Aggregate API with PPM

Mariana Raykova

What is MPC?

Secure Multiparty Computation (MPC)

Encryption, Signatures - protect data at rest and in transit

MPC - protect data while computing on it

➔ Enable the computation of a specific function without revealing more than the output

Aggregate Functionality

Differentially private histogram

Each client contributes a value to a bucket based on attributed impression and conversion data (**attribution done locally on device**)

Prio Secret Sharing Solution

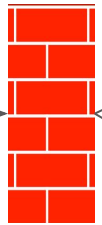
Private, Robust, and Scalable Computation of Aggregate Statistics,
Corrigan-Gibbs, Boneh, NSDI 2017

Ingestion Server Filter

1. Perform device attestation check on incoming data
2. Forward encrypted data to the two computation servers for reports with verified device attestation

Private Computation Server 2

AdTech



User's device

Input x_i

Encrypted shares and proofs

Encrypted shares and proofs

Encrypted shares and proofs

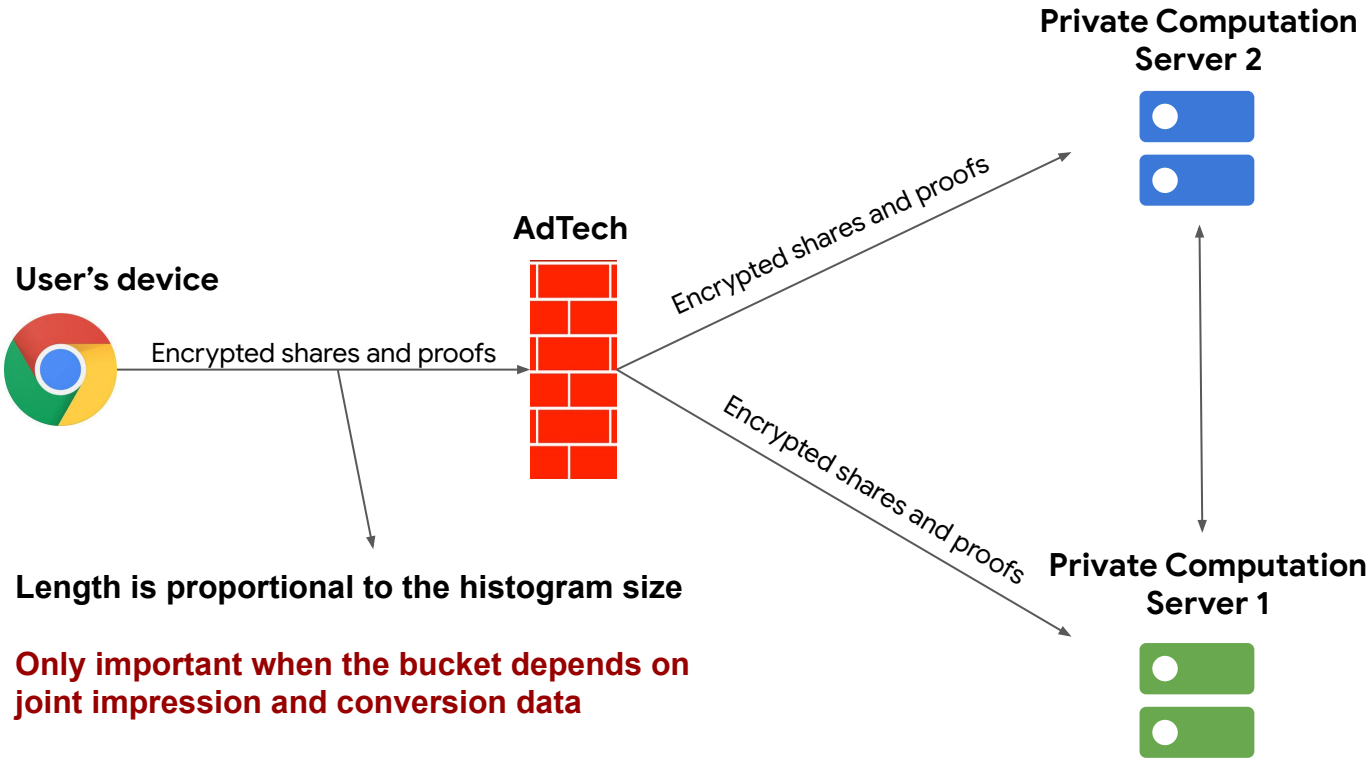
Distributed Private Computation between the Two Servers

