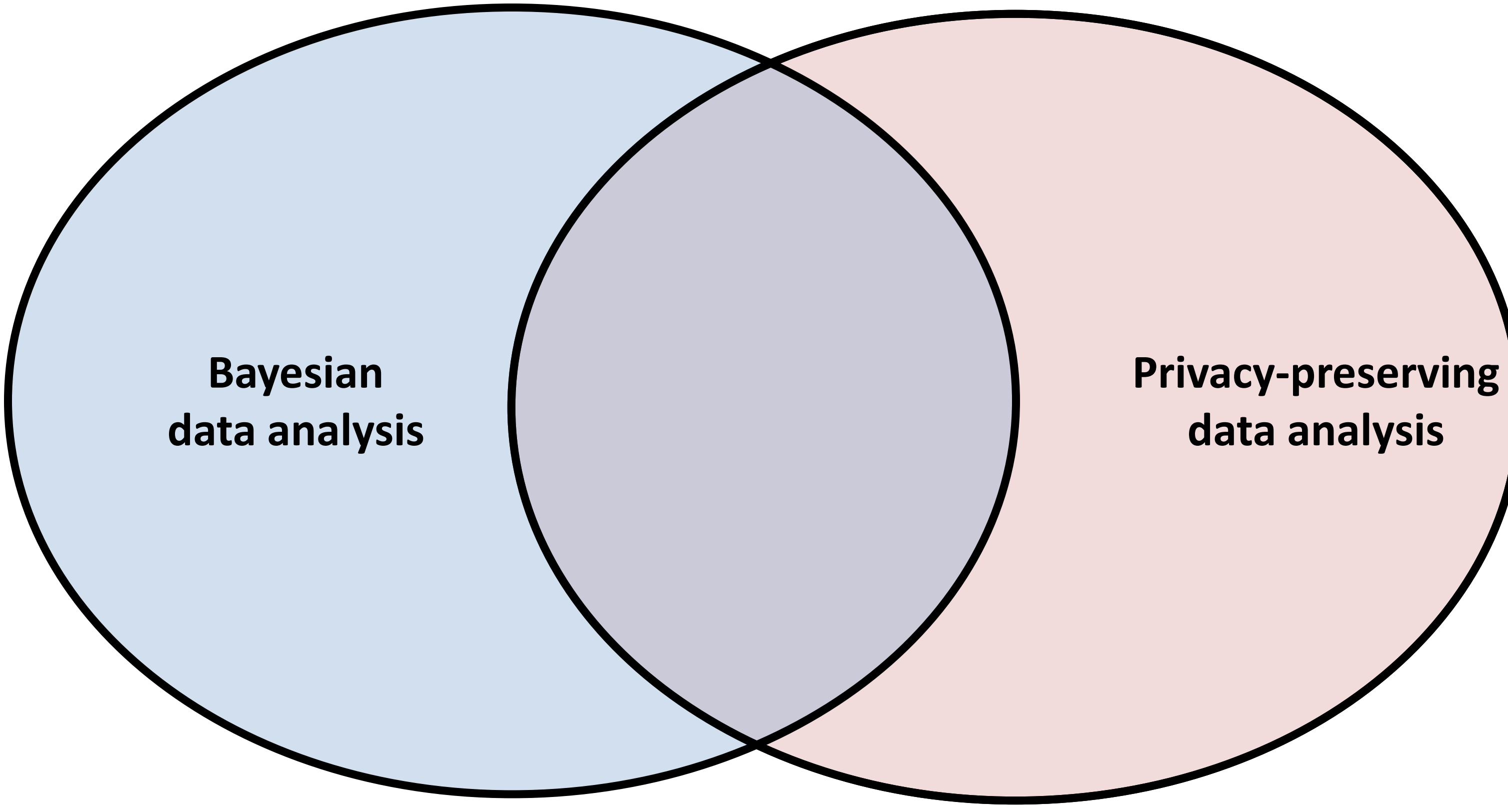
On the Theory and Practice of Privacy-Preserving Bayesian Data Analysis

