

На правах рукописи



Sign

Азимов Рустам Шухратуллович

**Алгоритмы выполнения навигационных запросов к
графам с использованием операций линейной алгебры**

Специальность **XX.XX.XX** —

**«Технология обработки, хранения и переработки злаковых,
бобовых культур, крупяных продуктов, плодоовощной продукции
и виноградарства»**

Автореферат

диссертации на соискание учёной степени
кандидата **физико-математических наук**

Город — 2020

Работа выполнена в учреждении с длинным длинным длинным длинным названием, в котором выполнялась данная диссертационная работа.

Научный руководитель: уч. степень, уч. звание
Фамилия Имя Отчество

Официальные оппоненты: **Фамилия Имя Отчество**,
доктор физико-математических наук, профессор,
Не очень длинное название для места работы,
старший научный сотрудник

Фамилия Имя Отчество,
кандидат физико-математических наук,
Основное место работы с длинным длинным
длинным длинным названием,
старший научный сотрудник

Ведущая организация: Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования с длинным длинным длинным длинным названием

Защита состоится DD mmmmmmmm YYYY г. в XX часов на заседании диссертационного совета Д123.456.78 при Название учреждения по адресу: Адрес.

С диссертацией можно ознакомиться в библиотеке Название библиотеки.

Отзывы на автореферат в двух экземплярах, заверенные печатью учреждения, просьба направлять по адресу: Адрес, ученому секретарю диссертационного совета Д123.456.78.

Автореферат разослан DD mmmmmmmm2020 года.
Телефон для справок: +7 (0000) 00-00-00.

Ученый секретарь
диссертационного совета
Д123.456.78,
д-р физ.-мат. наук

Sign
Фамилия Имя Отчество

Общая характеристика работы

Актуальность темы исследования. В современном мире происходит значительный рост данных, которые требуют анализа. Графы используются в качестве структуры данных для представления больших объемов информации в компактной и удобной для анализа форме во многих областях, например, в биоинформатике, в графовых базах данных, при статическом анализе программ. При этом оказывается необходимым выявлять сложные зависимости между вершинами графа, которые выражаются путями между ними. Поэтому очень важен широкий класс задач, посвященный поиску путей в графах. Чтобы описать свойства искомым путей в графе, необходимо задать определенные ограничения на них. Данные ограничения формулируются в виде запроса к графу, а ответом на запрос является информация о существовании путей, удовлетворяющих данным ограничениям между каждой парой вершин. Кроме того, в некоторых областях, в качестве доказательства существования таких путей необходимо предъявить все или хотя бы один из них.

Для описания ограничений на пути в помеченном графе естественно выделять пути с помощью формальных грамматик (регулярные выражения, контекстно-свободные грамматики) над алфавитом, содержащим метки на ребрах этого графа. В настоящее время активно исследуются ограничения в виде контекстно-свободных (КС) грамматик, так как они позволяют описывать более широкий класс запросов, чем регулярные выражения. Таким образом, актуальна задача поиска путей к графам с КС-ограничениями.

Ещё одной важной областью исследований является формулирование различных алгоритмов на графах в терминах операций линейной алгебры над матрицами и векторами. Для тех алгоритмов, которые позволяют найти такую формулировку, имеется возможность применить широкий класс оптимизаций, например, разреженное представление матриц и параллельное вычисление. Кроме того, данные алгоритмы зачастую просты в реализации, так как позволяют использовать существующие библиотеки для вычисления операций над матрицами. Таким образом, формулирование различных алгоритмов на графах в терминах операций линейной алгебры над матрицами и векторами является перспективным направлением.

Однако возможность использования операций линейной алгебры в задачах поиска путей с КС-ограничениями в графах в настоящее время не исследована. Большинство существующих алгоритмов в данной области имеют низкую производительность на больших графах, что затрудняет их применение на практике. Таким образом, использование операций линейной алгебры для поиска путей в графе с КС-ограничениями требует изучения.

Кроме того, в данной области КС-грамматика приводится в нормальную форму, что способствует её разрастанию и увеличивает время анализа. Таким образом, перспективным направлением является исследование возможности использования такой матричной операции, как произведение Кронекера. Использование данной операции позволит работать с КС-грамматиками произвольного вида без необходимости их преобразования. Таким образом, алгоритм поиска путей с КС-ограничениями в графах, использующий произведение Кронекера может позволить получить ещё больший прирост производительности за счёт отсутствия разрастания размеров грамматики.

Таким образом, для КС-ограничений необходима разработка алгоритмов поиска путей, использующих различные операции линейной алгебры и имеющих высокую производительность на больших графах.