1. Verify input validity proofs
2. Aggregate the verified inputs and add DP noise

Private Computation Server 1

Device Report Preparation

1. Split input into two cryptographic secret shares
   - $r_i^{HS} \leftarrow$ random; $r_i^{PHA} \leftarrow x_i - r_i^{HS}$
2. Compute a distributed validity proof for the input
   - Proof: π^{HS}, π^{PHA}
3. Encrypt shares and proofs under public keys of the computation servers
4. Send encrypted data together with device attestation information


Private Computation Server 2
AdTech
User's device
Encrypted shares and proofs
Encrypted shares and proofs
Encrypted shares and proofs
Private Computation Server 1
Length is proportional to the histogram size
Only important when the bucket depends on joint impression and conversion data

Privacy-preserving Firefox telemetry with Prio

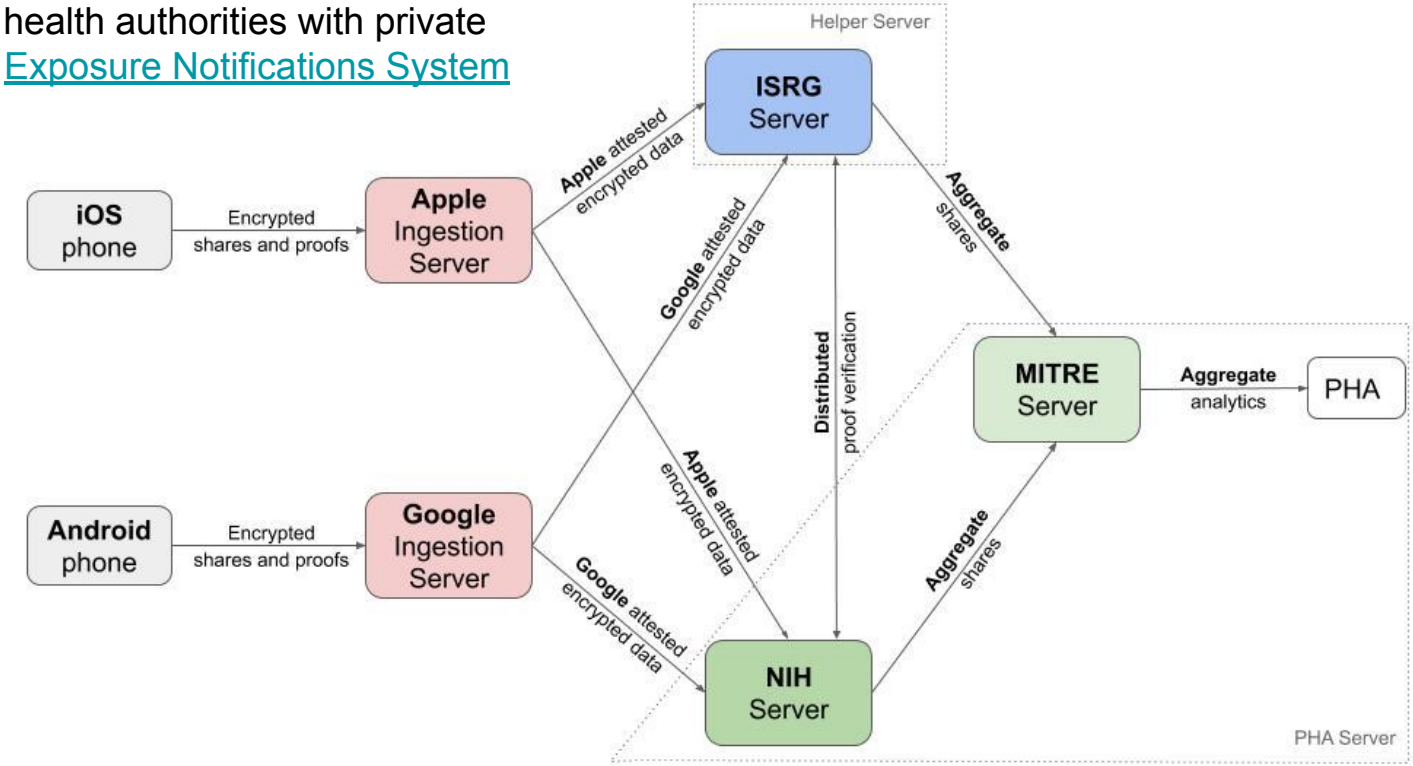
Henry Corrigan-Gibbs
(EPFL → MIT CSAIL)

