Differentially Private Policy Evaluation

Borja de Balle Pigem

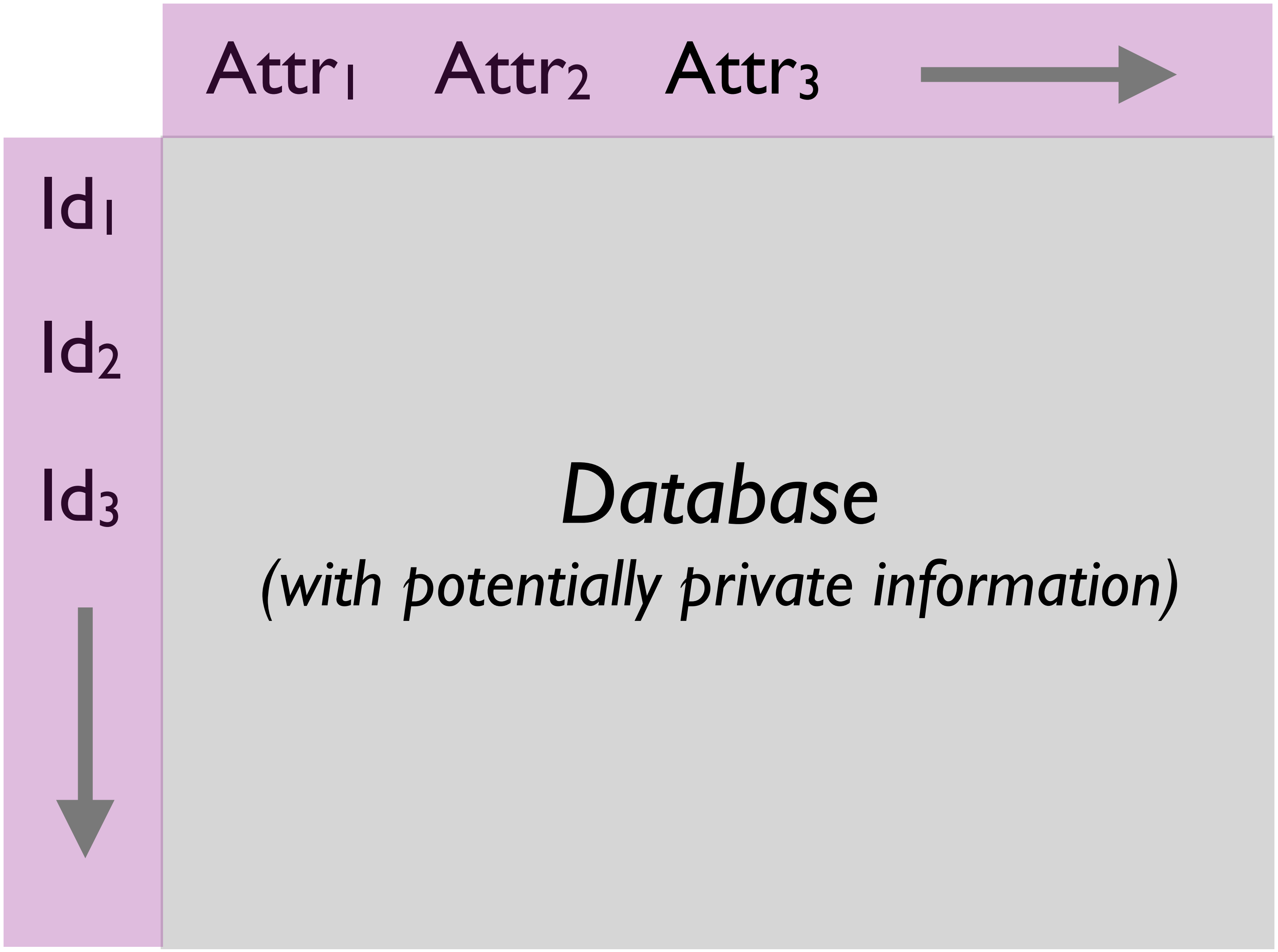
Data Science | Lancaster University

Mathematics & Statistics | Lancaster University

[CSML — Feb 18, 2016]

Part 1: Introduction to Differential Privacy

Data Science in the Big Data Era



The Privacy ~~Challenge~~ Nightmare

The Privacy ~~Challenge~~ Nightmare

The New York Times

Technology

A Face Is Exposed for AOL Searcher No. 4417749

By MICHAEL BARBARO and TOM ZELLER Jr.
Published: August 9, 2006

SIGN IN TO E-MAIL THIS
PRINT
SINGLE PAGE
REPRINTS

Buried in a list of 20 million Web search queries collected by AOL and recently released on the Internet is user No. 4417749. The number was assigned by the company to protect the searcher's anonymity, but it was not much of a shield.

Erik S. Lesser for The New York Times
Thelma Arnold's identity was betrayed by AOL records of her Web searches, like ones for her dog, Dudley, who clearly has a problem.

No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from "numb fingers" to "60 single men" to "dog that urinates on everything."

And search by search, click by click, the identity of AOL user No. 4417749 became easier to discern. There are queries for "landscapers in Lilburn, Ga," several people with the last name Arnold and "homes sold in shadow lake subdivision gwinnett county georgia."

It did not take much investigating to follow that data trail to Thelma Arnold, a 62-year-old widow who lives in Lilburn, Ga., frequently researches her friends' medical ailments and loves her three dogs. "Those are my searches," she said, after a reporter read part of the list to her.

The Privacy ~~Challenge~~ Nightmare

The New York Times

Technology

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION

AUTOS

CAMCORDERS CAMERAS CELLPHONES COMPUTERS HANDHELDS HOME VIDEO MUSIC PERIPHER

A Face Is Exposed for AOL Searcher No. 4417749

By MICHAEL BARBARO and TOM ZELLER Jr.
Published: August 9, 2006

SIGN IN TO E-MAIL THIS
PRINT
SINGLE PAGE
REPRINTS

Buried in a list of 20 million Web search queries collected by AOL and recently released on the Internet is user No. 4417749. The number was assigned by the company to protect the searcher's anonymity, but it was not much of a shield.

