РАЭК

Технологии повышения конфиденциальности как инструмент формирования доверия к обмену данными

Мария САЙКИНА | главный аналитик РАЭК

9-10/04/2024 —— москва

Задача

Поиск баланса между защитой данных и вовлечением в оборот данных, защищенных различными правовыми режимами (интеллектуальная собственность, коммерческая тайна, медицинская тайна, персональные данные и т.д.)

Дано

Обмен данными **необходим для цифровой трансформации, роста цифровой экономики** и развития инновационных технологий

Уровень доверия к обмену данными в обществе **невысокий**

52% россиян не готовы делится данными ни с кем, 56% считают, что данные, переданные коммерческим компаниям, плохо защищены, 46% не уверены в защите данных, предоставленных для получения госуслуг (по данным опроса ВЦИОМ и АНО «Национальные приоритеты»)

Организации **не до конца используют потенциал** данных из-за жёстких регуляторных требований, опасений относительно утечек и возможных санкций

порядка 60-73% данных, имеющихся в распоряжении организаций, не используются для аналитики (по оценкам Forrester)

Возможное решение

Privacy Enhancing Technologies (PETs) / Технологии повышения конфиденциальности

совокупность цифровых технологий и подходов, позволяющих осуществлять сбор, обработку, анализ и обмен информацией при одновременной защите конфиденциальности персональных данных

ОЭСР

программные и аппаратные решения, функционал которых включает технические процессы, методы или знания для достижения конфиденциальности, защиты данных или защиты от рисков, связанных с приватностью отдельного физического лица или группы лиц

Агентство ЕС по кибербезопасности (ENISA)

любые программные или аппаратные решения, технические процедуры, техники и т.д., предназначенные для снижения рисков приватности, возникающих в процессе обработки данных

Указ президента США от 30 октября 2023 года о безопасном и доверенном ИИ

Виды PETs

- K-анонимность
- Дифференциальная приватность
- Синтетические данные
- Многосторонние вычисления
- Гомоморфное шифрование
- Доказательства с нулевым разглашением
- Доверенные среды выполнения
- Федеративное обучение

Плюсы и минусы

Технология	Описание	Преимущество	Ограничения
K-анонимность	Преобразует заданный набор из k записей таким образом, что в опубликованной версии каждый субъект неотличим от других.	Снижает риск повторной идентификации.	Уязвимость для атаки повторной идентификации, если есть дополнительная общедоступная информация.
Дифференциальная приватность	Добавляет шум к исходным данным таким образом, что злоумышленник не может определить, были или не включены данные какого-либо человека в исходный набор данных.	Гарантирует конфиденциальность за счет снижения вероятности восстановления данных и атак связывания (linkage attacks).	Ограничена более простыми типами данных; Есть проблема в поиске компромисса между конфиденциальностью, точностью и полезностью данных.
Синтетические данные	Искусственно созданные данные как альтернатива реальным данным	Сохраняет общие свойства или характеристики исходного дата-сета	Может сохранять конфиденциальную информацию, содержащуюся в исходном наборе данных; Сложно точно отразить реальные данные
Многосторонние вычисления	Позволяет нескольким сторонам совместно выполнять согласованные вычисления над своими личными данными, при этом позволяя каждой стороне изучать только окончательные результаты вычислений.	Увеличивает возможности вычислений над распределенными наборами данных без раскрытия исходных данных.	Более высокие вычислительные и коммуникационные затраты/нагрузка; Трудности с масштабированием.

Источник: Национальная стратегия по развитию обмена данными и аналитики с сохранением конфиденциальности США

