



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ
(национальный исследовательский университет)»

Институт (Филиал) № 8 «Компьютерные науки и прикладная математика» Кафедра 806

Группа М8О-406Б-19 Направление подготовки 01.02.03 «Прикладная математика и информатика»

Профиль Информатика

Квалификация: бакалавр

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

на тему «Модификация базы данных отделения неотложной медицинской помощи с использованием PostgreSQL»

Автор ВКРБ: Алексеев Владислав Евгеньевич ()

Руководитель: Пивоваров Дмитрий Евгеньевич ()

Консультант: — ()

Консультант: — ()

Рецензент: — ()

К защите допустить

Заведующий кафедрой № 806 Крылов Сергей Сергеевич ()

____ мая 2023 года

Москва 2023

РЕФЕРАТ

Выпускная квалификационная работа бакалавра состоит из 48 страниц, 7 рисунков, 13 использованных источников.

БАЗА ДАННЫХ, АВТОМАТИЗАЦИЯ, БЕЗОПАСНОСТЬ, ШИФРОВАНИЕ, POSTGRESQL

Объектом исследования в данной выпускной квалификационной работе является деятельность отделения неотложной медицинской помощи.

Предметом исследования является процесс разработки и внедрения базы данных медицинских диагностических таблиц первичного осмотра пациента, таблиц дифференциальной диагностики и вспомогательных информационных таблиц.

Целью работы является автоматизация, ускорение и улучшение условий обработки данных для сотрудников отделения неотложной медицинской помощи.

Основное содержание работы состояло в разработке базы данных медицинских диагностических таблиц первичного осмотра пациента для сотрудников отделения неотложной медицинской помощи.

Для достижения поставленной цели был проведен анализ функциональных требований для разработки необходимой базы данных, а также анализ наиболее эффективных способов хранения взаимосвязанных атрибутов каждой сущности. Кроме того, была разработана и протестирована довольно обширная структура базы данных, объединяющая сильные стороны проанализированных методов хранения и обработки информации.

Основными результатами работы, полученными в процессе анализа и разработки являются: определение наиболее эффективных способов структурирования и хранения всей необходимой информации для работников скорой медицинской помощи, разработка и внедрение базы данных, включая вспомогательные информационные таблицы, оценка эффективной работы полученной базы данных и выявление сценариев ее оптимального использования.

Полученные результаты разработки представляют немаловажное значение для всех организаций здравоохранения, стремящихся оптимизировать работу своего персонала и минимизировать время и трудозатраты на рутинное заполнение бумажных документов.

СОДЕРЖАНИЕ

ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ	4
ПЕРЕЧЕНЬ СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ	5
ВВЕДЕНИЕ	6
1 ТЕОРЕТИЧЕСКОЕ ОБОСНОВАНИЕ РАЗРАБОТКИ БАЗЫ ДАННЫХ	9
1.1 Анализ предметной области	9
1.2 Назначение и возможности базы данных	12
2 ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА БАЗЫ ДАННЫХ	16
2.1 Выбор программных средств для реализации базы данных . . .	16
2.2 Проектирование базы данных	17
2.3 Проектирование логических моделей данных	20
2.4 Нормализация базы данных	21
2.5 Проектирование физических моделей данных	22
2.6 Готовые решения	23
3 ЗАЩИТА БАЗЫ ДАННЫХ ОТ НЕСАНКЦИОНИРОВАННОГО ДОСТУПА	31
3.1 Конфиденциальность персональных данных	31
3.2 Использование регистрационных имен, имен пользователей, ролей и разрешений	33
3.3 Шифрование базы данных	35
3.4 Порты	36
3.5 Брандмауэр	37
4 ТЕСТИРОВАНИЕ И ОПТИМИЗАЦИЯ ПРОИЗВОДИТЕЛЬНОСТИ БАЗЫ ДАННЫХ	39
4.1 Методы и инструменты тестирования	39
4.2 Нагрузочное тестирование и оптимизация производительности	39
ЗАКЛЮЧЕНИЕ	46
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	47

ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ

В настоящей выпускной квалификационной работе бакалавра применяют следующие термины с соответствующими определениями:

ACID — это набор свойств транзакций базы данных, предназначенных для гарантии достоверности данных, несмотря на ошибки, сбои питания и другие неудачи

BSON — это формат электронного обмена цифровыми данными, бинарная форма представления простых структур данных и ассоциативных массивов. Является надмножеством JSON, включая дополнительно регулярные выражения, двоичные данные и даты

JSON — это популярный формат текстовых данных, который используется для обмена данными в современных веб - и мобильных приложениях

SQL-инъекция — это уязвимость, которая позволяет атакующему использовать фрагмент вредоносного кода на языке структурированных запросов (SQL) для манипулирования базой данных и получения доступа к потенциально ценной информации

SSL — это криптографический протокол, который подразумевает более безопасную связь и использует асимметричную криптографию для аутентификации ключей обмена, симметричное шифрование для сохранения конфиденциальности, коды аутентификации сообщений для целостности сообщений

TCP — это важный протокол сети интернет, который позволяет двум хостам создать соединение и обмениваться потоками данных

UDP — это один из ключевых элементов набора сетевых протоколов для Интернета, с помощью которого компьютерные приложения могут посылать сообщения другим хостам по IP-сети без необходимости предварительного сообщения для установки специальных каналов передачи или путей данных

Модель OSI — это сетевая модель стека сетевых протоколов, посредством которой различные сетевые устройства могут взаимодействовать друг с другом

Наряд — это группа медицинских работников и транспортных средств, назначенных для выполнения определенной миссии или задачи, а также медицинское оборудование и инструменты

ПЕРЕЧЕНЬ СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ

В настоящей выпускной квалификационной работе бакалавра применяют следующие сокращения и обозначения:

ACID — атомарность, согласованность, изоляция и долговечность

ANSI — американский национальный институт стандартов

TCP — протокол управления передачей

UDP — протокол пользовательских датаграмм

XML — расширяемый язык разметки

БД — база данных

ИС — информационная система

МКБ — международная классификация болезней

ОНМП — отделение неотложной медицинской помощи

ОС — операционная система

СМП — скорая медицинская помощь

СУБД — система управления базами данных

ВВЕДЕНИЕ

Актуальность темы данной выпускной квалификационной работы бакалавра связана с большим значением проведения эффективных мероприятия по оказанию медицинской помощи пациентам со стороны медицинских работников, максимальное упрощение заполнения документов о выездах и действиях по ее оказанию. Кроме того, врачам иногда требуется некоторый вспомогательный материал в виде информационных данных в таблицах с симптомами диагнозов или дозировками того или иного препарата.

В наше время использование баз данных является одной из неотъемлемых частей работы многих компаний и организаций, так как это позволяет хранить, организовывать и управлять большими объемами данных. Базы данных позволяют быстро и эффективно хранить и обрабатывать данные, что является необходимым условием для принятия важных бизнес-решений.

Одним из наиболее популярных языков для работы с базами данных является PostgreSQL. Он обладает высокой степенью надежности, масштабируемости и производительности. PostgreSQL также поддерживает многие функции, которые делают его удобным для работы с различными типами данных и приложений [1].

Использование баз данных также позволяет легко и быстро находить необходимую информацию и управлять ею, а также обеспечивает сохранность и целостность данных, что немаловажно в наших современных реалиях, ведь большинство информации просто напросто может утечь в сеть и оказаться в открытом доступе, желаем мы того или нет. Базы данных позволяют использовать различные методы анализа и обработки данных, что является важным инструментом для повышения эффективности работы компаний и организаций.

Кроме того, использование баз данных является современным и необходимым подходом к работе с данными, который позволяет не только сохранять и управлять ими, но и анализировать и использовать в дальнейшем. Базы данных PostgreSQL являются открытым и свободно распространяемым решением, что делает их доступными и удобными для использования в различных проектах и приложениях.

В современном мире качество медицинской помощи и скорость ее

оказания имеют большое значение для общества. Скорая медицинская помощь играет важную роль в сохранении здоровья и жизни людей в случае непредвиденных ситуаций и аварий. Однако, существующая система скорой помощи имеет свои проблемы, связанные с низкой эффективностью и долгим временем ожидания при оказании медицинской помощи.

Внедрение цифровых технологий в работу скорой медицинской помощи может значительно улучшить качество и эффективность ее работы. В частности, организация электронных медицинских карт и использование баз данных на языке PostgreSQL позволят ускорить процесс диагностики и лечения пациентов, а также сократить время ожидания скорой медицинской помощи.

Данная выпускная квалификационная работа имеет практическое значение, так как внедрение цифровых технологий в работу скорой медицинской помощи может ускорить процесс оказания медицинских услуг и повысить их качество, что в свою очередь приведет к улучшению уровня здоровья и жизни населения в целом.

Сейчас врачам важно и нужно переходить на цифровое использование и заполнение медицинских карт при осмотре пациентов по нескольким причинам:

- цифровая медицинская карта позволяет быстро и удобно получить доступ к информации о пациенте, что повышает эффективность и точность диагностики и лечения,
- цифровая медицинская карта позволяет сохранять и обрабатывать большие объемы данных о пациентах, что является важным инструментом для проведения исследований и разработки новых методов лечения. Также это позволяет врачам получать доступ к общей истории лечения пациента и учитывать его предыдущие заболевания, что повышает качество и эффективность лечения,
- использование цифровых медицинских карт позволяет уменьшить вероятность ошибок и искажений при заполнении и хранении информации о пациентах, что обеспечивает сохранность и конфиденциальность медицинских данных.

Наконец, использование цифровых медицинских карт является современным и удобным подходом к организации работы медицинских учреждений и обслуживанию пациентов, что способствует повышению

уровня медицинской помощи и общей качества жизни населения.

В данной выпускной квалификационной работе бакалавра были использованы различные технологии и инструменты для облегчения исследования, реализации и анализа работы всей системы:

- PostgreSQL – мощная реляционная система управления базами данных (СУБД), которая обладает высокой производительностью, масштабируемостью, надежностью и расширяемостью. Она поддерживает широкий спектр функций, позволяет работать с различными типами данных и обеспечивает сохранность и целостность данных,
- PgAdmin – это графический инструмент для администрирования баз данных PostgreSQL. Он предоставляет удобный интерфейс для управления базами данных и объектами в них, такими как таблицы, индексы, пользователи и многое другое. PgAdmin поддерживает широкий диапазон функций, включая создание и редактирование объектов базы данных, выполнение SQL-запросов, экспорт и импорт данных, а также управление безопасностью и настройками базы данных,
- Erwin – это программное обеспечение для моделирования баз данных. Оно позволяет проектировать их визуально в удобном графическом интерфейсе. Erwin поддерживает множество популярных СУБД, включая PostgreSQL, и предоставляет широкий набор инструментов для работы с базами данных, таких как автоматическое генерирование SQL-кода, проверка целостности данных, анализ производительности и многое другое.

С научной точки зрения, данная работа вносит вклад в существующий массив знаний о разработке и проектировании баз данных и их эффективности в различных сценариях нагрузки и использования. С практической точки зрения, данное исследование предоставляет ценные идеи для медицинских работников, для которых так важно оптимизировать свои действия по заполнению медицинских карт, карт осмотра пациентов и многие другие рутинные задачи.

Таким образом, данная выпускная квалификационная работа бакалавра является актуальной и с научно-методической/теоретической, и с практической точек зрения.

1 ТЕОРЕТИЧЕСКОЕ ОБОСНОВАНИЕ РАЗРАБОТКИ БАЗЫ ДАННЫХ

1.1 Анализ предметной области

Анализ предметной области является первым этапом в проектировании базы данных, в ходе которого выделяются основные объекты и их свойства, определяются первоначальные требования и границы проекта, чтобы разработать эффективную и безопасную базу данных.