In collaboration with: Dan Boneh (Stanford),
Gary Chen, Steven Englehardt, Robert Helmer, Chris Hutten-Czapski,
Anthony Miyaguchi, Eric Rescorla, and Peter Saint-Andre (Mozilla)

Real World Crypto 2020
https://rwc.iacr.org/2020/slides/Gibbs.pdf

Exposure Notifications Private Analytics (ENPA)

ENPA: Provide health authorities with private analytics in the Exposure Notifications System



How to reduce communication?

Distributed Point Function

Can encode the contribution value v to bucket i among N possible buckets with O(log N) bits **while hiding the index i where the client is contributing**

Servers do additional work to expand the compact shares

Incremental DPFs

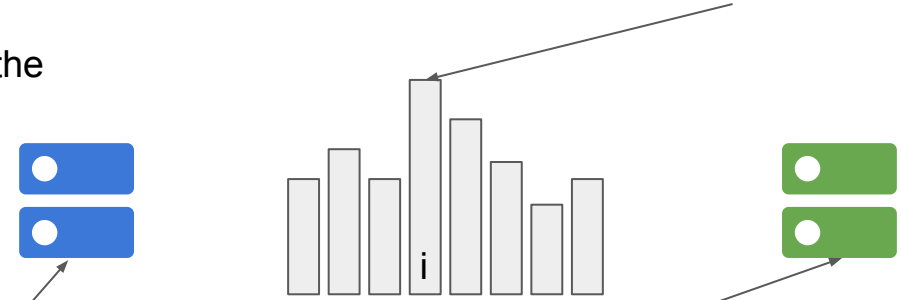
- Lightweight Techniques for Private Heavy Hitters, Boneh, Boyle, Corrigan-Gibbs, Gilboa, Ishai, IEEE S&P 2021

Construction

$\sum_m$ (Eval($K_{m,0}$, i) + Eval($K_{m,1}$, i)) = # clients with input i

To evaluate the share of the i-th location in the histogram compute

$\sum_m$ Eval($K_{m,0}$, i)



To evaluate the share of the i-th location in the histogram compute

$\sum_m$ Eval($K_{m,1}$, i)

$K_{m,0}$

$K_{m,1}$

Input i for m-th client : Construct $f_{i,x\_i}(x)$, generate DPF keys $K_{m,0}$, $K_{m,1}$

Some Benchmarks

https://github.com/google/distributed_point_functions

Domain space 2^{10}, 2^{15}, 2^{20} full evaluation

Key space size	2^{10}	2^{15}	2^{20}
Full Evaluation Cost Per Client	18.53µs	511.24µs	17.04ms

Domain space 2^{32} with 2^{20} non-zeros:

Distribution	PL1	PL5	PL10	PL50	Uniform
Hierarchical Evaluation Cost per Client	0.98s	1.36s	1.55s	2.42s	3.40s
Direct Evaluation Cost Per Client	0.66s	0.66s	0.67s	0.68s	0.70s

Privacy Preserving Measurement (PPM) at IETF

https://datatracker.ietf.org/doc/draft-gpew-priv-ppm/