СОДЕРЖАНИЕ

ОПРЕДЕЛЕНИЯ

В настоящей расчетно-пояснительной записке применяют следующие термины с соответствующими определениями.

Дифференциальная приватность — концепция, позволяющая анализировать наборы данных, основанных на личной информации, при этом защищая конфиденциальность отдельных лиц [1]

Параметр приватности — неотрицательное значение, которое определяет уровень приватности [2]

Шумоподобные механизмы — алгоритмы, которые генерируют шумы и добавляют их к данным

База данных — набор данных, относящихся к определённой области, структурированных по определённым правилам, обеспечивающим общие принципы описания, хранения и обработки данных, независимый от прикладных программ [3]

Фреймворк — это структура, представляющая собой набор готовых компонентов, библиотек, абстракций и правил разработки, предназначенных для облегчения процесса создания программного обеспечения [**framework**]

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

В настоящей расчетно-пояснительной записке применяют следующие сокращения и обозначения.

ДП — дифференциальная приватность

БД — база данных

ВВЕДЕНИЕ

В современном мире цифровые данные стали важной частью нашей повседневной жизни и вопросы конфиденциальности этих данных становятся все более актуальными. В контексте *bigdata* конфиденциальность может быть обеспечена различными способами, среди которых особое внимание уделяется дифференциальной приватности. Эта область обладает сильными математическими гарантиями и перспективами для дальнейшего развития.

Цель работы — провести обзор существующего программного обеспечения с возможностью дифференциальной приватности для баз данных. Для достижения этой цели требуется решить следующие задачи:

— провести анализ предметно области и обзор существующих решений;

— установить критерии сравнения существующих решений

— классифицировать существующие решения.

1 Аналитический раздел

ДП обеспечивает защиту личной информации, гарантируя, что результаты анализа не позволят точно определить информацию о конкретных пользователях [1].

В рамках ДП данные обрабатываются с добавлением шума, что минимизирует влияние индивидуальных записей на результаты статистических запросов [1].

Таким образом, даже если в набор данных добавляется или исключается информация о конкретном индивиде, это не влияет на общие выводы о группе [1].

Представим БД, содержащую медицинские записи 100 человек, из которых 20 страдают диабетом. Злоумышленник хочет узнать, есть ли у человека X диабет, и уже узнал, что 19 из остальных 99 человек в выборке являются диабетиками. Запрашивая базу данных и получая информацию о том, что в выборке 20 диабетиков, злоумышленник делает вывод, что человек X также страдает диабетом. Этот процесс извлечения дедуктивной информации называется дифференцированной атакой [2].

Теперь представим, что доступ к данным, хранящихся в БД, предоставляется через систему запросов с дифференциальной приватностью. Когда злоумышленник делает запрос, система возвращает истинное значение с некоторым случайным шумом. Даже если злоумышленник знает все о других 99 людях, то он не может быть уверен, есть ли у человека X диабет. Конфиденциальность человека X защищена, поскольку его риск нарушения конфиденциальности остается примерно одинаковым независимо от того, входит ли он в набор данных или нет. Это справедливо для любого отдельного человека и любого набора данных, независимо от того, насколько необычными или отличительными могут быть данные конкретного индивидуума.

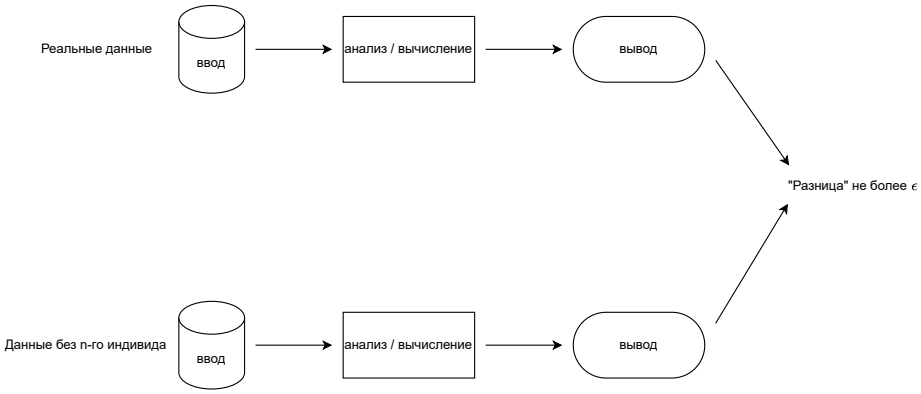

Рисунок 1.1 – Принцип ДП

1.1 Типы приватности

1.1.1 ϵ – дифференцированная приватность

Стандартная ϵ – дифференцированная приватность предоставляет строгую гарантию, что результаты анализа данных не изменятся сильно, даже если одна запись будет добавлена или удалена из базы данных. [4] Механизм M обеспечивает ϵ – дифференциальную приватность, если для любых двух соседних наборов данных D и D' (различающихся на одну запись):