James Foulds,* Joseph Geumlek,* Max Welling,+ Kamalika Chaudhuri*

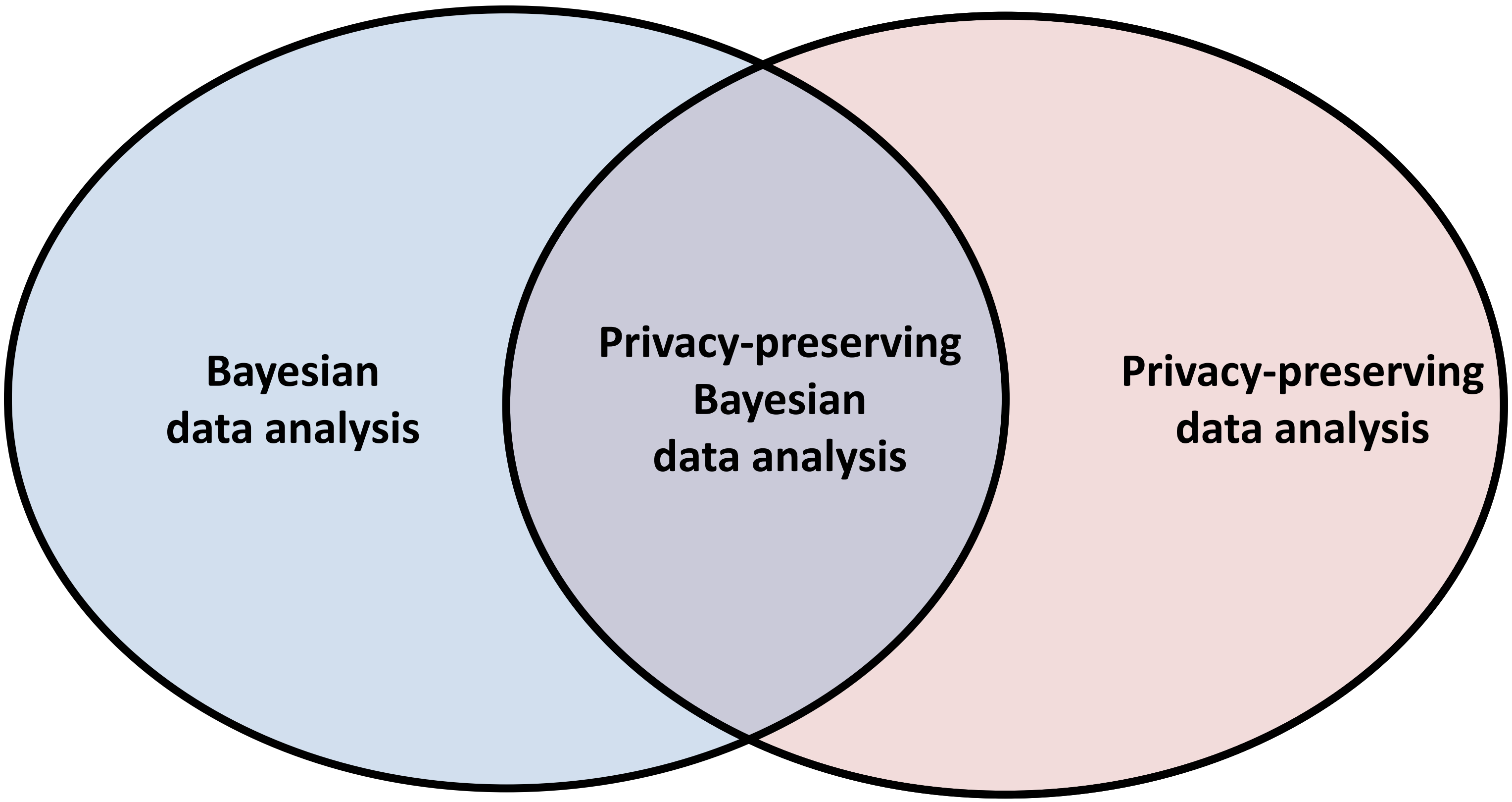
*University of California, San Diego

+University of Amsterdam

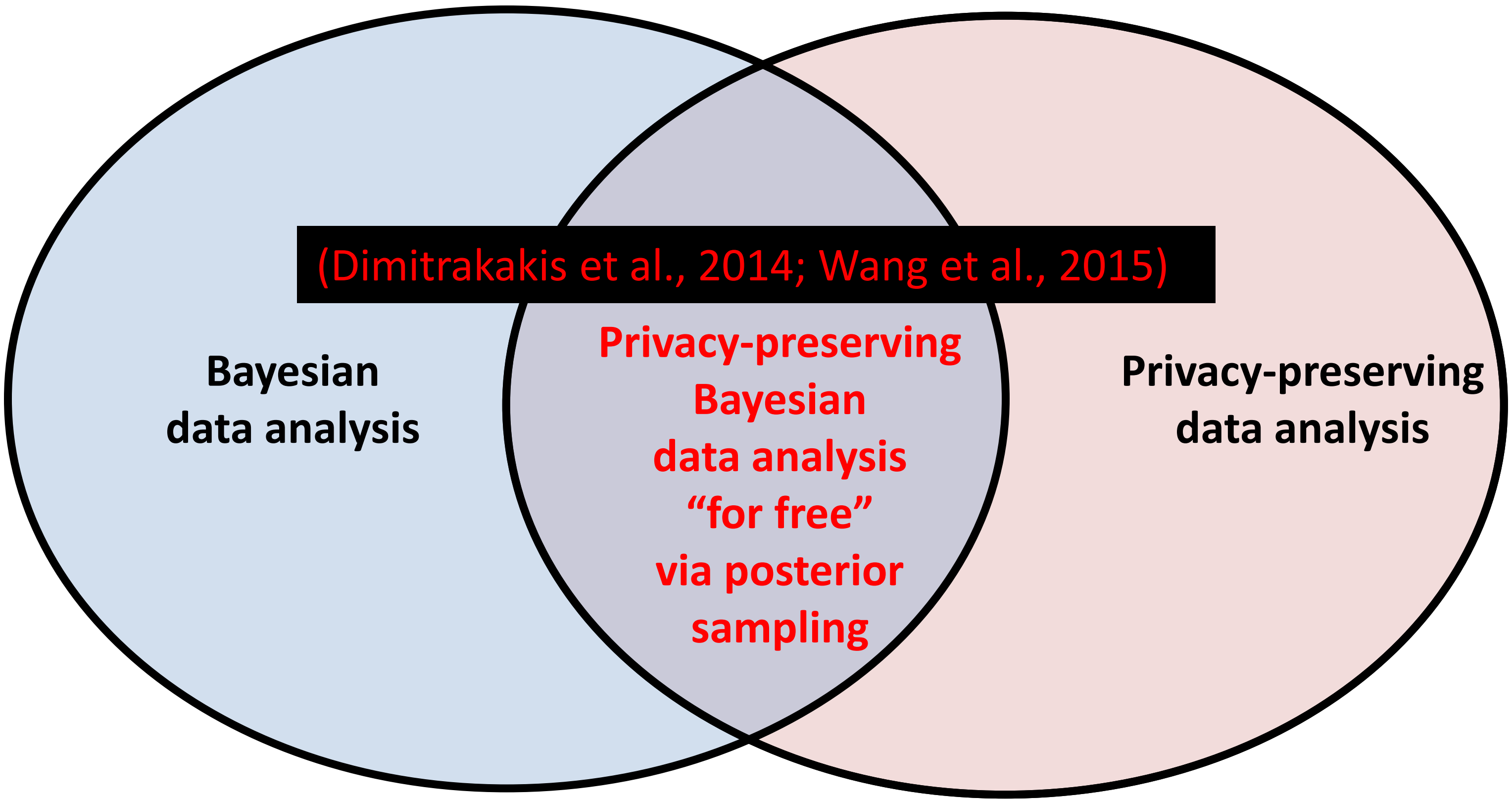
Overview



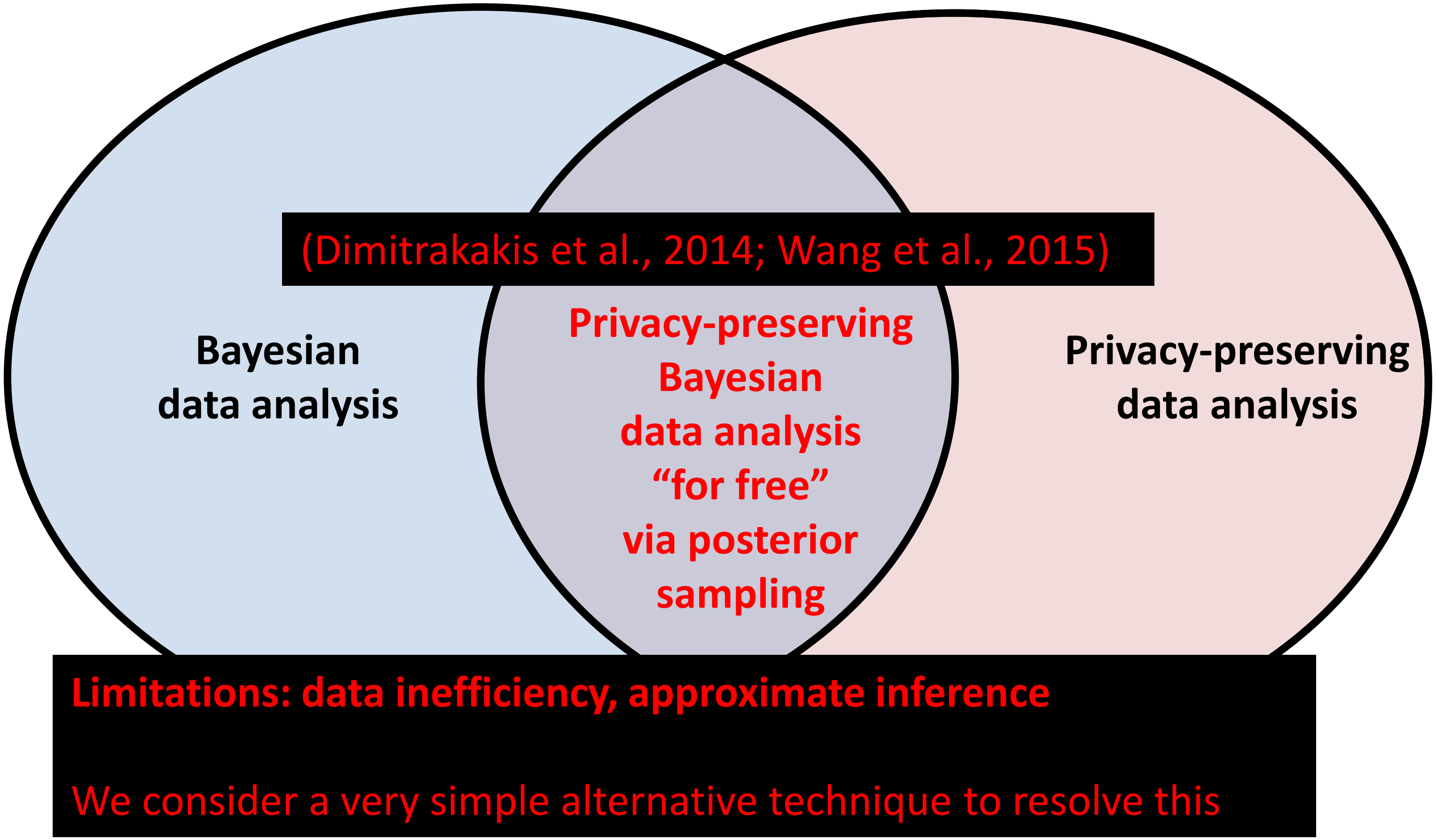
Overview


Bayesian
data analysis
Privacy-preserving
Bayesian
data analysis
Privacy-preserving
data analysis

Overview


(Dimitrakakis et al., 2014; Wang et al., 2015)
Bayesian data analysis
Privacy-preserving Bayesian data analysis "for free" via posterior sampling
Privacy-preserving data analysis

Overview


(Dimitrakakis et al., 2014; Wang et al., 2015)
Bayesian data analysis
Privacy-preserving Bayesian data analysis “for free” via posterior sampling
Privacy-preserving data analysis
Limitations: data inefficiency, approximate inference
We consider a very simple alternative technique to resolve this

Privacy and Machine Learning

- As individuals and consumers we benefit from ML systems trained on **OUR** data

Privacy and Machine Learning

- As individuals and consumers we benefit from ML systems trained on **OUR** data
  - **Internet search**

Privacy and Machine Learning

- As individuals and consumers we benefit from ML systems trained on **OUR** data
  - **Internet search**
  - **Recommendations**
    - products, movies, music, news, restaurants, email recipients

Privacy and Machine Learning

- As individuals and consumers we benefit from ML systems trained on **OUR** data
  - **Internet search**
  - **Recommendations**
    - products, movies, music, news, restaurants, email recipients
  - **Mobile phones**
    - Autocorrect, speech recognition, Siri, …

The cost is our privacy

Forbes / Tech

FEB 16, 2012 @ 11:02 AM 2,998,353 VIEWS

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Kashmir Hill
FORBES STAFF

Welcome to The Not-So Private Parts where technology & privacy collide

FULL BIO >

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

Charles Duhigg outlines in the New York Times how Target tries to hook parents-to-be at that crucial moment before they turn into rampant — and loyal — buyers of all things pastel, plastic, and miniature. He talked to Target statistician Andrew Pole — before Target freaked out and cut off all communications — about the clues to a customer's impending bundle of joy. Target assigns every customer a Guest ID number, tied to

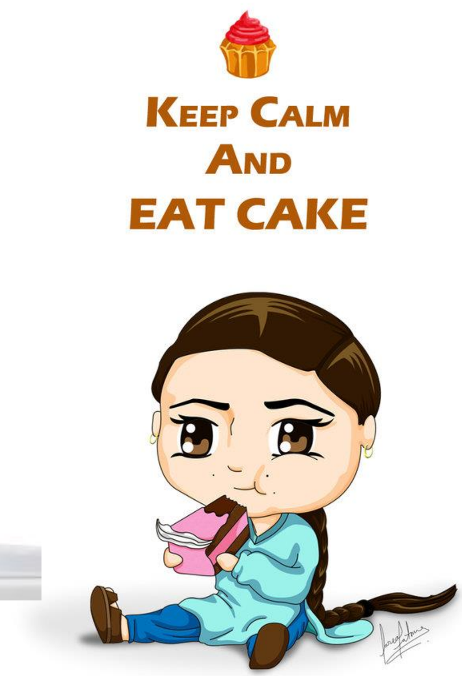
TARGET

Target has got you in its aim

http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#b228dae34c62 ,Retrieved 6/16/2016

Privacy and Machine Learning

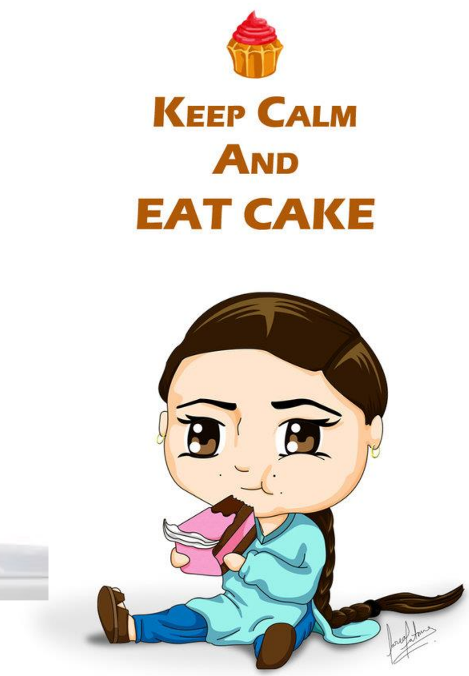
- Want the **benefits of sharing our data** while **protecting our privacy**
  - *Have your cake and eat it too!*



Privacy and Machine Learning

- Want the **benefits of sharing our data** while **protecting our privacy**
  - *Have your ~~cake~~* ***Apple*** *and eat it too!*



Privacy and Machine Learning

- Want the **benefits of sharing our data** while **protecting our privacy**
  - *Have your ~~cake~~* ***Apple*** *and eat it too!*

"We believe you should have

great features

and

***great privacy*.**

You demand it and we're dedicated to providing it."

Craig Federighi,
Apple senior vice president of Software Engineering.
June 13 2016, WWDC16

Quote from http://appleinsider.com/articles/16/06/15/inside-ios-10-apple-doubles-down-on-security-with-cutting-edge-differential-privacy , retrieved 6/16/2016

Statistical analysis of sensitive data

theguardian

election 2016 US world opinion sports soccer tech arts lifestyle fashion business travel environment ≡ browse all sections

home › world UK europe americas asia middle east africa australia cities development

Iraq: The war logs

Iraq war logs: disclosure condemned by Hillary Clinton and Nato

Officials say lives could be put at risk by WikiLeaks's release of 400,00 secret US army field reports

Amy Fallon

Friday 22 October 2010 17.53 EDT

This article is 5 years old

Shares

0

Save for later

The Pentagon has condemned the release of the secret US army field reports by WikiLeaks saying that lives will be put at risk. Photograph: Charles Dharapak/AP

[the Wikileaks disclosure] ***"puts the lives of United States and its partners' service members and civilians at risk."***

- Hillary Clinton

Most popular in US

USA 2-1 Ecuador: Copa América quarter-final - as it happened

Bayesian analysis of sensitive data

- Bayesian inference widely and successfully used in application domains where privacy is invaluable
  - Text analysis (Blei et al., 2003; Goldwater and Griffiths, 2007)
  - Personalized recommender systems (Salakhutdinov and Mnih, 2008)
  - Medical informatics (Husmeier et al., 2006)
  - MOOCs (Piech et al., 2013).

