИДЕНТИФИКАТОРОВ» посвящен проблеме отсутствия системы автоматической генерации идентификаторов для добавляемых сущностей в графовой СУБД Neo4j и способу её решения. [21]

Подраздел «РЕАЛИЗАЦИЯ ПОДТИПОВ СУЩНОСТЕЙ» содержит описание преимуществ и недостатков различных способов представления иерархии наследования в реляционных и графовых базах данных.

Подраздел «РЕАЛИЗАЦИЯ ДВУНАПРАВЛЕННОЙ СВЯЗИ» описывает отличия подходов, используемых в реляционных и графовых базах данных для реализации представления двунаправленных связей, например, отношения дружбы. [22]

В подразделе «РЕАЛИЗАЦИЯ РЕКОМЕНДАЦИЙ» рассматривается реализация одного из приложений графовых баз данных – системы рекомендаций. Подраздел содержит примеры запросов для получения списка возможных друзей пользователя социальной сети на языках SQL и Cypher.

Подраздел «СКОРОСТЬ ВЫПОЛНЕНИЯ ЗАПРОСОВ» содержит результаты измерения скорости выполнения запросов в реляционной СУБД SQL Server и графовой СУБД Neo4j. Измерения скорости запросов проводились при различном количестве сущностей и связей в базах данных. Были рассмотрены запросы на получение, добавление и изменения записей. Результаты представлены в виде таблиц и графиков.

## ЗАКЛЮЧЕНИЕ

В ходе данной работы было проведено сравнение эффективности использования реляционной и графовой баз данных для обработки информации о пользователях социальной сети. На основе анализа структуры данных пользователей социальных сетей было создано веб-приложение, в котором реализованы наиболее часто используемые элементы и функции существующих в этой области решений. Изучены основные принципы работы с графовыми базами данных. При разработке веб-приложения было реализовано хранение, добавление, изменение, удаление информации и выполнение запросов в реляционной и графовой базах данных. Проведено сравнение графового и реляционного подхода по основным действиям работы с базой данных.

В ходе сравнения было выяснено, что графовые базы данных являются удобным средством хранения структуры систем, модель которых представляет собой множество узлов с разнородными данными и большим количеством связей между ними. Одним из примеров таких систем может являться социальная сеть.

Если в реляционных базах данных для представления отношений между сущностями или реализации иерархии требуется особым образом организовывать структуру таблиц, прибегая к созданию дополнительных таблиц и атрибутов, то графовые базы данных позволяют описывать графовую модель естественным образом: узлы и отношения являются объектами языка запросов. Это позволяет не только упростить моделирование самого графа, но и дает возможность формулировать лаконичные запросы. В языке запросов рассмотренной графовой СУБД данные и отношения между ними описываются в виде шаблонов, что позволяет значительно сократить объем кода по сравнению с запросом на языке SQL в случае работы с графовой структурой, например, при реализации системы рекомендаций. Однако по скорости работы запросов Neo4j в рассмотренных случаях уступает реляционной СУБД SQL Server. При этом можно наблюдать, что на некоторых запросах при увеличении количества связей графовая СУБД практически не снижает скорость работы.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Лекция «Модели организации баз данных» [Электронный ресурс]. URL: <http://www.intuit.ru/studies/courses/3439/681/lecture/14023> (дата обращения: 04.05.2017).
2. Базовые понятия реляционной модели данных [Электронный ресурс]. URL: <http://citforum.ru/database/dblearn/dblearn02.shtml> (дата обращения: 04.05.2017).
3. Ограничения первичных и внешних ключей [Электронный ресурс]. URL: <https://msdn.microsoft.com/ru-ru/library/ms179610.aspx> (дата обращения: 05.05.2017).
4. Связь объектно-ориентированных СУБД с общими понятиями объектно-ориентированного подхода [Электронный ресурс]. URL: [http://citforum.ru/database/osbd/glava\\_111.shtml](http://citforum.ru/database/osbd/glava_111.shtml) (дата обращения: 07.05.2017).
5. М. Фаулер, Дж. Садаладж. NoSQL: новая методология разработки нереляционных баз данных. – М.: «Вильямс», 2013 – 48 с.
6. Я. Робинсон, Дж. Вебер, Э. Эифрем. Графовые базы данных. Новые возможности для работы. – М.: «ДМК Пресс», 2016 – 20 с.
7. I. Robinson, J. Webber, E. Eifrem. Graph Databases: New Opportunities for Connected Data. – «O'Reilly», 2014 – 105 с.
8. DB-Engines Ranking of Graph DBMS [Электронный ресурс]. URL: <https://db-engines.com/en/ranking/graph+dbms> (дата обращения: 03.05.2017).
9. Бартенев М.В., Вишняков И.Э. Использование графовых баз данных в целях оптимизации анализа биллинговой информации. Инженерный журнал: наука и инновации, 2013, вып. 11 [Электронный ресурс]. URL: <http://engjournal.ru/catalog/it/hidden/1058.html> (дата обращения: 06.05.2017).
10. Обзор графовых баз данных. Молодежный научно-технический вестник [Электронный ресурс]. URL: <http://sntbul.bmstu.ru/doc/791943.html> (дата обращения: 07.05.2017)

