Next level MPC

What comes after DAP

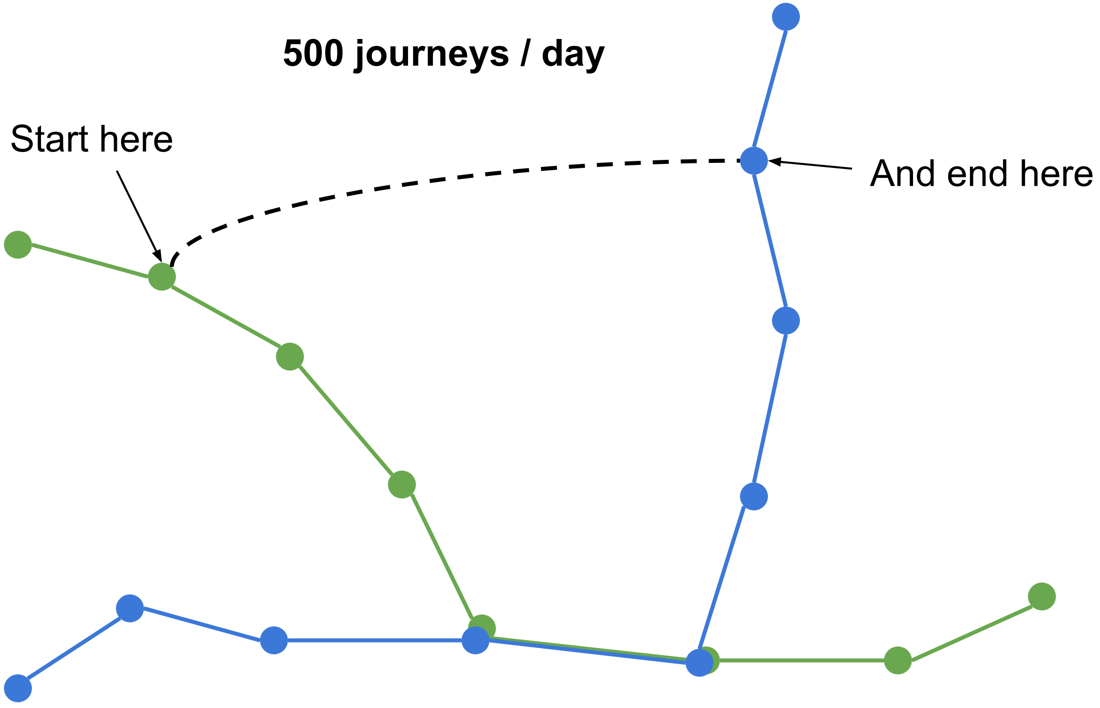
Here’s a use-case

Ad attribution

Website where I buy ads	Cost of ads on site	# Conversions attributed to these ads	Cost per conversion
foo.com	$571.5	127	$4.5
bar.com	$284.2	98	$2.9
baz.com	$220.8	32	$6.9
example.com	$255.2	58	$4.4
foo.example	$86.4	24	$3.6

Here’s a use-case (actually the same thing)

Bus Route Planning



Here’s a use-case (actually the same thing)

Foodborne Illness

Restaurant	**Percentage of customers who were diagnosed with food-poisoning within 2 days of a visit**
Han’s Chinese	2%
Joshua’s Steakhouse	1%
Bill’s Hamburgers	18%
Karina’s Tacos	3%
Guiseppe’s Pizza	2%

DAP is not enough

We've considered DAP as an option for this

DAP is best suited to asking simple questions of the data, and the zero-knowledge proofs get expensive quickly as you do more complex queries. What we are aiming to support:

- Some amount of adaptive querying
- Enable the training of machine learning models