- Data scientists must balance benefits and potential insights vs privacy concerns (Daries et al., 2014).

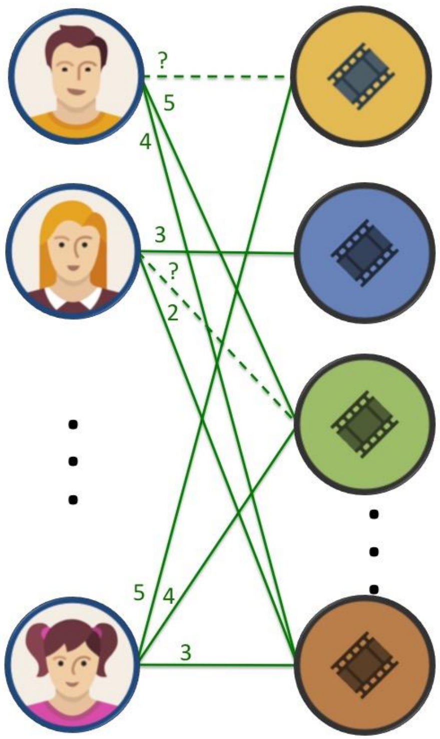
Anonymization?

Alice

Bob

Claire

....



Anonymized Netflix data

(Narayanan and Shmatikov, 2008)

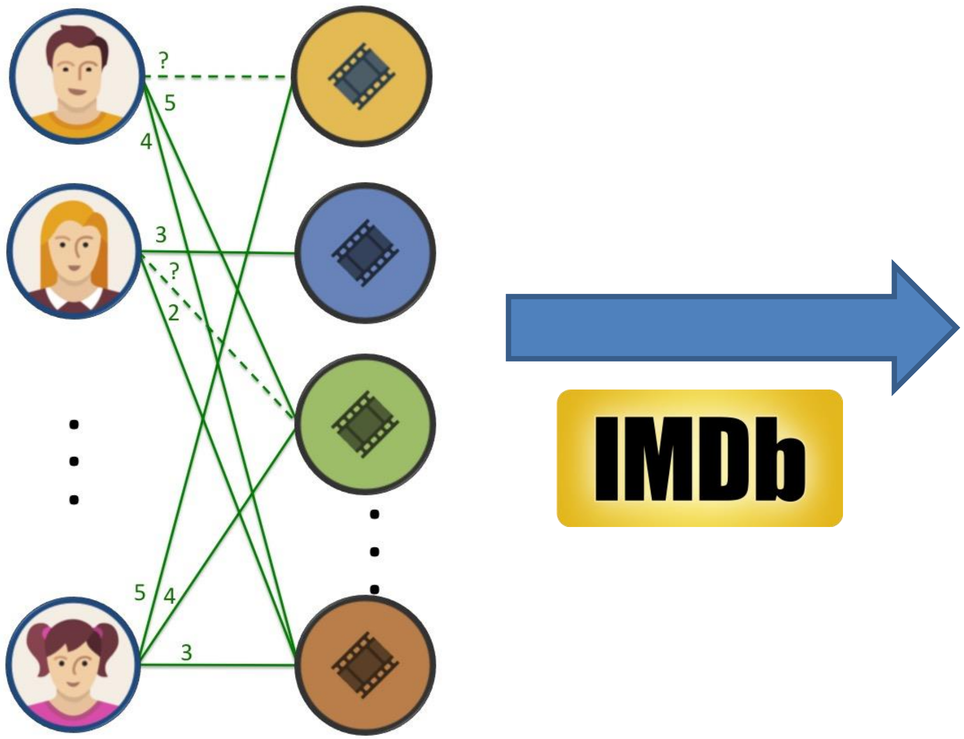
Anonymization?

Alice

Bob

Claire

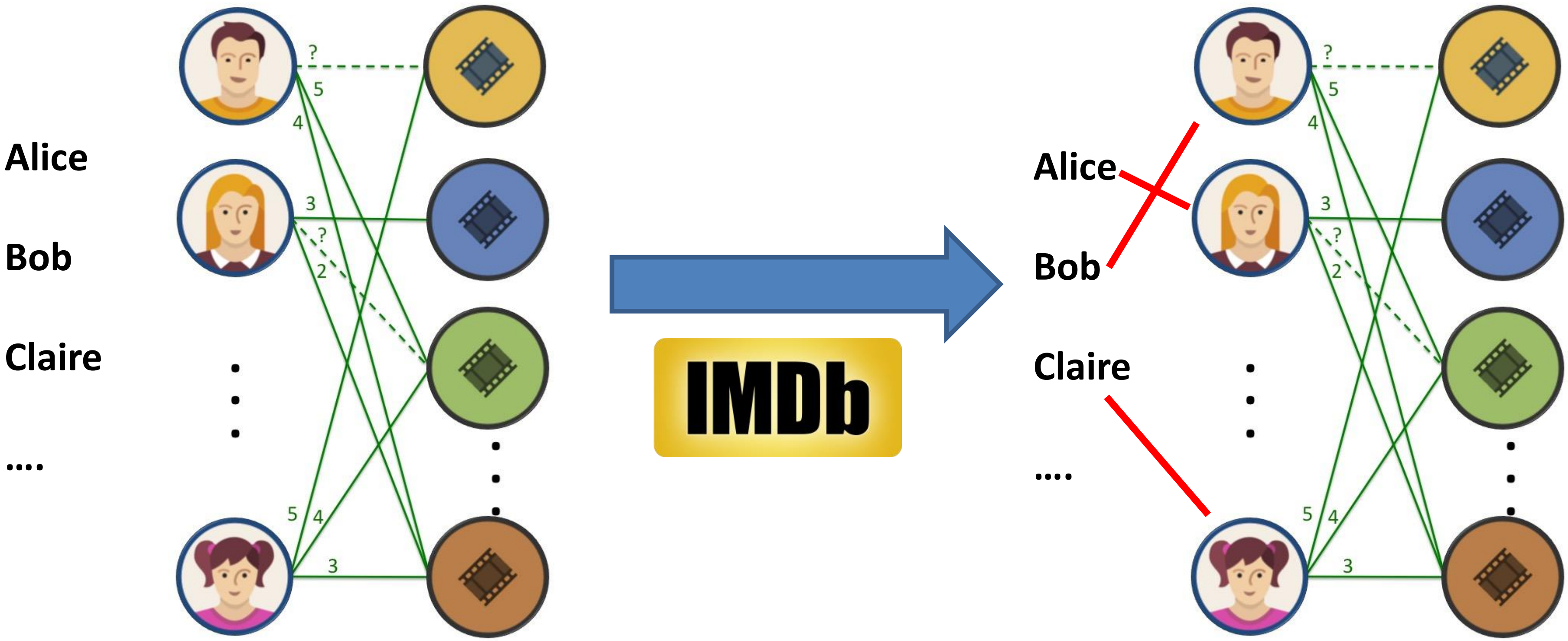
....



Anonymized Netflix data + public IMDB data

(Narayanan and Shmatikov, 2008)

Anonymization?



Anonymized Netflix data + public IMDB data = identified Netflix data

(Narayanan and Shmatikov, 2008)

Aggregation?

BuzzFeed

LOL win omg cute fail wtf

Can You Find All 10 People Hiding In This Crowd?

Do you have the vision of a majestic eagle?

posted on Apr. 20, 2016, at 2:00 p.m.

https://www.buzzfeed.com/nathanwpyle/can-you-spot-all-26-letters-in-this-messy-room-369?utm_term=.gyRdVVvV5#.kkovLL1LE
Retrieved 6/16/2016

Hiding in the crowd

- Only release statistics aggregated over many individuals. Does this ensure privacy?

Hiding in the crowd

- Only release statistics aggregated over many individuals. Does this ensure privacy?

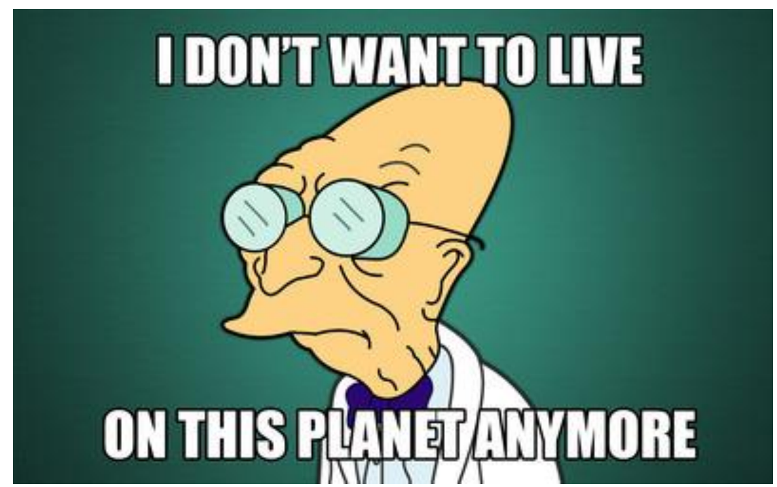
- Report average salary in CS dept. 

Hiding in the crowd

- Only release statistics aggregated over many individuals. Does this ensure privacy?

- Report average salary in CS dept.

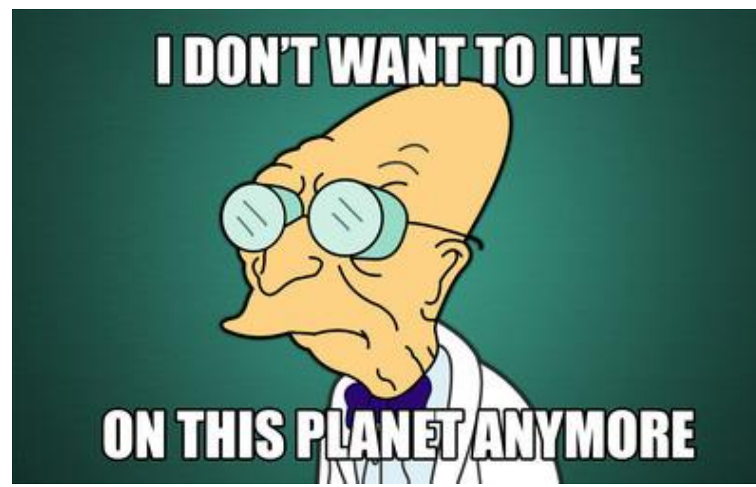
- Prof. X leaves.



Hiding in the crowd

- Only release statistics aggregated over many individuals. Does this ensure privacy?

- Report average salary in CS dept. 

