ГАЗПРОМБАНК QApp ФИНТЕХ АССОЦИАЦИЯ

Конфиденциальные вычисления

Первая встреча рабочей группы
31.10.2023

gazprombank.ru

QApp ГАЗПРОМБАНК

Антон Гугля

Генеральный директор QApp

Программа встречи

	№	Длительность	Задачи и темы	Докладчик / Модератор
Открытая часть	1	15:00 - 15:10	Сбор участников	
	2	15:10 - 15:30	Приветствие	Гугля А.П.
	3	15:30 - 15:40	Технологические и бизнес-вызовы Банка ГПБ АО в которых технологии конфиденциальных вычислений могут быть ответом	Пузырникова Н.В.
	4	15:40 - 16:00	Вступительное слово от представителя Ассоциации ФинТех	Григорьев М.А.
	5	16:00 - 16:50	Технологии конфиденциальных вычислений	Гугля А.П. Кот М.А.
Закрытая часть	6	16:50 - 17:00	Кофе-брейк	
	7	17:00 - 17:45	Дискуссия участников рабочей группы	Гугля А.П
	8	17:45 - 18:00	Выводы встречи	Гугля А.П

Конфиденциальные вычисления

QApp ГАЗПРОМБАНК

Существующее шифрование

Данные в покое

Защита ценных данных в процессе хранения

Данные в процессе передачи

Защита данных, передаваемых между общедоступными или частными сетями

Новое

Конфиденциальные вычисления

Данные в использовании

Защита данных от раскрытия владельцу системы выполняющей вычисления над данными

Кейсы применения конфиденциальных вычислений

QApp ГАЗПРОМБАНК

MPC ALLIANCE

MPC Альянс включает в себя около 60 компаний, среди которых Meta, Bosch, Acronis, Binance и др. Альянс был создан с целью расширения осведомленности, продвижения и внедрения технологий MPC в различных отраслях

RBC Royal Bank of Canada

Работники банка могут брать данные о покупках от продавцов и сочетать их со своей собственной информацией о транзакциях по кредитным картам потребителей, чтобы получить полное представление о транзакциях, не нарушая конфиденциальности информации потребителей или продавцов

МАГНИТ

В 2019 году ритейлер «Магнит» и компания по управлению цифровыми лицензиями Aggregion запустили первую в России доверенную среду совместной работы с данными

ВТБ

Технология позволяет компаниям совместно работать с массивами данных для улучшения качества сервисов, повышения эффективности бизнеса и решения других задач

Альфа Банк

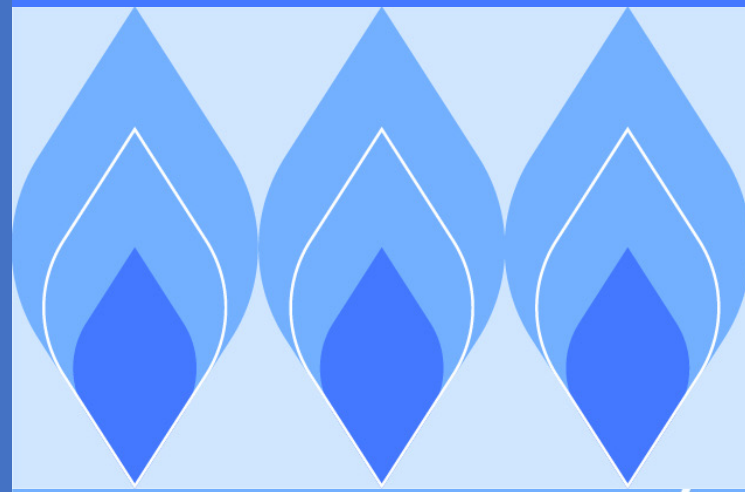
Научно-исследовательские и пилотные проекты по конфиденциальным вычислениям

ГАЗПРОМБАНК QApp

01

Конфиденциальные вычисления в Банке ГПБ

Технологические и бизнес-вызовы Банка ГПБ АО
в которых технологии конфиденциальных вычислений
могут быть ответом



Наталия Пузырникова

Заместитель председателя правления Банка ГПБ АО

Области применения конфиденциальных вычислений в Банке ГПБ

QApp ГАЗПРОМБАНК

Вычисления над зашифрованными данными

- Данные — ключевая бизнес-ценность сторон
- Передача данных в открытом виде не отвечает бизнес-целям сторон
- Для передачи данных в незашифрованном виде есть законодательные ограничения

