DATA FUSION

FIRST RUSSIAN DATA FORUM

Скоринговая межбанковская инфраструктура

И конфиденциальные вычисления

ЕМЕЛЬЯНОВ ПЕТР

ООО Блуттех, Генеральный Директор

О компании

Bloomtech – российская, коммерческая, независимая компания, не аффилированная ни с одной кредитной организацией. Наша цель: взаимовыгодное сотрудничество банков друг с другом и с нами.

01. РАЗРАБАТЫВАЕМ РЕШЕНИЯ ДЛЯ БАНКОВ

Изучаем и развиваем технологии совместных конфиденциальных вычислений, разрабатываем новые протоколы, оптимизируем существующие, применяем на практике в Fintech (и не только).

01. КАПИТАЛИЗИРУЕМ HANDS-ON ОПЫТ В BIG DATA

Знаем, что такое Big Data, Artificial Intelligence, Machine Learning. Умеем внедрять сложные программные продукты в крупнейшие компании и государственные структуры Российской Федерации.

Агрегация банковских данных

Данные каждого отдельного банка не передаются другим банкам и не консолидируются у третьей стороны. Единственное, что раскрывается, – сумма метрик по всем банкам сразу.

КРЕДИТНЫЙ СКОРИНГ

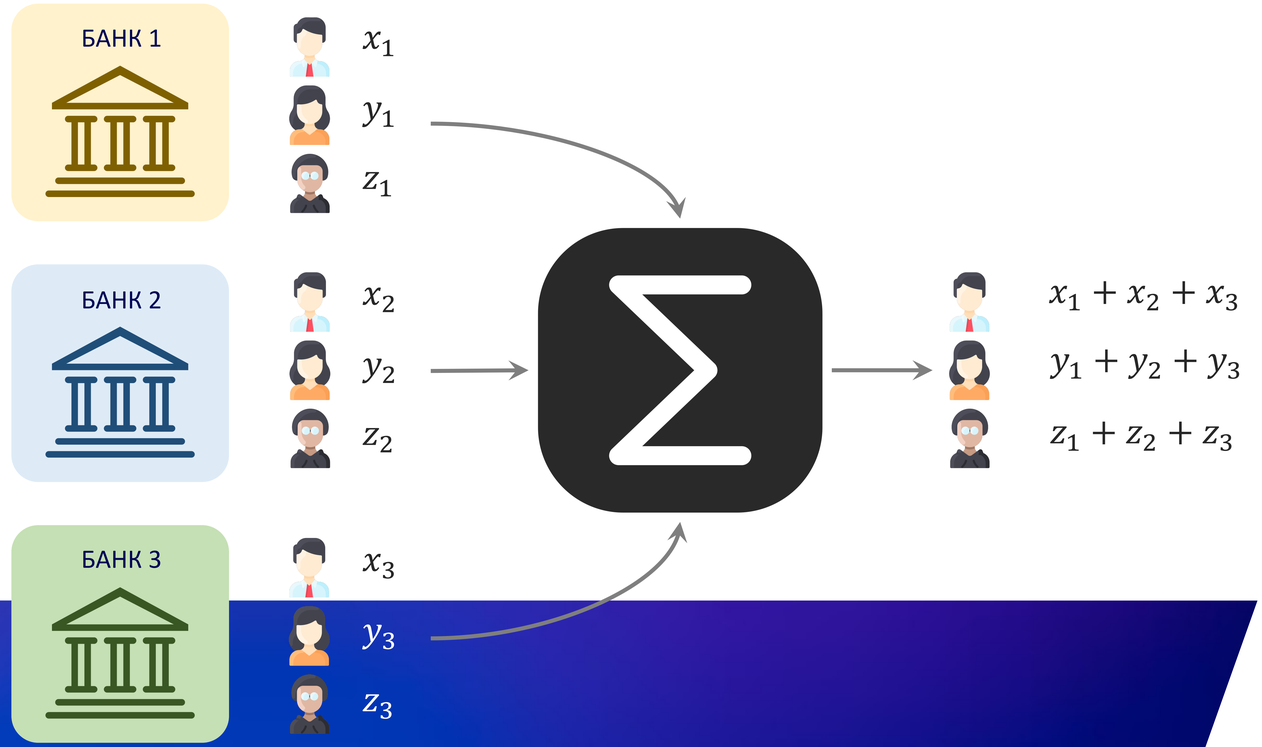
Объединенные межбанковские данные – обогащение скоринговых моделей;

ОЦЕНКА ДОХОДА/PTI

Верификация дохода в заявке, оценка дохода по сумме денежных поступлений на дебетовые счета;

AML/KYC

Новые данные о клиентах, выявление подозрительных операций, маркировка дроперов (мулов), etc;



Проще простого!