Степень разработанности темы исследования. Существует классическое исследование Вэлианта, посвященное синтаксическому анализу КС-грамматик с использованием матричных операций. Однако изложенный в нём алгоритм позволяет производить анализ только над строками, что эквивалентно анализу частного случая графов — линейных помеченных графов. Данный алгоритм не применим к анализу произвольных графов.

Хорошо изучена задача поиска путей в графах с ограничениями в виде регулярных выражений. Этой задаче посвящены работы грq грq грq грq. Однако алгоритмы, представленные в данных работах не позволяют формулировать такие КС-ограничения, как, например, язык правильных последовательностей вложенных скобок. Анализ графов с ограничениями в виде данного языка применим, например, в статическом анализе программ.

Поиску путей с КС-ограничениями в произвольных графах посвящены работы Семёна Григорьева, Джелле Хеллингса, Сяованга Чжана. В данных исследованиях используются подходы, основанные на динамическом программировании. Однако, в отличие от алгоритмов, использующих операции линейной алгебры, данные алгоритмы сложно поддаются применению таких вычислительных техник, как параллельные вычисления, что затрудняет анализ больших графов.

Таким образом, актуальной является задача дальнейшего исследования анализа графов с ограничениями в виде КС-грамматик в терминах операций линейной алгебры.

Целью данной работы является разработка алгоритмов поиска путей в графе с КС-ограничениями, использующих операции линейной алгебры и имеющих высокую производительность на больших графах.

Достижение поставленной цели обеспечивается решением следующих **задач**:

1. Разработать подход к вычислению КС-запросов к графам, использующий операции линейной алгебры.

2. Разработать семейство алгоритмов вычисления КС-запросов к графам, использующих предложенный подход и позволяющих предоставлять искомые пути.
3. Разработать алгоритм вычисления КС-запросов к графам, использующий предложенный подход и произведение Кронекера в качестве основной операции линейной алгебры.

Научная новизна:

1. Подход, предложенный в диссертации, отличается от аналогов (работы Семёна Григорьева, Джелле Хеллингса) активным использованием матричных операций в процессе вычисления запросов. Это позволяет применять широкий класс оптимизаций для вычисления матричных операций и получать высокую производительность при работе с большими графами.
2. Семейство алгоритмов, предложенное в диссертации, отличается от аналогов (работы Семёна Григорьева, Джелле Хеллингса, Сяованга Чжана) возможностью автоматически распараллеливать вычисления за счёт существующих решений для матричных операций. Кроме того, предложенные алгоритмы просты в разработке, так как могут использовать существующие библиотеки для вычисления матричных операций.
3. Алгоритм, предложенный в диссертации, отличается от аналогов (работы Семёна Григорьева, Джелле Хеллингса, Сяованга Чжана) возможностью работать с произвольными КС-грамматиками без необходимости приводить их в нормальную форму. Таким образом, удаётся улучшить производительность алгоритма за счёт отсутствия разрастания размеров грамматики.

Теоретическая и практическая значимость работы.

Теоретическая значимость диссертационного исследования заключается в разработке подхода к вычислению КС-запросов к графам, использующего операции линейной алгебры, в разработке формальных алгоритмов, использующих данный подход, а также в формальном доказательстве завершаемости и корректности разработанных алгоритмов.

Полученные в ходе исследования реализации могут быть использованы для интеграции с графовой базой данных RedisGraph. Это позволит расширить выразительность языка запросов с которым работает RedisGraph, добавив возможность вычислять КС-запросы.

Методология и методы исследования. TODO

Основные положения, выносимые на защиту:

1. Разработан подход к вычислению КС-запросов к графам, использующий операции линейной алгебры.
2. Разработано семейство алгоритмов вычисления КС-запросов к графам, использующих предложенный подход и позволяющих

предоставлять искомые пути. Доказана завершаемость и корректность предложенных алгоритмов. Проведено экспериментальное исследование.

3. Разработан алгоритм вычисления КС-запросов к графам, использующий предложенный подход и произведение Кронекера в качестве основной операции линейной алгебры. Доказана завершаемость и корректность предложенного алгоритма. Проведено экспериментальное исследование.

Степень достоверности и апробация результатов.

Достоверность и обоснованность результатов исследования опирается на использование формальных методов исследуемой области, выполнение формальных доказательств и инженерные эксперименты.

Основные результаты работы были доложены на ряде международных научных конференций: GRADES'18, GRADES'20, ADBIS'20, SIGMOD'21(еще не приняли). Дополнительной апробацией является то, что разработка предложенных алгоритмов была поддержана грантом РФФ №18-11-00100 и грантом РФФИ №19-37-90101.