Бизнес-применение конфиденциальных вычислений

- Расчет кредитный рисков
- Профилирование клиентов
- Антифрод

Концепция решения Банка ГПБ

Банк ГПБ уже сегодня проводит прикладные исследования и пилотные проекты по направлению конфиденциальных вычислений и готов к диалогу с отраслью

Вызовы отечественной индустрии в сфере конфиденциальных вычислений

QApp ГАЗПРОМБАНК

Нет зрелого рынка (недостаток технологий, кейсов, обмена опытом)

Недостаток на 100% отечественных решений (ПАК – аналог Intel SGX)

Нет четких рекомендаций от регулятора и законодателя

Максим Григорьев

Генеральный директор Ассоциации ФинТех

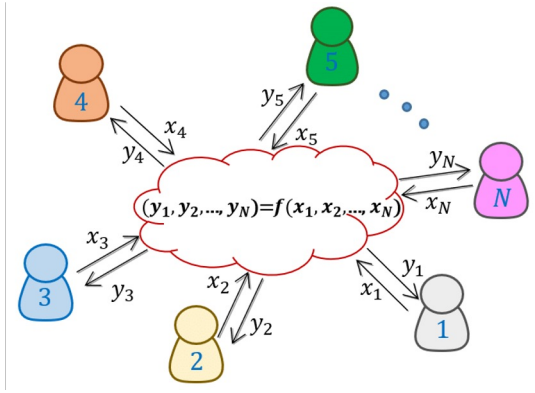
ГАЗПРОМБАНК QApp

02

Технологии конфиденциальных вычислений

Семинар

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ



Технология

Криптографический протокол, позволяющий нескольким участникам вычислить некоторую функцию от набора данных, сохраняемых от каждого из них в секрете, таким образом, чтобы ни один участник не смог получить никакой информации о чужих входных данных.

Пример: два миллионера хотят узнать, кто богаче, но не хотят раскрывать размеры своего достатка.

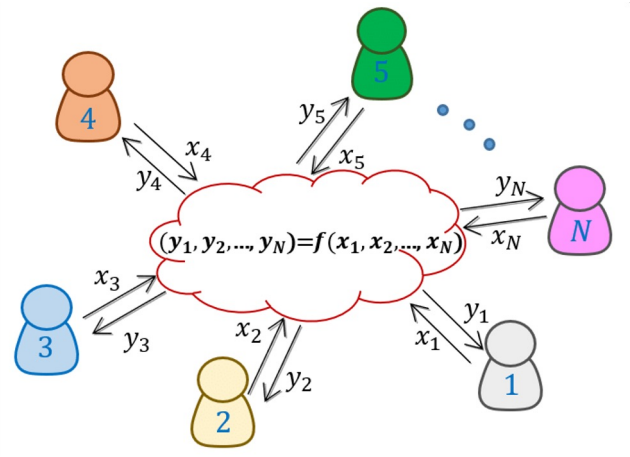
Подробнее

Сферы применения

- Конфиденциальный поиск в базе данных
- Электронное голосование
- Электронные аукционы
- Статистическая обработка конфиденциальных данных
- Распределённый центр сертификации

Пример реализации: ABY

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ



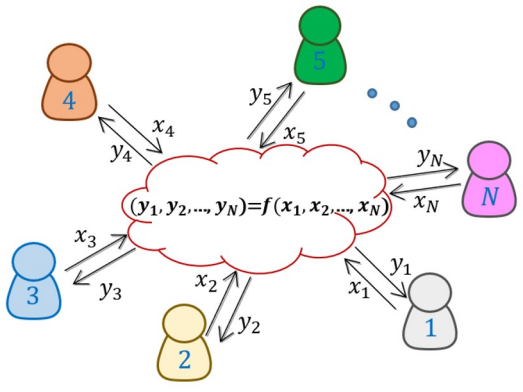
- В протоколах конфиденциального вычисления (secure multiparty computation, MPC) участвует ряд различных, но связанных, вычислительных устройств (или участников).
- Задачей такого протокола является проведение совместного вычисления некоторой функции безопасным способом. Обычно выход каждого участника одинаковый, но в общем случае выходы могут быть различными
- Участники выбирают функцию для вычисления, а затем используют MPC-протокол для совместного вычисления выхода функции на их секретных данных без разглашения этих данных.