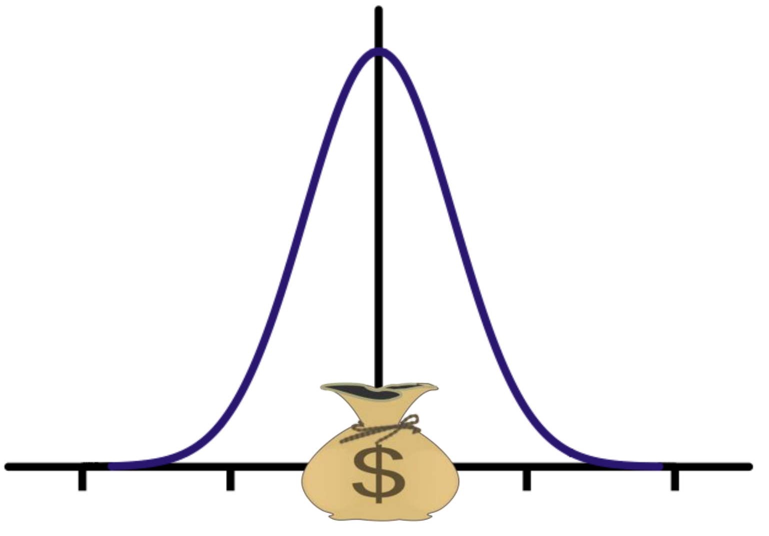


- Prof. X leaves.

- Report avg salary again. 
  - We can identify Prof. X's salary

Noise / data corruption

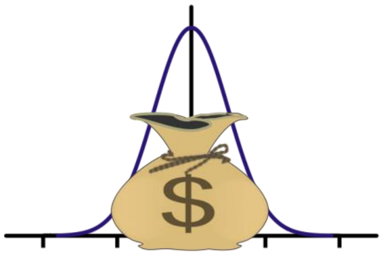
- Release Prof. X's salary + noise



- Once we sufficiently obfuscate Prof. X's salary, it is no longer useful

Noise + crowd

- Release **mean salary** + noise



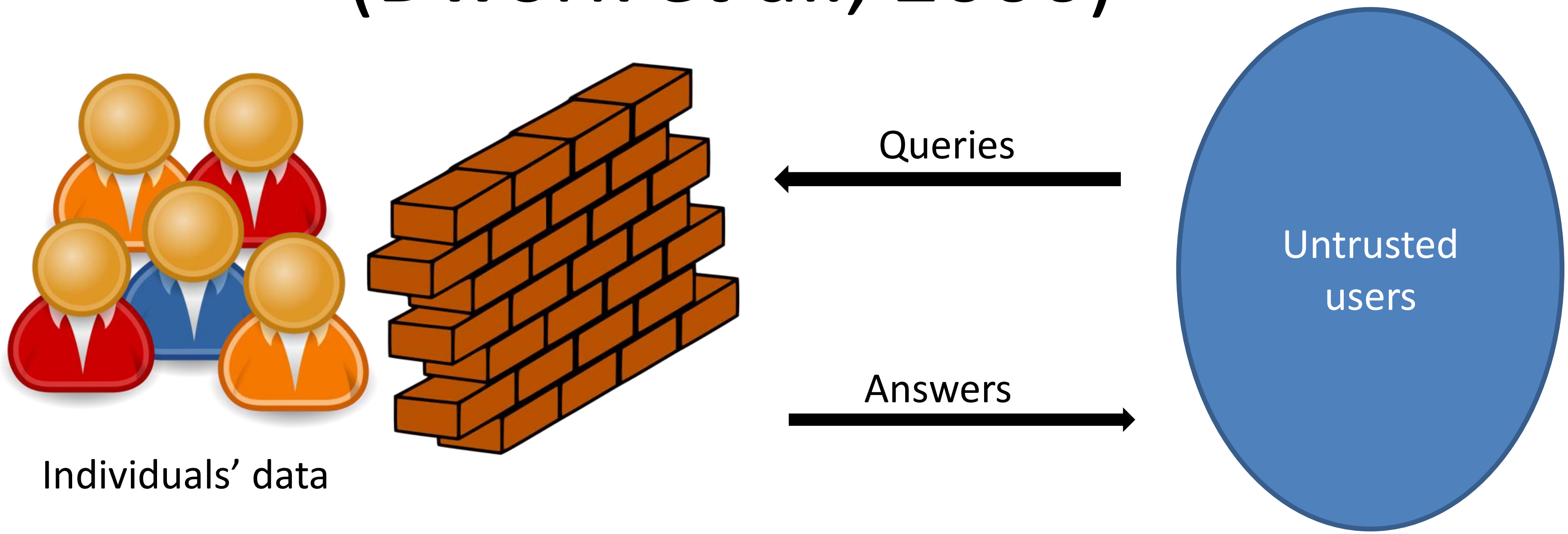
- Need much less noise to protect Prof. X's salary

Solution

- "Noise + crowds" can provide both **individual-level privacy**, and accurate **population-level queries**

- How to quantify privacy loss?
  - Answer: **Differential privacy**

Differential privacy (Dwork et al., 2006)



Privacy-preserving interface: randomized algorithms

- DP is a promise:
  - "If you add your data to the database, you will not be affected much"

Differential privacy (Dwork et al., 2006)

- Consider randomized algorithm $\mathcal{M}(\mathbf{X})$
- DP guarantees that the likely output of $\mathcal{M}(\mathbf{X})$ is not greatly affected by **any one data point**
- In particular, the **distribution over the outputs** of the algorithm will not change too much

Individuals' data

Differential privacy (Dwork et al., 2006)

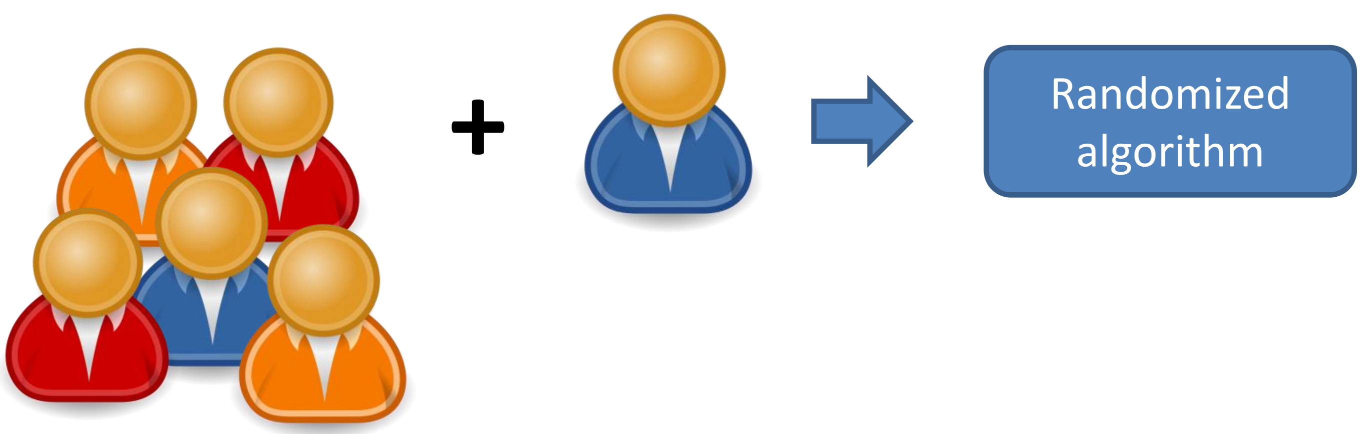
- Consider randomized algorithm $\mathcal{M}(\mathbf{X})$
- DP guarantees that the likely output of $\mathcal{M}(\mathbf{X})$ is not greatly affected by **any one data point**
- In particular, the **distribution over the outputs** of the algorithm will not change too much

+

Individuals' data

Differential privacy (Dwork et al., 2006)

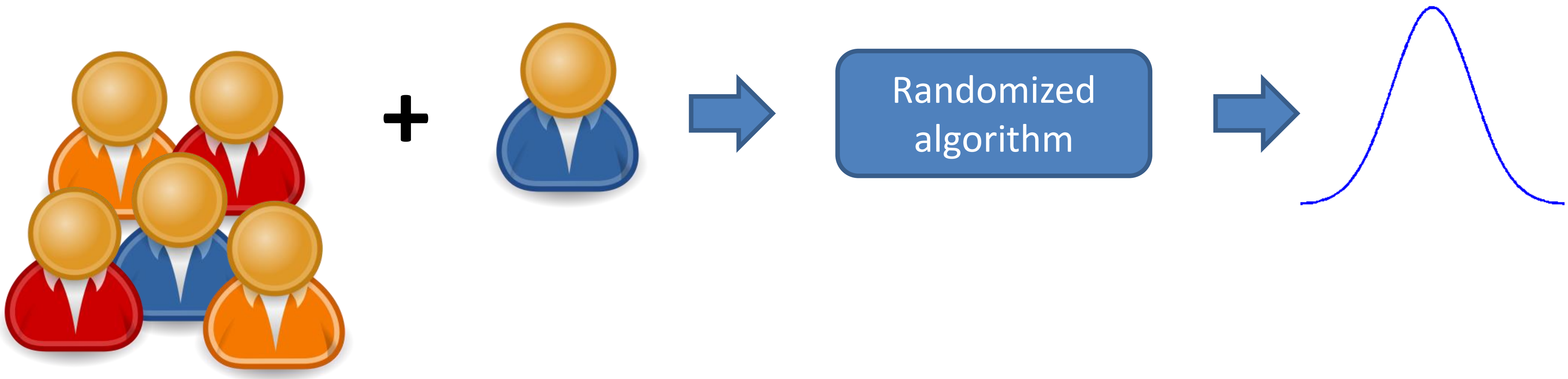
- Consider randomized algorithm $\mathcal{M}(\mathbf{X})$
- DP guarantees that the likely output of $\mathcal{M}(\mathbf{X})$ is not greatly affected by **any one data point**
- In particular, the **distribution over the outputs** of the algorithm will not change too much



Individuals' data

Differential privacy (Dwork et al., 2006)