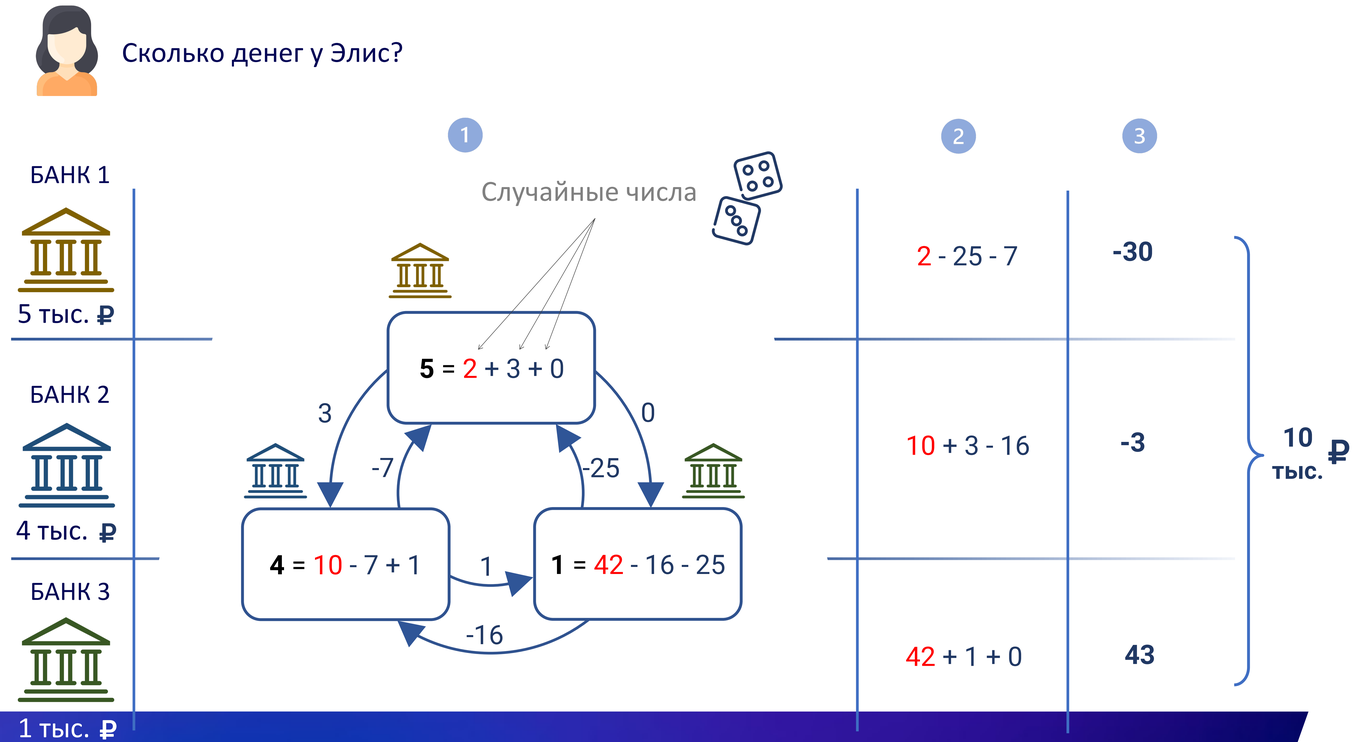
От перестановки мест слагаемых сумма не изменяется. Мы просто делаем слагаемые случайными.

1. РАЗДЕЛЕНИЕ СЕКРЕТА

Участники вычислений представляют секретные значения в виде суммы случайных слагаемых и обмениваются случайными числами;

2. НЕПОСРЕДСТВЕННО ВЫЧИСЛЕНИЯ

Участники вычислений складывают случайные числа: одно свое и два, полученные от соседей;



ВОССТАНОВЛЕНИЕ РЕЗУЛЬТАТА

Участники вычислений обмениваются полученными слагаемыми (тоже случайными), складывают их и узнают оригинальную сумму. Профит!

Не тут то было!

БАНКОВСКАЯ ТАЙНА!

Криптографический протокол

Рассмотрим задачу агрегации банковских данных как протокол, в котором участвуют банки и оператор. Банки владеют данными и совместно используют данные друг друга, а оператор не владеет и не получает никаких подлежащих регуляции данных.