Предметной областью разрабатываемой модели данных является система заполнения электронных медицинских карт для отделения неотложной медицинской помощи.

Скорая медицинская помощь (СМП) – вид медицинской помощи, оказываемой гражданам при заболеваниях, несчастных случаях, травмах, отравлениях и других состояниях, требующих срочного медицинского вмешательства.

Отделение неотложной (экстренной) медицинской помощи больницы является структурным подразделением многопрофильной больницы, которое в круглосуточном режиме оказывает экстренную (неотложную) медицинскую помощь.

Медицинская помощь включает в себя процедуру проведения осмотра, непосредственно манипуляции по оказанию медицинской помощи, а также консультирование пациента по с целью определения наиболее эффективного, безопасного и экономически оправданного курса лечения.

ОНМП осуществляет следующие функции:

- прием пациентов с острыми заболеваниями, несчастными случаями, травмами, отравлениями и другими состояниями, требующими немедленной медицинской помощи,
- проведение первичной медицинской диагностики и оценки состояния пациента, осуществление мер, направленных на стабилизацию состояния пациента,
- предоставление неотложной медицинской помощи, включая проведение лечебных манипуляций, инъекций, переливание крови, а также оказание психологической помощи,
- организация и координация работы других специалистов и служб в медицинском учреждении, если необходимо,

- подготовка пациента к транспортировке в стационар для дальнейшего лечения,
- соблюдение всех медицинских стандартов и требований, направленных на обеспечение безопасности пациентов и медицинского персонала,
- организация транспортировки пациентов в случае необходимости,
- обеспечение работы необходимого медицинского оборудования, инструментов, материалов и медикаментов, необходимых для оказания неотложной медицинской помощи,
- проведение медицинской документации, включая учет медицинских случаев, регистрацию историй болезни и медицинских записей,
- обеспечение мониторинга и контроля за состоянием пациентов, находящихся на лечении в отделении неотложной (экстренной) медицинской помощи,
- обучение медицинского персонала и сотрудников отделения неотложной медицинской помощи новым методам лечения и диагностики,
- сотрудничество и консультации с другими медицинскими учреждениями, специалистами и службами для повышения качества оказываемой медицинской помощи.

Отделение неотложной (экстренной) медицинской помощи является ключевым звеном в системе оказания медицинской помощи населению и обеспечивает быстрое и эффективное лечение при острых заболеваниях и травмах.

Базы данных для неотложной медицинской помощи являются критически важным элементом в работе современных медицинских учреждений. Они позволяют эффективно организовывать, хранить и обрабатывать медицинскую информацию, ускоряя процессы принятия решений и повышая качество медицинской помощи.

Основным требованием к базе данных для неотложной медицинской помощи является ее способность оперативно и точно хранить медицинскую информацию о пациентах. Эта информация может включать данные о медицинской истории пациента, диагнозе, принятых мероприятиях, результатах обследований и лекарственном лечении.

Для эффективного управления медицинской информацией в базе

данных неотложной медицинской помощи необходима специализированная система управления базами данных (СУБД). Существует множество СУБД, которые могут использоваться для этой цели, включая Oracle, MySQL, Microsoft SQL Server, PostgreSQL и др.

Также необходимо учитывать специфику работы медицинских учреждений, которые могут иметь различные требования к хранению и обработке медицинской информации. Например, больницы могут иметь разные отделения, каждое из которых может иметь свои особенности в обработке информации. Кроме того, необходимо учитывать возможность интеграции с другими системами, такими как системы управления ресурсами и планирования процессов.

Важно также отметить, что база данных для неотложной медицинской помощи должна соответствовать требованиям законодательства в области защиты персональных данных, таких как GDPR в Европейском союзе и HIPAA в США. Регулярное обновление и адаптация базы данных к изменениям в законодательстве помогут поддерживать соответствие нормам и предотвращать правовые проблемы.

Исследования показывают, что эффективное использование баз данных в медицинских учреждениях может значительно улучшить качество медицинской помощи и снизить затраты на ее оказание. Например, использование баз данных может уменьшить количество ошибок в принятии решений, сократить время, необходимое для оказания медицинской помощи, и повысить точность диагностики.

На текущий момент существует множество различных подходов к проектированию баз данных для медицинских учреждений. Некоторые из них ориентированы на сохранение структурированной информации, другие на работу с полуструктурированными и неструктурированными данными.

Одним из наиболее распространенных подходов к проектированию баз данных для медицинских учреждений является использование реляционной модели данных. Реляционная модель данных позволяет хранить информацию в виде таблиц, которые могут быть связаны друг с другом через ключи. Это позволяет обрабатывать большие объемы структурированных данных и осуществлять запросы на выборку данных, используя язык SQL.

Тем не менее, в последние годы набирают популярность нереляционные базы данных, такие как MongoDB, Cassandra, Couchbase и др. Они хранят

данные в более гибкой форме, позволяют более легко масштабировать базы данных и работать с полуструктурированными и неструктурированными данными, такими как тексты медицинских записей, изображения и видео.

Важным аспектом разработки баз данных для медицинских учреждений является также обеспечение защиты медицинской информации от несанкционированного доступа. Для этого применяются различные меры безопасности, такие как шифрование данных, многофакторная аутентификация и разграничение прав доступа к информации в зависимости от роли пользователя.

Наконец, важно отметить, что разработка баз данных для медицинских учреждений является сложным и многогранным процессом, требующим учета множества факторов, таких как специфика медицинских процедур, законодательство, требования к безопасности и другие. Поэтому необходимо проводить тщательный анализ предметной области и разрабатывать индивидуальные решения для каждого конкретного медицинского учреждения.

1.2 Назначение и возможности базы данных

В предметной области системы заполнения электронных медицинских карт для отделения неотложной медицинской помощи, основные объекты и свойства, которые следует рассмотреть, могут включать:

- пациенты: информация о каждом пациенте, включая персональные данные (имя, дата рождения, пол и контактная информация), медицинскую историю, диагнозы, принятые мероприятия, результаты обследований и лекарственное лечение,
- медицинские работники: данные о врачах, медсестрах и других медицинских специалистах, включая их идентификационные данные, специализацию, график работы и доступные привилегии,
- медицинские процедуры: информация о проводимых процедурах, включая коды процедур, описания, стоимость, требуемое оборудование и прочие детали,
- отделения: данные о различных отделениях неотложной медицинской помощи в больнице, их назначение, доступный персонал и оборудование,
- ресурсы: информация о доступных ресурсах, таких как медицинское

- оборудование, лекарства, материалы и другие необходимые средства,
- расписание: график работы медицинского персонала и расписание доступности отделений и ресурсов,
- системы управления и интеграция: необходимо учесть возможность интеграции с другими системами, такими как системы управления ресурсами и планирования процессов, чтобы обеспечить эффективное взаимодействие и координацию деятельности медицинских учреждений.

Для лучшего понимания предметной области, рассмотрим конкретный пример - разработку базы данных для отделения неотложной медицинской помощи.

Медицинские работники оказывают медицинскую помощь пациентам с острыми заболеваниями и травмами, которые требуют немедленного вмешательства, также необходимо быстро и точно определить диагноз, назначить лечение и принять меры по сохранению жизни пациента.

Одной из основных задач базы данных для отделения неотложной медицинской помощи является хранение и обработка медицинских данных пациентов, включая информацию о симптомах, диагнозах, назначенных лекарствах, процедурах и т.д. Кроме того, необходимо учитывать, что пациенты могут обращаться за медицинской помощью в нескольких отделениях неотложной медицинской помощи, поэтому база данных должна позволять обмениваться информацией между различными медицинскими учреждениями.

Для решения этих задач можно использовать реляционную базу данных, в которой каждый пациент будет представлен в виде отдельной записи в таблице, содержащей данные о пациенте, диагнозах, лекарствах и т.д. Ключами в таблицах могут быть номера пациента, номера записи и т.д. Это позволит производить выборку данных о конкретном пациенте, обращаться к истории его болезни, проводить анализ данных и выявлять тенденции в заболеваемости и лечении.

Однако, реляционная модель может столкнуться с проблемами при работе с полуструктурированными и неструктурированными данными, такими как медицинские изображения, видео и тексты медицинских записей. Для работы с этими данными может использоваться нереляционная база данных, такая как MongoDB. В MongoDB данные могут быть храниться в

более гибкой форме, используя форматы, такие как JSON и BSON. MongoDB также позволяет хранить и обрабатывать файлы в формате BLOB (binary large object), что делает его идеальным инструментом для хранения медицинских изображений и других неструктурированных данных.

Еще одной важной задачей при разработке базы данных для отделения неотложной медицинской помощи является обеспечение безопасности хранения и доступа к медицинским данным. Для этого может использоваться различные меры, такие как шифрование данных, авторизация пользователей и аудит доступа.

Также при проектировании базы данных для отделения неотложной медицинской помощи необходимо учитывать требования к ее масштабируемости, отказоустойчивости и производительности. В случае большого количества пациентов и медицинских записей может потребоваться использование кластерной архитектуры, репликации данных и других технологий, позволяющих обеспечить высокую доступность и производительность базы данных. Также необходимо обрабатывать большие объемы данных и поддерживать быстрый доступ к ним.

На сегодняшний день существует множество различных систем управления базами данных, которые могут быть использованы для разработки базы данных для отделения неотложной медицинской помощи, включая MySQL, PostgreSQL, Oracle Database, Microsoft SQL Server, MongoDB и др. Выбор конкретной системы управления базами данных зависит от требований к производительности, масштабируемости, отказоустойчивости и других факторов.

В заключение, разработка базы данных для отделения неотложной медицинской помощи является сложной задачей, которая требует учета множества факторов, таких как безопасность, масштабируемость, отказоустойчивость и производительность. Решение этих задач может потребовать использования различных технологий и систем управления базами данных, а также тщательного анализа требований и потребностей пользователей.

Анализируя все вышеперечисленные требования к организации базы данных, можно сделать вывод, что основным субъектом данной базы данных является врач, который будет осуществлять выезд и непосредственное оказание всех необходимых медицинских услуг.

Врач может иметь доступ к данным о препаратах, находящихся в использовании у данного наряда, он может осуществить поиск по названию лекарственного средства для оперативного предоставления ответа по запросу клиента.

Основные реализуемые функции:

- аутентификация и авторизация,
- разграничение ролей пользователей,
- добавление и удаление данных,
- поиск данных по нескольким критериям,
- вывод вспомогательных информационных таблиц.

Выводы по разделу

В конечном итоге, проектирование базы данных для системы заполнения электронных медицинских карт является сложным процессом, требующим тщательного анализа предметной области, определения требований и границ проекта, выбора подходящей модели данных и СУБД, обеспечения безопасности данных, обучения пользователей и поддержки системы. Это комплексный процесс, который должен выполняться с участием экспертов в области медицины и баз данных для достижения оптимального результата. Результатом успешной разработки будет эффективная система управления медицинской информацией, способствующая повышению качества медицинской помощи, оптимизации процессов и улучшению результатов лечения пациентов.

2 ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА БАЗЫ ДАННЫХ

2.1 Выбор программных средств для реализации базы данных

Выбор программных средств играет большое значение для проектируемой базы данных, поскольку определенные на этом этапе программные продукты, средства разработки и технологии непосредственным образом окажут влияние на распространение разработанной продукта: чем более распространенными будут используемые при разработке и требуемые для функционирования базы данных технологии, тем шире круг потенциальных потребителей, имеющих возможность ей воспользоваться. Кроме того, выбор среды разработки непосредственно влияет на стоимость и простоту процесса разработки в целом.