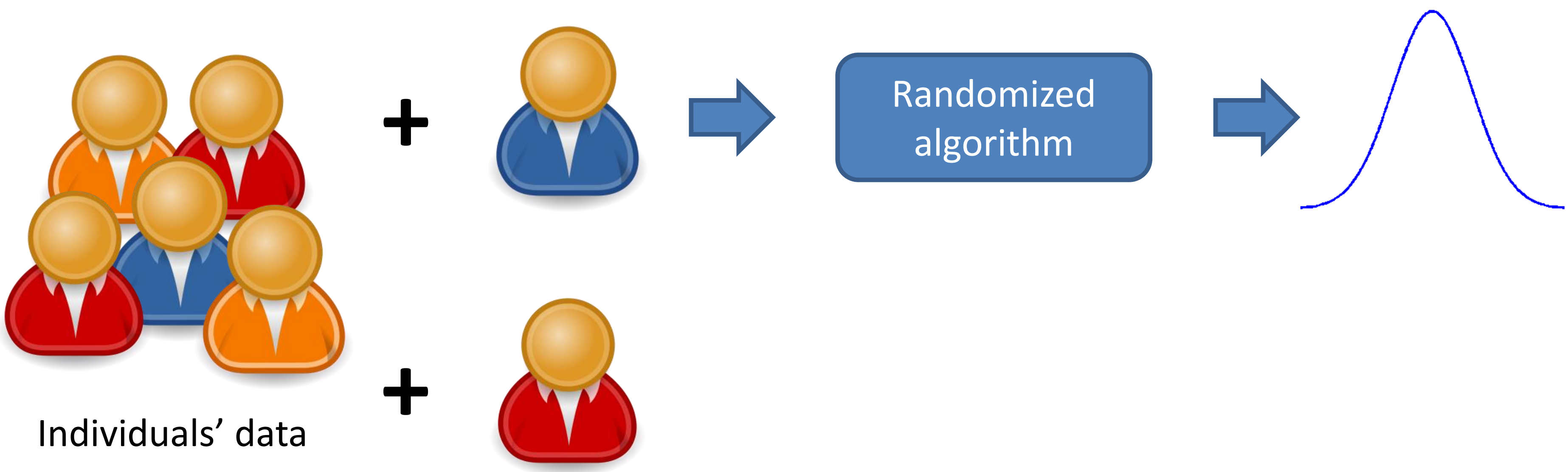
- Consider randomized algorithm $\mathcal{M}(\mathbf{X})$
- DP guarantees that the likely output of $\mathcal{M}(\mathbf{X})$ is not greatly affected by **any one data point**
- In particular, the **distribution over the outputs** of the algorithm will not change too much



Individuals’ data

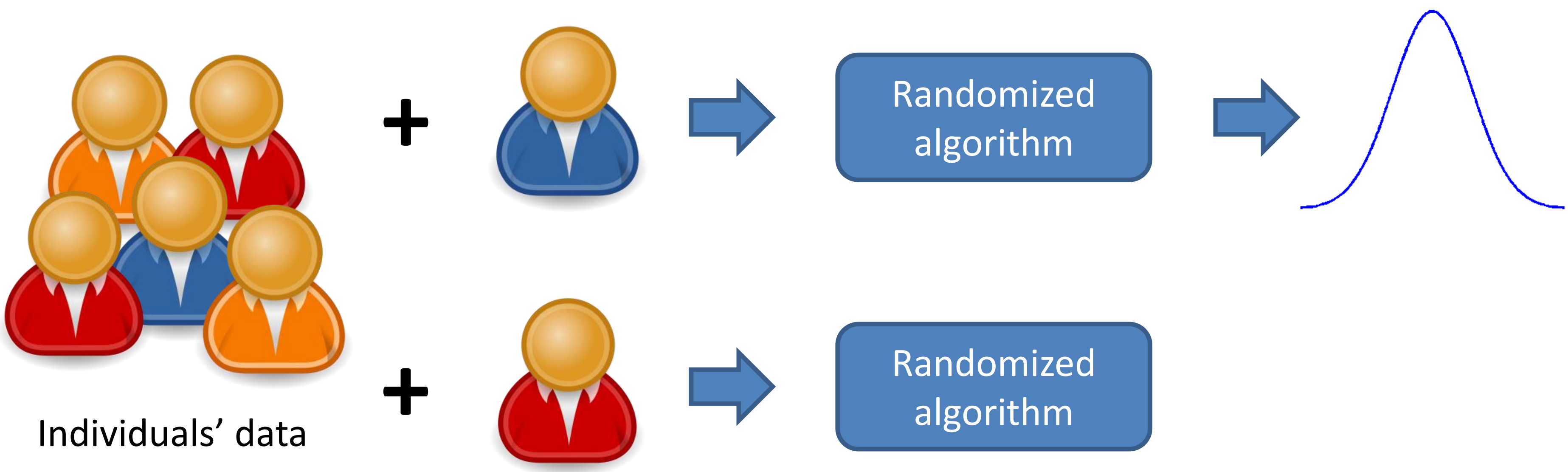
Differential privacy (Dwork et al., 2006)

- Consider randomized algorithm $\mathcal{M}(\mathbf{X})$
- DP guarantees that the likely output of $\mathcal{M}(\mathbf{X})$ is not greatly affected by **any one data point**
- In particular, the **distribution over the outputs** of the algorithm will not change too much



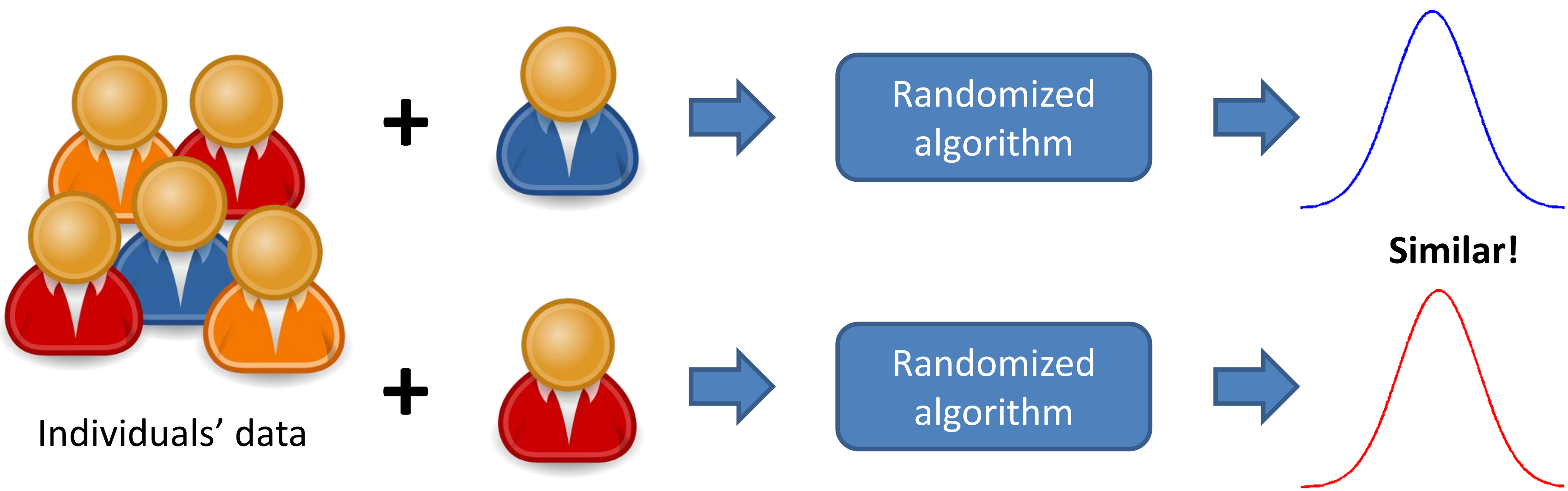
Differential privacy (Dwork et al., 2006)

- Consider randomized algorithm $\mathcal{M}(\mathbf{X})$
- DP guarantees that the likely output of $\mathcal{M}(\mathbf{X})$ is not greatly affected by **any one data point**
- In particular, the **distribution over the outputs** of the algorithm will not change too much



Differential privacy (Dwork et al., 2006)

- Consider randomized algorithm $\mathcal{M}(\mathbf{X})$
- DP guarantees that the likely output of $\mathcal{M}(\mathbf{X})$ is not greatly affected by **any one data point**
- In particular, the **distribution over the outputs** of the algorithm will not change too much

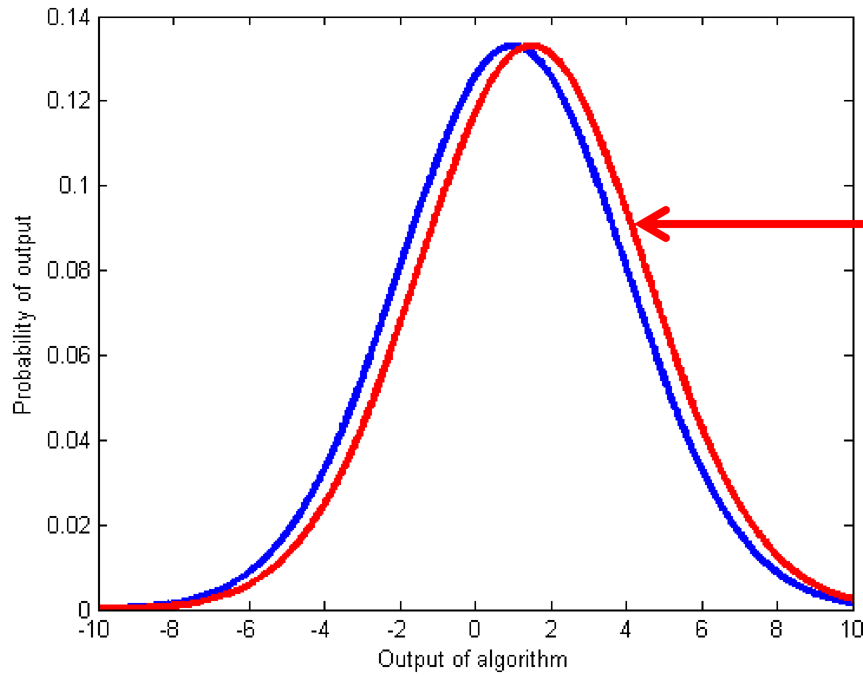


Differential privacy (Dwork et al., 2006)

Definition: $\mathcal{M}(\mathbf{X})$ is ϵ-differentially private if

$$\frac{Pr(\mathcal{M}(\mathbf{X}) \in \mathcal{S})}{Pr(\mathcal{M}(\mathbf{X}') \in \mathcal{S})} \leq e^{\epsilon}$$

for all outcomes $\mathcal{S}$, and pairs of databases $\mathbf{X}$, $\mathbf{X}'$ differing in a single element.



