РЕФЕРАТ

Расчетно — пояснительная записка Объектом исследования являются программного обеспечения с возможностью дифференциальной приватности для баз данных.

Цель работы — провести обзор существующего программного обеспечения с возможностью дифференциальной приватности для баз данных.

В ходе исследования проанализирована предметная область применения дифференциальной приватности для баз данных и изучены алгоритмы ее реализации.

В результате исследования предложены критерии сравнения программного обеспечения с возможностью дифференциальной приватности для баз данных, определен критерии их сравнения и классифицированы 4 существующего ПО.

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**]

Библиотека —

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

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

ДП — дифференциальная приватность

БД — база данных

СУБД — система управлением базами данных

ВВЕДЕНИЕ

В 2023 году ЭАЦ зарегистрировал 11549 утечек информации, произошедших в мире, что на 61,5% больше, чем в 2022 году, когда в различных источниках были найдены сведения о 7149 утечках, как показано на рисунке 1. Впервые за многолетнюю историю зафиксировано пятизначное количество утечек данных. [**utechki**].

Также, с ссылкой на источник [**forbes**], объем слитых персональных данных в России в 2023 году составил 1,12 млрд записей. Всего за отчетный период из российских компаний утекло 95 крупных баз данных.

Рисунок 1 – Количество утечек

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;

— провести обзор существующего ПО;

— определить критерии сравнения существующего ПО;

— классифицировать существующее ПО.

1 Аналитический раздел

Дифференциальная приватность обеспечивает защиту личной информации, гарантируя, что результаты анализа не позволят точно определить информацию о конкретных индивидах. В рамках ДП данные обрабатываются с добавлением шума, что минимизирует влияние индивидуальных записей на результаты статистических запросов. 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. [2]

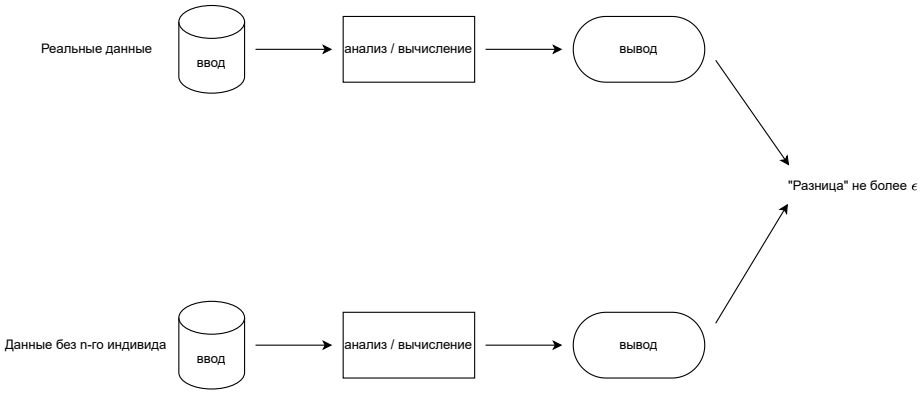

Рисунок 1.1 – Принцип ДП

1.1 Типы приватности

1.1.1 ϵ – дифференцированная приватность

Данная ϵ – дифференцированная приватность гарантирует, что результаты анализа данных не изменятся, если одна запись будет добавлена или удалена. [4]

Механизм M обеспечивает ϵ – дифференциальную приватность, если для любых двух соседних наборов данных D и D' (различающихся на одну запись) выполняется условие:

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

, где ϵ — параметр приватности

1.1.2 (ϵ, δ) – дифференцированная приватность

Модель (ϵ, δ) – дифференцированной приватности вводит дополнительный параметр δ, который позволяет немного ослабить строгие гарантии приватности. Параметр δ предоставляет вероятность того, что гарантия приватности будет нарушена в некоторых случаях. Это дает гибкость в балансе между приватностью и точностью результатов. [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, δ — параметр вероятности нарушения приватности

1.2 Типы механизмов

Параметр чувствительности — это максимальное изменение результата механизма при изменении одного элемента в базе данных. Чувствительностью механизма M является максимум разности значений механизма для любых двух соседних наборов данных D и D' (различающихся на одну запись) и вычисляется по формуле 1.4:

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

, где M — параметр чувствительности.

1.2.1 Механизм Лапласа

Лапласов механизм добавляет к результату запроса шум, сгенерированный из распределения Лапласа [4]. Работа Лапласова механизм следует 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 и D' 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

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

, где ϵ — параметр приватности, $u(r, D)$ — оценка полезности

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]