1. МНОЖЕСТВО БАНКОВ

Протокол чувствителен к изменению состава участников, поэтому множество банков фиксируется для каждой вычислительной сессии.

2. СЕССИОННЫЕ ДАННЫЕ

Банки генерируют необходимый криптографический материал, обмениваются им друг с другом и устанавливают новую вычислительную сессию.

3. ВЫЧИСЛЕНИЕ МЕТРИК

В рамках установленной сессии банки выполняют запросы на вычисление агрегированных метрик, используя сессионную ключевую информацию.

Модель угроз

Банковская тайна ~ 5 свойств безопасности

01. КОНФИДЕНЦИАЛЬНОСТЬ МЕТРИК

Никто не может узнать, какие операции по счетам совершало физическое лицо, а также ненулевой остаток на указанных счетах.

02. АНОНИМНОСТЬ КЛИЕНТОВ

Никто не может узнать, в каком конкретно банке у физического лица открыты счета (заключены иные договоры с банком).

03. АНОНИМНОСТЬ ИНИЦИАТОРА

Никто (кроме оператора) не может определить, какой банк-участник направил запрос информации о конкретном физическом лице.

04. АНОНИМНОСТЬ КЛИЕНТА В ЗАПРОСЕ

Оператор не может определить, в отношении каких физических лиц (персональные данные) участники направляют свои запросы.

05. КОНФИДЕНЦИАЛЬНОСТЬ РЕЗУЛЬТАТА

Никто (кроме банка-инициатора запроса) не может получить значения агрегированных метрик физического лица, по которому другие банки не выполняли запросов на получение тех же метрик.

Банковская тайна – юридическое понятие. Свойства безопасности – конкретные требования, выполнение которых можно проверить.

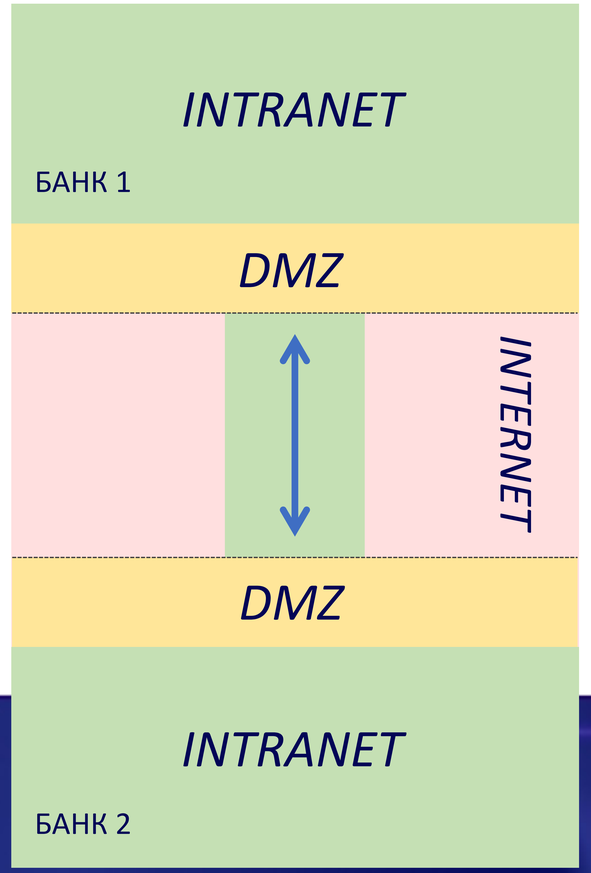
Модель нарушителя

ВНУТРЕННИЙ НАРУШИТЕЛЬ

- Отклоняется от протокола, отправляет недостоверную информацию, прерывает протокол;
- Использует произвольные входные параметры;
- Навязывает честным участникам значения их параметров;
- Вступает в сговор с другими участниками;

ВНЕШНИЙ НАРУШИТЕЛЬ

- Авторизованная зона;
- Защита каналов связи;



Защититься от внешнего нарушителя можно вне средств протокола. Основная угроза – внутренний нарушитель, защиту от которого должен обеспечивать протокол.

Инженерные ограничения

МАСШТАБИРУЕМОСТЬ

Банков много. Очень много. Производительность протокола не должна зависеть от количества участвующих в системе банков.

СТРОГИЙ SLA

Банки предъявляют строгие требования к производительности внешних сервисов. Консолидированное мнение: 1 секунда на ответ.