Примеры функций:

- математическая функция (элементарная функция или их комбинация)
- строковая операция
- обращения к полям распределенной базы данных без раскрытия как результатов, так и запросов

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ

Формальная модель: основные положения



Предположения

- Между двумя каждыми участниками существует безопасный канал передачи информации
- Злоумышленник контролирует подмножество участников, которые называются испорченными (corrupted)

Модель безопасности (модель злоумышленника)

- **Semi-honest (passive) security** - испорченные участники не отклоняются от спецификации протокола, однако злоумышленник получает всю информацию, доступную им. Используя эти данные, злоумышленник пытается раскрыть секретные данные честных участников
- **Malicious (active) security** - контролируемые злоумышленником участники могут произвольно отклоняться от спецификации протокола. В данной модели любые активные атаки не нарушают свойство конфиденциальности

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ

Формальная модель: основные положения

Свойства безопасности (требования к безопасности)

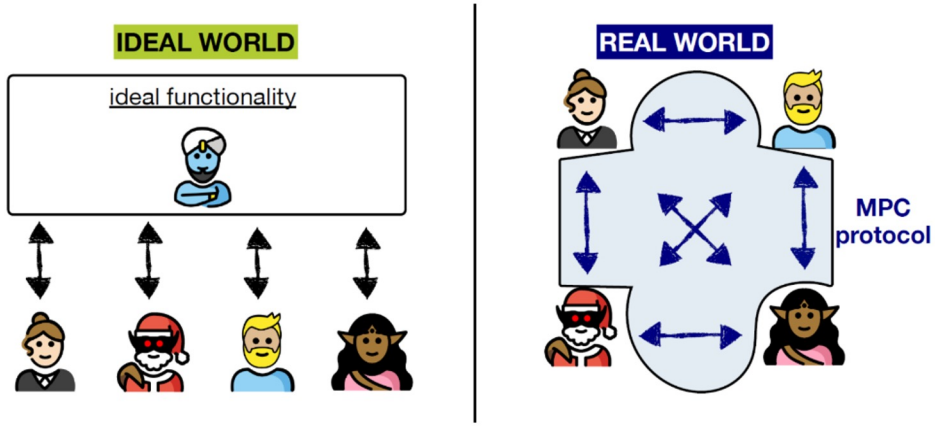
- **Privacy**: информация каждого участника ограничена тем, что он может посчитать на основе своего входа и предназначенного ему результата
- **Correctness**: гарантируется, что результат вычислений каждого участника правильный
- **Guaranteed Output Delivery**: испорченные участники не могут предотвратить получение честными участниками предназначенного для них результата
- **Fairness**: испорченные участники должны получить результат тогда и только тогда, когда честные участники получают свой результат

При использовании модели активного злоумышленника обычно рассматриваются более слабые свойства, например как:

- **Security with abort**: злоумышленник получает результат раньше честных участников и решает, разглашать результат или нет

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ

Формальная модель: основные положения



Для доказательства безопасности Secure-MPC протокола (с заданным набором предположений) используется парадигма Реальный-Идеальный мир. В идеальном мире предполагается, что существует доверенная третья сторона, которая принимая входные данные каждого участника, вычисляет требуемую функциональность и отправляет результат участникам.

Протокол реального мира считается стойким (в некоторых предположениях) при условии, что если в реальном мире нарушаются свойства протокола, то существует некий злоумышленник и в идеальном мире (где все примитивы стойкие), который нарушит те же свойства

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ

Центральным примитивом MPC для реализации арифметических операций является схема разделения секрета

Линейная схема разделение секрета:

- Для разделения входного секретного значения **x** при наличии **n** участников, участник генерирует **n** случайных чисел (называемых долями) так, чтобы их сумма равнялась **x** (сумма и числа определены по модулю большого числа)
- Каждому участнику отправляется предназначенная ему доля
- Участники могут решить раскрыть секрет, опубликовав свои доли (в общем случае требуется **t ≤ n** долей для восстановления секрета)
- Локальное сложение двух долей приводит к получению доли, являющейся разделением секрета для суммы изначальных секретов
- Для выполнения операции умножения используется дополнительный подпротокол

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ

Пример вычисления суммы

Вычисления производятся по модулю $q = 11$

Участник 1 имеет секретный вход **2**, Участник 2 – **3**

Задача: вычислить сумму (известно, что входы и результат меньше q)

$[2] = (5, 8)$ $[3] = (4, 10)$

8 →

← 4

Вычисление с использованием локальных долей:

$5 + 4 = 9$ $8 + 10 = 7$

Результат: $[5] = (9, 7)$

Для получения результата в открытом виде необходимо опубликовать доли

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ

Пример вычисления линейной функции

Вычисления производятся по модулю $q = 101$

Участник 1 имеет секретный вход $x_1 = 5$, Участник 2 – $x_2 = 3$

Задача: вычислить $4 \cdot x_1 + 7 \cdot x_2$

$[5] = (1, 4)$ $[3] = (41, 63)$

4 →

← 41

Вычисление с использованием локальных долей:

$4 \cdot 1 + 7 \cdot 41 = 89$ $4 \cdot 4 + 7 \cdot 63 = 53$

Результат: $[4 \cdot 5 + 7 \cdot 3] = [41] = (89, 53)$

Для получения результата в открытом виде необходимо опубликовать доли

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ

На практике MPC протокол разбивается на два этапа:

- **Offline(preprocessing) фаза** - предварительная фаза, в которой не используются входные данные участников. Данная фаза предназначена для ускорения онлайн фазы и обеспечения безопасности протокола. В ходе этой фазы участники (или доверенная третья сторона) совместно генерируют случайные данные, удовлетворяющие некоторым соотношениям;

- **Online фаза** - фаза, в которой участники непосредственно вычисляют требуемую функцию с использованием своих входных данных и данных, сгенерированных в оффлайн фазе.

ТЕХНОЛОГИЧЕСКИЕ ПОДХОДЫ К РЕШЕНИЮ ЗАДАЧ КОНФИДЕНЦИАЛЬНЫХ ВЫЧИСЛЕНИЙ

Подход №1

Secure-MPC (SMPC)
Программные решения

Подход №2

ПАК
Программно-аппаратные решения

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ

Подход №1. Secure-MPC (программный подход)

MPC как общая модель

Модель MPC — весьма сильная и общая. Она позволяет описать и другие схемы, в том числе: гомоморфное шифрование и разделение секрета, при этом парадоксальным образом опирается на некоторые из них.

используемые в MPC криптографические примитивы:

- Защищенные каналы
- Гомоморфное шифрование
- Разделение секрета
- Привязка к биту (bit commitment)
- Протоколы с нулевым разглашением (zero-knowledge proofs)

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ

Secure-MPC подход

Достоинства

- Корень доверия — криптография
- Лучшее масштабирование в долгосрочной перспективе
- Независимость от процессорной архитектуры
- Отсутствие зависимости от конкретного производителя оборудования
- Строгое доказательство конфиденциальности

Недостатки

- Низкая производительность
- Отсутствие практики сертификации решений регулятором
- Малый промежуток времени практического криптоанализа решений
- Дешёвое масштабирование атаки на всю сеть

ПРОГРАММНО-АППАРАТНЫЙ ПОДХОД

ПАК (программно-аппаратный комплекс) связан с обработкой данных в защищенном периметре и опирается на доверие к производителю оборудования.
Подход гарантируется выполнением кода в аппаратно-защищенной среде (примеры: Intel SGX, ARM TrustZone). Таким образом, доверие обеспечивается доверием к корневому сертификату производителя вычислительной аппаратуры. Передача данных от поставщика к вычислительному ядру должна осуществляться по защищенному каналу, затем данные расшифровываются и обрабатываются внутри защищенного периметра.

Достоинства

- Высокая производительность
- Скорость введения в эксплуатацию
- Проверенность временем модели обеспечения конфиденциальности, прецеденты сертификации
- Дорогое масштабирование атаки на всю сеть
- Отработанные механизмы аттестации узла доверия

Недостатки

- Корень доверия — сертифицирующая лаборатория регулятора
- Зависимость от производителей аппаратной части
- Возможные ограничения по месту размещения аппаратуры
- Нельзя построить математическую модель стойкости решения

ПРОГРАММНО-АППАРАТНЫЙ ПОДХОД

Основные понятия

TPM (Trusted platform module) — компонент программно-аппаратной системы, обеспечивающий аттестацию системы и хранение криптографической информации.

TEE (Trusted execution environment) — частный случай TPM — изолированная от основной операционной системы (таких как например: Android и Linux) вычислительная среда, выполняющая функции корня доверия: хранение конфиденциальной информации и контроль доступа к этой информации

Доверенная загрузка (secureboot) — общее название набора технологий позволяющих реализовать аттестацию системы на этапе загрузки.

