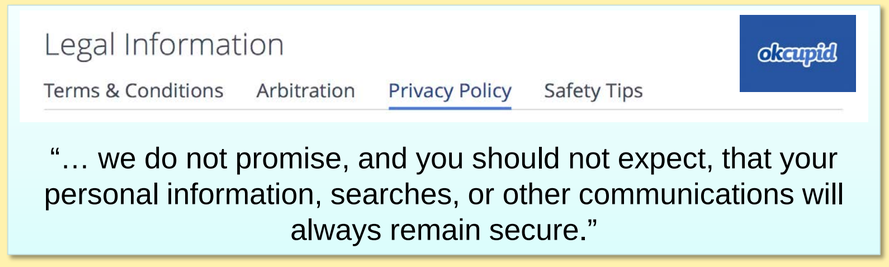
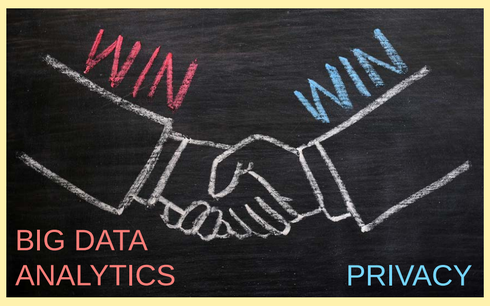
TWC SBE: Small: Towards an Economic Foundation of Privacy-Preserving Data Analytics: Incentive Mechanisms and Fundamental Limits (Award #: 1618768)

PIs: Lei Ying and Junshan Zhang, Arizona State University

http://inlab.lab.asu.edu/data-privacy/

Big Data versus Privacy: The commoditization of private data has been trending up. It is becoming increasingly difficult to know how data may be used, or to retain control over data about oneself. The common practices of collecting private data are becoming untenable, with vague privacy policies and a behind-the-scenes data brokerage market becoming the norm.





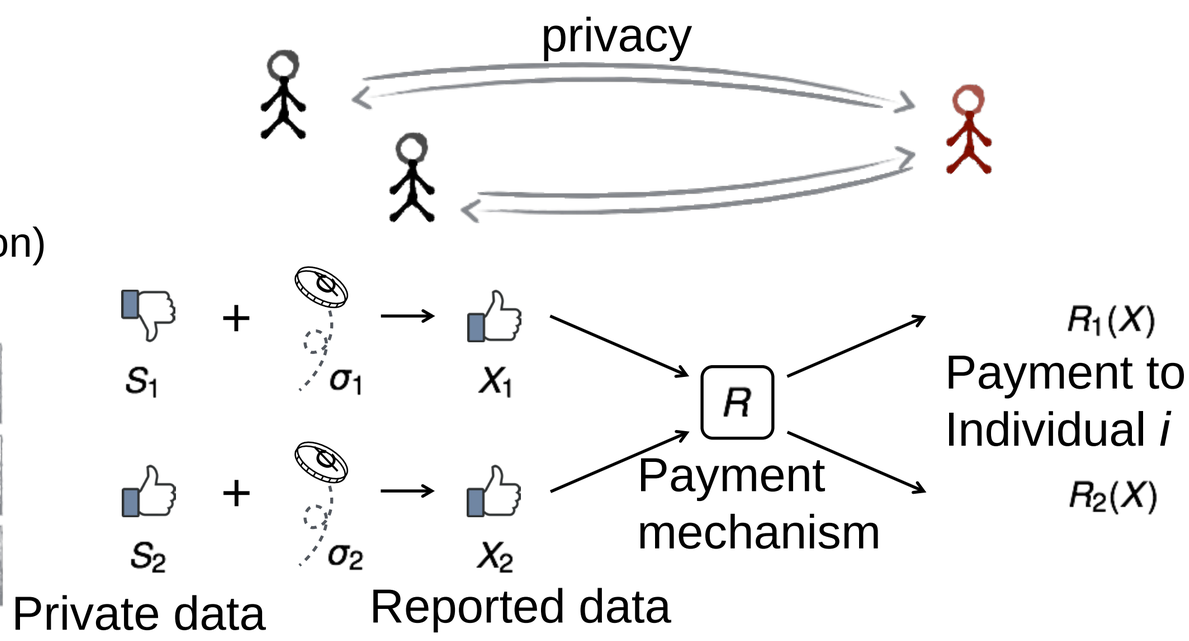
Proposed approach: A market model for private data analytics