База данных «оптр» является реляционной. Реляционная база данных представляет собой набор таблиц, однако достаточно часто в состав базы данных входят и другие элементы, позволяющие дополнительно влиять на организацию и структуру данных в соответствии с определенным набором требований.

Создание и развитие динамических веб-страниц требует использования различных технологий. Разработка динамических веб-страниц включает три основных компонента: веб-сервер, язык программирования сценариев, исполняемых на стороне сервера, и базу данных.

Язык структурированных запросов (Structured Query Language, SQL) – самый распространенный язык, предназначенный для записи, извлечения, обновления и удаления информации в системах управления реляционными базами данных.

PostgreSQL (полное название "PostgreSQL: The world's most advanced open source relational database") - это мощная объектно-реляционная система управления базами данных (СУБД), которая поддерживает SQL-запросы и соответствует многим стандартам ANSI SQL. PostgreSQL является свободным и открытым программным обеспечением, доступным для использования и модификации бесплатно.

PostgreSQL обладает рядом преимуществ по сравнению с другими СУБД [2; 3] :

- надежность и целостность данных: PostgreSQL имеет высокую

степень надежности и обеспечивает высокую степень целостности данных благодаря поддержке транзакций, атомарности, согласованности и изоляции (ACID),

- безопасность: PostgreSQL обеспечивает множество функций безопасности, включая возможность управления правами доступа на уровне таблиц, столбцов, функций и процедур, проверку подлинности и защиту от SQL-инъекций,
- масштабируемость: PostgreSQL обладает высокой степенью масштабируемости и может обрабатывать большие объемы данных. Он поддерживает репликацию данных и многоуровневую архитектуру серверов,
- поддержка SQL: PostgreSQL поддерживает стандарт SQL и соответствует многим стандартам ANSI SQL, что облегчает разработку и поддержку приложений,
- расширяемость: PostgreSQL поддерживает расширения, которые позволяют разработчикам создавать свои собственные функции, типы данных и языки программирования, расширяя тем самым функциональность базы данных,
- мощный функционал: PostgreSQL имеет широкий функционал, включая поддержку полнотекстового поиска, гео-пространственных запросов, JSON- и XML-обработки, а также многопоточности,
- открытость: PostgreSQL является свободным и открытым программным обеспечением, доступным для использования и модификации бесплатно, что обеспечивает независимость от поставщика и поддержку со стороны большого сообщества разработчиков и пользователей.

2.2 Проектирование базы данных

Проектирование базы данных для любой автоматизированной системы разделяется на следующие этапы:

- определение требований к базе данных,
- концептуальное проектирование базы данных,
- выбор средств реализации базы данных,
- логическое проектирование,
- физическое проектирование.

Объекты, которые хранятся в базе данных, имеют некую логическую структуру, то есть описываются некоторой моделью представления данных. К числу классических моделей относят следующие: иерархическая, сетевая и реляционная.

Иерархическая модель – представляет собой упорядоченную совокупность экземпляров данных типа «дерево», содержащих экземпляры типа «запись». Тип «дерево» является составным. Он включает в себя подтипы («поддерева»), каждый из которых, в свою очередь, является типом «дерево». Каждый из типов «дерево» состоит из одного «корневого» типа и упорядоченного набора (возможно, пустого) подчиненных типов, тем самым устанавливая связь «предок-потомок».

Сетевая модель – отображает разнообразные взаимосвязи данных в виде произвольного графа, обобщая тем самым иерархическую модель данных. Для описания этой модели используется два типа «запись» и «связь». Тип «связь» определяется для двух типов «запись», то есть для предка и потомка. Если в иерархической модели потомок может иметь только одного предка, то тут он может иметь произвольное число предков.

Реляционная модель – основывается на понятии отношение, представляя собой множеством элементов, называемых кортежами. Наглядной формой представления отношения является привычная для восприятия двумерная таблица. Данная модель состоит из строк и столбцов. Каждая строка имеет одинаковую структуру и состоит из полей. Строкам таблицы соответствуют кортежи, а столбцам – атрибуты. С помощью таблицы удобно описывать простейший вид связи между объектами, а именно деление одного на множество подобъектов, каждому из которых соответствует строка таблицы. Поскольку в рамках одной таблицы не удастся описать более сложные логические структуры данных из предметной области, применяют связывание таблиц.

Для реализации данной медицинской информационной системы будем использовать реляционную модель.

Концептуальная модель базы данных – это некая наглядная диаграмма, изображаемая в принятых обозначениях и подробно показывающая связь между объектами и их характеристиками. Концептуальная модель создается для дальнейшего проектирования базы данных и перевода ее, например, в реляционную базу данных. В концептуальной модели в визуальном

виде прописываются связи между объектами данных и их характеристиками [4].

ER-модель – модель данных, позволяющая описывать концептуальные схемы на основе диаграмм сущность-связь (ER-диаграмм).

Модель базы данных описывают с помощью одной или нескольких ER-диаграмм, содержащих сущности, атрибуты и связи.

Сущность определяется как объект, событие или концепция, информация о котором должна сохраниться. Сущности имеют наименование, несущее четкое смысловое значение. Каждый экземпляр сущности на диаграмме уникален.

Атрибут хранит информацию об определенном свойстве сущности и имеет четкое смысловое значение. Атрибут или группа атрибутов, которые однозначно идентифицируют экземпляры сущности, называются первичным ключом (англ. primary key).

Связь описывает логическое соотношение между сущностями. Связь сущности с другими сущностями определяет ее тип: различают два типа сущностей – зависимые и независимые.

Можно установить следующие связи между сущностями: идентифицирующая связь «Один–ко–Многим», связь «Многие–ко–Многим», неидентифицирующая связь «Один–ко–Многим» и связь «Один–к–Одному».

Связь «Многие–ко–Многим» существует только на логическом уровне. При переходе на физический уровень это отношение должно быть преобразовано за счет добавления новой зависимой сущности, связанной идентифицирующими связями «Один–ко–Многим» с сущностями, находящимися в исходном отношении.

Сущности, имеющие связь «Один–к–Одному», можно объединить в одну. При физической реализации базы данных две таблицы могут использоваться вместо одной по соображениям конфиденциальности, для удобства, для экономии дискового пространства, из семантических соображений.

Идентифицирующая связь устанавливается между независимой и зависимой сущностями, при этом зависимая сущность не может существовать самостоятельно - экземпляр зависимой сущности определяется только через отношение к сущности, которая его идентифицирует. При установлении идентифицирующей связи атрибуты первичного ключа родительской

сущности автоматически переносятся в состав ключевых атрибутов дочерней сущности. В дочерней сущности новые атрибуты помечаются как внешний ключ. В случае неидентифицирующей связи внешний ключ не входит в состав первичного ключа дочерней сущности.

2.3 Проектирование логических моделей данных

Проектирование модели данных состоит из двух уровней представления данных - логического и физического. Такое разделение на модели позволяет разделять задачу на более мелкие элементы.

Логический уровень – это некоторая абстрактная модель данных, позволяющая описать исследуемый объект, подчеркивая в нем необходимые свойства. Преимущество этого уровня заключается в том, что он является универсальным, поэтому реализовав его, можно создать БД, используя всевозможные для этого инструменты.

Различают три уровня логической модели, отличающихся по глубине представления информации о данных:

- диаграмма сущность-связь (Entity Relationship Diagram, ERD),
- модель данных, основанная на ключах (Key Based model, KB),
- полная атрибутивная модель (Fully Attributed model, FA).

Диаграмма сущность-связь - этот тип логической модели используется при первоначальной разработке БД. Он позволяет описать основные объекты или процессы, необходимые для реализации в БД. Для этого вводятся 3 понятия: сущность, связь, атрибут. Сущность – это объект, находящийся в БД. Используя сущности, можно описать основные таблицы в будущей БД. Связь – это соединение или введение некоторого взаимодействия между сущностями. Если интерпретировать связь в БД, то связь выступает в роли внешних ключей. Атрибут – это ключевые элементы сущности, необходимые для его описания. В БД атрибут можно сравнить со столбцами, хранящимися в таблице. Этот тип моделирования позволяет опускать описание атрибутов, чтобы не загромождать диаграмму.

Модель данных, основанная на ключах - это дополненная ER-диаграмма. Ее главное отличие от первой: обязательное наличие некоторых атрибутов. Этими атрибутами являются первичные ключи, необходимые для соблюдения уникальности данных, и внешние ключи, обеспечивающие целостность данных. Стоит отметить, что внешние ключи

никак не изображаются, они лишь являются следствием связей между сущностями. Причем внешний ключ находится у слабой сущности, а сам этот ключ ссылается на первичный ключ сильной сущности.

Первичные ключи принято обозначать как РК – Primary Key, графически они обозначаются двумя подчеркнутыми линиями, также допускается использование жирного шрифта. Внешние ключи принято обозначать как FK – Foreign Key, графически обозначаются одной подчеркнутой линией, допускается использование курсивного шрифта.

Полная атрибутивная модель - самая детализированная модель на логическом уровне. В ней исключается лишняя или дублирующая информация – этот процесс называется нормализацией БД. Помимо минимизации избыточной информации, в этой модели представляются все существующие сущности, связи и атрибуты.

Для построения полной атрибутивной модели используется IDEF1X нотация. В ней основными объектами также являются сущности, связи и атрибуты. Сущности называются в единственном числе и имеют четкое смысловое значение. Атрибуты должны именоваться в единственном числе и иметь четкое смысловое значение. Каждая связь должна именоваться глаголом или глагольной фразой, причем глагол ставится от 3 лица.

2.4 Нормализация базы данных

Для того, чтобы оптимизировать модель и избавиться от избыточных данных, проводят нормализацию данных. Нормализация – разбиение таблицы на две или более, обладающие лучшими свойствами при добавлении, изменении и удалении данных. Нормализация осуществляется с целью оптимизации объема БД и быстродействия запросов. Всего существует пять нормальных форм, но реально, на практике, используются лишь третья, а именно база данных в третьей нормальной форме (3НФ). Процесс нормализации отношений осуществляется пошагово и заключается в последовательном переводе отношения от первой нормальной формы к нормальным формам более высокого порядка. При этом каждая следующая нормальная форма сохраняет все свойства предыдущих.

Нужно проверить, находится ли полученная схема отношений в третьей нормальной форме (3НФ). Если схема отношений не находится в 3НФ, то ее нужно нормализовать для минимизации избыточности данных и устранения

потенциальной противоречивости данных.

Отношение находится в первой нормальной форме (1НФ), если значения всех атрибутов атомарные, то есть значение атрибута не должно быть множеством или повторяющейся группой.

Отношение находится во второй нормальной форме (2НФ), если оно находится в 1НФ, и нет частичной функциональной зависимости неключевых атрибутов от ключа (зависимость не ключевых атрибутов от части ключа). Из схемы отношений видно, что ни один из неключевых атрибутов функционально полно не зависит от ключа, следовательно, схема отношений находится в 2НФ.

Отношение находится в 3НФ, если оно находится во 2НФ, и отсутствуют транзитивные зависимости неключевых атрибутов от ключа. Между атрибутами А и С есть транзитивная зависимость, если выполняется совокупность условий: если хотя бы одно из условий не выполняется, то транзитивной зависимости между атрибутами А, В, С нет. Причем атрибуты А, В, С могут быть составными.

Анализируя атрибуты, можно сделать вывод, что транзитивная зависимость отсутствует, то есть отношение находится в 3НФ. Следовательно, все схемы отношений являются окончательными схемами отношений.