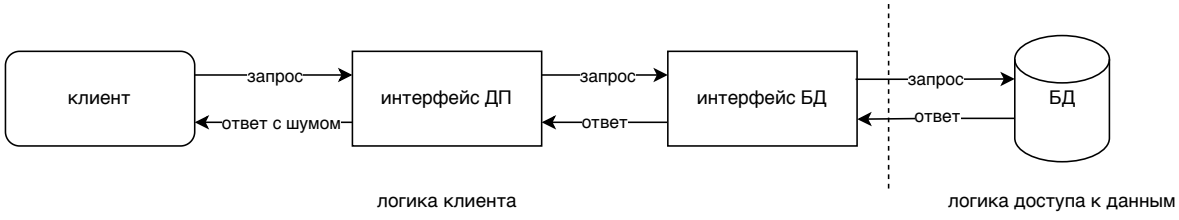

Рисунок 1.2 – Многоуровневая реализация

1.3.2 Монолитная реализаиця

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

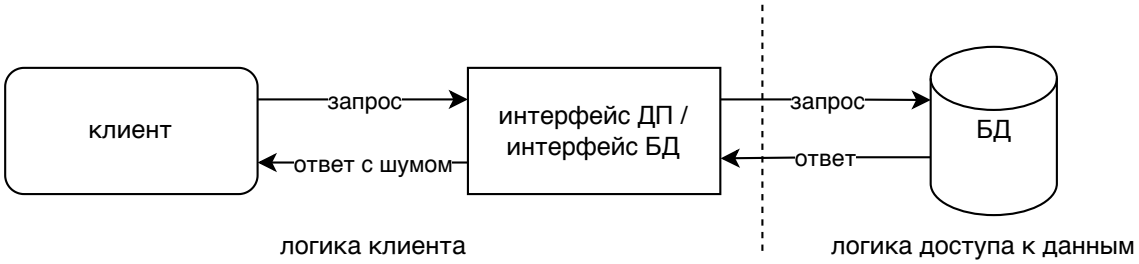

Рисунок 1.3 – Монолитная реализация

1.3.3 Реализация на уровне СУБД

Идея данной реализации представлена на рисунке **??** и заключается в добавлении шума к результатам агрегированных запросов, не требуя вмешательства в данные на стороне клиента.

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

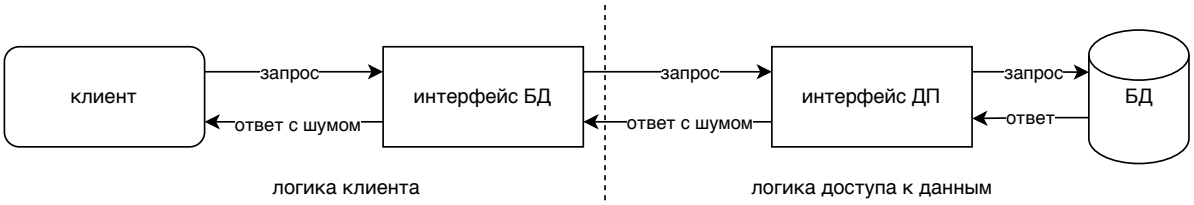

Рисунок 1.4 – Архитектура на уровне БД

1.4 Критерии сравнений

Из выше приведенного анализа предметной области были выделены следующе критерии сравнения ПО с возможностью дифференциальной приватности для баз данных:

— тип приватности;

— поддерживаемые механизмы;

— тип реализации;

Первый критерий важен, т. к. является ключевым значением в реализации ДП основой всех механизмов приватности в целом. Поддерживаемые механизмы были выделены как критерий, потому что в большей мере влияют на функционал ПО, предлагаемого пользователю. Тип реализации был выбран из–за разных подходов к взаимодействию с базами данных.

2 Обзор существующих решений

2.1 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 значениям, таким как среднее, сумма, дисперсия и медиана. [**diffprivlibibmdifferentialprivacy**] Для работы с базами данных через IBM Diffprivlib необходимо использовать промежуточный метод для подключения к БД, т. к. IBM Diffprivlib изначально лишен этой возможности.

2.2 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]

2.3 GoogleSQL

GoogleSQL — это разновидность SQL, используемая в экосистеме Google Cloud. Google Cloud — это облачная платформа, предоставляемая компанией Google, которая включает в себя набор сервисов для хранения данных, вычислений, машинного обучения, аналитики и множества других функций для разработки и развертывания приложений. [**googlecloud2023**]

GoogleSQL представляет собой расширение стандартного SQL и включает в себя некоторые уникальные функции и синтаксис, специфичные для работы с большими объемами данных. Одной из такой функции является возможность дифференциальной приватности: в GoogleSQL существует возможность добавлять механизмы шума и анонимизации для защиты данных в запросах. [**googlecloud2023**]

2.4 DP SQL

Библиотека DP SQL добавляет непосредственно в СУБД PostgreSQL функционал для выполнения дифференциально приватных агрегатных запросов. Данная библиотека реализована на внутреннем SQL языке. [**DPSQL**]

3 Сравнение

В таблице приведена классификация ранее рассмотренныего ПО с дифференциальной приватности для баз данных. По горизонтали расположены критерии сравнения методов функционала ПО.

1. тип приватности;
2. поддерживаемые механизмы приватности;
3. тип реализации;

ПО	1	2	3
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

Выявлено, что каждое ПО имеет ряд своих преимуществ и ограничений. Библиотека IBM Diffprivlib имеет широкий функционал реализации ДП, благодаря поддержке большого спектра возможностей в виде механизмов приватности и типов приватности. К особенности можно отнести работу с базами данных через стороннее ПО. Данная библиотеку можно использовать в направлении Data Science.

Библиотека PINQ обладает меньшим функционалом, чем IBM Diffprivlib, но имеет встроенную возможность взаимодействия с БД, что упрощает использование. Является самой популярной библиотекой для ДП на языке C#, поэтому подойдет для использования в проектах на языке C#.

GoogleSQL также обладает широким функционалом реализации ДП, но уже на уровне СУБД, что упрощает и усикоряет работу с конкретной БД. Но GoogleSQL является проприетарным решением компании Google на платформе Google Cloud, поэтому данное ПО подойдет для коммерческих проектов на базе Google Cloud.

DP SQL обладает менее широким функционалом, чем GoogleSQL, но является свободным распростроняемым расширением для PostrgeSQL. Поэтому данный вариант подойдет при работе с PostgreSQL

ЗАКЛЮЧЕНИЕ

Во ходе выполнения работы достигнута цель — проведен обзор существующего программного обеспечения с возможностью дифференциальной приватности для баз данных.

Определены 3 критерия сравнения ПО с возможностью дифференциальной приватности для баз данных.

Предложена классификация 4 существующих решений:

— IBM Diffprivlib

— PINQ

— GoogleSQL

— DP SQL

Для каждого из существующего решения были выявлены направления применения.

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

Презентация к научно–исследовательской работе состоит из .