$$Pr[M(D)] \leqslant \exp(\epsilon) \times Pr[M(D')] \tag{1.1}$$

, где ϵ — параметр приватности

1.1.2 (ϵ, δ) – дифференцированная приватность

Модель (ϵ, δ) – дифференцированной приватности вводит дополнительный параметр δ, который позволяет немного ослабить эти строгие гарантии. Взамен δ предоставляет вероятность того, что гарантия приватности будет нарушена в некоторых случаях. Это даёт большую гибкость в балансе между приватностью и точностью результатов, но с оговоркой, что такая утечка будет происходить с очень маленькой вероятность [4] Механизм M обеспечивает (ϵ, δ) – дифференциальную приватность, если для любых двух соседних наборов данных D и D' (различающихся на одну запись):

$$Pr[M(D)] \leqslant \exp(\epsilon) \times Pr[M(D')] + \delta \tag{1.2}$$

, где ϵ — параметр приватности

1.2 Типы механизмов

Параметр чувствительности — это максимальное изменение результата механизма при изменении одного элемента в базе данных. Чувствительность механизма ΔM — это максимум разности значений механизма M при изменении одного элемента и высчитывается по формуле

$$\Delta M = \max(|M(D) - M(D')|) \tag{1.3}$$

, где D и D' — две соседние базы данных, которые отличаются на один элемент.

1.2.1 Механизм Лапласа

Лапласов механизм добавляет к результату запроса шум, сгенерированный из распределения Лапласа [4]. Результат Лапласовского механизма выяисляется по формуле 1.4:

$$M'(D) = M(D) + \mathrm{Lap}(0, \frac{\Delta M}{\epsilon}) \tag{1.4}$$

где ΔM — параметр чувствительности, ϵ — параметр приватности

1.2.2 Механизм Гаусса

Гауссов механизм похож на Лапласов, но шум добавляется на основе нормального (Гауссова) распределения. К результатам запроса добавляется шум по формуле 1.5:

$$M'(D) = M(D) + \mathcal{N}(0, \sigma^2) \tag{1.5}$$

, где $\mathcal{N}(0, \sigma^2)$ — нормальное распределение.

Параметр σ вычисляется по формуле 1.6.

$$\sigma = \frac{\Delta M \sqrt{2 \ln(1.25/\delta)}}{\epsilon} \tag{1.6}$$

1.2.3 Экспоненциальный механизм

Экспоненциальный механизм работает путем добавления случайности в процесс выбора элементов, где вероятность выбора каждого элемента зависит от его "качества"(оценки полезности) и параметра приватности.

Для двух соседних наборов данных я D_1 и D_2 экспоненциальный механизм M обеспечивает ϵ – дифференциальную приватность, если вероятность выбора результата r зависит от его полезности $u(r, D)$, и вычисляется по формуле 1.7

$$P[M(D)] \propto \exp(\epsilon u(r, D)) \tag{1.7}$$

1.3 Типы реализаций

Реализации ДП при работе базами данных классифицируют по 3 типам:

— многоуровневая реализация;

— монолитная реализация;

— реализация на уровне СУБД;

1.3.1 Многоуровневая реализация

Данная реализация представлена на рисунке 1.2 и предполагает использование нескольких фреймворков на стороне клиента для реализации ДП.

Наиболее частым решением данной в реализации является использование фреймворка БД для получения данных и фреймворка ДП для обработки данных [4]

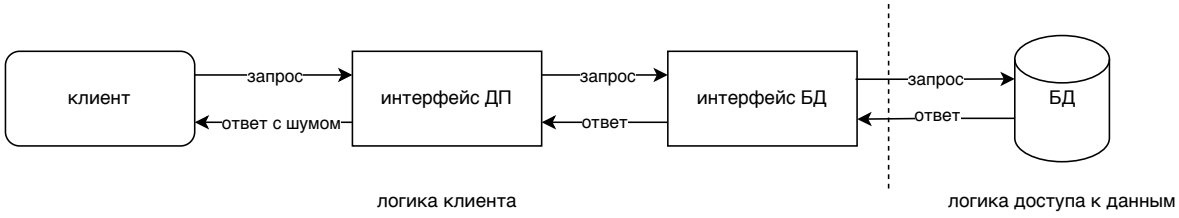

Рисунок 1.2 – Многоуровневая реализация

1.3.2 Монолитная реализаия

В данной реализации, как показано на рсиунке 1.3 один фреймворк выполняет роль интейрфейса как для функционала ДП, так и для интеграции с БД.

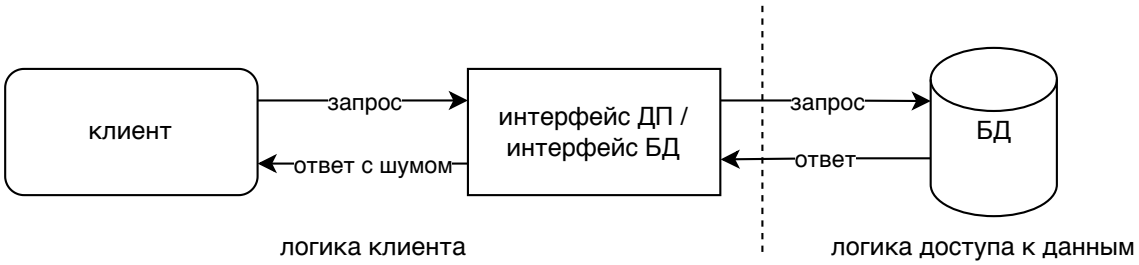

Рисунок 1.3 – Монолитная реализация

1.3.3 Реализация на уровне СУБД

Идея данной реализации представлена на рисунке **??** и заключается в добавлении шума к результатам агрегированных запросов, не требуя модификации данных на сторон клиента.

