Differential Privacy for DAP

Junye Chen, Christopher Patton, Shan Wang

IETF 117 – PPM
Differential Privacy (DP) in a slide

- Keeping the measurements private (as DAP does) may not be enough: the aggregate result may leak (bits of) an individual measurement
  - Motivating example: Average height of a group of people with or without an especially tall (or short) individual

- Differential Privacy (DP): the aggregate result (or, more generally, the adversary's view) should not change significantly if any one measurement is replaced by another
  - Achieved by adding noise to:
    - the measurements by the Clients;
    - the aggregate shares by the Aggregators; and/or
    - the aggregate result by the Collector
DP is simple in principle, but…

- Mechanisms (e.g., sampling from a discrete Gaussian distribution) are easy to implement incorrectly
  - Implementation errors can violate DP [Mir’12, CSVW’22, JMRO’22]
  - Side channels leak the noise [JMRO’22]
- For a given VDAF, choosing the optimal DP policy requires careful analysis
- Protocol details (of DAP) can impact analysis
  - auxiliary data, metadata, intersection attacks…
- DP literature provides little guidance on tuning privacy parameters (e.g., $\epsilon$, $\delta$)
  - Roughly speaking, more noise $\Rightarrow$ better privacy, but worse utility: how much noise is enough? Requires a lot of expertise.

Proposal: draft-wang-ppm-differential-privacy

- Plan: Write the draft, then recommend changes to DAP/VDAF drafts as needed
- Status of draft: early stages (we missed the draft deadline). We're currently looking for contributors.
- Open questions:
  - Working (or research) group
  - Scope
Scope of the draft

- Overview of relevant DP literature (and constraints imposed by DAP)
  - Remove- versus Substitution-DP (DAP leaks the report count to Aggregators/Collector)
  - User-level DP (no mechanism in DAP prevents Client from uploading the same measurement multiple times)
  - Metadata (report timestamp, client IP)
- Mechanisms (for building DP "policies")
  - Guidance for mitigating side channels
  - Extensive test vectors
- Policies (for concrete VDAFs)
- DAP integration
Out-of-scope (for now?)

- Privacy parameter tuning (requires domain expertise)
- More sophisticated DP mechanisms [DKM+'06, KKL+'23]
- Non-VDAF applications (though the low-level mechanisms should be useful in different contexts)

DAP integration

```python
class DpPolicy:
    Measurement
    AggregateShare
    AggregateResult

def add_noise_to_measurement(self, meas: Measurement,)
    -> Measurement:
        """Add noise to measurement, if required."
        pass

def add_noise_to_agg_share(self, agg_share: AggregateShare,
    meas_count: Unsigned, min_batch_size: Unsigned,
    ) -> AggregateShare:
        """Add noise to aggregate share, if required."
        pass

def debias_agg_result(self, agg_result: AggregateResult,
    meas_count: Unsigned, min_batch_size: Unsigned,
    num_aggregators: Unsigned,
    ) -> AggregateResult:
        """De-bias the aggregate result, if required."
        pass
```

- **Protocol changes**
  - New Client, Aggregator, Collector behavior for executing policy
  - (Maybe?) Minimizing report metadata (i.e., removing the timestamp) may help support security analysis
  - (Maybe?) Early collection: Some DP policies allow collecting an aggregate result before min batch size is met **without violating DP**

- **Requirements**
  - One DP mechanism per task
  - Aggregators noise an aggregate share at most once
  - Clients enforce privacy budget (how many times they will upload a measurement)
Fin.

- Questions or thoughts about scope?
- **Question:** which working (or research) group?
  - PPM (we have a vested interest, but perhaps not the expertise)
  - CFRG (DP is closely related to cryptography, but often viewed as a separate discipline)
  - DISPATCH (perhaps there are other groups at IETF working on the same problem)
  - Anywhere else?