Ratios of probabilities bounded by e^{ϵ}

Properties of differential privacy

- **Immune to post-processing**
  - Resists attacks using side information, as in the Netflix Prize linkage attack

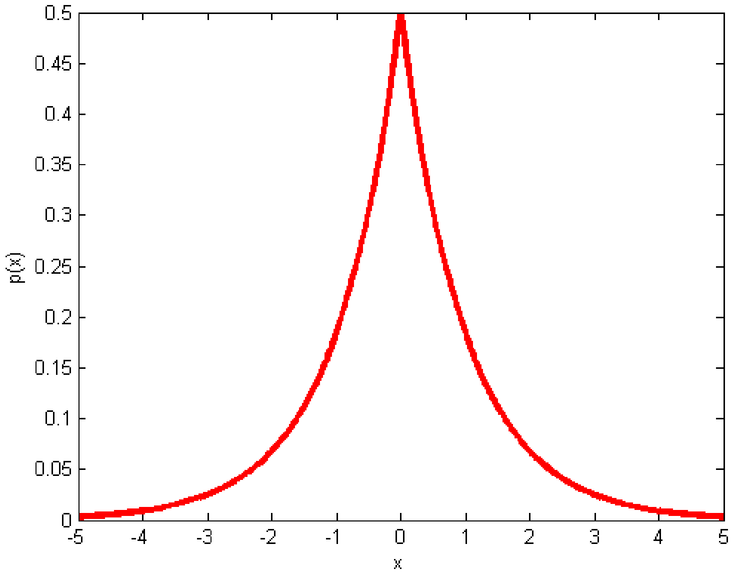
Properties of differential privacy

- **Immune to post-processing**
  - Resists attacks using side information, as in the Netflix Prize linkage attack
- **Composition**
  - If you run multiple DP queries, their epsilons add up.
  - Can think of this as a "privacy budget" we spend over all queries

Laplace mechanism (Dwork et al., 2006)

- Adding Laplace noise is sufficient to achieve differential privacy

- The Laplace distribution is two exponential distributions, back-to-back



- The noise level depends on a quantity called the **L1 sensitivity** of the query *h*:

$$\triangle h = \max_{\mathbf{X},\mathbf{X}'} \|h(\mathbf{X}) - h(\mathbf{X}')\|_1$$

$$\text{Add Laplace}(\triangle h/\epsilon) \text{ noise to each dimension of } h(\mathbf{X}).$$

Exponential mechanism (McSherry and Talwar, 2007)

- Aims to output responses of high utility
- Given real-valued utility function $u(\mathbf{X}, \mathbf{r})$, the exponential mechanism selects outputs r via

$$Pr(\mathcal{M}_E(\mathbf{X}, u, \epsilon) = \mathbf{r}) \propto \exp\left(\frac{u(\mathbf{X}, \mathbf{r})}{T}\right) , \qquad T = \frac{2\triangle u}{\epsilon}$$

Temperature depends on sensitivity, epsilon

$$\text{Sensitivity: } \triangle u \triangleq \max_{r,(\mathbf{X},\mathbf{X}')} \|u(\mathbf{X}, r) - u(\mathbf{X}', r)\|_1$$

Privacy-preserving Bayesian inference via the exponential mechanism (OPS)

(Dimitrakakis et al., 2014; Wang et al., 2015)

- **Privacy cost of drawing a sample from posterior**
  - Interpret as exponential mechanism with the log joint probability $u(\mathbf{X}, \theta) = \log Pr(\theta, \mathbf{X})$ as the utility function:

$$f(\theta; \mathbf{X}, \epsilon) \propto \exp\left(\frac{\log Pr(\theta, \mathbf{X})}{T}\right) = Pr(\theta, \mathbf{X})^{1/T} \ , T = \frac{2\triangle \log Pr(\theta, \mathbf{X})}{\epsilon}$$

$$\text{where } \triangle \log Pr(\theta, \mathbf{X}) \triangleq \max_{\theta, (\mathbf{X}^{(1)}, \mathbf{X}^{(2)})} \| \log Pr(\theta, \mathbf{X}^{(1)}) - \log Pr(\theta, \mathbf{X}^{(2)}) \|_1$$

Privacy-preserving Bayesian inference via the exponential mechanism (OPS)

(Dimitrakakis et al., 2014; Wang et al., 2015)

- **Privacy cost of drawing a sample from posterior**
  - Interpret as exponential mechanism with the log joint probability $u(\mathbf{X}, \theta) = \log Pr(\theta, \mathbf{X})$ as the utility function:

$$f(\theta; \mathbf{X}, \epsilon) \propto \exp\left(\frac{\log Pr(\theta, \mathbf{X})}{T}\right) = Pr(\theta, \mathbf{X})^{1/T} \ , T = \frac{2\triangle \log Pr(\theta, \mathbf{X})}{\epsilon}$$

$$\text{where } \triangle \log Pr(\theta, \mathbf{X}) \triangleq \max_{\theta, (\mathbf{X}^{(1)}, \mathbf{X}^{(2)})} \| \log Pr(\theta, \mathbf{X}^{(1)}) - \log Pr(\theta, \mathbf{X}^{(2)}) \|_1$$

  - Setting $\epsilon = 2\triangle \log Pr(\theta, \mathbf{X})$ gives the **privacy we get “for free”** from posterior sampling
  - For smaller ϵ , flatten posterior by increasing the temperature

Privacy for exponential families

- Consider an exponential family likelihood with conjugate prior

$$Pr(\mathbf{X}|\theta) = \Big(\prod_{i=1}^{N} h(\mathbf{x}^{(i)})\Big) g(\theta)^N \exp\Big(\theta^\intercal \sum_{i=1}^{N} S(\mathbf{x}^{(i)})\Big)$$

$$Pr(\theta|\chi, \alpha) = f(\chi, \alpha) g(\theta)^\alpha \exp\Big(\alpha\theta^\intercal \chi\Big)$$

Privacy for exponential families

- Consider an exponential family likelihood with conjugate prior

$$Pr(\mathbf{X}|\theta) = \Big(\prod_{i=1}^{N} h(\mathbf{x}^{(i)})\Big) g(\theta)^N \exp\Big(\theta^\intercal \sum_{i=1}^{N} S(\mathbf{x}^{(i)})\Big)$$

$$Pr(\theta|\chi,\alpha) = f(\chi,\alpha) g(\theta)^\alpha \exp\Big(\alpha\theta^\intercal \chi\Big)$$

- The posterior is

$$Pr(\theta|\mathbf{X},\chi,\alpha) \propto g(\theta)^{N+\alpha} \exp\Big(\theta^\intercal \big(\sum_{i=1}^{N} S(\mathbf{x}^{(i)}) + \alpha\chi\big)\Big)$$

Privacy for exponential families: Exponential mechanism

- Sample from temperature-adjusted posterior

$$f(\theta; \mathbf{X}, \chi, \alpha, \epsilon) \propto g(\theta)^{\frac{N+\alpha}{T}} \exp\left(\theta^{\intercal} \frac{\sum_{i=1}^{N} S(\mathbf{x}^{(i)}) + \alpha\chi}{T}\right) , T = \frac{2\triangle \log p(\theta, X)}{\epsilon}$$

Privacy for exponential families via the Laplace mechanism

$$Pr(\theta|\mathbf{X},\chi,\alpha) \propto g(\theta)^{N+\alpha} \exp\Big(\theta^\intercal\big(\sum_{i=1}^{N} S(\mathbf{x}^{(i)}) + \alpha\chi\big)\Big)$$

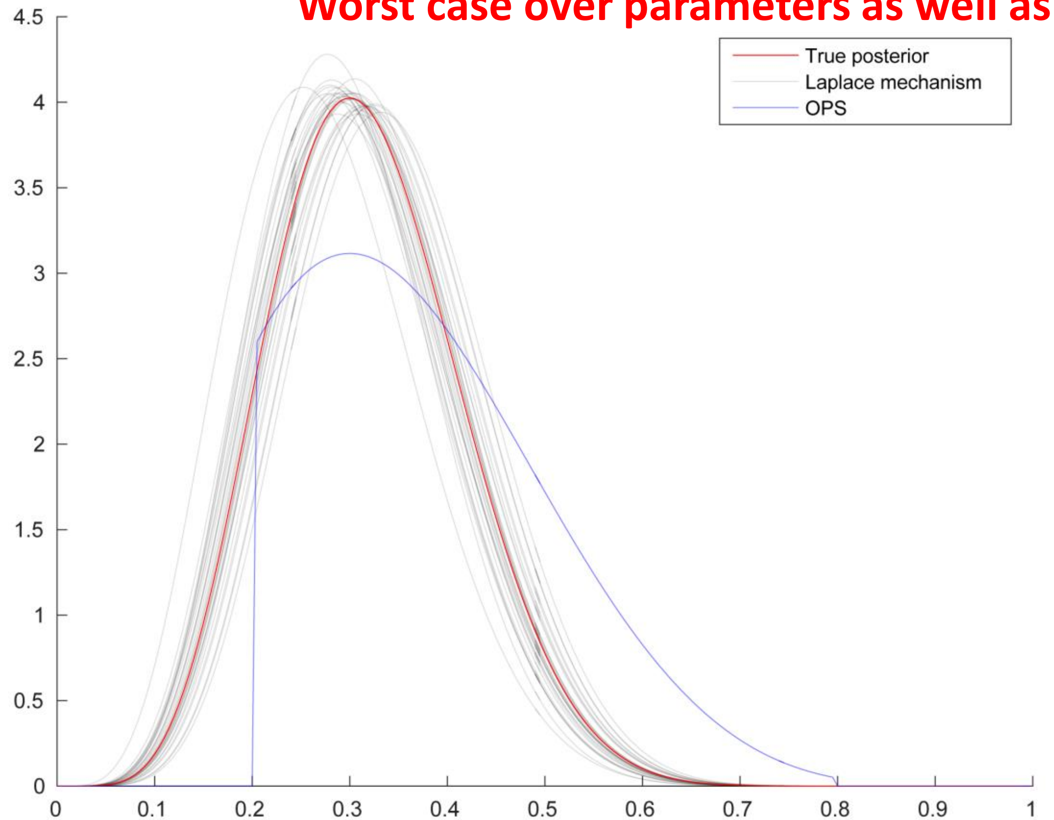
- Only interacts with the data via the aggregate sufficient statistics, $S(\mathbf{X}) = \sum_{i=1}^{N} S(\mathbf{x}^{(i)})$
- Add Laplace noise to $S(\mathbf{X})$.
  Releases privatized posterior, not just a sample!