2.5 Проектирование физических моделей данных

В отличие от логической модели, физическая модель строится в зависимости от выбранной СУБД. Также этот вид модели позволяет напрямую создавать команды для системного каталога СУБД. Главное отличие физической от логической модели состоит в том, что в первой уделяется внимание на типы объектов. Так, атрибуты могут иметь разные форматы данных: числовой или строчный. Помимо формата данных, физическая модель допускает использование физических объектов СУБД: например, процедуры.

Различают два уровня физической модели:

- трансформационная модель (Transformation Model),
- модель СУБД (DBMS Model).

Трансформационная модель содержит информацию для реализации отдельного проекта, который может быть частью общей ИС и описывать подмножество предметной области. Трансформационная модель позволяет

проектировщикам и администраторам БД лучше представлять, какие объекты БД хранятся в словаре данных, и проверить, насколько физическая модель данных удовлетворяет требованиям к ИС.

Модель СУБД является точным описанием того, как выглядит будущая БД в системном каталоге СУБД. Именно эта модель физического уровня будет использоваться в данной работе, потому что она имеет практическое влияние.

Самым главным объектом в БД является таблица. Если интерпретировать физическую модель через логическую, то получим такой вывод, что таблицы являются сущностями, столбцы - атрибутами, а связи образуются через внешние ключи. Таблица содержит столбцы, именно поэтому им стоит уделить достойное внимание.

Первое свойство столбца, про которое стоит сказать, может ли оно принимать нулевые значения, в случае если может, нужно писать этому столбцу предложение «NULL», иначе «NOT NULL». Причем первичные ключи не могут быть нулевыми значениями, также не рекомендуется использовать «NULL» для внешних ключей. Второе свойство столбца есть сам формат данных, для целочисленных (причем 0 в старшем бите, не являющемся значимым) стоит использовать «INTEGER», для строковых значений можно использовать «VARCHAR(n)», а для использования даты применяется «DATE», а если нужна дата с временем, то «DATETIME» [5].

2.6 Готовые решения

Проанализировав программные средства для реализации баз данных, рассмотрев все основные принципы построения таблиц, создания первичных и внешних ключей и нормализацию данных, приведем ER-диаграммы основных структур [6].

Логично предположить, что для начала нам нужно будет создать аккаунт и пройти аутентификацию в приложении, пройдя верификацию по логину и паролю. Здесь возможно два исхода:

- верификация прошла успешно,
- отказано в доступе.

Как уже говорилось ранее, у нас возможны некоторые роли:

- администратор,
- пользователь.

В зависимости от этого у нас будут формироваться данные входа для

разных ролей. Соответственно для этого нам нужно сделать структуру таблиц под эти требования. На рисунке 1 показана ER-диаграмма организации структуры хранения данных для хранения данных о пользователе с определенными правами, ролями и группами.

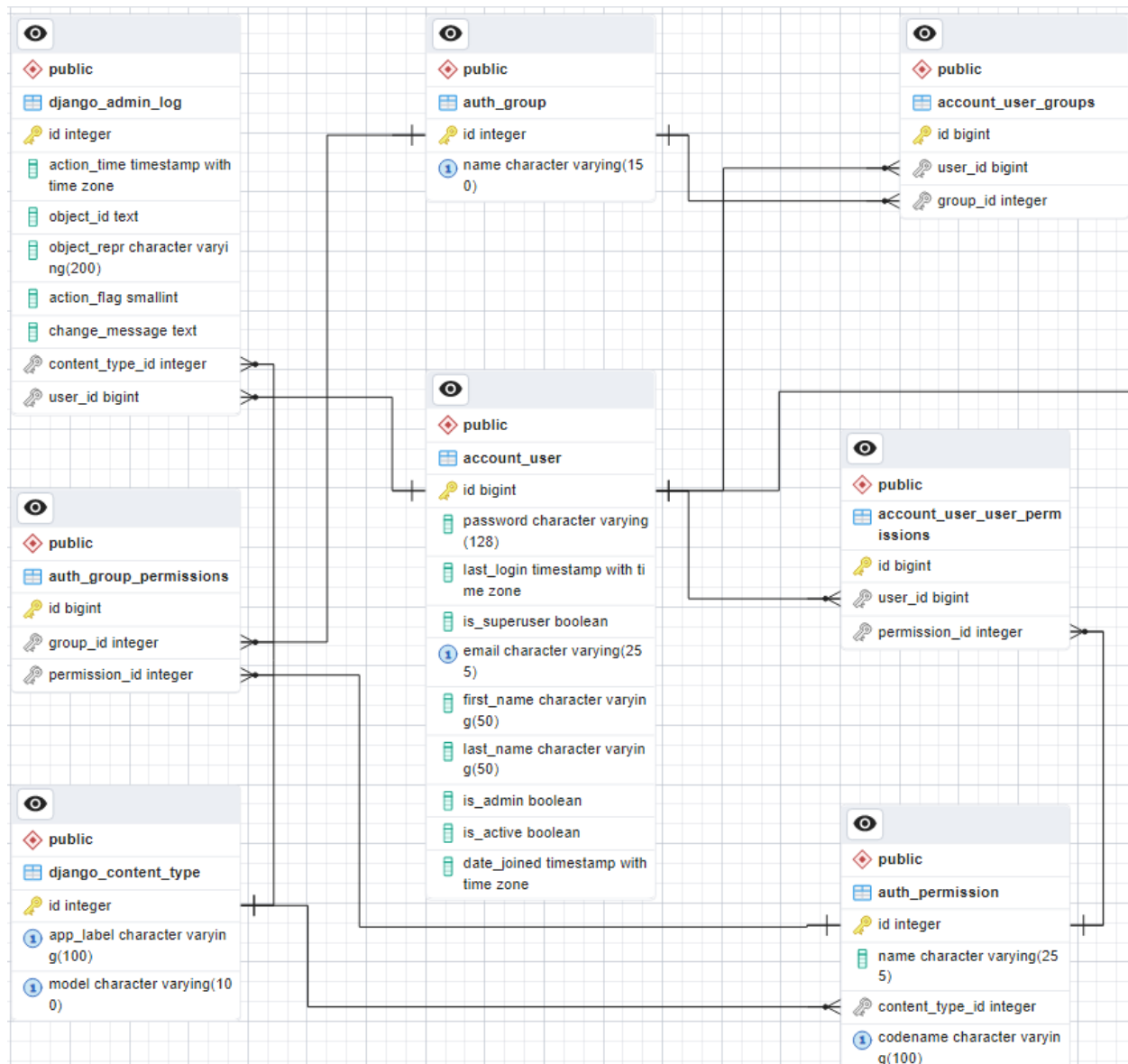


Рисунок 1 – Данные с определенными правами, ролями и группами

Как можно заметить, на рисунке 1 видна еще одна связь, которая уходит за пределы изображения. Это не случайно, сейчас объясним этот момент.

На рисунке 2 показана показана ER-диаграмма организации структуры хранения данных для связи нашей БД с взаимодействиями на стороне back-end'а. Есть таблица, которая заполняется при первичной регистрации пользователя и служит для информации на стороне back-end'а. Кроме того,

есть таблица пользователя, которая нужна только в БД и служит непосредственно для связи пользователя с медицинскими картами пациентов, которые заполняет медицинский работник при выезде.

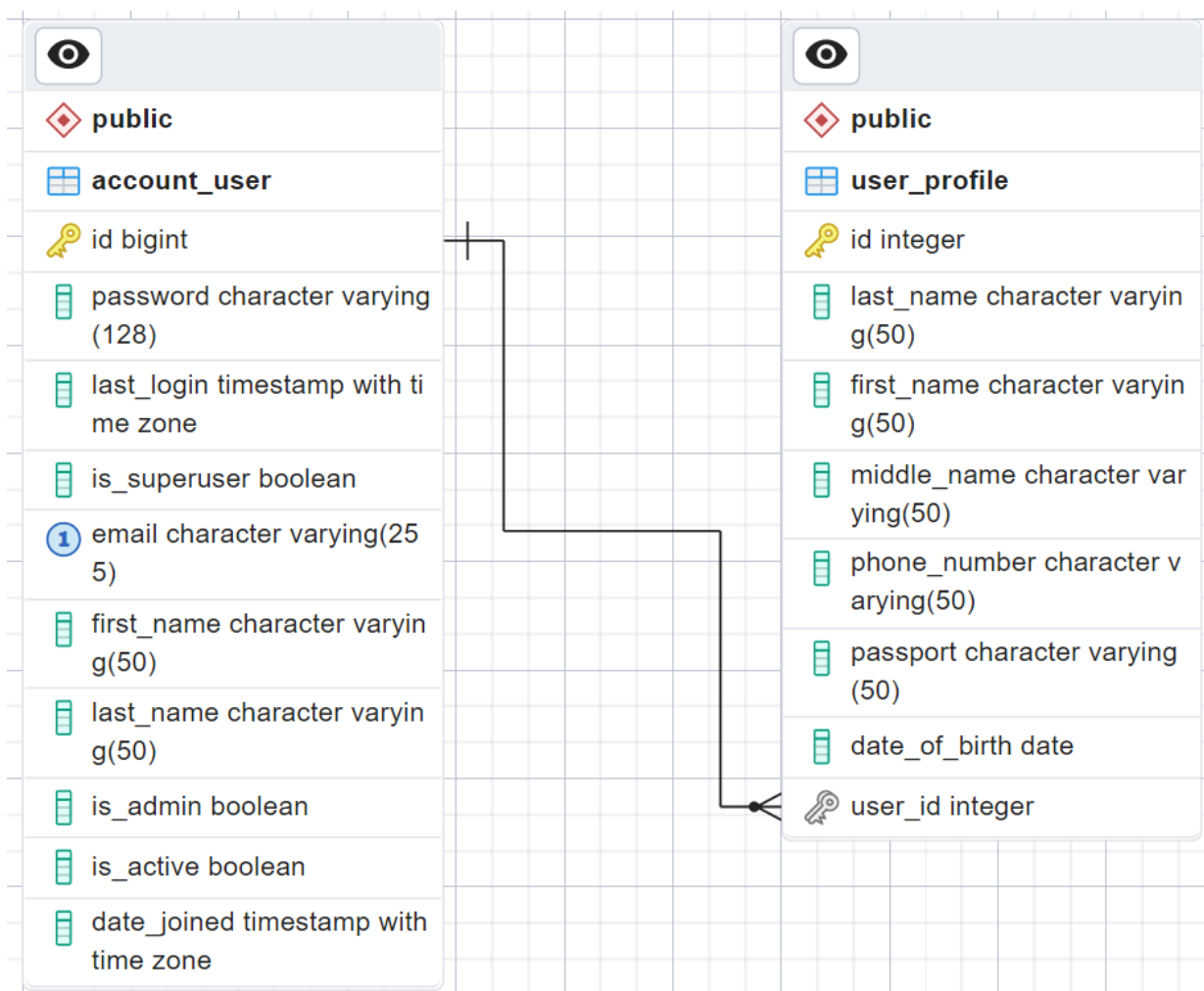


Рисунок 2 – Связи взаимодействия

На рисунке 2 была схематично показана связь двух таблиц. Теперь приведем описание таблицы **user_profile** и объясним ее назначение в нашей БД. На рисунке 3 показана показана ER-диаграмма организации структуры хранения данных для пользователей нашего приложения.