We don't think DAP will address this use case in a sufficiently efficient manner

What's happened in the past 7 years

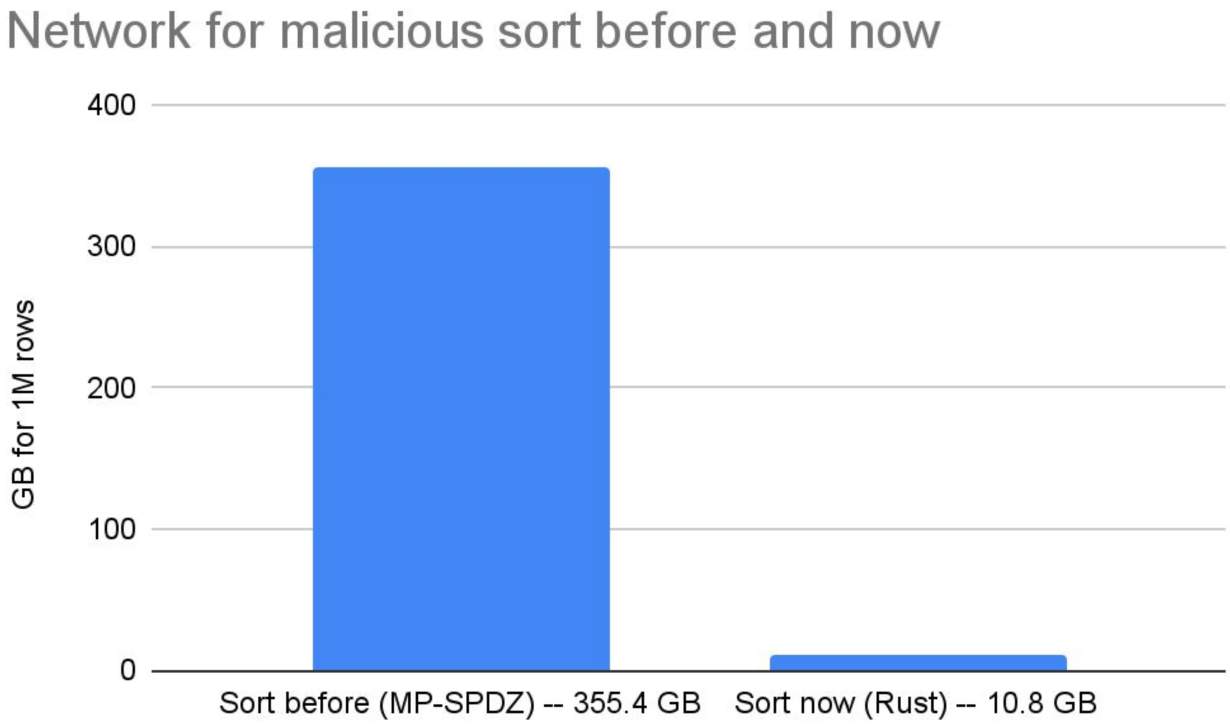
We've seen huge advances in MPC technology in cost and latency

- In particular **3-party, honest-majority** MPC using replicated secret-sharing
- Boolean circuit acceleration

Timeline

- 2016: High-Throughput Semi-Honest Secure Three-Party Computation with an Honest Majority https://dl.acm.org/doi/10.1145/2976749.2978331 (multiplication!)
- 2017: Prio https://crypto.stanford.edu/prio/paper.pdf
- 2018: High-throughput secure AES computation https://dl.acm.org/doi/10.1145/3267973.3267977 (faster)
- 2018: Efficient Bit-Decomposition and Modulus Conversion Protocols with an Honest Majority https://eprint.iacr.org/2018/387.pdf (faster)
- 2019: Malicious Security with Distributed Zero-Knowledge Proofs https://eprint.iacr.org/2019/188.pdf (sub-linear malicious security)
- 2019: Practical Fully Secure Three-Party Computation via Sublinear
- Distributed Zero-Knowledge Proofs https://eprint.iacr.org/2019/1390.pdf refined the approach (faster)
- 2023: Efficient 3PC for Binary Circuits with Application to Maliciously-Secure DNN Inference https://eprint.iacr.org/2023/909.pdf (boolean, faster)