11. OpenNews: Представлена СУБД ArangoDB 2.4. [Электронный ресурс]. URL: <https://www.opennet.ru/opennews/art.shtml?num=41494.html> (дата обращения: 08.05.2017)
12. AllegroGraph. Semantic Graph Database. [Электронный ресурс]. URL: <https://allegrograph.com/allegrograph/> (дата обращения: 20.04.2017)
13. What is Cypher? – The Neo4j Manual. [Электронный ресурс]. URL: <http://neo4j.com/docs/stable/cypher-introduction.html> (дата обращения: 11.05.2017)
14. Patterns. Syntax. – The Neo4j Manual. [Электронный ресурс]. URL: <http://neo4j.com/docs/developer-manual/current/cypher/syntax/patterns/> (дата обращения: 11.05.2017)
15. Clauses – The Neo4j Manual. [Электронный ресурс]. URL: <http://neo4j.com/docs/developer-manual/current/cypher/clauses/> (дата обращения: 11.05.2017)
16. Многоуровневая архитектура в ASP.NET MVC 5 [Электронный ресурс]. URL: <https://metanit.com/sharp/mvc5/23.5.php> (дата обращения: 28.04.2017).
17. А. Фримен. ASP.NET MVC 5 с примерами на C# 5.0 для профессионалов, 5-е издание – М.: «Вильямс», 2014 – 8 с.
18. Маршрутизация ASP.NET [Электронный ресурс]. URL: [https://msdn.microsoft.com/ru-ru/library/cc668201\(v=vs.100\).aspx](https://msdn.microsoft.com/ru-ru/library/cc668201(v=vs.100).aspx) (дата обращения: 28.04.2017).
19. Веб-API ASP.NET [Электронный ресурс]. URL: [https://msdn.microsoft.com/ru-ru/library/hh833994\(v=vs.108\).aspx](https://msdn.microsoft.com/ru-ru/library/hh833994(v=vs.108).aspx) (дата обращения: 29.04.2017).
20. Создание веб-сайта с использованием синтаксиса Razor. [Электронный ресурс]. URL: [https://msdn.microsoft.com/ru-ru/library/gg606533\(v=vs.100\).aspx](https://msdn.microsoft.com/ru-ru/library/gg606533(v=vs.100).aspx) (дата обращения: 03.05.2017).

21.Node (Neo4j 3.2.0 API) [Электронный ресурс]. URL:  
<http://neo4j.com/docs/java-reference/current/javadocs/org/neo4j/graphdb/Node.html> (дата обращения: 12.05.2017).

22.Selecting friends – EXPLAIN EXTENDED. URL:  
<https://explainextended.com/2009/03/07/selecting-friends/> (дата обращения: 15.05.2017).