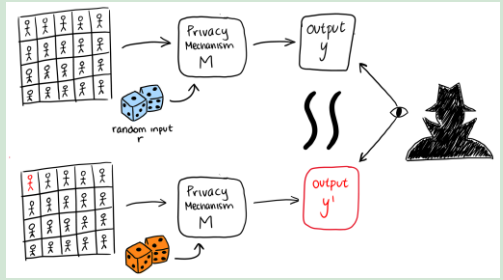
Bounded-Leakage Differential Privacy

Katrina Ligett, Charlotte Peale, Omer Reingold
katrina@cs.huji.ac.il cpeale@stanford.edu reingold@stanford.edu

Differential Privacy

A mathematically rigorous definition of privacy for individuals included in statistical computations [Dwork et al., 2006]



Intuitively, the output of a computation should not depend too much on the data of one particular individual.

DP + Auxiliary Information

DP guarantees are *relative*—don't depend on existing background information.

But, arbitrary auxiliary information combined with differentially private results can result in substantial privacy harms [Dwork and Naor, 2010]

Would it be useful to have a variant of DP that could deal with some auxiliary information?

Goals

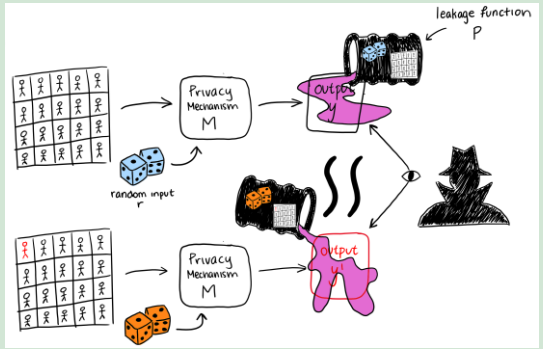
Define a relaxed variant of differential privacy that can give further insights about privacy when we know an upper bound on potential auxiliary (leaked) information.

Ideally should satisfy useful properties such as composability, post-processing, and meaningful conversions to standard differential privacy.

Bounded-leakage Differential Privacy

We introduce an additional *leakage function* to the standard definition of differential privacy, which defines the type of auxiliary information being considered.

The leakage function takes in the same database and random input as the privacy mechanism.



Regardless of external knowledge, after seeing the leakage, an adversary who sees the output of the privacy mechanism draws the same conclusions whether or not a particular individual's data was included in the original database.

A privacy mechanism $M: X^n \times R \to O_M$ is

(ϵ, δ)-*bounded-leakage* differentially private

with respect to leakage $P: X^n \times R \to O_P$ if for all neighboring databases $x \sim x' \in X^n$, subsets $S \subseteq O_M$, and leakage $o \in O_P$ such that $\Pr_{r \in R}[P(x,r) = o] \cdot \Pr_{r \in R}[P(x',r) = o] \neq 0$, we have

$$\Pr_{r \in R}[M(x,r) \in S | P(x,r) = o] \leq e^{\epsilon} \Pr_{r \in R}[M(x',r) \in S | P(x',r) = o] + \delta$$

Properties

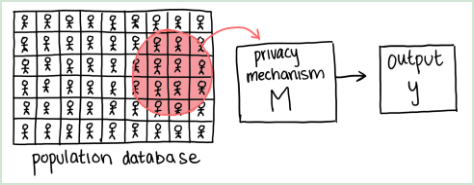
Bounded-leakage privacy satisfies post-processing, group privacy, and composability properties analogous to standard differential privacy.

Explicitly "leaking" the value of the leakage function does not affect bounded-leakage differential privacy.

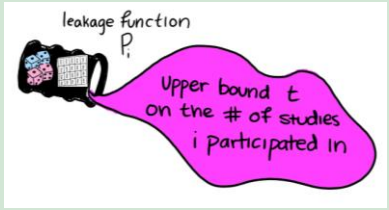
Scenario 2: Non-participation

Can my privacy be degraded by studies that I *didn't* participate in?

If we model this question by assuming a database contains an entire population, differential privacy tells us that individuals will incur a privacy loss for *every* study run on the population.



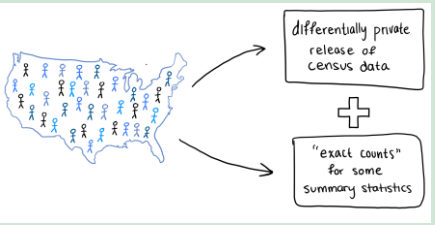
Introducing a leakage function that gives an upper bound on the number of studies that an individual participated in gives tighter privacy guarantees.



Privacy loss can now be expressed in terms of the upper bound rather than the total number of studies over the entire population.

Scenario 1: 2020 Census

How would the release of "exact counts" affect the privacy of census participants? [Garfinkel, 2018]



When the leakage and mechanism output are independent, bounded-leakage privacy preserves the same bounds as the differentially private mechanism.

So, releasing exact counts won't degrade census participants' privacy in unexpected ways.

Future Directions

It would be interesting to develop additional mechanisms that enjoy bounded-leakage privacy and to apply the notion in new domains.

Variations on the definition of bounded-leakage privacy deserve further exploration.

One possibility: a variant that allows weakened privacy on low-probability leakage outputs.

Acknowledgements

We are grateful to Daniel Kifer, Ashwin Machanavajjhala, and Adam Smith for assisting us in gaining a better understanding of their related work and existing privacy literature.

Part of this work was done while the first and third authors were visiting the Simons Institute for the Theory of Computing.

This work was supported in part by NSF Award IIS-1908774, the VMWare fellowship, Israel Science Foundation (ISF) grant #1044/16, United States Air Force and DARPA under contracts FA8750-16-C-0022 and FA8750-19- 2-0222, and the Federmann Cyber Security Center in conjunction with the Israel national cyber directorate. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Air Force and DARPA: Supported in part by NSF Award IIS-1908774 and by VMWare fellowship.