Удалённая аттестация (RA, remote attestation) — набор технологий, позволяющих проводить аттестацию оборудования удаленно, в том числе используя ключи заинтересованных сторон.

Защита памяти (memory protection) — набор методик, использующих аппаратные возможности и шифрование памяти для обеспечения изоляции участка памяти от других компонент системы.

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ

Сравнение подходов

ПАК-подход

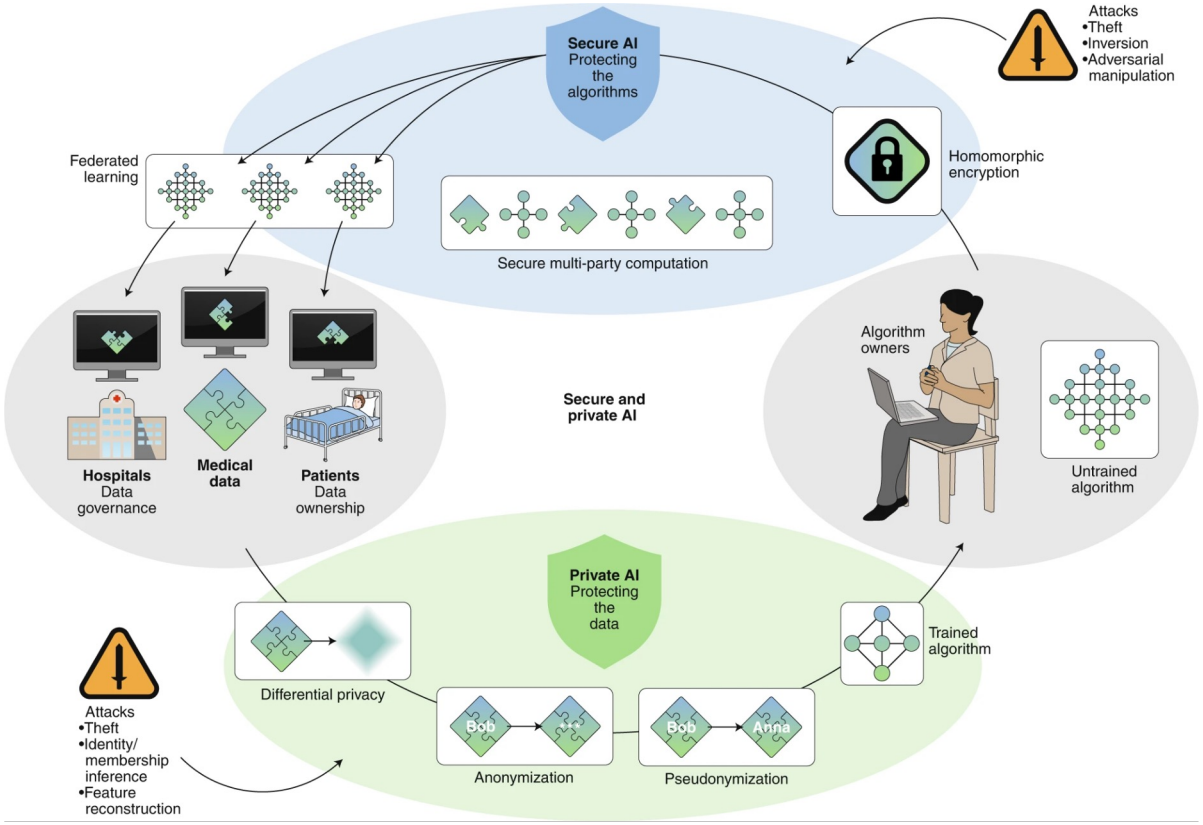
- может быть реализован в относительно короткие сроки при условии поддержки со стороны владельца аппаратной платформы
- защита данных при передаче может быть построена на сертифицированных СКЗИ
- защита данных в вычислительном ядре обеспечивается аппаратными методами, доступ к данным обеспечивается уровнями доверия, установленными производителем, при условии поддержки корневого сертификата производителя
- накладные расходы, в том числе падение скорости обработки данных, незначительны
- решение может быть сертифицировано регулятором на общих основаниях

Secure-MPC (SMPC) подход

- внедряемые методы защиты данных могут гарантировать квантовую устойчивость и защиту на долгий срок
- независимость от аппаратной платформы и от вендора гарантирует легкую переносимость и большой жизненный цикл решения (например, Intel отказалась от поддержки технологии SGX)
- интересен в долгосрочной перспективе

КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ

SMPC и Machine Learning



Конфиденциальные вычисления позволяют решать задачу обучения модели машинного обучения среди нескольких участников таким образом, что они получают модель, но используемые данные остаются в секрете.