ИЗМЕНЧИВАЯ СРЕДА

Множество банков непостоянно. Банки могут отключаться, подключаться и “моргать”.

СЕТЬ

Сетевое взаимодействие всех со всеми требует квадратичного (в зависимости от числа банков) количества обменов сообщениями ☹

Можно сконструировать безопасный протокол, который не будет работать (или будет, но очень, очень долго) в реальном мире. Инженерные ограничения – важны!

В результате

ПРОТОКОЛ

Вместе с компанией КриптоПРО разработали протокол, который удовлетворяем всем пяти требованиям к безопасности.

РЕАЛИЗАЦИЯ

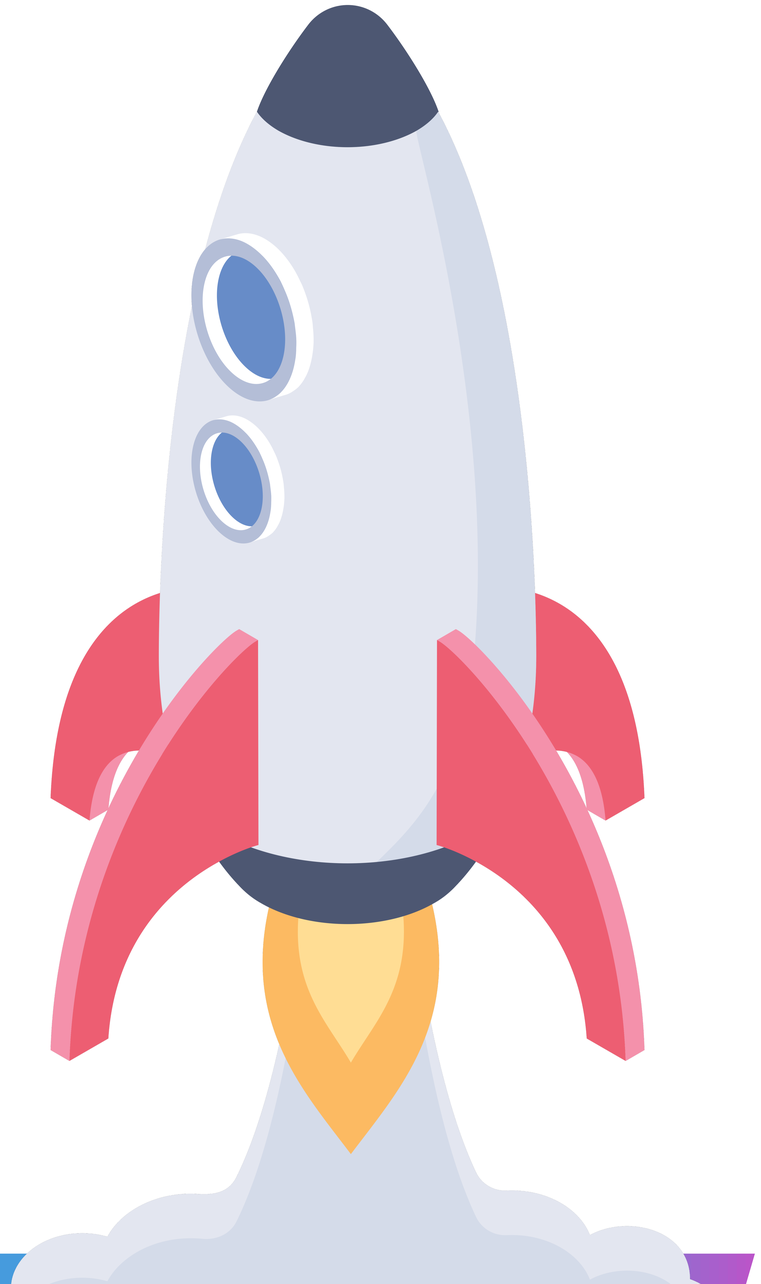
Реализовали программную инфраструктуру, которая учитывает все инженерные ограничения и делает протокол практически применимым.

ЭКСПЛУАТАЦИЯ

Вышли в прод. В настоящий момент 5 банков подключены к системе, еще 5+ находятся на разных фазах интеграционных работ.

СТАНДАРТИЗАЦИЯ

Вместе с компаниями КриптоПРО, Актив, СПБ, АФТ и QApp работаем над стандартом по конфиденциальным вычислениям в ТК26.