Erik S. Lesser for The New York Times

Thelma Arnold's identity was betrayed by AOL records of her Web searches, like ones for her dog, Dudley, who clearly has a problem.

No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from "numb fingers" to "60 single men" to "dog that urinates on everything."

And search by search, click by click, the identity of AOL user No. 4417749 became easier to discern. There are queries for "landscapers in Lilburn, Ga," several people with the last name Arnold and "homes sold in shadow lake subdivision gwinnett county georgia."

It did not take much investigating to follow that data trail to Thelma Arnold, a 62-year-old widow who lives in Lilburn, Ga., frequently researches her friends' medical ailments and loves her three dogs. "Those are my searches," she said, after a reporter read part of the list to her.

Robust De-anonymization of Large Sparse Datasets

Arvind Narayanan and Vitaly Shmatikov
The University of Texas at Austin

Abstract

We present a new class of statistical de-anonymization attacks against high-dimensional micro-data, such as individual preferences, recommendations, transaction records and so on. Our techniques are robust to perturbation in the data and tolerate some mistakes in the adversary's background knowledge.

We apply our de-anonymization methodology to the Netflix Prize dataset, which contains anonymous movie ratings of 500,000 subscribers of Netflix, the world's largest online movie rental service. We demonstrate that an adversary who knows only a little bit about an individual subscriber can easily identify this subscriber's record in the dataset. Using the Internet Movie Database as the source of background knowledge, we successfully identified the Netflix records of known users, uncovering their apparent political preferences and other potentially sensitive information.

The Privacy ~~Challenge~~ Nightmare

The New York Times

Technology

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION

AUTOS

CAMCORDERS CAMERAS CELLPHONES COMPUTERS HANDHELDS HOME VIDEO MUSIC PERIPHER

A Face Is Exposed for AOL Searcher No. 4417749

By MICHAEL BARBARO and TOM ZELLER Jr.
Published: August 9, 2006

SIGN IN TO E-MAIL THIS
PRINT
SINGLE PAGE
REPRINTS

Buried in a list of 20 million Web search queries collected by AOL and recently released on the Internet is user No. 4417749. The number was assigned by the company to protect the searcher's anonymity, but it was not much of a shield.

Erik S. Lesser for The New York Times

Thelma Arnold's identity was betrayed by AOL records of her Web searches, like ones for her dog, Dudley, who clearly has a problem.

No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from "numb fingers" to "60 single men" to "dog that urinates on everything."

And search by search, click by click, the identity of AOL user No. 4417749 became easier to discern. There are queries for "landscapers in Lilburn, Ga," several people with the last name Arnold and "homes sold in shadow lake subdivision gwinnett county georgia."

It did not take much investigating to follow that data trail to Thelma Arnold, a 62-year-old widow who lives in Lilburn, Ga., frequently researches her friends' medical ailments and loves her three dogs. "Those are my searches," she said, after a reporter read part of the list to her.

Robust De-anonymization of Large Sparse Datasets

Arvind Narayanan and Vitaly Shmatikov
The University of Texas at Austin

Abstract

We present a new class of statistical de-anonymization attacks against high-dimensional micro-data, such as individual preferences, recommendations, transaction records and so on. Our techniques are robust to perturbation in the data and tolerate some mistakes in the adversary's background knowledge.

We apply our de-anonymization methodology to the Netflix Prize dataset, which contains anonymous movie ratings of 500,000 subscribers of Netflix, the world's largest online movie rental service. We demonstrate that an adversary who knows only a little bit about an individual subscriber can easily identify this subscriber's record in the dataset. Using the Internet Movie Database as the source of background knowledge, we successfully identified the Netflix records of known users, uncovering their apparent political preferences and other potentially sensitive information.

Yahoo Releases Largest Ever Machine Learning Dataset To Researchers (tumblr.com) 41

Posted by samzenpus on Thursday January 14, 2016

An anonymous reader writes:

> Yahoo Labs has released a record-breaking dataset containing 110 billion interactions from 20 million Yahoo News users in 1.5TB of zipped data. The anonymized data is intended for research initiatives in artificial intelligence, including user-behavior modeling, collaborative filtering techniques and unsupervised learning methods.

Moritz Hardt
@mrtz
Follow

Dear privacy researchers, I'm sure deanonymizing that new Yahoo data set is a homework exercise, but can we just NOT this time.

What Makes Privacy Difficult?

High-dimensional data is essentially unique

What Makes Privacy Difficult?

High-dimensional data is essentially unique

Example 1: LU Employee Database

Position	Department	Gender	Year Joined	Nationality	Salary
Lecturer	Math & Stats	Male	2015	Catalan	-

Only one employee fits the description ;-)

What Makes Privacy Difficult?

High-dimensional data is essentially unique

Example 1: LU Employee Database

Position	Department	Gender	Year Joined	Nationality	Salary
Lecturer	Math & Stats	Male	2015	Catalan	-

Only one employee fits the description ;-)

Example 2: Netflix Prize Dataset

"For the vast majority of records, there isn't a single record with similarity score over 0.5 in the entire 500K-record dataset, even if we consider only the sets of movies rated without taking into account numerical ratings or dates."

Differential Privacy: Definition

A *randomized* algorithm $\mathcal{A}$ is ε-differentially private if for every pair of *neighbouring* databases $X \sim X'$ and every possible output y we have

$$\frac{\mathbb{P}[\mathcal{A}(X) = y]}{\mathbb{P}[\mathcal{A}(X') = y]} \leqslant e^{\varepsilon} \quad (\approx 1 + \varepsilon)$$

[Dwork '06]

Key Properties of DP

Key Properties of DP

- Provides privacy against attackers with **side-knowledge**

Key Properties of DP

- Provides privacy against attackers with **side-knowledge**
- Is preserved by any **post-processing** on the output

Key Properties of DP

- Provides privacy against attackers with **side-knowledge**
- Is preserved by any **post-processing** on the output
- Users get **bound on privacy loss** when contributing data

