Optimal Privacy-Constrained Mechanisms

Ran Eilat, Kfir Eliaz and Xiaosheng Mu*

AEA Meetings

Jan. 3, 2020

Motivation

Introduce and analyze a Bayesian measure of privacy loss.

- Review of "ϵ-differential privacy" (Dwork et al. '06, Pai-Roth '13): outcome does not reveal much about any individual's private info.

Motivation

Introduce and analyze a Bayesian measure of privacy loss.

- Review of "ϵ-differential privacy" (Dwork et al. '06, Pai-Roth '13): outcome does not reveal much about any individual's private info.

  $$\mathbb{P}\{x \text{ is chosen given } t\} \leq \exp(\epsilon) \cdot \mathbb{P}\{x \text{ is chosen given } t'\}$$

  for all outcomes x and type profiles t and t' differing in a single agent.

Motivation

Introduce and analyze a Bayesian measure of privacy loss.

- Review of "ϵ-differential privacy" (Dwork et al. '06, Pai-Roth '13): outcome does not reveal much about any individual's private info.

$$\mathbb{P}\{x \text{ is chosen given } t\} \leq \exp(\epsilon) \cdot \mathbb{P}\{x \text{ is chosen given } t'\}$$

  for all outcomes x and type profiles t and t' differing in a single agent.

- DP uses outsider's perspective and is *prior-free*.

- But to implement this, types have to be reported.
  We might worry about the principal knowing too much.

Motivation

Introduce and analyze a Bayesian measure of privacy loss.

- Review of "ϵ-differential privacy" (Dwork et al. '06, Pai-Roth '13): outcome does not reveal much about any individual's private info.

$$\mathbb{P}\{x \text{ is chosen given } t\} \leq \exp(\epsilon) \cdot \mathbb{P}\{x \text{ is chosen given } t'\}$$

  for all outcomes x and type profiles t and t' differing in a single agent.

- DP uses outsider's perspective and is *prior-free*.

- But to implement this, types have to be reported.
  We might worry about the principal knowing too much.

- Our approach: mechanism design under a privacy constraint that *limits how much information the principal can collect from the agents*.

Screening Environment

Focus on the single-agent screening model of Mussa-Rosen '78.

- A seller sells some quantity/quality $q \geq 0$ to a buyer for payment p.
- Buyer type $\theta \in [\underline{\theta}, \overline{\theta}]$ distributed as F with positive density.
- Buyer utility $q \cdot \theta - p$.
- Production cost $\frac{q^2}{2}$; seller profit $p - \frac{q^2}{2}$.

Screening Environment

Focus on the single-agent screening model of Mussa-Rosen '78.

- A seller sells some quantity/quality $q \geq 0$ to a buyer for payment p.
- Buyer type $\theta \in [\underline{\theta}, \overline{\theta}]$ distributed as F with positive density.
- Buyer utility $q \cdot \theta - p$.
- Production cost $\frac{q^2}{2}$; seller profit $p - \frac{q^2}{2}$.
- Assume increasing and positive virtual values $v(\theta) := \theta - \frac{1 - F(\theta)}{f(\theta)}$.
  - positive ensures participation; can be relaxed
- Mussa-Rosen showed that optimal mechanism perfectly separates types:
  - type θ receives quantity $v(\theta)$
  - payment given by envelope theorem

Privacy Measure

We depart by adding a (privacy) constraint to seller's problem:

1. Seller has *prior belief* F about buyer type θ.
2. He offers general (potentially indirect) mechanism with message set M, allocation function $q : M \to \mathbb{R}^+$ and payment function $p : M \to \mathbb{R}$.
3. Each buyer θ sends message $m(\theta)$ to maximize EU given $q(\cdot)$, $p(\cdot)$.
4. Observing message m, seller forms *posterior belief* $F(\theta \mid m)$ about θ.
5. Will put a constraint on how posterior changes relative to prior.

Constrained Problem

- Privacy loss of a mechanism $\mathbb{M}$ defined as maximum (across messages) KL-divergence between posterior and prior beliefs:

  $$I(\mathbb{M}) = \max_{m} \ D(F(\cdot \mid m) \,||\, F),$$

  where $D(P \,||\, Q) = \displaystyle\int \log\left(\frac{dP}{dQ}\right) \, dP.$

  - results extend to general divergences

- Maximize profit among mechanisms s.t. $I(\mathbb{M}) \leq \kappa$ (exogenously given).

Interpretation

- *Paternalistic view*: a regulator imposes a constraint of this form to protect consumer privacy.