СУБД предоставляет интерфейс для использования методов ДП и гарантирует возврат результата с учетом принципов дифференциальной приватности. [4]

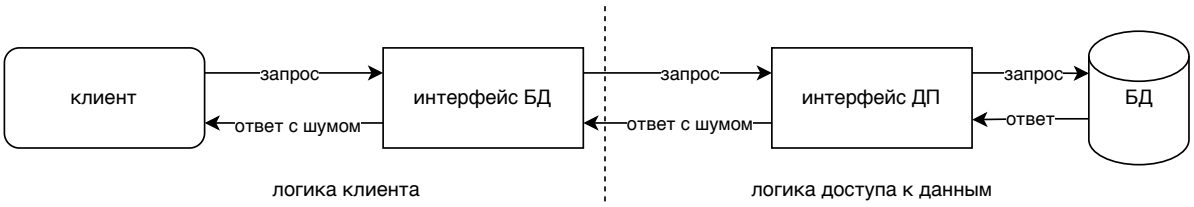

Рисунок 1.4 – Архитектура на уровне БД

1.4 Критерии сравнений

Из выше приведенного анализа предметной области были выведены следующе критерии сравнений:

— тип приватности;

— поддерживаемые механизмы;

— тип реализации;

2 Обзор существующих решений

2.1 PINQ

Privacy Integrated Queries (PINQ) — библиотека для обеспечения дифференциальной приватности при работе с данными, написанная на языке C# и основа на Language Integrated Queries (LINQ), который является .NET фреймворком. [5]

2.2 OpenDP

OpenDP — это открытый фреймворк для реализации дифференцированной приватности (Differential Privacy, DP), разработанный для упрощения применения концепций дифференцированной приватности в реальных вычислениях и статистическом анализе данных. Цель OpenDP — предложить удобные инструменты для безопасного и приватного анализа данных, при этом сохраняя высокую точность результатов.

OpenDP предоставляет программные интерфейсы, которые позволяют пользователям применять методы дифференцированной приватности к различным операциям, таким как агрегация, статистический анализ, машинное обучение и публикация данных. [6]

2.3 GoogleSQL

GoogleSQL for BigQuery uses differential privacy to protect the privacy of individuals when they query data from BigQuery. When you query a dataset with differential privacy, GoogleSQL for BigQuery will:

2.4 IBM Diffprivlib

Данная библиотека предназначена для реализации дифференциальной приватности на языке *Python* и интегрирации с разными фреймворками как для машинного обучения, так и для баз данных. Библиотека включает функции для добавления шума к базовым статистическим меткам, таким как среднее, сумма, дисперсия и медиана. [**diffprivlibibmdifferentialprivacy**]

2.5 DP-SQL

2.6 San Francisco DP Lib

3 Сравнение

ПО	Тип приватности	Механизм	Реализация
IBM Diffprivlib	ϵ, (ϵ, δ)	Лапласа, Гаусса, Экспоненциальный	интеграция
PINQ	ϵ	Лапласа	фреймворк
GoogleSQL	ϵ, (ϵ, δ)	Лапласа, Гаусса, Экспоненциальный	уровень СУБД
DP-SQL	ϵ	Лапласа, Гаусса	уровень СУБД

Таблица 3.1 – Пример таблицы

4 Исследовательский раздел

ЗАКЛЮЧЕНИЕ

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Differential Privacy: A Primer for a Non-Technical Audience / A. Wood [и др.] // ChemRN: Computational Materials Science (Topic). — 2018. — URL: https://api.semanticscholar.org/CorpusID:3352963.
2. *Gandhi R.* Technology Factsheet: Differential Privacy // Belfer Center for Science and International Affairs, Harvard Kennedy School. — 2020.
3. *Кришпопова Е. А.* Защита удаленной пользовательской статистики с помощью механизмов дифференциальной приватности // Технические средства защиты информации : тезисы докладов XVII Белорусско-российской научно – технической конференции. — 2019.
4. *Jain P. Gyanchandani M. Khare N.* Differential privacy: its technological prescriptive using big data. — 2018.
5. *Lee D. G.-Y.* Protecting patient data confidentiality using differential privacy // Scholar Archive. — 2008. — № 392.
6. *Gaboardi M.*, *Hay M.*, *Vadhan S.* A Programming Framework for OpenDP. — 2020. — Май. — Available at: https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/34954895/c3450e16-c497-4f55-8081-c9df3cc1c471/paste-2.txt.

ПРИЛОЖЕНИЕ А