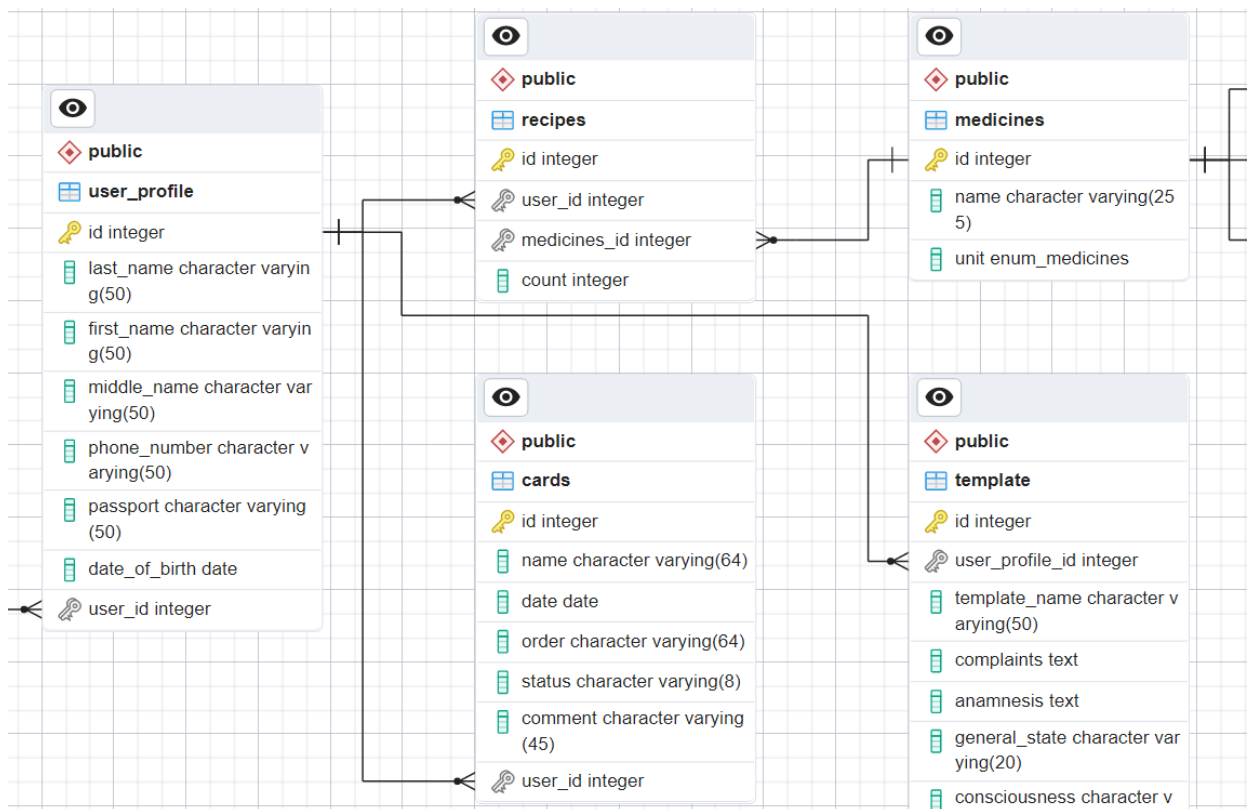


Рисунок 3 – Пользователь

Как можно заметить, в таблицк medicines есть еще три связи, которые не попали на данный рисунок. Эти связи будут продемонстрированы на рисунке 7.

На рисунке 4 показана ER-диаграмма организации структуры хранения данных полного взаимодействия с момента регистрации и аутентификации пользователя в приложении до связи определенного медицинского работника с шаблоном для заполнения медицинской карты при выезде к пациенту.

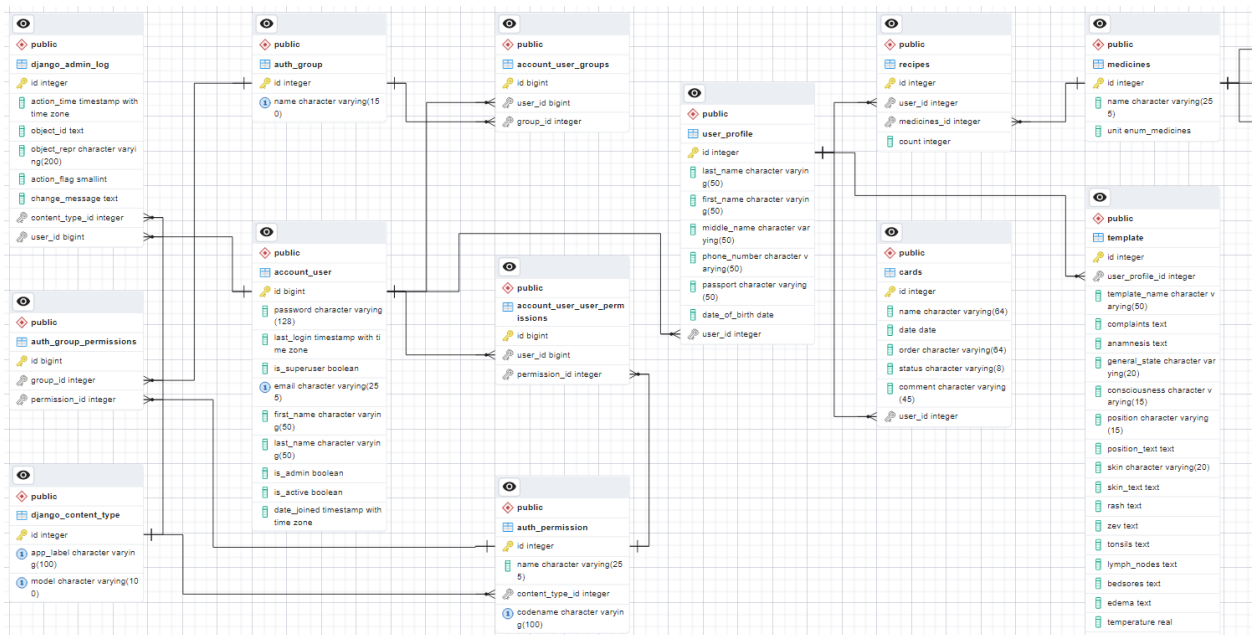


Рисунок 4 – Полная структура взаимодействия

На рисунке 5 показана ER-диаграмма организации структуры хранения данных для заболеваний. Каждое заболевание относится к определённой категории в медицине - так называемый тег. У заболеваний есть свои определённые симптомы. Кроме того, заболевание имеет свой собственные формы и относящиеся уже к этой форме симптомы.

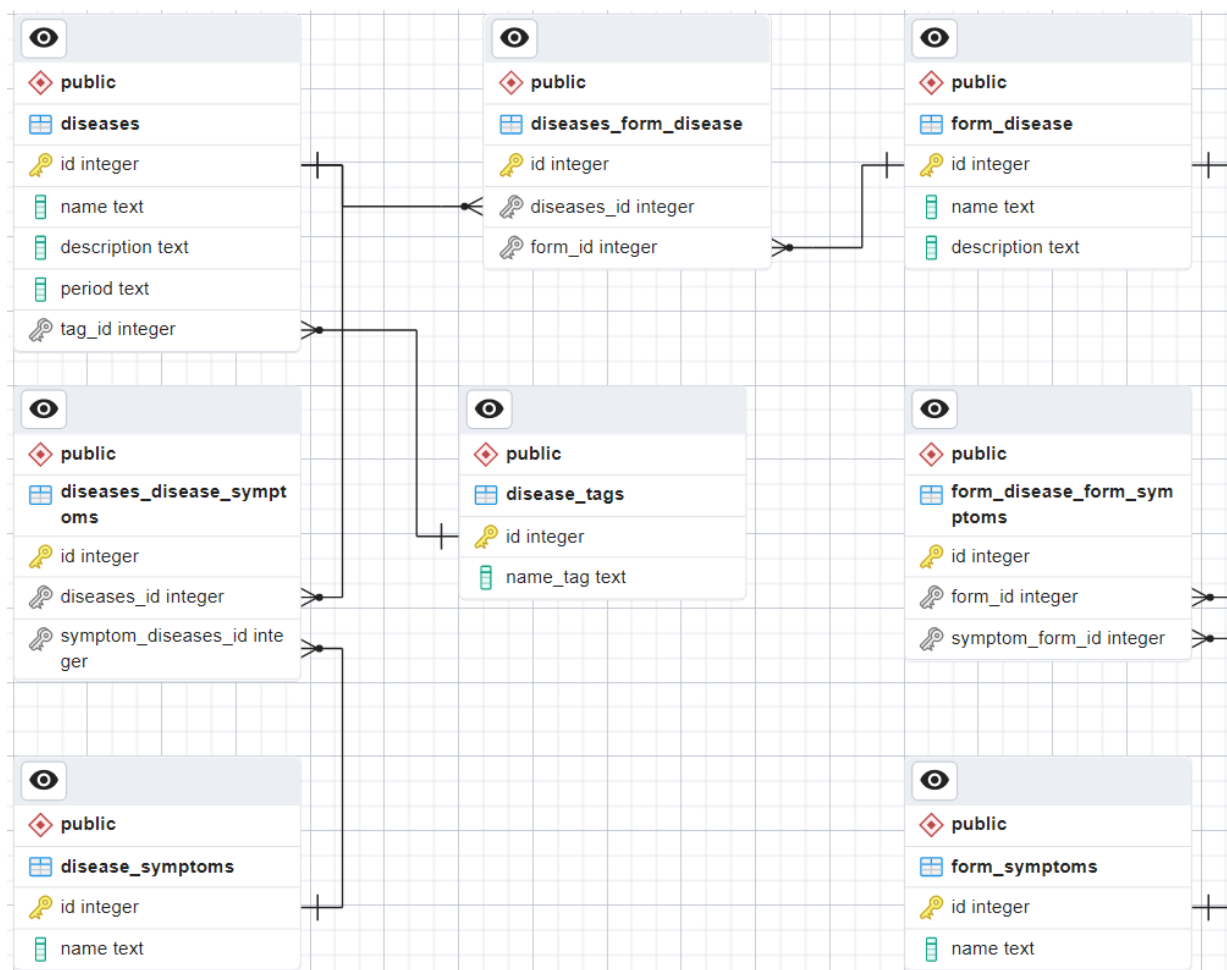


Рисунок 5 – Заболевания

На рисунке 6 показана показана ER-диаграмма организации структуры хранения данных для диагнозов. Каждый диагноз относится к определённому направлению в медицине - так называемый тег. Каждый диагноз имеет свой определенный код МКБ. У диагноза есть свой определенный перечень оказания объема необходимой медицинской помощи и тактика выполнения действий. Кроме того, диагноз имеет свой так называемые формы(поддиагнозы) и относящийся уже к ним объем необходимой медицинской помощи.

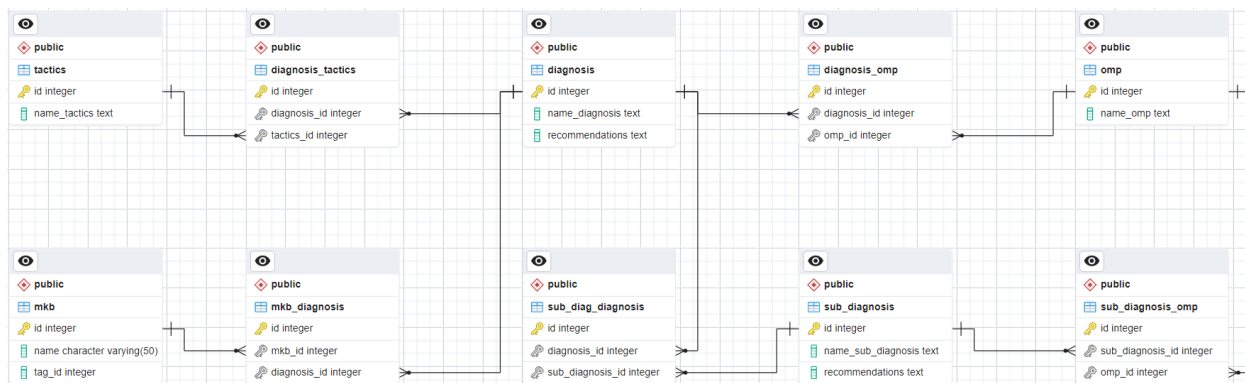


Рисунок 6 – Диагнозы

На рисунке 7 показана показана ER-диаграмма организации структуры хранения данных для медикаментов. Каждый препарат имеет определенные противопоказания, взрослую и детскую дозировки, которые, в свою очередь зависят от конкретного диагноза.