DP101: The Laplace Mechanism

Deterministic function

$$f : \mathbb{X} \to \mathbb{R}$$

DP101: The Laplace Mechanism

Deterministic function

$$f : \mathbb{X} \rightarrow \mathbb{R}$$

Global sensitivity

$$GS_f = \sup_{X \sim X'} |f(X) - f(X')|$$

DP101: The Laplace Mechanism

$$\mathcal{A}(X) = f(X) + \mathrm{Lap}\left(\frac{GS_f}{\varepsilon}\right)$$

Deterministic function

$f : \mathbb{X} \to \mathbb{R}$

Global sensitivity

$GS_f = \sup_{X \sim X'} |f(X) - f(X')|$

DP101: The Laplace Mechanism

$$\mathcal{A}(X) = f(X) + \mathrm{Lap}\left(\frac{GS_f}{\varepsilon}\right)$$

Deterministic function

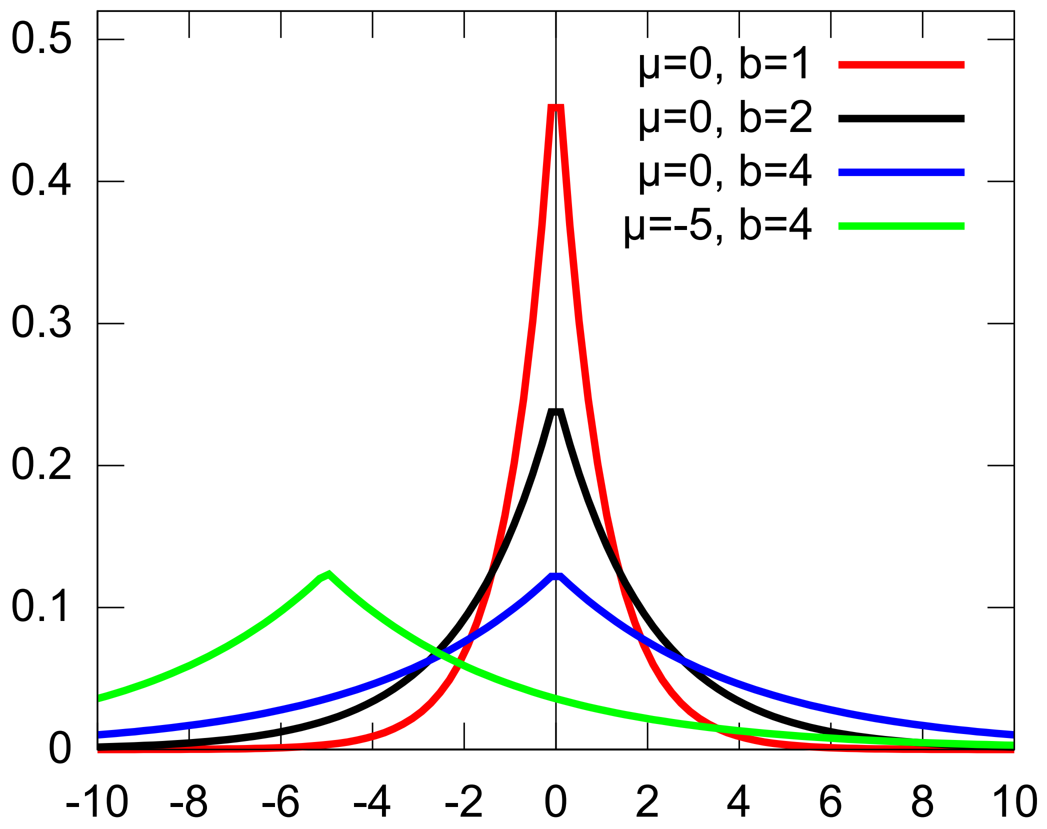
$$f : \mathbb{X} \to \mathbb{R}$$

Global sensitivity

$$GS_f = \sup_{X \sim X'} |f(X) - f(X')|$$

Laplace distribution

$$p_{\mathrm{Lap}(b)}(y) = \frac{1}{2b} \exp\left(-\frac{|y|}{b}\right)$$



Proof: Laplace Mechanism is DP

$$\frac{p_{\mathcal{A}(X)}(y)}{p_{\mathcal{A}(X')}(y)} = \frac{\frac{\varepsilon}{2GS_f}\exp\left(-\frac{\varepsilon|y-f(X)|}{GS_f}\right)}{\frac{\varepsilon}{2GS_f}\exp\left(-\frac{\varepsilon|y-f(X')|}{GS_f}\right)}$$

$$= \exp\left(\frac{\varepsilon(|y-f(X')|-|y-f(X)|)}{GS_f}\right)$$

$$\leqslant \exp\left(\frac{\varepsilon|f(X)-f(X')|}{GS_f}\right) \leqslant e^{\varepsilon}$$

More Queries, More Data

More Queries, More Data

More Queries, Less Privacy

ε-DP $\mathcal{A}_1, \ldots, \mathcal{A}_k \implies (\mathcal{A}_1(X), \ldots, \mathcal{A}_k(X))$ is $(k\varepsilon)$-DP

More Queries, More Data

More Queries, Less Privacy

ε-DP $\mathcal{A}_1, \dots, \mathcal{A}_k \implies (\mathcal{A}_1(X), \dots, \mathcal{A}_k(X))$ is $(k\varepsilon)$-DP

More Data, Less Noise

$$X \sim X'$$
$$X = (x_1, \dots, x_{n-1}, x_n)$$
$$X' = (x_1, \dots, x_{n-1}, x_n')$$

Linear Query

$$f(X) = \frac{1}{n} \sum_{i=1}^{n} g(x_i)$$

Global Sensitivity

$$GS_f = \frac{GS_g}{n}$$

Two Basic Questions

Two Basic Questions

- How to maximize access to information with a fixed privacy budget?
  - ✦ **Interactive DP:** e.g. perturbation dependent on correlation between queries
  - ✦ **Release "Synthetic" DP Data:** e.g. output noisy histograms

Two Basic Questions

- How to maximize access to information with a fixed privacy budget?
  - **Interactive DP:** e.g. perturbation dependent on correlation between queries
  - **Release "Synthetic" DP Data:** e.g. output noisy histograms
- How to design DP mechanisms for more complex queries?
  - **DP Machine Learning:** e.g. train logistic regression / SVM / DNN with DP guarantee on model parameters

Diff. Priv. Machine Learning

[see book by Dwork & Roth ’14]