Interpretation

- *Paternalistic view*: a regulator imposes a constraint of this form to protect consumer privacy.

- *Participation constraint*: each buyer type tolerates privacy loss up to KL-distance of κ.

Interpretation

- *Paternalistic view*: a regulator imposes a constraint of this form to protect consumer privacy.

- *Participation constraint*: each buyer type tolerates privacy loss up to KL-distance of κ.

- We use KL as a reduced-form measure of seller's information gain.
  - prior works Taylor (2004), Calzolari and Pavan (2006) model agents who value privacy due to specific future interactions with principal
  - our approach is applicable if *future interactions are unknown* ("context-free")

Ex-post vs. Ex-ante

- Above definition $I(\mathbb{M}) = \max_m D(F(\cdot \mid m) \,||\, F)$ considers worst-case privacy loss across all messages (thus types).

- Alternatively, may require average loss $\mathbb{E}_m[D(F(\cdot \mid m) \,||\, F)] \leq \kappa$.
  - relates to rational inattention since average KL is equal to MI

- Ex-post criterion is stricter and fits better with above interpretations. But similar results hold for the ex-ante model (see paper)

Main Result

Theorem (Coarse Revelation)

Given $0 < \kappa < \infty$. There exists an optimal privacy-constrained mechanism $\mathbb{M}$, where the set of types $[\underline{\theta}, \overline{\theta}]$ is partitioned into finitely many intervals, and in equilibrium each type truthfully reports its interval.

Why Intervals?

Several papers (Bergemann et al. '11, Kos '12) derived optimality of intervals by assuming an upper bound on *number* of messages.

We put upper bound on *informational content* of each message.

Proof of interval partition structure:

Why Intervals?

Several papers (Bergemann et al. ’11, Kos ’12) derived optimality of intervals by assuming an upper bound on *number* of messages.

We put upper bound on *informational content* of each message.

Proof of interval partition structure:

1. First remove “redundant” messages. If two messages lead to same outcome, combine them into a single message.

   $\Longrightarrow$ posterior belief is *averaged*, hence smaller privacy loss by convexity

Why Intervals?

Several papers (Bergemann et al. '11, Kos '12) derived optimality of intervals by assuming an upper bound on *number* of messages.

We put upper bound on *informational content* of each message.

Proof of interval partition structure:

1. First remove "redundant" messages. If two messages lead to same outcome, combine them into a single message.
   $\Longrightarrow$ posterior belief is *averaged*, hence smaller privacy loss by convexity
2. Messages are ranked by the quantities they induce.

Why Intervals?

Several papers (Bergemann et al. '11, Kos '12) derived optimality of intervals by assuming an upper bound on *number* of messages.

We put upper bound on *informational content* of each message.

Proof of interval partition structure:

1. First remove "redundant" messages. If two messages lead to same outcome, combine them into a single message.
   $\Longrightarrow$ posterior belief is *averaged*, hence smaller privacy loss by convexity
2. Messages are ranked by the quantities they induce.
3. By single-cross property of buyer preference, types that choose a particular quantity (and associated price) form an interval.

Why Intervals?

Several papers (Bergemann et al. '11, Kos '12) derived optimality of intervals by assuming an upper bound on *number* of messages.

We put upper bound on *informational content* of each message.

Proof of interval partition structure:

1. First remove "redundant" messages. If two messages lead to same outcome, combine them into a single message.
   $\implies$ posterior belief is *averaged*, hence smaller privacy loss by convexity
2. Messages are ranked by the quantities they induce.
3. By single-cross property of buyer preference, types that choose a particular quantity (and associated price) form an interval.
4. Distinct intervals can only intersect at the boundary.

Why Intervals?

Several papers (Bergemann et al. '11, Kos '12) derived optimality of intervals by assuming an upper bound on *number* of messages.

We put upper bound on *informational content* of each message.

Proof of interval partition structure:

1. First remove "redundant" messages. If two messages lead to same outcome, combine them into a single message.
   $\Longrightarrow$ posterior belief is *averaged*, hence smaller privacy loss by convexity
2. Messages are ranked by the quantities they induce.
3. By single-cross property of buyer preference, types that choose a particular quantity (and associated price) form an interval.
4. Distinct intervals can only intersect at the boundary.
5. Thus interval partition — this only uses convexity of privacy measure. Extends also to multiple agents with one-dimensional types.

Reformulation

Recall KL-divergence defined as $D(P \,||\, Q) = \int \log\left(\frac{dP}{dQ}\right) \, dP$.