Публикации. Все результаты диссертации изложены в 5 [1; 2; 3; 4; 5] научных работах, которые содержат основные результаты работы и индексируются Scopus. Работы [1; 2; 3; 4] написаны в соавторстве. В [1] Р. Азимову принадлежит разработка алгоритма, доказательство его корректности и завершаемости, реализация алгоритма, работа над текстом. В [2] Р. Азимову принадлежит разработка алгоритма, доказательство его корректности и завершаемости, работа над текстом. В [3; 4] Р. Азимову принадлежит работа над доказательствами корректности и завершаемости алгоритма, работа над текстом.

Работа [5] еще не написана.

Работы, которые можно включить в список: Публикация [6] ВАК и SCOPUS матричный алгоритм для конъюнктивных грамматик. Статья [7] ВАК конъюнктивные труды ИСП РАН. Статья [8] РИНЦ про конъюнктивные и выступление на конференции PLC'17.

Список литературы

- [1] Azimov R. Context-Free Path Querying by Matrix Multiplication / Azimov R., Grigorev S. // In Proceedings of the 1st Joint International Workshop on Graph Data Management Experiences & Systems (GRADES) and Network Data Analytics (NDA) (GRADES-NDA'18)
- [2] Azimov R. Context-Free Path Querying with Single-Path Semantics by Matrix Multiplication / Terekhov A., Khoroshev A., Azimov R., Grigorev S. // In Proceedings of the 3rd Joint International Workshop on Graph

- [3] Azimov R. Context-Free Path Querying by Kronecker Product / Orachev E., Epelbaum I., Azimov R., Grigorev S. // In Proceedings of the 24th European Conference on Advances in Databases and Information Systems (ADBIS'20)
- [4] Azimov R. Context-Free Path Querying by Kronecker Product большая версия / Orachev E., Epelbaum I., Azimov R., Grigorev S. // In Proceedings of the (SIGMOD'21)
- [5] Azimov R. Ненаписанная работа матричный алгоритм по всем путям
- [6] Azimov R. Path Querying with Conjunctive Grammars by Matrix Multiplication / Azimov R., Grigorev S. // Programming and Computer Software. – 2019. – Vol. 45. – №. 7. – pp. 357-364.
- [7] Азимов Р. Ш. Синтаксический анализ графов с использованием конъюнктивных грамматик / Азимов Р., Григорьев С. // Труды ИСП РАН, 2018, том 1 вып. 2, с. 3-4.
- [8] Азимов Р. Ш. Синтаксический анализ графов и задача генерации строк с ограничениями / Азимов Р., Григорьев С. // Сборник трудов конференции PLC 2017, с. 24-27.

Содержание работы

Во введении обосновывается актуальность исследований, проводимых в рамках данной диссертационной работы, приводится обзор научной литературы по изучаемой проблеме, формулируется цель, ставятся задачи работы, излагается научная новизна и практическая значимость представляемой работы. В последующих главах сначала описывается общий принцип, позволяющий ..., а потом идёт апробация на частных примерах: ... и

Первая глава посвящена ...

картинку можно добавить так:

Формулы в строку без номера добавляются так:

$$\lambda_{T_s} = K_x \frac{dx}{dT_s}, \quad \lambda_{q_s} = K_x \frac{dx}{dq_s},$$

Вторая глава посвящена исследованию

Третья глава посвящена исследованию

L^AT_EX

а) L^AT_EX



б) Knuth

Рис. 1 — Подпись к картинке.

Можно сослаться на свои работы в автореферате. Для этого в файле `Synopsis/setup.tex` необходимо присвоить положительное значение счётчику `\setcounter{usefootcite}{1}`. В таком случае ссылки на работы других авторов будут подстрочными. Изложенные в третьей главе результаты опубликованы в `[vakbib1; vakbib2]`. Использование подстрочных ссылок внутри таблиц может вызывать проблемы.

В четвертой главе приведено описание

В заключении приведены основные результаты работы, которые заключаются в следующем:

1. На основе анализа ...
2. Численные исследования показали, что ...
3. Математическое моделирование показало ...
4. Для выполнения поставленных задач был создан ...

При использовании пакета `biblatex` список публикаций автора по теме диссертации формируется в разделе «Публикации.» файла `common/characteristic.tex` при помощи команды `\nocite`

Азимов Рустам Шухратулович

Алгоритмы выполнения навигационных запросов к графам с использованием
операций линейной алгебры

Автореф. дис. на соискание ученой степени **канд. физ.-мат. наук**

Подписано в печать _____._____._____. Заказ № _____

Формат 60×90/16. Усл. печ. л. 1. Тираж 100 экз.

Типография _____