Diff. Priv. Machine Learning

- **Output Perturbation** (generalize Laplace Mechanism): multivariate outputs, data-dependent noise

[see book by Dwork & Roth '14]

Diff. Priv. Machine Learning

- **Output Perturbation** (generalize Laplace Mechanism): multivariate outputs, data-dependent noise
- **Objective Perturbation:** ERM with random regularization

[see book by Dwork & Roth ’14]

Diff. Priv. Machine Learning

- **Output Perturbation** (generalize Laplace Mechanism): multivariate outputs, data-dependent noise
- **Objective Perturbation:** ERM with random regularization
- Private **Bayesian ML:** privacy inducing priors

[see book by Dwork & Roth '14]

Diff. Priv. Machine Learning

- **Output Perturbation** (generalize Laplace Mechanism): multivariate outputs, data-dependent noise
- **Objective Perturbation:** ERM with random regularization
- Private **Bayesian ML:** privacy inducing priors
- Deep connections between **privacy and generalization**

[see book by Dwork & Roth ’14]

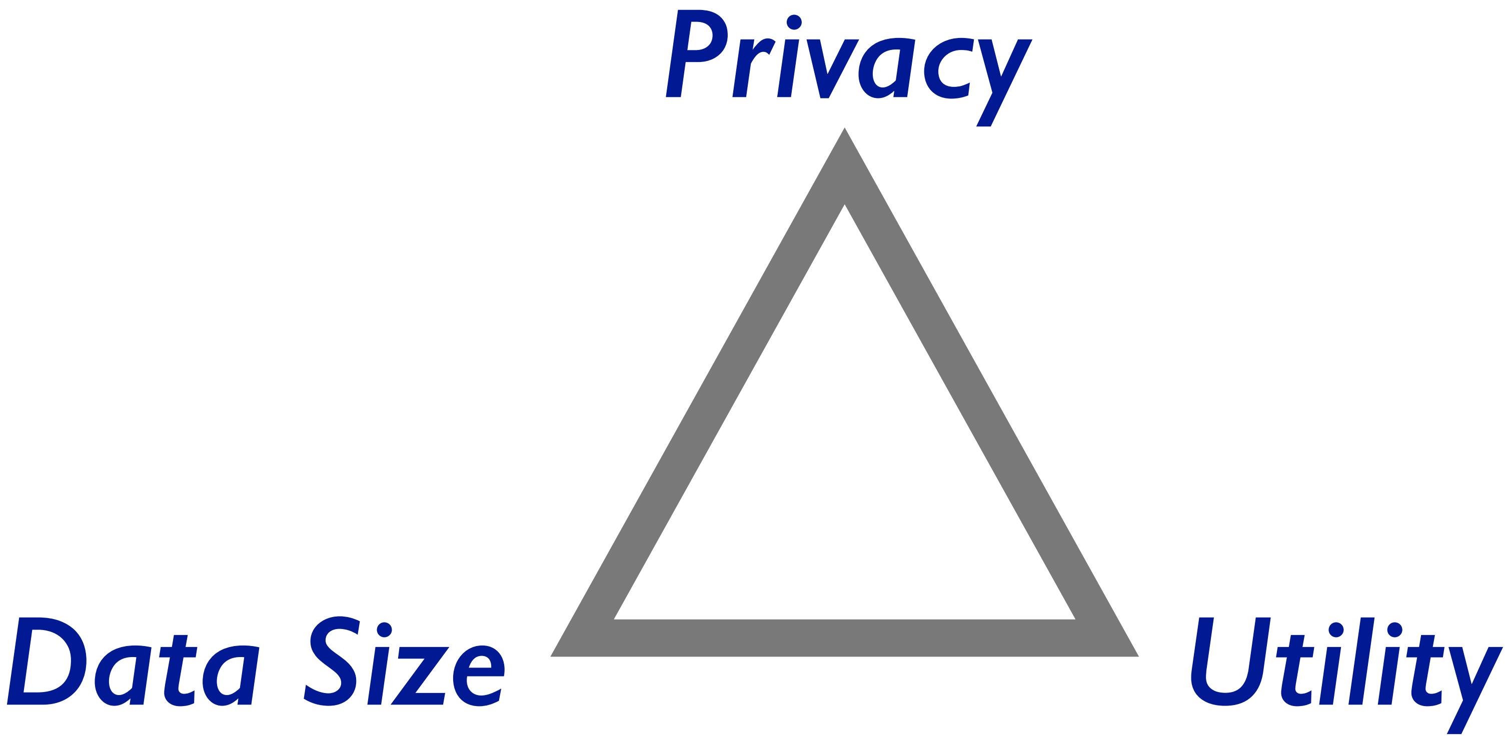
Diff. Priv. Machine Learning

- **Output Perturbation** (generalize Laplace Mechanism): multivariate outputs, data-dependent noise
- **Objective Perturbation:** ERM with random regularization
- Private **Bayesian ML:** privacy inducing priors
- Deep connections between **privacy and generalization**
- DP for **On-line Algorithms**: bandits, on-line optimization, reinforcement learning

[see book by Dwork & Roth '14]