- When P is given by Q conditional on an interval $[\theta_1, \theta_2]$, we have

$$D(P \,||\, Q) = -\log\left(Q([\theta_1, \theta_2])\right).$$

Reformulation

Recall KL-divergence defined as $D(P \,||\, Q) = \int \log\left(\frac{dP}{dQ}\right)\, dP$.

- When P is given by Q conditional on an interval $[\theta_1, \theta_2]$, we have

$$D(P \,||\, Q) = -\log\left(Q([\theta_1, \theta_2])\right).$$

- Maximize profit among partitions s.t. each interval has mass $\geq e^{-\kappa}$.

- Corollaries:
  - For each κ, optimal mechanism exists (by compactness).

Reformulation

Recall KL-divergence defined as $D(P \,||\, Q) = \int \log\left(\frac{dP}{dQ}\right)\, dP$.

- When P is given by Q conditional on an interval $[\theta_1, \theta_2]$, we have

$$D(P \,||\, Q) = -\log\left(Q([\theta_1, \theta_2])\right).$$

- Maximize profit among partitions s.t. each interval has mass $\geq e^{-\kappa}$.

- Corollaries:
  - For each κ, optimal mechanism exists (by compactness).
  - $0 \leq \kappa < \log(2) \qquad \Longrightarrow$ one interval, no screening

Reformulation

Recall KL-divergence defined as $D(P \,||\, Q) = \int \log\left(\frac{dP}{dQ}\right) \, dP$.

- When P is given by Q conditional on an interval $[\theta_1, \theta_2]$, we have

$$D(P \,||\, Q) = -\log\left(Q([\theta_1, \theta_2])\right).$$

- Maximize profit among partitions s.t. each interval has mass $\geq e^{-\kappa}$.

- Corollaries:
  - For each κ, optimal mechanism exists (by compactness).
  - $0 \leq \kappa < \log(2) \implies$ one interval, no screening
  - $\log(2) \leq \kappa < \log(3) \implies$ two intervals

Reformulation

Recall KL-divergence defined as $D(P \,||\, Q) = \int \log\left(\frac{dP}{dQ}\right)\, dP$.

- When P is given by Q conditional on an interval $[\theta_1, \theta_2]$, we have

$$D(P \,||\, Q) = -\log\left(Q([\theta_1, \theta_2])\right).$$

- Maximize profit among partitions s.t. each interval has mass $\geq e^{-\kappa}$.

- Corollaries:
  - For each κ, optimal mechanism exists (by compactness).
  - $0 \leq \kappa < \log(2) \implies$ one interval, no screening
  - $\log(2) \leq \kappa < \log(3) \implies$ two intervals
  - Privacy constraint does not in general bind

Uniform Case

Consider special case with *uniform* prior F.

Characterization

With uniform prior, for any $\log(n) \leq \kappa < \log(n+1)$, the optimal privacy-constrained mechanism partitions $[\underline{\theta}, \overline{\theta}]$ into n equally long intervals.

Uniform Case

Consider special case with *uniform* prior F.

Characterization

With uniform prior, for any $\log(n) \leq \kappa < \log(n+1)$, the optimal privacy-constrained mechanism partitions $[\underline{\theta}, \overline{\theta}]$ into n equally long intervals.

Proof:

1. Since $\kappa < \log(n+1)$, each interval has mass at least $e^{-\kappa} > \dfrac{1}{n+1}$.
2. There can be at most n intervals.
3. Equal partition maximizes profit among *all* partitions of size n.

Welfare Analysis

Comparative Statics w.r.t. κ

- Profit from a κ-constrained optimal mechanism (weakly) increases in κ.

Welfare Analysis

Comparative Statics w.r.t. κ

- Profit from a κ-constrained optimal mechanism (weakly) increases in κ.
- Buyer surplus is maximized with "full privacy" $\kappa = 0$, and minimized with "no privacy" $\kappa = \infty$.

Welfare Analysis

Comparative Statics w.r.t. κ

- Profit from a κ-constrained optimal mechanism (weakly) increases in κ.
- Buyer surplus is maximized with "full privacy" $\kappa = 0$, and minimized with "no privacy" $\kappa = \infty$.
- If prior density $f(\theta)$ decreases, $\kappa = \infty$ maximizes total welfare.

Future Work

- Further properties of optimal interval partition for general prior F:
  - Is the optimal number of intervals increasing in κ?
  - Is buyer surplus decreasing in κ?
- Regulation: how to elicit seller's prior and choose κ accordingly?
- Multiple agents: how to aggregate privacy?

Thank You!