В результате работы протокола возникает возможность получить более точную модель (по сравнению с моделью, которую может создать каждый из участников) благодаря большему набору данных.

Участники могут использовать SMPC протоколы в случаях, когда они не могут (банки, медицинские организации) или не хотят (коммерческие компании) делиться данными в открытом виде.

Источник: https://www.nature.com/articles/s42256-020-0186-1

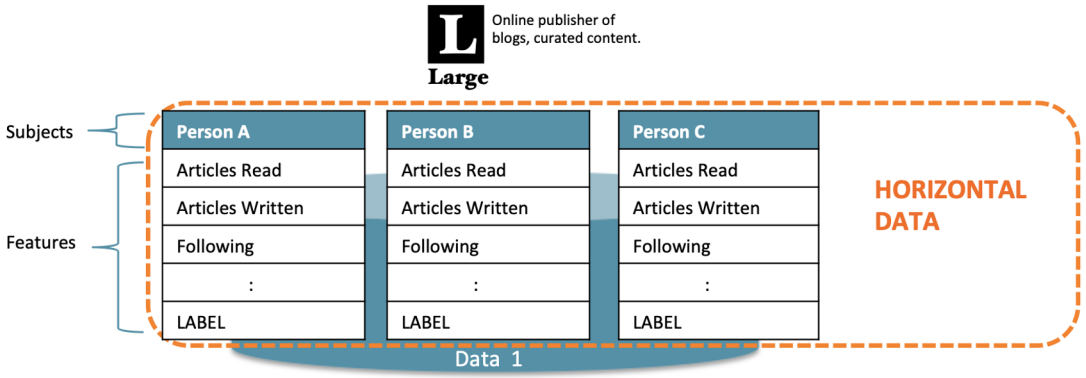
КОНФИДЕНЦИАЛЬНЫЕ ВЫЧИСЛЕНИЯ

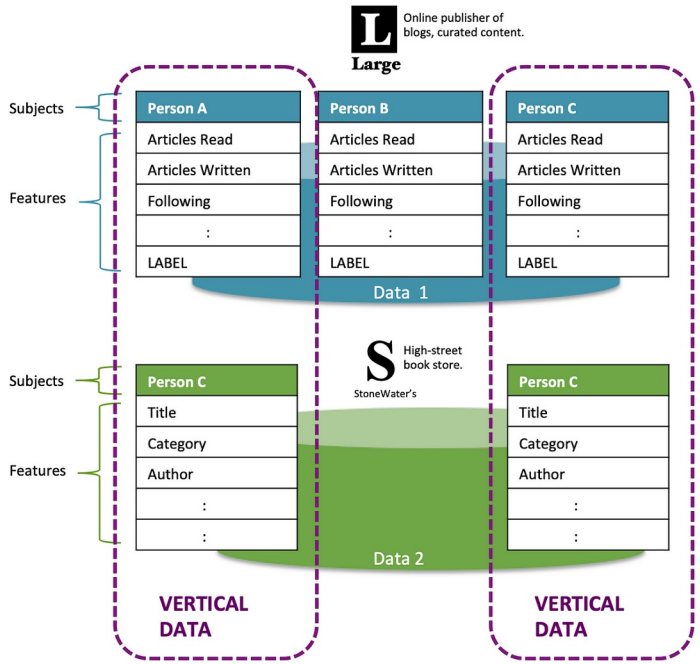
SMPC и Machine Learning

При обучении общей модели существует два типа разделения данных:

- **Горизонтальное разделение**: формат данных для каждого участника одинаков. Можно представить, что единую базу данных разделили по строкам среди участников.
- **Вертикальное разделение**: каждый участник имеет некоторые общие признаки (колонки) данных, а другие признаки уникальны для каждого участника.

В общем случае протоколы для горизонтального разделения данных быстрее и проще, чем для вертикального.





Источник: https://blog.openmined.org/federated-learning-types/

ГАЗПРОМБАНК QApp ФИНТЕХ АССОЦИАЦИЯ

Конфиденциальные вычисления

Кофе-брейк

gazprombank.ru

ГАЗПРОМБАНК QApp

03

Дискуссия и выводы встречи