Плюсы и минусы

Технология	Описание	Преимущество	Ограничения
Гомоморфное шифрование	Позволяет выполнять вычисления над зашифрованными данными и получать результаты в зашифрованной форме.	Видеть исходные и/или полученные данные могут только авторизованные пользователи	Высокие временные и вычислительные затраты
Доказательства с нулевым разглашением	Позволяет одной стороне доказать другой стороне, что конкретное утверждение верно, не раскрывая конфиденциальную информацию.	Повышает способность проверять информацию без раскрытия конфиденциальной информации.	Высокие затраты; сложность с масштабированием
Доверенные среды исполнения	Создает безопасную, изолированную среду выполнения, параллельную основной операционной системе, для обработки конфиденциальных данных.	Обеспечивает более быстрый и безопасный анализ данных по сравнению с методами, основанными на шифровании.	Создает другие возможности для утечки конфиденциальных данных.
Федеративное обучение	Обеспечивает возможность сотрудничества при построении модели машинного обучения на основе распределенных данных без совместного использования исходных данных.	Минимизирует обмен данными при обучении комбинированной модели.	Сохраняется возможность различных атак по реконструкции данных (data reconstruction attack) или атак логического вывода (inference attack); требует согласованности между наборами данных, хранящимися у нескольких объектов.

Источник: Национальная стратегия по развитию обмена данными и аналитики с сохранением конфиденциальности США

Мнение бизнеса

2.4 млрд $

оценка объема рынка PETs в 2023 году

↓

25.8 млрд $

оценка объема рынка PETs к 2033 году

26.6%

среднегодовые темпы роста

Источник: Future Market Insights

Мнение регуляторов

апрель 2023: Национальная стратегия по развитию обмена данными и аналитики с сохранением конфиденциальности
// National Strategy to Advance Privacy-Preserving Data Sharing and Analytics

сентябрь 2023: проект «Закона об исследованиях в области технологий, повышающих конфиденциальность»
// Privacy Enhancing Technology Research Act

октябрь 2023: Указ президента США о безопасном и доверенном ИИ
// Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence (отдельный раздел посвящен PETs)

январь 2022: отчет ENISA о техниках проектирования защиты данных
// Data Protection Engineering. From Theory to Practice

май 2022: проект «Закона о создании Европейского пространства медицинских данных»
// Regulation of the European Parliament and of the Council on the European Health Data Space (в рамках создания пространства мед. данных через программы Horizon Europe и EU4Health предоставляются гранты на исследования в области PETs)

сентябрь 2023: «Закон об управлении данными»
// Data Governance Act (в п. 7 Преамбулы рассматриваются техники, позволяющие анализировать базы данных, содержащие персональные данные)

Мнение регуляторов

июнь 2023: Управление уполномоченного по информации Великобритании (ICO), свод руководящих принципов по использованию PETs
// Privacy-enhancing technologies (PETs) Guidance

ноябрь 2017: Управление Уполномоченного по вопросам конфиденциальности Канады (OPC), Обзор инструментов и технологий PETs
// Privacy Enhancing Technologies — A Review of Tools and Techniques Report

---

2022 год: Управление по развитию инфокоммуникационных средств массовой информации Сингапура (IMDA) запустило «песочницу» для тестирования пилотных проектов использования PETs в безопасной среде и инициировало работу по изучению использования отдельных PETs

---

G7

июнь 2023 года: объединенная позиция регуляторов в области защиты данных и конфиденциальности (Data Protection and Privacy Agencies) стран «большой семерки» по развитию свободного обмена данными, основанного на доверии
// Data Free Flow with Trust (отмечается значимая роль PETs как инструмента интеграции принципов защиты данных в процессы их обработки)

PETs - решение всех проблем?

PETs являются инструментом, позволяющим обеспечить надлежащую защиту персональных данных при сохранении полезных качеств данных для их последующего анализа

PETs обладают значительным потенциалом для повышения доверия к обмену данными за счет реализации принципа проектируемой приватности и существенному снижению рисков повторной идентификации данных

Ни одна из PETs не является универсальной, каждая имеет свои преимущества и недостатки, и целесообразность их использования должна оцениваться в каждом конкретном случае

Все аспекты применения PETs пока еще недостаточно изучены, для их надлежащего использования требуются соответствующие руководства и стандарты

Использование PETs должно быть гармонизировано с принципами регулирования персональных данных

РАЭК

Следите за новостями Рунета в Telegram-канале РАЭК

Слушайте подкаст «А за окном Россия» о российском ИТ-бизнесе и технологиях в нашей жизни

Спасибо!

Мария САЙКИНА | главный аналитик РАЭК