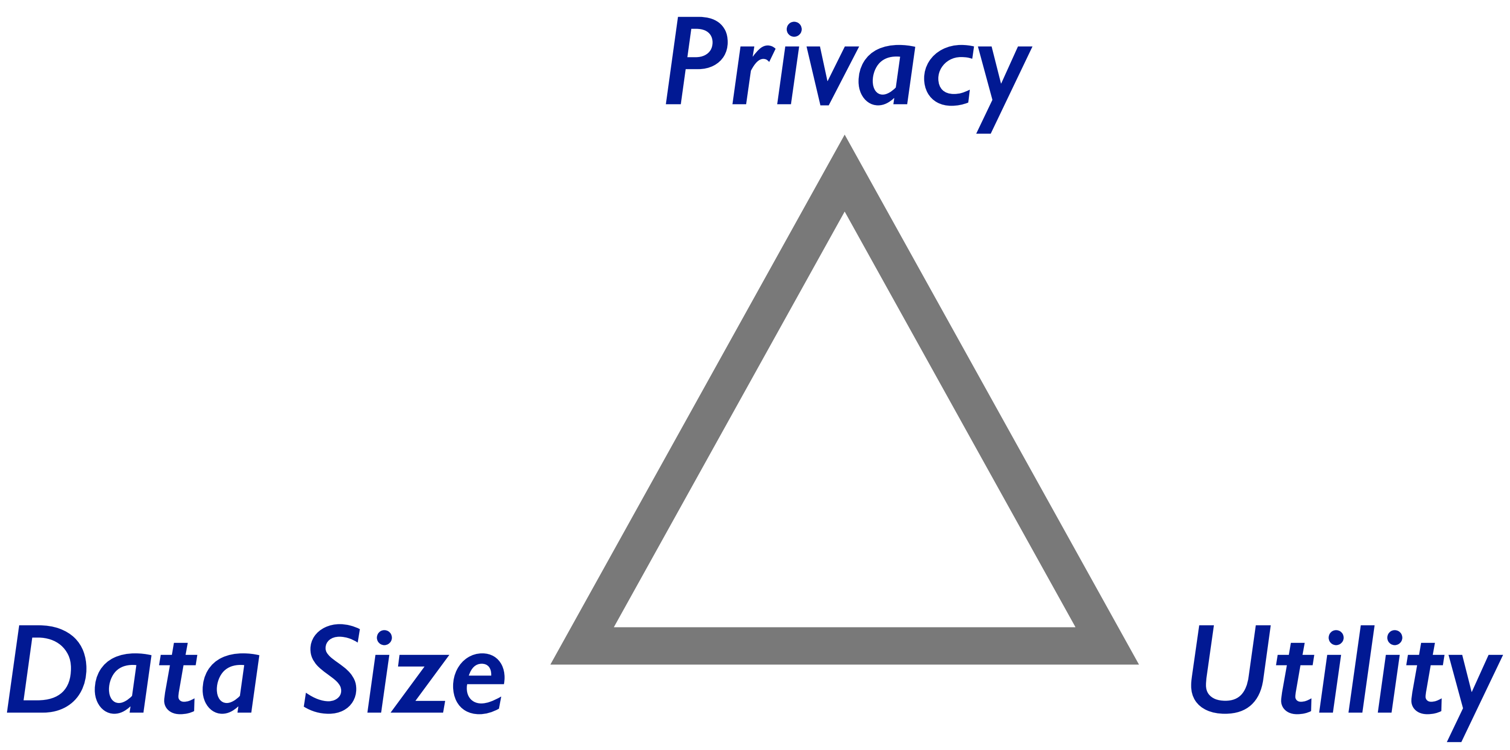
Fundamental Trade-offs

The Golden Triangle of Private Data Analysis



Fundamental Trade-offs

The Golden Triangle of Private Data Analysis



<u>*Example*</u>*: DP Mean estimation with Laplace Mechanism*

$$|\mu - \hat{\mu}_{\mathrm{Lap}}| = O\left(\frac{1}{\sqrt{n}} + \frac{1}{\varepsilon n}\right)$$

Part 2: Differentially Private Policy Evaluation

Markov Decision Processes

- State space $\mathcal{S}$
- Action space $\mathcal{A}$
- Transition kernel $P(s' \mid s, a)$
- Reward function/distribution $0 \leqslant R(s, a) \leqslant R_{\max}$

MDP

$$M = \langle \mathcal{S}, \mathcal{A}, P, R \rangle$$

Dynamics: Observe state, choose action, get reward, transition state, … (s_t, a_t, r_t, s_{t+1})

Learning the Value Function

- Behavior specified by a policy $\pi : \mathcal{S} \rightarrow \mathcal{A}$
- Each state has a value given by the expected discounted cumulative reward collected by the policy

$$V^{\pi}(s) = \mathbb{E}_{M,\pi}[\sum_{t \geqslant 0} \gamma^t r_t \mid s_0 = s]$$

- Policy Evaluation: use data to learn a value function $\hat{V}^{\pi} \approx V^{\pi}$
- Challenge: large or continuous state spaces require function approximation (eg. linear representation)

$$\phi : \mathcal{S} \rightarrow \mathbb{R}^d \qquad \hat{V}^{\pi}(s) = \langle \phi(s), \theta \rangle$$

Example: Evaluating Medical Treatments

- States are symptoms observed in a patient
- Actions are possible drug + dosage combinations
- Rewards reflect outcome of treatment
- Can observe states, actions, and rewards, but transition structure is unknown
- Can compare two treatments from estimated value functions

Batch First-Visit Monte Carlo

Input: $X = (x_1, \ldots, x_m)$ batch of trajectories collected from fixed policy and MDP

Batch First-Visit Monte Carlo

Input: $X = (x_1, \ldots, x_m)$

batch of trajectories collected from fixed policy and MDP

$$x = (s_0 \neq s, r_0, s_1 \neq s, r_1, s_2 = s, r_2, \ldots, s_T, r_T)$$

FVMC Value Estimates:

Batch First-Visit Monte Carlo

Input: $X = (x_1, \ldots, x_m)$ batch of trajectories collected from fixed policy and MDP

$$x = (s_0 \neq s, r_0, s_1 \neq s, r_1, s_2 = s, r_2, \ldots, s_T, r_T)$$

FVMC Value Estimates:

$$F_x(s) = \sum_{t \geqslant 2} \gamma^{t-2} r_t$$

unbiased: $\mathbb{E}_{x \sim \pi}[F_x(s)] = V^\pi(s)$

Batch First-Visit Monte Carlo

Input: $X = (x_1, \ldots, x_m)$ batch of trajectories collected from fixed policy and MDP

$x = (s_0 \neq s, r_0, s_1 \neq s, r_1, s_2 = s, r_2, \ldots, s_T, r_T)$

FVMC Value Estimates:

$F_x(s) = \sum_{t \geqslant 2} \gamma^{t-2} r_t$

unbiased: $\mathbb{E}_{x \sim \pi}[F_x(s)] = V^\pi(s)$

Learn a value function by regressing on these estimated values

Two FVMC Regression Algorithms

$$\theta^{\bullet}_{X} = f^{\bullet}(X) = \operatorname*{argmin}_{\theta \in \mathbb{R}^d} J^{\bullet}_{X}(\theta)$$