Individuals (Data Subjects)

Maximize utility (non-convex optimization)

Best responses

- prior distribution
- privacy cost model
- payment mechanism



Data Collector

Minimize cost (payment) s.t. mechanisms that have the desired NE

Mechanism design

- prior distribution
- privacy cost model

Untrustworthy

- Individuals control their own privacy
- Data collector cannot observe strategies

Game-theoretic approach

- Individuals are self-interested players
- Strategy: how to perturb private data
- Utility: payment – privacy cost
- Strategy profile σ is a Nash equilibrium (NE) of the payment mechanism R if for any individual i and any strategy σ_i'

$$\mathbb{E}_{\sigma}[R_i(X) - g(\zeta(\sigma_i))] \geq \mathbb{E}_{(\sigma_i', \sigma_{-i})}[R_i(X) - g(\zeta(\sigma_i'))].$$

- A payment mechanism obtains ϵ privacy from individual i if σ_i at a NE has privacy level ϵ.
  - Denote the set of such mechanisms by $\mathcal{R}(i;\epsilon)$.

Fundamental questions

What is the minimum payment to obtain ϵ privacy from an individual?

The value of privacy: $V(\epsilon) = \inf_{R \in \mathcal{R}(i;\epsilon)} \mathbb{E}_{\sigma(R;\epsilon)}[R_i(X)]$

- Tradeoff between privacy and cost
- Characterizes the balance point of the market

Which payment mechanism can achieve the above minimum cost? Optimal mechanism

Payment–accuracy tradeoff [1]

- The data collector is interested in learning the underlying state W
- Hypothesis testing $H_0: W = 0,$ $H_1: W = 1.$

Goal: min Total expected payment $\rightarrow F(\tau)$ subject to Error $\leq \tau$

Theorem

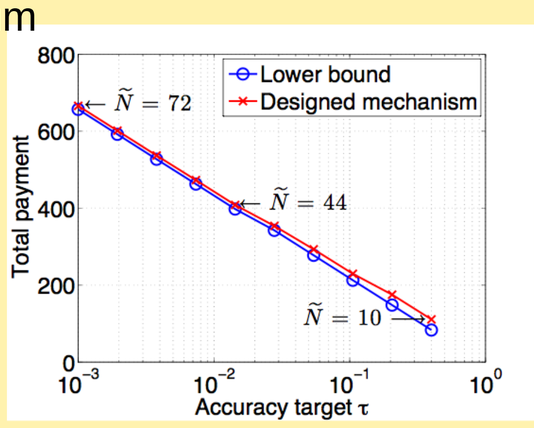
The optimal payment in the payment–accuracy problem satisfies

$$(\tilde{N} - 1) V_{\mathrm{LB}}(\tilde{\epsilon}) \leq F(\tau) \leq \tilde{N} V_{\mathrm{LB}}(\tilde{\epsilon}) + O(\tau \ln(1/\tau)).$$

$$V_{\mathrm{LB}}(\epsilon) = g'(\epsilon) \frac{e^{\epsilon} + 1}{e^{\epsilon}} \left(\frac{\theta}{2\theta - 1}(e^{\epsilon} + 1) - 1 \right).$$

with properly chosen $\tilde{N}$ and $\tilde{\epsilon}$

- At most one individual's payment away from the minimum



[1] W. Wang, L. Ying, and J. Zhang. The Value of Privacy: Strategic Data Subjects, Incentive Mechanisms and Fundamental Limits. Proc. ACM SIGMETRICS, Antibes Juan-les-Pins, France, June, 2016. (**the Kenneth C. Sevcik Outstanding Student Paper Award**)

Interested in meeting the PIs? Attach post-it note below!