Выводы

МОДЕЛИ

Безопасность протокола в отношении подлежащих регуляции данных может быть доказана только в модели.

КОМПРОМИССЫ

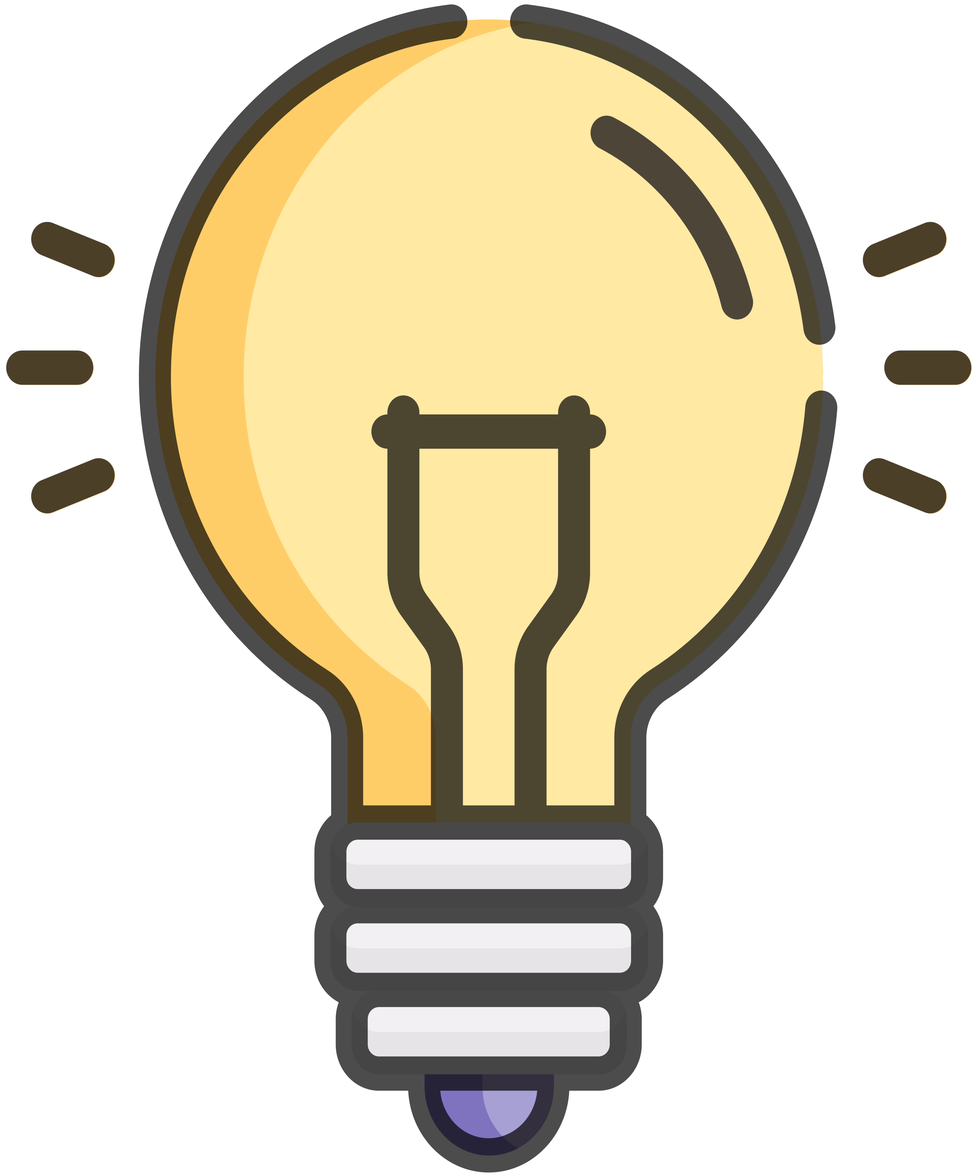
Строгость моделей нарушителя/угроз определяется компромиссом между потенциальной выгодой и потенциальным ущербом.

СПРОС

Рынок (причем не только банковский) демонстрирует запрос на системы конфиденциальных вычислений.

ПРЕДУБЕЖДЕНИЯ

”Новизна” таких систем – основной ограничивающий фактор их применения. Сотрудничество и открытый диалог нивелируют его влияние.



DATA FUSION

FIRST RUSSIAN DATA FORUM

Спасибо за внимание!

Обменивайтесь данными, не обмениваясь ими.

ЕМЕЛЬЯНОВ ПЕТР

ООО Блуттех, Генеральный Директор