Two FVMC Regression Algorithms

$$\theta_X^{\bullet} = f^{\bullet}(X) = \operatorname*{argmin}_{\theta \in \mathbb{R}^d} J_X^{\bullet}(\theta)$$

$$J_X^{l}(\theta) = \frac{1}{m} \sum_{i=1}^{m} \sum_{s \in \mathcal{S}_{x_i}} \rho_s \left(F_{x_i}(s) - \langle \phi(s), \theta \rangle \right)^2 + \frac{\lambda}{2m} \|\theta\|_2^2$$

Two FVMC Regression Algorithms

$$\theta_X^\bullet = f^\bullet(X) = \operatorname*{argmin}_{\theta \in \mathbb{R}^d} J_X^\bullet(\theta)$$

$$J_X^l(\theta) = \frac{1}{m} \sum_{i=1}^{m} \sum_{s \in \mathcal{S}_{x_i}} \rho_s \left(F_{x_i}(s) - \langle \phi(s), \theta \rangle\right)^2 + \frac{\lambda}{2m} \|\theta\|_2^2$$

$$J_X^w(\theta) = \sum_{s \in \mathcal{S}} w_s \left(\sum_{x \in X_s} \frac{F_x(s)}{|X_s|} - \langle \phi(s), \theta \rangle \right)^2$$

Data-dependent Output Perturbation

Output perturbation: $\hat{\theta}^{\bullet}_{X} = f^{\bullet}(X) + \eta_X$ $\quad Var[\eta_X] \propto (\text{sensitivity})^2$

Data-dependent Output Perturbation

Output perturbation: $\hat{\theta}^{\bullet}_{X} = f^{\bullet}(X) + \eta_X$ $\quad Var[\eta_X] \propto (\text{sensitivity})^2$

Global sensitivity: $$GS_{f^{\bullet}} = \sup_{X, X', X \sim X'} \|f^{\bullet}(X) - f^{\bullet}(X')\|_p$$

Data-dependent Output Perturbation

Output perturbation: $\hat{\theta}^{\bullet}_{X} = f^{\bullet}(X) + \eta_X$ $\quad Var[\eta_X] \propto (\text{sensitivity})^2$

Global sensitivity: $\mathrm{GS}_{f^\bullet} = \sup_{X, X', X \sim X'} \|f^\bullet(X) - f^\bullet(X')\|_p$

too pessimistic

Data-dependent Output Perturbation

Output perturbation: $\hat{\theta}^{\bullet}_X = f^{\bullet}(X) + \eta_X$ $\quad Var[\eta_X] \propto (\text{sensitivity})^2$

Global sensitivity: $GS_{f^\bullet} = \sup_{X, X', X \sim X'} \|f^{\bullet}(X) - f^{\bullet}(X')\|_p$

too pessimistic

Local sensitivity: $LS_{f^\bullet}(X) = \sup_{X', X \sim X'} \|f^{\bullet}(X) - f^{\bullet}(X')\|_p$

Data-dependent Output Perturbation

Output perturbation: $\hat{\theta}^{\bullet}_{X} = f^{\bullet}(X) + \eta_X$ $\qquad Var[\eta_X] \propto (\text{sensitivity})^2$

Global sensitivity: $GS_{f^\bullet} = \sup_{X, X', X \sim X'} \|f^\bullet(X) - f^\bullet(X')\|_p$

too pessimistic

Local sensitivity: $LS_{f^\bullet}(X) = \sup_{X', X \sim X'} \|f^\bullet(X) - f^\bullet(X')\|_p$

not private

Data-dependent Output Perturbation

Output perturbation: $\hat{\theta}^{\bullet}_{X} = f^{\bullet}(X) + \eta_X$ $\quad Var[\eta_X] \propto (\text{sensitivity})^2$

Global sensitivity: $GS_{f^{\bullet}} = \sup_{X,X',X \sim X'} \|f^{\bullet}(X) - f^{\bullet}(X')\|_p$

too pessimistic

Local sensitivity: $LS_{f^{\bullet}}(X) = \sup_{X',X \sim X'} \|f^{\bullet}(X) - f^{\bullet}(X')\|_p$

not private

Smoothed sensitivity:

$$SS_{f^{\bullet}}(X) \geqslant LS_{f^{\bullet}}(X)$$

$$X \sim X' \Rightarrow |\ln SS_{f^{\bullet}}(X) - \ln SS_{f^{\bullet}}(X')| \leqslant \beta$$

Data-dependent Output Perturbation

Output perturbation: $\hat{\theta}^{\bullet}_{X} = f^{\bullet}(X) + \eta_X$ $\quad \mathrm{Var}[\eta_X] \propto (\text{sensitivity})^2$

Global sensitivity: $\mathrm{GS}_{f^\bullet} = \sup_{X, X', X \sim X'} \|f^\bullet(X) - f^\bullet(X')\|_p$

too pessimistic

Local sensitivity: $\mathrm{LS}_{f^\bullet}(X) = \sup_{X', X \sim X'} \|f^\bullet(X) - f^\bullet(X')\|_p$

not private

Smoothed sensitivity: $\mathrm{SS}_{f^\bullet}(X) \geqslant \mathrm{LS}_{f^\bullet}(X)$

hard to compute? $X \sim X' \Rightarrow |\ln \mathrm{SS}_{f^\bullet}(X) - \ln \mathrm{SS}_{f^\bullet}(X')| \leqslant \beta$

The NRS Lemma

- The optimal smoothed sensitivity is given by:

$$\mathrm{SS}_{f^\bullet}(X) = \sup_{k \geqslant 0} \left(e^{-k\beta} \sup_{X', X \underset{k}{\sim} X'} \mathrm{LS}_{f^\bullet}(X') \right)$$

[Nissim, Raskhodnikova, Smith '07]

The NRS Lemma

- The <u>optimal smoothed sensitivity</u> is given by:

$$\mathrm{SS}_{f^\bullet}(X) = \sup_{k \geqslant 0} \left(e^{-k\beta} \sup_{X', X \underset{k}{\sim} X'} \mathrm{LS}_{f^\bullet}(X') \right)$$

- <u>Problem:</u> LS and the second "sup" involve uncountably many data sets (because rewards are reals). In general SS is NP-hard to compute