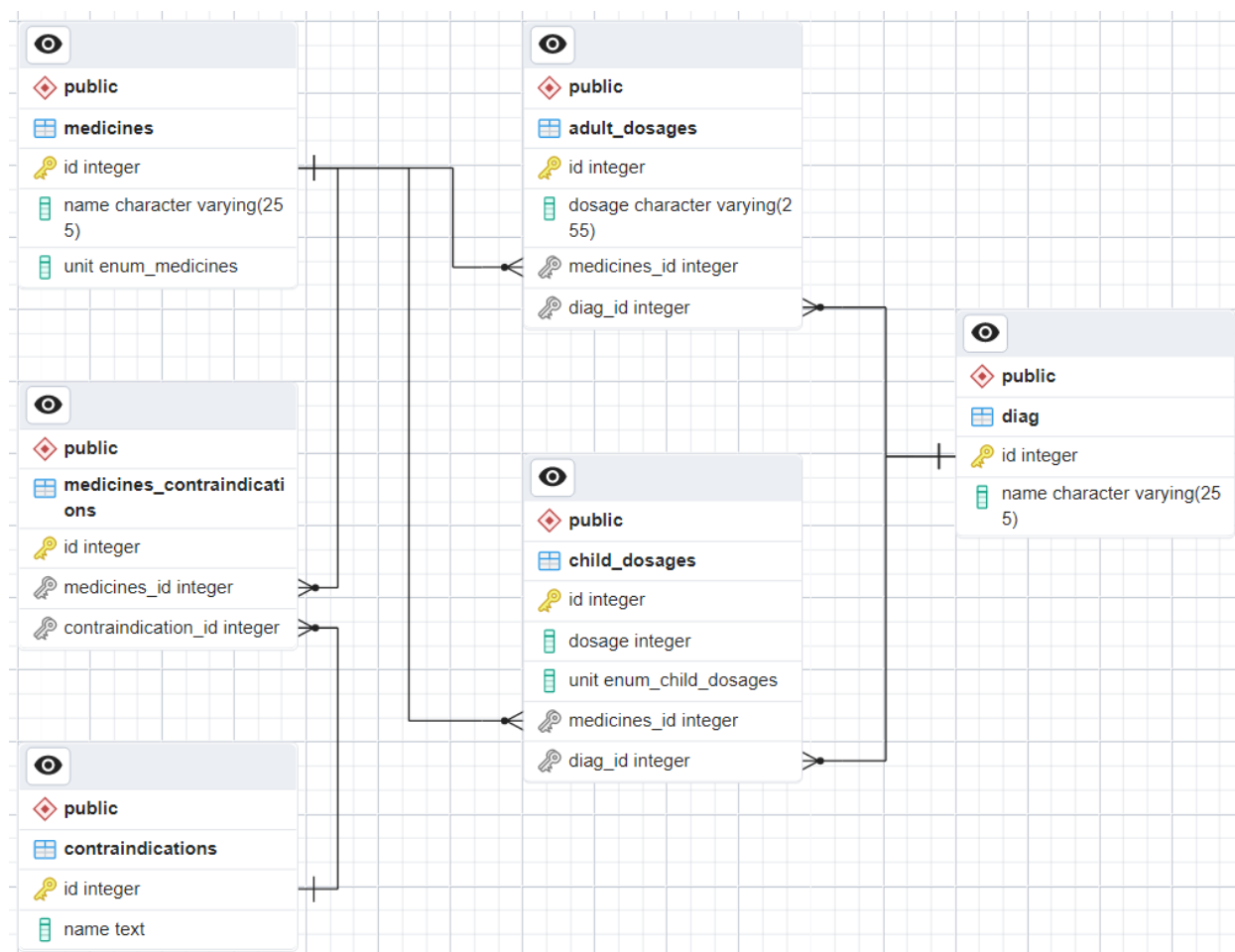


Рисунок 7 – Медикаменты

Конечно, это не все таблицы, требуемые для функционала целого приложения. Медицинскому работнику иногда требуется какая-то

незначительная вспомогательная информация и какой-то сложной организационной структуры эти таблицы не требуют. Подробное описание этих таблиц приведено в приложении А.

Таким образом, были объявлены основные таблицы БД, были рассмотрены столбцы, а именно проанализированы типы данных, а также определены ограничения для внесения конкретных значений. Этого достаточно, чтобы между таблицами сохранялась целостность данных, а вставка значений была корректной [7; 8; 9].

Важно понимать, что перед внедрением базы данных необходимо провести тестирование и проверку ее функциональности, производительности и безопасности. Это поможет выявить потенциальные проблемы и ошибки, которые могут повлиять на работу системы. Регулярное тестирование после внедрения также является важной частью обслуживания базы данных.

Выводы по разделу

В этом разделе разбирались теоретические сведения, связанные с созданием БД. С помощью анализа организации и метода моделирования мы смогли решить задачи, связанные с определением предметной области ОНПМ и разработкой логических и физических моделей БД. Также была выполнена практическая задача, связанная с формированием системного каталога структуры БД.

Важно отметить, что разработка базы данных не является заключительным шагом. Необходимо предусмотреть систему поддержки и обслуживания, которая будет обеспечивать регулярные обновления базы данных, исправление ошибок, мониторинг производительности и резервное копирование данных. Регулярное техническое обслуживание поможет сохранить работоспособность и безопасность базы данных на протяжении всего ее существования.

3 ЗАЩИТА БАЗЫ ДАННЫХ ОТ НЕСАНКЦИОНИРОВАННОГО ДОСТУПА

3.1 Конфиденциальность персональных данных

При разработке базы данных важно учитывать требования безопасности и конфиденциальности медицинской информации. Для этого могут быть приняты следующие меры:

- шифрование данных: чувствительная медицинская информация, такая как персональные данные пациентов и медицинская история, должна быть зашифрована при хранении и передаче. Использование сильных шифровальных алгоритмов поможет защитить данные от несанкционированного доступа,
- многофакторная аутентификация: для обеспечения безопасности доступа к базе данных, особенно для медицинского персонала, следует реализовать многофакторную аутентификацию. Это требует предоставления нескольких форм идентификации, таких как пароль, биометрические данные или специальные токены, для подтверждения легитимности пользователя,
- управление доступом: реализация гибкой системы управления доступом поможет ограничить доступ к конфиденциальной информации только уполномоченному персоналу. Ролевая модель доступа может быть использована для определения прав доступа на основе ролей и ответственностей сотрудников,
- аудит и мониторинг: важно вести аудит и мониторинг доступа к базе данных для обнаружения и предотвращения несанкционированного доступа или неправомерной активности. Журналы доступа и системы мониторинга позволят выявить подозрительную активность и принять соответствующие меры,
- соответствие нормам и законодательству: разработка базы данных должна соответствовать применимым нормам и законодательству в области защиты персональных данных. Это включает в себя соблюдение правил по сбору, хранению, использованию и передаче медицинской информации,
- восстановление после катастрофы и резервное копирование:

необходимо реализовать надежный план восстановления после катастрофы и резервного копирования, чтобы обеспечить доступность и целостность медицинской базы данных. Регулярные резервные копии должны выполняться, чтобы защитить от потери данных в случае сбоев системы, стихийных бедствий или кибератак. Также важно периодически тестировать процесс восстановления, чтобы убедиться в возможности точного и эффективного восстановления данных.

Анализируя вышеперечисленные пункты, можно сделать вывод, что обеспечение конфиденциальности хранения персональных данных является одним из ключевых вопросов, связанных с использованием цифровых медицинских карт и баз данных [10]..

Для обеспечения конфиденциальности данных необходимо проводить работу по защите информации и соблюдению законодательных требований. Организация должна предпринимать меры для защиты персональных данных пациентов, такие как использование паролей и шифрования данных. Кроме того, необходимо обучать медицинских работников основам информационной безопасности и контролировать доступ к информации.

Конфиденциальность хранения персональных данных пациентов - это важный аспект использования цифровых медицинских карт и организации должны принимать соответствующие меры для обеспечения безопасности информации и соблюдения законодательных требований.

PostgreSQL поддерживает множество функций для обеспечения безопасности данных, включая конфиденциальность хранения персональных данных.

Одним из ключевых механизмов для обеспечения конфиденциальности данных в PostgreSQL является использование различных методов шифрования, включая шифрование данных в пути и в покое, а также использование SSL для защиты соединений.

PostgreSQL также обеспечивает механизмы авторизации и аутентификации, которые позволяют контролировать доступ к базе данных и ее объектам, таким как таблицы и представления. Например, администраторы могут назначать различные роли и права доступа к объектам базы данных, что позволяет управлять доступом к конфиденциальным данным.

PostgreSQL также поддерживает аудиторскую функциональность,

которая позволяет записывать действия пользователей в базе данных, такие как входы и выходы из системы, выполнение запросов и изменение данных. Это позволяет отслеживать и анализировать действия пользователей, чтобы обеспечить безопасность данных и защитить их от несанкционированного доступа.

В целом, PostgreSQL - это надежная и безопасная реляционная база данных, которая обеспечивает множество механизмов для защиты конфиденциальности данных, включая персональные данные пациентов.

3.2 Использование регистрационных имен, имен пользователей, ролей и разрешений

PostgreSQL обладает базовым механизмом защиты, включающим в себя идентификацию, аутентификацию и авторизацию. Идентификация – проверка на то, существует ли данный субъект, желающий воспользоваться данным ресурсом. Следующим шагом после идентификации следует аутентификация – проверка подлинности субъекта, в основном для этого используется паролевый метод. Конечным шагом является авторизация – предоставление необходимых прав субъекту или отказ в доступе к нужным ресурсам. Для реализации описанного механизма применяются регистрационные имена, которые инициализируются следующим образом: «CREATE ROLE login_name WITH LOGIN PASSWORD 'login_password' CREATEDB DEFAULT_DATABASE db_name;». Стоит остановиться на параметрах «CHECK_EXPIRATION» и «CHECK_POLICY». «CHECK_EXPIRATION» отвечает за то, есть ли у пароля срок истечения, а «CHECK_POLICY» определяет, должна ли использоваться паролевая политика. По умолчанию, проверка срока действия пароля и политики паролей включена в PostgreSQL, поэтому опции CHECK_EXPIRATION и CHECK_POLICY не нужно указывать.

После создания регистрационных имен для них требуется создать имена пользователей. Это требуется из-за того, что регистрационные имена нужны лишь для получения доступа на сервер, а для взаимодействия внутри сервера нужны уже вторые. Чтобы создать имя пользователя и привязать его к регистрационному имени можно применить следующий шаблон: «GRANT login_name TO user_name;».

