DATA FUSION

FIRST RUSSIAN DATA FORUM

Технологии защищенной обработки данных

Григорий Маршалко

АК РФ

Развитие технологий обработки — развитие технологий защиты

ЗАЩИТА ПРИ ПЕРЕДАЧЕ

Защита каналов при передаче информации между контролируемыми зонами

ЗАЩИТА ПРИ ХРАНЕНИИ

Хранение больщих объемов информации после ее передачи и обработки в контролируемой/ неконтролируемой зоне

ЗАЩИТА ПРИ ОБРАБОТКЕ

Защита информации при обработке недоверенным пользователем

Спектр технологий защищенной обработки данных

Организационно-технические

Разграничение доступа к данным

Обезличивания

Преобразование данных или алгоритмов их обработки для невозможности повторной идентификации

Криптографические

Вычисления, запросы с использованием криптографических методов

Федеративное обучение

Совместное распределенное обучение на нескольких устройствах без непосредственного обмена данными

Доверенные анклавы

Вычисления в доверенной среде с аппаратным контролем доступа

Спектр технологий защищенной обработки данных

Организационно-технические

Разграничение доступа к данным

Обезличивания

Преобразование данных или алгоритмов их обработки для невозможности повторной идентификации

Криптографические

Вычисления, запросы с использованием криптографических методов

Защищенная публикация данных

Табличные данные преобразуются оператором (обобщаются, подавляются, зашумляются, перемешиваются) с последующей публикацией в закрытом или открытом контуре

Защищенное машинное обучение

Оператор реализует API с зашумляющими алгоритмами анализа данных, данные не передаются, передаются результаты их обработки

Спектр технологий защищенной обработки данных

Организационно-технические

Разграничение доступа к данным

Обезличивания

Преобразование данных или алгоритмов их обработки для невозможности повторной идентификации

Криптографические

Вычисления, запросы с использованием криптографических методов

Протоколы безопасных вычислений

Совместные распределенные вычисления в условиях наличия нарушителя

Гомоморфное шифрование

Делегирование вычислений недоверенному серверу через возможность вычислений над зашифрованными данными

Безопасное извлечение информации

Возможность обращения к записи в базе данных без разглашения оператору базы факта обращения к записи (протоколы передачи с забыванием)

Поиск по шифрованным данным

Алгоритмы шифрования баз данных, допускающие поиск по ключевым словам

Как выбрать подходящую технологию?

Кто обрабатывает?

Кто участвует в обработке:
— пользователь,
— владелец данных,
— аналитик…?
Какие у них:
— права доступа,
— возможности?

Что обрабатывает?

Какие данные обрабатываются:
— категориальные,
— числовые
  — целочисленные,
  — вещественные,
  — даты…?
Какие форматы используются?
У каких участников какие данные?

С какой целью?

Какая цель обработки:
— вычисление статистик,
— обучение моделей…?
Какая нужна точность?

Какие угрозы?

Кто является потенциальным нарушителем?
Какие у него цели?
Что является угрозой безопасности данных:
— утечка одной конкретной записи,
— произвольной записи,
— ее части?