[Nissim, Raskhodnikova, Smith '07]

The NRS Lemma

- The optimal smoothed sensitivity is given by:

$$SS_{f^\bullet}(X) = \sup_{k \geqslant 0} \left(e^{-k\beta} \sup_{X', X \underset{k}{\sim} X'} LS_{f^\bullet}(X') \right)$$

- Problem: LS and the second "sup" involve uncountably many data sets (because rewards are reals). In general SS is NP-hard to compute
- Solution: use SS of a simple upper bound of LS

[Nissim, Raskhodnikova, Smith '07]

From Trajectories to Signatures

Visit signature of a dataset: $\langle X \rangle = (|X_s|)_{s \in \mathcal{S}} \in \mathbb{N}^{\mathcal{S}}$

From Trajectories to Signatures

Visit signature of a dataset: $\langle X \rangle = (|X_s|)_{s \in \mathcal{S}} \in \mathbb{N}^{\mathcal{S}}$

$$\mathrm{LS}_{f^l}(X) \leqslant \varphi^l(\langle X \rangle) = C_l \sqrt{\sum_{s \in \mathcal{S}} \rho_s |X_s|}$$

$$\mathrm{LS}_{f^w}(X) \leqslant \varphi^w(\langle X \rangle) = C_w \sqrt{\sum_{s \in \mathcal{S}} \frac{w_s}{\max\{|X_s|, 1\}^2}}$$

From Trajectories to Signatures

Visit signature of a dataset: $\langle X \rangle = (|X_s|)_{s \in \mathcal{S}} \in \mathbb{N}^{\mathcal{S}}$

$$LS_{f^l}(X) \leqslant \varphi^l(\langle X \rangle) = C_l \sqrt{\sum_{s \in \mathcal{S}} \rho_s |X_s|}$$

$$LS_{f^w}(X) \leqslant \varphi^w(\langle X \rangle) = C_w \sqrt{\sum_{s \in \mathcal{S}} \frac{w_s}{\max\{|X_s|, 1\}^2}}$$

The smooth sensitivity of these functions is easy to compute

DP Policy Evaluation Algorithms

1. Compute $\theta_X^\bullet = \mathrm{argmin}_\theta\, J_X^\bullet(\theta)$
2. Compute $\psi_X^\bullet = \mathrm{SS}_{\varphi^\bullet}(\langle X \rangle)$
3. Sample $\eta_X \sim \mathcal{N}(0, C{\psi_W^\bullet}^2 I)$
4. Output $\hat{\theta}_X^\bullet = \theta_X^\bullet + \eta_X$

Utility Analysis

- Assume trajectories in dataset are i.i.d.
- Bound the empirical excess risk: how worse is the private estimate on the target task versus the non-private estimate
- Ideally this should vanish as the size of the dataset grows (it is easier to satisfy privacy of a user among many)

Utility Analysis

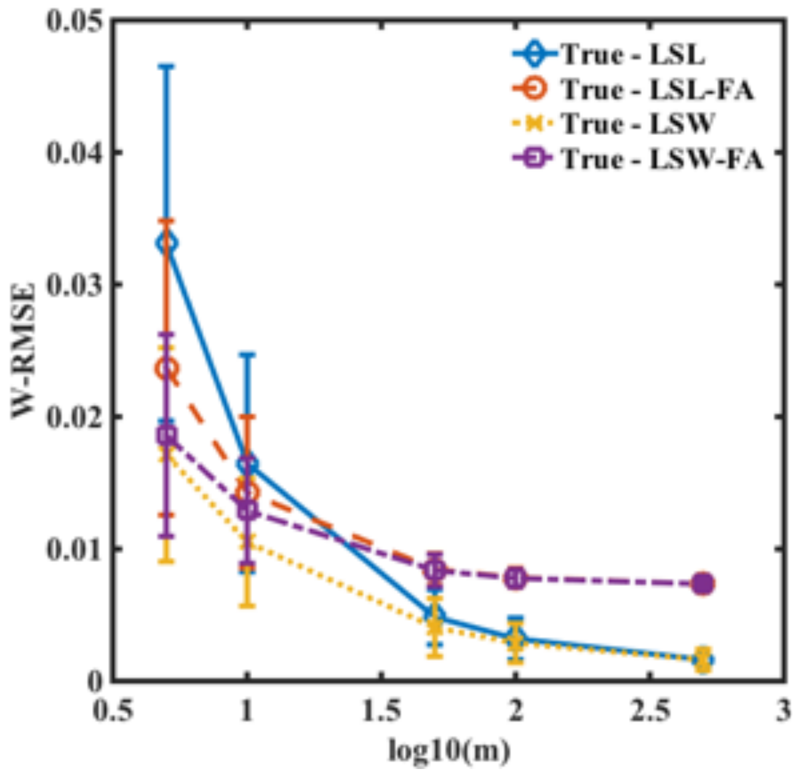
- Assume trajectories in dataset are i.i.d.
- Bound the empirical excess risk: how worse is the private estimate on the target task versus the non-private estimate
- Ideally this should vanish as the size of the dataset grows (it is easier to satisfy privacy of a user among many)