Определив регистрационные имена и их пользователей, остается лишь

предоставить им необходимые разрешения, позволяющие взаимодействовать с БД. Можно отдельно каждому пользователю прописывать разрешения, но в данном случае это нецелесообразно – потому что у всех медиков должны быть одинаковые права. Но стоит еще упомянуть такое понятие как роль. Роль – объект БД, позволяющий объединять пользователей в единую группу с целью эффективного администрирования. Для создания ролей используется следующий шаблон: «CREATE ROLE role_name;». Для добавления пользователя в данную роль применяют следующий синтаксис: «GRANT role_name TO user_name;». Также в PostgreSQL есть возможность создавать ролевые комбинации на уровне базы данных, чтобы сделать связь логин-роль доступной во всех базах данных. Для этого нужно использовать команду CREATE ROLE с опцией LOGIN и указать логин, который должен быть связан с ролью: «CREATE ROLE user_name LOGIN LOGIN_NAME login_name;».

Создав роли для пользователей, нужно определить некоторые права, позволяющие манипулировать БД. Для этого в PostgreSQL предусмотрена возможность выдачи прав как отдельным пользователям, так и ролям. Для выдачи права используется конструкция: «GRANT action_name ON object_name TO user_name/role_name;», в эту конструкцию можно добавить свойство «WITH GRANT OPTION», позволяющее предоставлять права другим пользователям или ролям. Здесь action_name - право доступа, которое нужно предоставить (например, SELECT, INSERT, UPDATE или DELETE), object_name - имя объекта базы данных, к которому нужно предоставить доступ, user_name или role_name - имя пользователя или роли, которым нужно предоставить доступ. Для запрета на определенное право используется конструкция: «REVOKE action_name ON object_name FROM user_name/role_name;», при запрете права также может использоваться «CASCADE», это свойство наоборот запрещает некоторые права тем субъектам, которым оно выдавалось. Помимо выдачи и запрета прав есть третья возможность – отмена прав, она имеет следующий синтаксис: «REVOKE action_name ON object_name FROM user_name/role_name [CASCADE];», свойство «CASCADE» используется для отмены прав у тех субъектов, которым данный пользователь выдал. Стоит обратить внимание, что для использования команды REVOKE в PostgreSQL необходимо иметь соответствующие привилегии администратора базы данных.

3.3 Шифрование базы данных

Вторым базовым механизмом защиты, реализующимся в SQL Server, от НСД является шифрование БД. Шифрование – это преобразование исходной информации в данные, которые теряют иной смысл, не зная секретного слова или алгоритма, преобразовывающие данные обратно в смысловую информацию. Самый популярный способ зашифровать информацию, хранящуюся в БД – это прозрачное шифрование данных. Преимущество такого вида шифрования заключается в том, что конечные пользователи могут даже и не знать, что применяется защита информации. Перед тем, как перейти к шифрованию, следует, на всякий случай, сделать резервную копию БД. Для этого следует включить pgAgent.

Затем выбираем нужную БД, вызываем контекстное меню, в нем выбираем опцию «создать резервную копию».

Создав резервную копию БД, можно приступать к процессу шифрования. Этот процесс происходит в несколько этапов:

- создание главного ключа БД,
- создание сертификата,
- создание ключа шифрования,
- запуск процесса шифрования,
- проверка состояния шифрования.

Главный ключ БД – это ключ, который применяется для шифрования других ключей, использующихся для реализации шифрования в БД. Для его создания применяется следующая конструкция: «CREATE MASTER KEY ENCRYPTION BY PASSWORD 'password';».

Следующим шагом будет создание сертификата. Сертификат – это объект безопасности, имеющий подпись. Этот объект хранит в себе ключи шифрования. Для создания данного объекта применяется следующая конструкция: «CREATE CERTIFICATE name_db_and_cert WITH SUBJECT 'description';».

Последним подготовительным этапом перед началом шифрования БД является создание ключа шифрования. Именно этим ключом будет происходить шифрование БД, а сам ключ будет находиться в сертификате, созданном заранее. Для его создания требуется применять следующий шаблон: «CREATE DATABASE ENCRYPTION KEY WITH ALGORITHM =

AES_128 ENCRYPTION BY SERVER CERTIFICATE cert_name;».

Следующий шаг является необязательным, но его желательно выполнять. При создании главного ключа, сертификата и ключа шифрования рекомендуется создание резервных копий этих объектов безопасности. Для создания резервной копии ключа шифрования можно использовать команду `pg_dump`, которая создаст резервную копию всей базы данных, включая ключи шифрования: «`pg_dump dbname > backup_file_name;`».

Все объекты безопасности проинициализированы, также для критических объектов были созданы резервные копии. Теперь можно приступать к шифрованию БД. Для начала данного процесса используется следующий пакет команд: «`UPDATE table_name SET column_name = pgp_sym_encrypt(column_name, 'key_password');`».

Осталось лишь проверить, работает ли ключ шифрования для БД. Для этого используются системные БД «`pg_database_encryption`». Нас интересует столбец `is_encrypted`, если шифрование завершилось успешно, то столбец принимает значение 1.

3.4 Порты

Порт – это некоторое идентифицирующее число от 1 до 65535, позволяющее протоколам взаимодействовать друг с другом на транспортном уровне модели OSI.

На самом деле, у PostgreSQL сервера по умолчанию установлен только один порт, и это TCP порт.

PostgreSQL использует TCP/IP для обмена данными между клиентами и сервером. По умолчанию, порт TCP для PostgreSQL установлен на 5432. TCP 5432: database engine, используется для подключения к СУБД. Этот порт можно изменить при настройке сервера [11]..

Несмотря на то, что UDP не используется для коммуникации между клиентами и сервером PostgreSQL, есть несколько расширений, таких как `pgpool-II`, которые могут использовать UDP для взаимодействия между серверами PostgreSQL.

Таким образом, можно сказать, что PostgreSQL сервер по умолчанию использует только один порт TCP для обмена данными между клиентами и сервером.

Хорошей практикой является смена портов на другие, чтобы при случае

атаки порты пришлось перебирать методом грубого подбора, а не использовать установленные по умолчанию. По данным IANA, порты 834-846 являются свободными, поэтому мы их можем использовать.

Для изменения порта, используемого PostgreSQL, можно внести изменения в конфигурационный файл `postgresql.conf`, который находится в каталоге данных PostgreSQL.

В файле `postgresql.conf` необходимо изменить параметр `port` на желаемое значение порта. Например, если вы хотите использовать порт 840 вместо стандартного порта 5432, необходимо изменить строку: `port = 5432` на `port = 840`. После этого нужно перезапустить PostgreSQL, чтобы изменения вступили в силу.

Также можно использовать параметр командной строки `-p` при запуске сервера PostgreSQL для указания порта. Например, чтобы запустить PostgreSQL на порту 840, необходимо использовать команду: `postgres -p 840`, но в этом случае необходимо убедиться, что порт 840 свободен и не используется другим приложением на сервере.

Помимо смены порта, слушающий входящие соединения, можно отключить PostgreSQL Server. PostgreSQL Server – это одна из служб PostgreSQL, отвечающая за прослушивание запросов. Если эту службу отменить, то для установления соединения придется явно указывать номер порта, что обеспечивает дополнительную безопасность. Для остановки этой службы требуется просто ввести `sudo systemctl stop postgresql` в командной строке.

3.5 Брандмауэр

Брандмауэр – это программный межсетевой экран, имеющийся в ОС Windows. В свою очередь межсетевой экран – это элемент локальной сети, осуществляющий мониторинг активности в данной сети, а именно фильтрацию сетевого трафика по заранее описанным правилам. PostgreSQL Server функционирует на ОС Windows, поэтому рекомендуется использовать Брандмауэр вместе с программно-аппаратным межсетевым экраном. Такое сочетание позволяет повысить уровень защищенности за счёт использования дополнительных механизмов защиты информации. Также Брандмауэр обеспечивает не только защищенность от НСД, но и от поступающего вредоносного трафика в принципе.

Как было описано выше, межсетевой экран анализирует сетевой трафик и отклоняет его, если он считается подозрительным. Выше мы поменяли стандартный TCP/UDP порт 5432 на 840, теперь нужно объявить правило в межсетевом экране, что публичная сеть не может ссылаться на этот порт. Для этого заходим в Монитор Брандмауэра Защитника Windows в режиме повышенной безопасности и создадим новое правило. Далее определяем, что правило создается для порта, затем вписываем наш нужный порт, после этого разрешаем подключение, после этого исключаем публичный доступ, и в конце сохраняем наше правило.

Выводы по разделу

В этом разделе разбирались аспекты, связанные с безопасностью. При помощи прозрачного шифрования файлы БД были защищены от кражи физических носителей. Создание и управление регистрационных имен позволило осуществить легитимный доступ к информации. А изменение портов и использование Брандмауэра позволило защититься на сетевом уровне. В совокупности была выполнена задача определения защитных мер БД.

4 ТЕСТИРОВАНИЕ И ОПТИМИЗАЦИЯ ПРОИЗВОДИТЕЛЬНОСТИ БАЗЫ ДАННЫХ

4.1 Методы и инструменты тестирования

Предзаключительным этапом является тестирование базы данных и ее оптимизация. На этом этапе производятся проверки корректной и эффективной работы базы данных. Кроме того, выполняются тесты производительности, определяются проблемы с ней и соответственно решение этих проблемы путем оптимизации запросов, индексов и настройки СУБД PostgreSQL.

Для тестирования базы данных можно использовать различные методы и инструменты, такие как:

- модульное тестирование: это тестирование отдельных компонентов базы данных, таких как хранимые процедуры, триггеры, функции и т.д,
- интеграционное тестирование: это тестирование совместной работы базы данных и приложения, которое использует эту базу данных,
- нагрузочное тестирование: это тестирование производительности базы данных в условиях высокой нагрузки,
- тестирование безопасности: это тестирование уязвимостей базы данных и проверка соответствия базы данных требованиям безопасности,
- ручное тестирование: это тестирование, при котором тестировщик проверяет базу данных на соответствие требованиям и наличие ошибок вручную.

4.2 Нагрузочное тестирование и оптимизация производительности

Нагрузочное тестирование БД в PostgreSQL - это процесс измерения производительности и стабильности базы данных при большой нагрузке. Это необходимо для определения максимального количества пользователей и запросов, которые могут быть обработаны базой данных, без ухудшения ее производительности.

Для проведения нагрузочного тестирования БД в PostgreSQL,

необходимо выполнить следующие шаги:

- определить сценарии использования БД: определить типы запросов, которые будут отправлены к БД, и количество пользователей, которые будут использовать приложение одновременно,
- подготовить тестовые данные: создать тестовые данные, которые будут использоваться при выполнении запросов в тестовом окружении. Важно создать реалистичные тестовые данные, чтобы результаты нагрузочного тестирования были максимально близки к реальной нагрузке,
- настроить окружение для нагрузочного тестирования: убедиться, что база данных и приложение настроены правильно, чтобы обеспечить максимальную производительность и стабильность,
- запустить нагрузочное тестирование: выполнить тестовые сценарии использования БД в тестовом окружении и измерить производительность базы данных, используя метрики, такие как скорость ответа, время выполнения запросов, количество ошибок и использование ресурсов,
- анализ результатов: проанализировать результаты нагрузочного тестирования, чтобы определить максимальную нагрузку, которую может выдержать база данных, и определить узкие места, которые могут быть оптимизированы.

При тестировании в PostgreSQL можно использовать специальные инструменты, такие как `pgbench` и `Apache JMeter`, чтобы автоматизировать процесс выполнения тестовых сценариев и сбора метрик производительности. Важно понимать, что результаты нагрузочного тестирования могут изменяться в зависимости от настроек и конфигурации сервера и базы данных, поэтому рекомендуется проводить тестирование на реальном оборудовании, которое будет использоваться в итоговой среде.