Summary

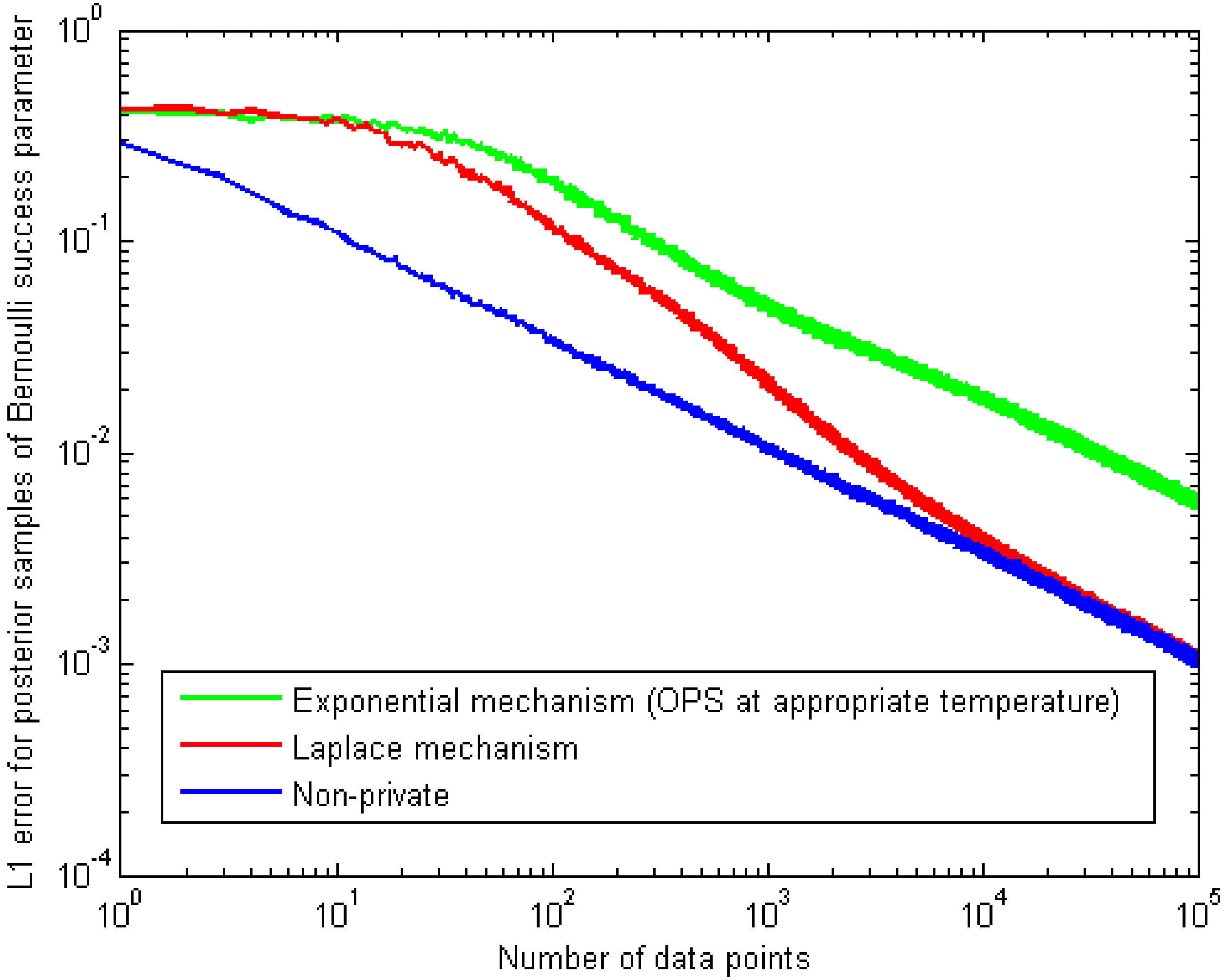
Mechanism	Sufficient statistics $S(\mathbf{X})$ are:	Release	Sensitivity
Laplace	Noised additively	Statistics	$\sup_{\mathbf{x},\mathbf{x}'} \lVert S(\mathbf{x}') - S(\mathbf{x}) \rVert_1$
Exponential	Rescaled multiplicatively	One sample	$\sup_{\mathbf{x},\mathbf{x}' \in \chi, \theta \in \Theta} \lvert \theta^{\mathsf{T}} \Big(S(\mathbf{x}') - S(\mathbf{x}) \Big) + \log h(\mathbf{x}') - \log h(\mathbf{x}) \rvert$

Worst case over parameters as well as data

Example: Beta-Bernoulli model



Data (in)efficiency in beta-Bernoulli model



Asymptotic relative efficiency

- **ARE** = ratio between variance of estimator and optimal variance achieved by posterior mean in the limit

- **Exponential mechanism**: **ARE = 1 + T**
  Temperature T >= 1 (Wang et al., 2015)

  Our results: under general conditions,
- **Laplace mechanism** (one sample): **ARE = 2**

- **Laplace mechanism** (posterior mean): **ARE = 1**

Assumptions for ARE result

- Laplace regularity conditions, and posterior satisfies asymptotic normality as in Bernstein-von Mises theorem:

1. The data $\mathbf{X}$ comes i.i.d. from a minimal exponential family distribution with natural parameter $\theta_0 \in \Theta$
2. θ_0 is in the interior of Θ
3. The function $A(\theta)$ has all derivatives for θ in the interior of Θ
4. $cov_{Pr(\mathbf{x}|\theta)}(S(\mathbf{x}))$ is finite for $\theta \in \mathcal{B}(\theta_0, \delta)$
5. $\exists w > 0$ s.t. $\det(cov_{Pr(\mathbf{x}|\theta)}(S(\mathbf{x}))) > w$ for $\theta \in \mathcal{B}(\theta_0, \delta)$
6. The prior $Pr(\theta|\chi, \alpha)$ is integrable and has support on a neighborhood of θ^*

Corollary 1. *The Laplace mechanism on an exponential family satisfies the noise distribution requirements above when the sensitivity of the sufficient statistics is finite and either the exponential family is minimal, or if the exponential family parameters θ are identifiable.*

Privacy of approximate sampling

- Posterior sampling in general intractable
  - exponential mechanism typically must be approximated.
- Approximate sampler is “close” to true posterior
  - Privacy cost will be close to that of a true posterior sample (Wang et al., 2015). However, cannot typically verify MCMC convergence
- Wang et al. also proposed an approximate sampling scheme via stochastic gradient Langevin dynamics.

Privacy of Gibbs sampling: Exponential mechanism

- We can interpret Gibbs updates as an instance of the exponential mechanism:

$$T^{(Gibbs,l,\epsilon)}(\theta,\theta') \propto Pr(\theta'_l, \theta_{\neg l}, \mathbf{X})^{\frac{\epsilon}{2\triangle \log Pr(\theta'_l, \theta_{\neg l}, \mathbf{X})}} ,$$

with utility function $u(\mathbf{X}, \theta'_l; \theta_{\neg l}) = \log Pr(\theta'_l, \theta_{\neg l}, \mathbf{X})$, over the space of possible assignments to θ_l, holding $\theta_{\neg l}$ fixed.

Privacy of Gibbs sampling: Exponential mechanism

- We can interpret Gibbs updates as an instance of the exponential mechanism:

$$T^{(Gibbs,l,\epsilon)}(\theta, \theta') \propto Pr(\theta'_l, \theta_{\neg l}, \mathbf{X})^{\frac{\epsilon}{2\triangle \log Pr(\theta'_l, \theta_{\neg l}, \mathbf{X})}} \ ,$$

with utility function $u(\mathbf{X}, \theta'_l; \theta_{\neg l}) = \log Pr(\theta'_l, \theta_{\neg l}, \mathbf{X})$, over the space of possible assignments to θ_l, holding $\theta_{\neg l}$ fixed.

- A Gibbs update is therefore $\epsilon = 2\triangle \log Pr(\theta'_l, \theta_{\neg l}, \mathbf{X})$-DP

Privacy of Gibbs sampling: Exponential mechanism

- We can interpret Gibbs updates as an instance of the exponential mechanism:

$$T^{(Gibbs,l,\epsilon)}(\theta,\theta') \propto Pr(\theta'_l,\theta_{\neg l},\mathbf{X})^{\frac{\epsilon}{2\triangle \log Pr(\theta'_l,\theta_{\neg l},\mathbf{X})}} \;,$$

with utility function $u(\mathbf{X},\theta'_l;\theta_{\neg l}) = \log Pr(\theta'_l,\theta_{\neg l},\mathbf{X})$, over the space of possible assignments to θ_l, holding $\theta_{\neg l}$ fixed.

- A Gibbs update is therefore $\epsilon = 2\triangle \log Pr(\theta'_l,\theta_{\neg l},\mathbf{X})$-DP

- Since worst case is computed over a strictly smaller set of outcomes,

$$\triangle \log Pr(\theta'_l,\theta_{\neg l},\mathbf{X}) \leq \triangle \log Pr(\theta,\mathbf{X})$$

Privacy of Gibbs sampling: Laplace mechanism

- If the Gibbs update interacts with the data via an exponential family likelihood, only need to privatize the sufficient statistics

- Can do this once at the beginning of the algorithm, and run as many iterations as we'd like!

- Unlike the exponential mechanism, the sampler does not need to converge to get verifiable privacy guarantees

- For this to work well, we need aggregate sufficient statistics to be large relative to Laplace noise, e.g. multiple observations per latent variable

Case study: Wikileaks war logs

- We investigate the performance of our technique on sensitive military data:
  - US military war logs from the wars in Iraq and Afghanistan disclosed by the Wikileaks organization.

- January 2004 - December 2009,
- Afghanistan: 75,000 log entries
- Iraq: 390,000 log entries

Wikileaks features

- Coarse-grained label ***"Type"**:*
  - *friendly action, explosive hazard, …*

- Fine-grained label ***"Category"***:
  - *mine found/cleared, show of force, …*

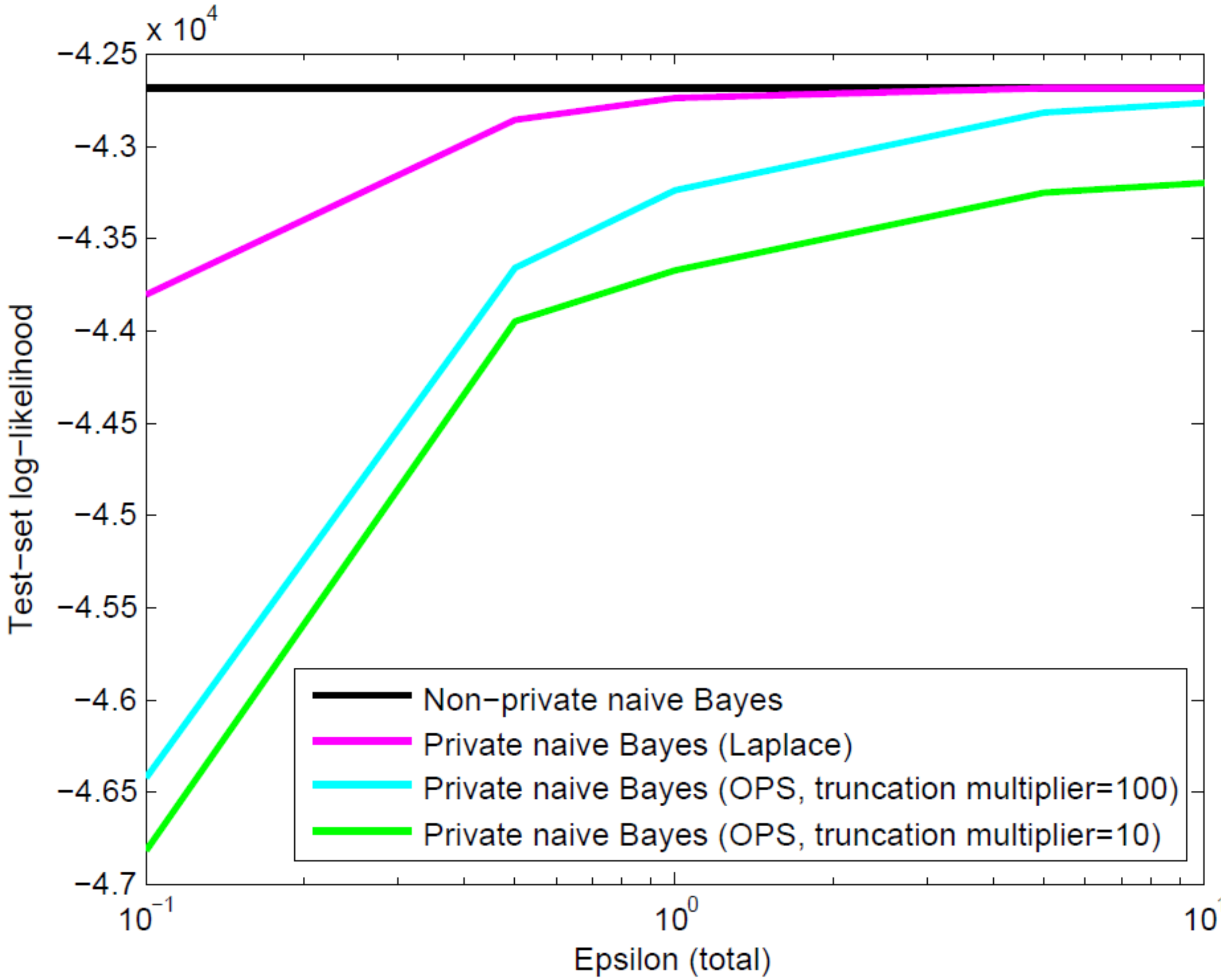
- **Casualties** for different factions:
  - Friendly/HostNation, Civilian, Enemy
    (names relative to US military perspective)
    1 IFF > 0 killed/wounded/captured/detained