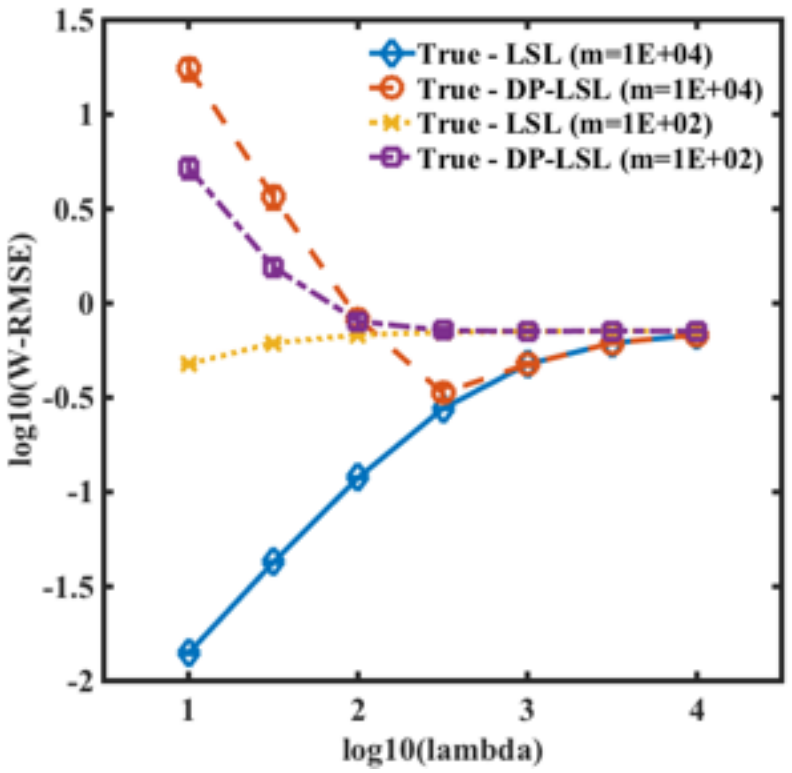
$$\mathbb{E}_{X,\eta_X}[J_X^w(\hat{\theta}_X^w) - J_X^w(\theta_X^w)] = O\left(\frac{1}{m^2}\right)$$

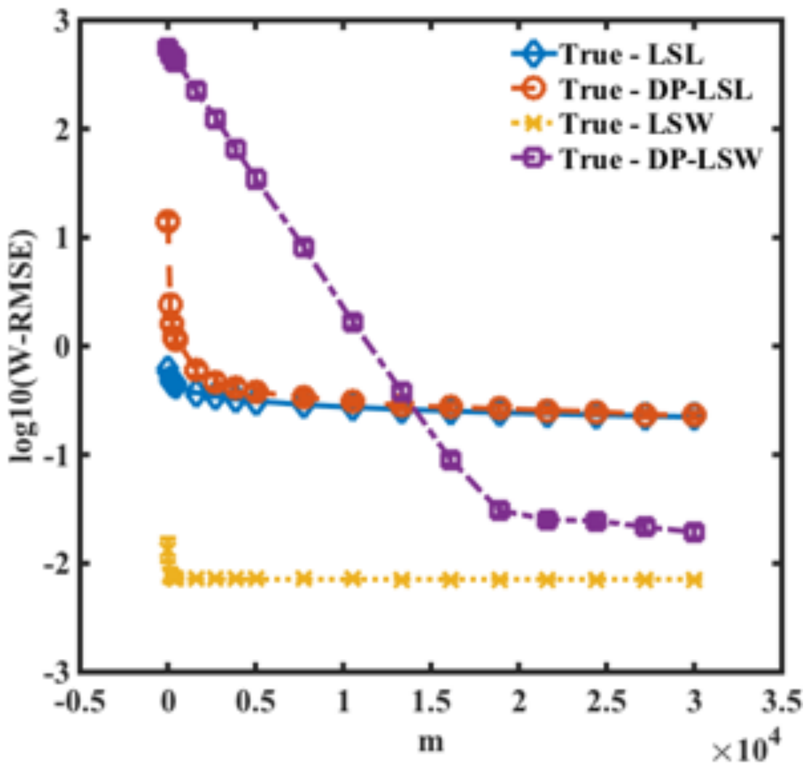
$$\mathbb{E}_{X,\eta_X}[J_X^l(\hat{\theta}_X^l) - J_X^l(\theta_X^l)] = O\left(\frac{1}{\lambda m} + \frac{1}{\lambda^2} + \frac{m}{\lambda^3}\right)$$

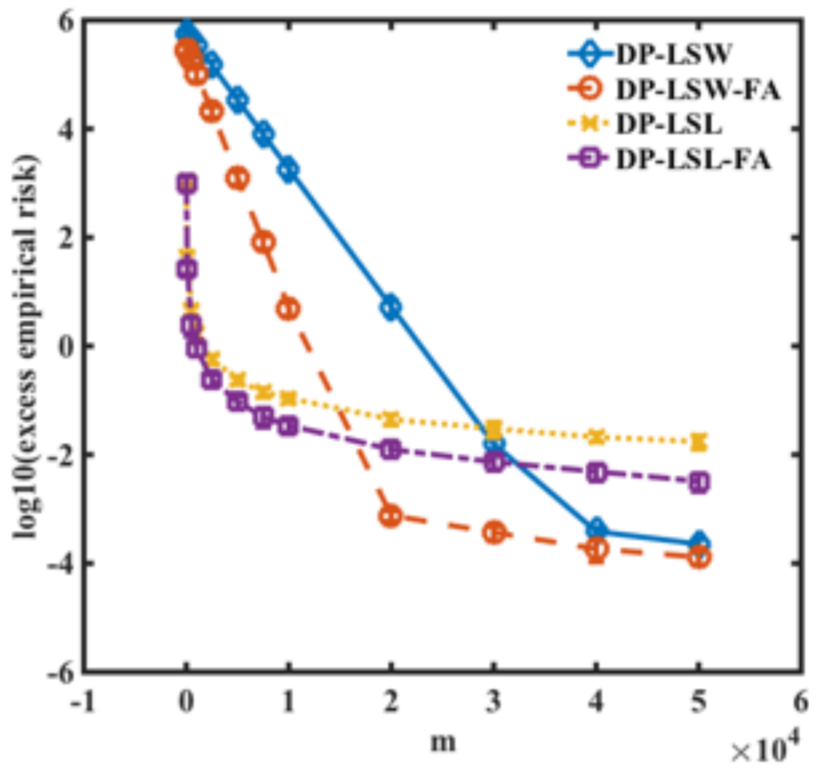
Experimental Results

- 40 state chain MDP with reward in last state, advance with prob. 0.5
- Function approximation aggregates adjacent states
- Test effect of regularization, function approximation, and privacy









Conclusion and Future Work

- Two DP algorithms for policy evaluation in the batch setting
- DP-LSL better with small data, DP-LSW better with large data
- Function approximation helps privacy
- Tighter approximation to the smooth sensitivity would yield less noise
- Alternative approaches: objective perturbation and LSTD

Differentially Private Policy Evaluation

Borja de Balle Pigem

Data Science | Lancaster University

Mathematics & Statistics | Lancaster University

[CSML — Feb 18, 2016]