Our Research Findings



<- Feb 2023

“Private Set Intersection” style solution with OPRF

Cost of processing 1 billion (10^9) rows:

- 760 Gb per MPC helper, 2.1 Tb total
- 40 min latency for 5 Gpbs network utilized at 50%
- Highly parallelizable approach - with 20 shards, ingress traffic per host is 40Gb

Malicious Security

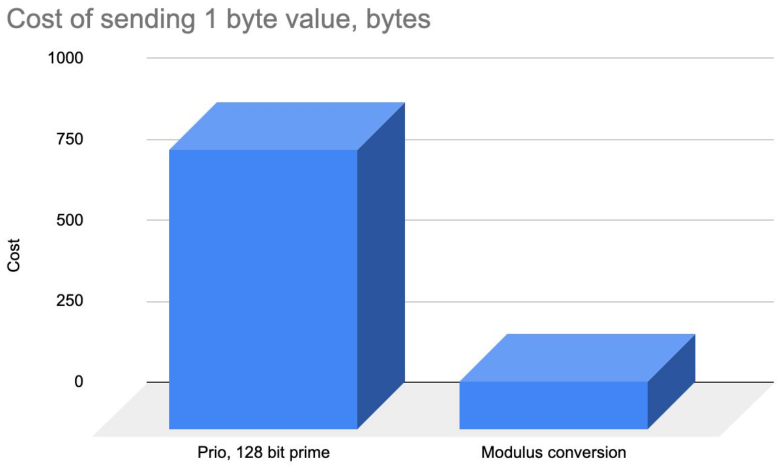
Before: 2.5x communication cost over semi-honest

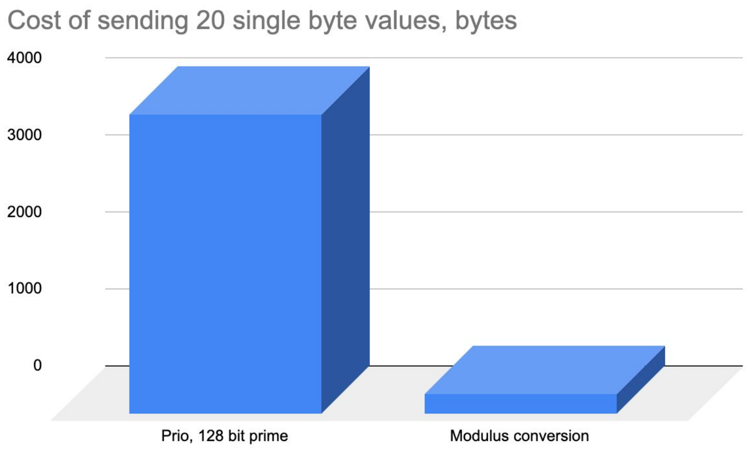
New: close to free with distributed fully-linear zero-knowledge proofs: O(log(N))

Comparison Point with DAP

Eliminates the need to round-trip so much data through the client

Using “modulus conversion” protocol allows us to do away with costly ZKPs





Status

Presently still a bit of a research project, but we have a working prototype

- With a few pieces yet to build
- Working on the details for differential privacy protections in MPC

IETF, we would like to work on this

- The protocol:
  - Private-Set-Intersection inspired ORPF-based protocol
  - Boolean MPC circuits evaluated per matching group
- Reusable protocol elements:
  - Helper Party initialization / setup
  - MPC Shuffle protocol
  - MPC-OPRF protocol
  - Additive DP noise generation
  - Share consistency check
- Foundational components:
  - Pseudo-Random Secret-Sharing (PRSS)
  - MPC Multiplication Protocol
  - Malicious Security Upgrade

Resources:

Github Repo: https://github.com/private-attribution/ipa/

An example of some DP analysis: Google Slides