- `pgbench`: это инструмент, входящий в стандартный комплект поставки PostgreSQL. Он предназначен для проведения нагрузочного тестирования на уровне SQL-запросов,
- `pgBadger`: это инструмент для анализа логов PostgreSQL. Он позволяет анализировать логи запросов и генерировать отчеты о производительности базы данных,
- `HammerDB`: это инструмент для тестирования производительности

баз данных. Он поддерживает PostgreSQL и позволяет тестировать производительность базы данных на различных уровнях нагрузки,

- Sysbench: это инструмент, который может быть использован для тестирования производительности базы данных PostgreSQL. Он позволяет тестировать производительность базы данных на уровне SQL-запросов и многопоточности.

Важно помнить, что результаты нагрузочного тестирования БД PostgreSQL могут зависеть от конкретной конфигурации сервера и структуры данных, а также от условий использования базы данных в реальном мире. Поэтому результаты тестирования должны быть интерпретированы с осторожностью и использованы для оптимизации конкретной базы данных.

Оптимизация производительности базы данных PostgreSQL может включать в себя ряд мероприятий, чтобы ускорить выполнение SQL запросов и обеспечить более эффективную работу БД. Вот основные подходы, которые можно эффективно использовать для оптимизации БД PostgreSQL [12; 13]:

- создание правильных индексов: индексы ускоряют выполнение запросов, позволяя PostgreSQL быстро находить нужные данные. Убедитесь, что у вас есть индексы на часто используемые столбцы в запросах, а также на столбцы, используемые в условиях WHERE и JOIN и фильтрации данных, может значительно ускорить выполнение запросов,
- оптимизация запросов: оптимизация запросов, такая как использование подзапросов, объединения и группировки данных, может улучшить производительность запросов и уменьшить количество запросов, необходимых для выполнения операции,
- разделение данных на отдельные таблицы: если у нас есть большие таблицы, то не следует хранить большие объемы данных в одной таблице. Разделение этих данных на отдельные таблицы или партиционирование может помочь ускорить выполнение запросов. Это позволяет более эффективно управлять объемом данных, используемых при выполнении запросов,
- оптимизация схемы базы данных: иногда изменение структуры базы данных может привести к улучшению производительности. Например, использование более эффективных типов данных, уменьшение количества NULL значений или устранение

избыточных таблиц и связей,

- использование кэширования: если важна скорость выполнения, то можно использовать кэширование для часто запрашиваемых данных. Использование кэширования запросов может значительно улучшить производительность, позволяя избежать выполнения сложных запросов, возвращая результаты из кэша. Это особенно полезно для запросов, которые выполняются часто и имеют статические результаты,
- оптимизация запросов на запись: если есть интенсивная нагрузка на запись данных, то различные техники, такие как пакетная вставка (batch insert), использование транзакций и пакетных обновлений, могут помочь ускорить процесс записи данных,
- предварительная компиляция запросов: в PostgreSQL есть возможность предварительной компиляции запросов (prepared statements), что позволяет повторно использовать выполненные запросы и сократить накладные расходы на компиляцию. Предварительную компиляцию стоит использовать для запросов, которые выполняются многократно с различными параметрами,
- настройка конфигурации PostgreSQL: немаловажное значение имеет изучение и настройка параметров конфигурации PostgreSQL в соответствии с требованиями конкретной системы. Некоторые параметры, которые можно настроить, включают размер буферов, параллелизм выполнения запросов и максимальное количество одновременных соединений,
- обновление до последней версии PostgreSQL: немаловажное значение имеет актуальность поддерживаемого ПО, поэтому рекомендуется использовать последнюю стабильную версию PostgreSQL. Каждое новое обновление может содержать оптимизации и улучшения производительности, которые могут существенно повысить скорость выполнения запросов,
- настройка памяти и дискового пространства: правильная настройка памяти и дискового пространства может существенно повлиять на производительность PostgreSQL. Следует выделять достаточное количество памяти для работы с базой данных. Кроме того, система должна иметь достаточное дисковое пространство для хранения

данных и временных файлов,

- использование материализованных представлений: материализованные представления - это предварительно вычисленные результаты запросов, сохраняемые в виде таблиц. Они могут быть особенно полезны, если имеются сложные запросы с большими объемами данных, которые выполняются часто. Материализованные представления позволяют значительно сократить время выполнения запросов,
- настройка параллелизма: PostgreSQL поддерживает параллельное выполнение запросов, что может значительно ускорить обработку больших объемов данных. Можно настроить параметры параллелизма в зависимости от характеристик системы и требований к производительности,
- мониторинг и профилирование: регулярный мониторинг производительности базы данных может помочь идентифицировать проблемные запросы или узкие места производительности. Рекомендуется использовать инструменты мониторинга и профилирования, такие как `pg_stat_statements` и `EXPLAIN ANALYZE`, чтобы анализировать и оптимизировать запросы,
- горизонтальное масштабирование: если даже после оптимизации базы данных путем проведения вышеперечисленных манипуляций все еще возникают проблемы с производительностью, можно рассмотреть вариант горизонтального масштабирования. Распределение данных на несколько серверов может помочь справиться с высокой нагрузкой и улучшить производительность,
- оптимизация настройки сервера: настройка параметров сервера PostgreSQL, таких как размер буфера и количество параллельных запросов, может улучшить производительность базы данных.

Важно проводить тестирование на реалистичных тестовых данных и сценариях использования, чтобы получить наиболее точные результаты.

Также при проведении нагрузочного тестирования БД PostgreSQL необходимо учитывать следующие факторы:

- объем данных: чем больше объем данных в базе данных, тем больше ресурсов требуется для обработки запросов. Поэтому важно учитывать объем данных при выборе конфигурации сервера и

оптимизации производительности,

- конфигурация сервера: настройка сервера PostgreSQL имеет большое значение для производительности базы данных. Например, увеличение размера буфера может ускорить выполнение запросов, но при этом может потребовать больше памяти,
- структура данных: структура таблиц и связей между ними может влиять на производительность базы данных. Например, использование ненормализованных таблиц может привести к медленной работе базы данных,
- распределение запросов: распределение запросов между разными серверами или узлами может увеличить производительность и снизить нагрузку на базу данных,
- утилизация ресурсов: нагрузочное тестирование может потребовать большое количество ресурсов, таких как память, процессорное время и дисковое пространство. Поэтому необходимо убедиться, что ресурсы доступны и достаточны для проведения тестирования.

Тестирование БД PostgreSQL является важным этапом при разработке и оптимизации базы данных. Оно позволяет выявить узкие места в работе базы данных и произвести необходимую оптимизацию, чтобы обеспечить максимальную производительность и стабильность приложения.

Стоит отметить, что после внедрения базы данных необходимо проводить анализ ее использования и результата. Это позволит выявить потенциальные улучшения и оптимизации процессов работы с медицинской информацией. На основе анализа можно внести изменения в структуру базы данных, внедрить новые функциональные возможности или улучшить существующие.

Выводы по разделу

Тестирование базы данных является критически важным этапом в разработке и поддержке приложений, использующих базы данных. Тестирование базы данных позволяет убедиться в корректности и надежности ее работы, а также обеспечить защиту от ошибок и нарушений безопасности.

Использование различных методов и инструментов позволяет выявить проблемы в базе данных еще на ранних стадиях и устранить их до того, как они приведут к серьезным проблемам в работе приложения. Кроме того, тестирование базы данных помогает повысить уровень доверия пользователей к приложению и защитить данные от нарушений и утечек. В целом,

проведение тестирования базы данных является критически важным для обеспечения надежности, безопасности и производительности приложений.

ЗАКЛЮЧЕНИЕ

В выпускной квалификационной работе бакалавра были рассмотрены особенности внедрения цифровых технологий в работу отделения неотложной медицинской помощи, а также разработана система электронных медицинских карт на базе PostgreSQL, которая позволяет улучшить качество и эффективность работы скорой медицинской помощи.

В работе использовался PostgreSQL для хранения всех возможных данных пациентов. Это позволило создать систему, которая может автоматизировать процесс сбора и обработки данных, а также обеспечивать быстрый и безопасный доступ к истории болезней и личным данным пациентов.

Кроме того, был проведен анализ системы медицинских карт и баз данных на языке PostgreSQL, а также их применение в работе скорой медицинской помощи. Также были рассмотрены вопросы обеспечения безопасности хранения медицинских данных и конфиденциальности пациентов при использовании электронных медицинских карт.

Помимо БД, имеющей удобную структуру, была проведена работа по её защите. Защита строилась как внутренними механизмами и функционалом PostgreSQL Server, так и посредством использования встроенной опциональности ОС.

БД сгенерирована и защищена, поэтому её можно вводить в эксплуатацию. По желанию в неё можно добавить дополнительный функционал, состоящий из: представлений, индексов, функций и процедур. Все эти объекты БД может настроить администратор или лицо, обслуживающее её, по желанию сотрудников, работающих с ней. Основной функционал и защита были выполнены исходя из целей, а именно:

- изучения особенностей предметной области,
- проектирования логических и физических моделей данных,
- проектирования БД при помощи СУБД,
- определения и настройки защитных мер.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. *Joshua D. Drake J. C. W.* Practical PostgreSQL. — New York, USA : O'Reilly Media, Inc., 2002. — 640 с.
2. *Regina O. Obe L. S. H.* PostgreSQL: Up and Running. — New York, USA : O'Reilly Media, Inc., 2014. — 232 с.
3. *Ibrar Ahmed Gregory Smith E. P.* PostgreSQL High Performance. — Birmingham : Packt Publishing Ltd, 2018. — 508 с.
4. *studfile.* Концепция баз данных. — 2022. — URL: <https://studfile.net/preview/1504046/> (дата обращения 03.03.2022).
5. *SQLZOO.* SQLZOO. — 2019. — URL: https://sqlzoo.net/wiki/SQL_Tutorial (дата обращения 09.03.2022).
6. *Советов Б.Я. Цехановский В.В. Ч. В.* Базы данных. Теория и практика. — Москва : Юрайт, 2014. — 464 с.
7. *PostgreSQLn.* PostgreSQL: Documentation: 14: PostgreSQL 14.8 Documentation. — 2022. — URL: <https://www.postgresql.org/docs/14/index.html> (дата обращения 20.02.2023).
8. *Профессиональный П.* Документация к Postgres Pro Standard 14. — 2016. — URL: <https://postgrespro.ru/docs/postgrespro/14/index> (дата обращения 24.02.2022).
9. *Tutorial P.* PostgreSQL Tutorial – Comprehensive Postgresql Tutorial. — 2022. — URL: <https://www.postgresqltutorial.com/>.
10. *Stedihabr.* Обеспечение безопасности базы данных PostgreSQL. — 2021. — URL: <https://habr.com/ru/articles/550882/> (дата обращения 11.04.2022).
11. *8host.* Защита PostgreSQL от автоматизированных хакерских атак. — 2017. — URL: <https://www.8host.com/blog/zashhita-postgresql-ot-avtomatizirovannykh-hakerskix-atak/> (дата обращения 20.03.2023).

12. *LikaChernigo*. Как ускорить работу PostgreSQL с помощью конфигурации базы и оптимизации запросов. — 2022. — URL: <https://habr.com/ru/companies/southbridge/articles/684826/> (дата обращения 01.05.2023).

13. *Овсянкин А.* Разбор оптимизаций запросов PostgreSQL на живых примерах. — 2019. — URL: <https://infostart.ru/1c/articles/1196217/> (дата обращения 30.04.2023).