Hidden Markov model for Wikileaks

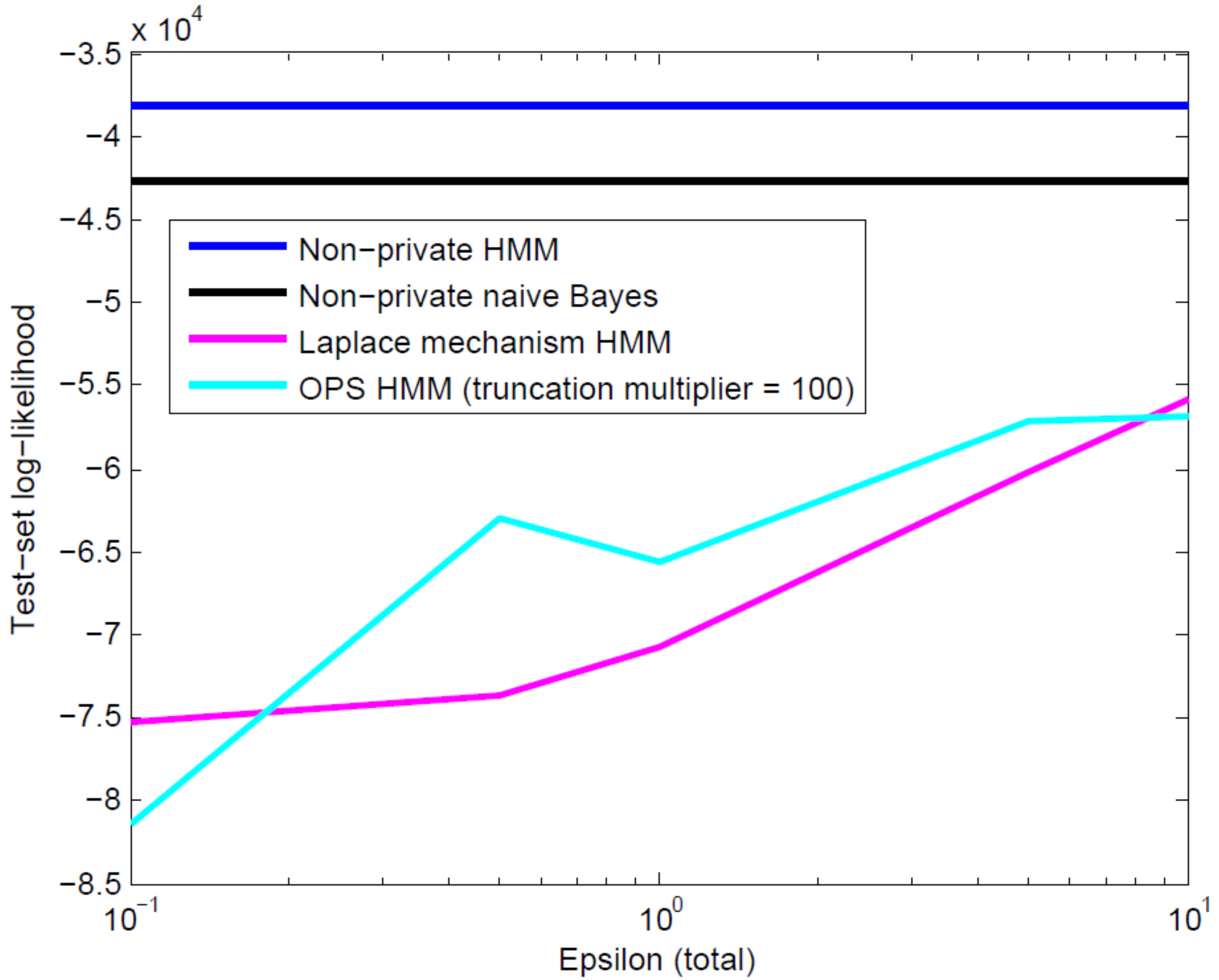
- An HMM chain of latent states for each region, with a timestep per month
  - Multiple emissions per timestep (all logs in that month)
- Naïve Bayes multinomial emissions
- 2 states for Iraq, 3 states for Afghanistan
- MCMC with a partially collapsed Gibbs sampler
- Total privacy budget epsilon = 5 for visualization results, varied from 10^{-1} to 10 for held-out log-likelihood experiments

  (10% timestep/region pairs held out, 10 train/test splits)

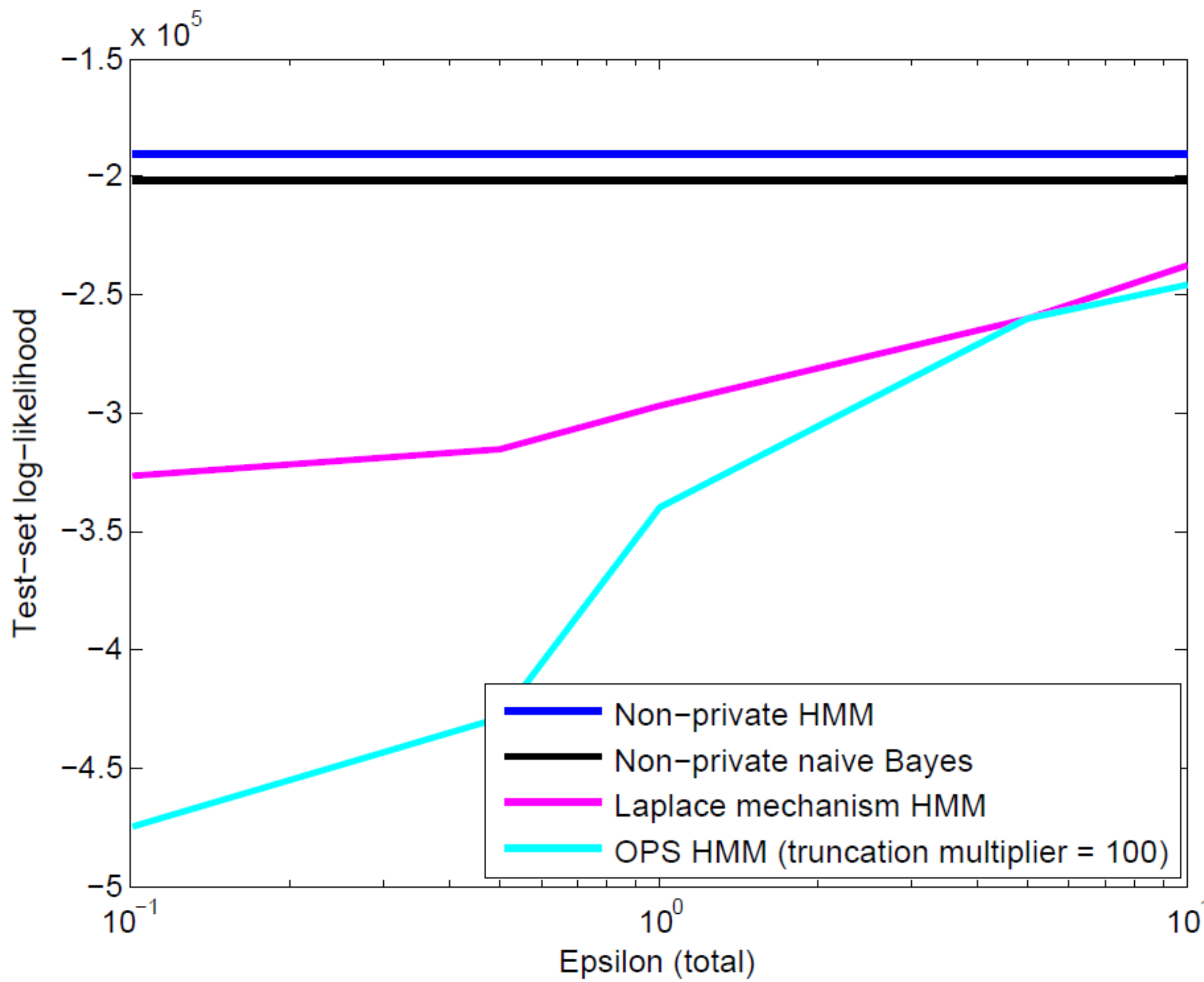
Held-out log-likelihood: Naïve Bayes (Afghanistan)



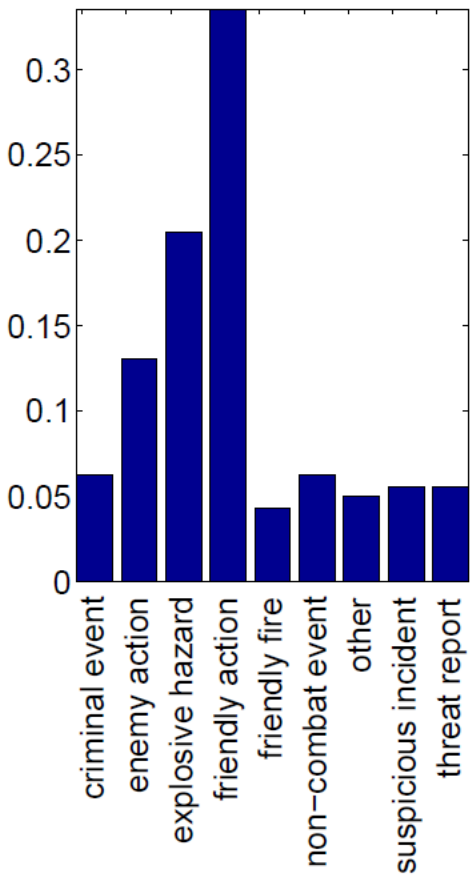
Held-out log-likelihood: Afghanistan



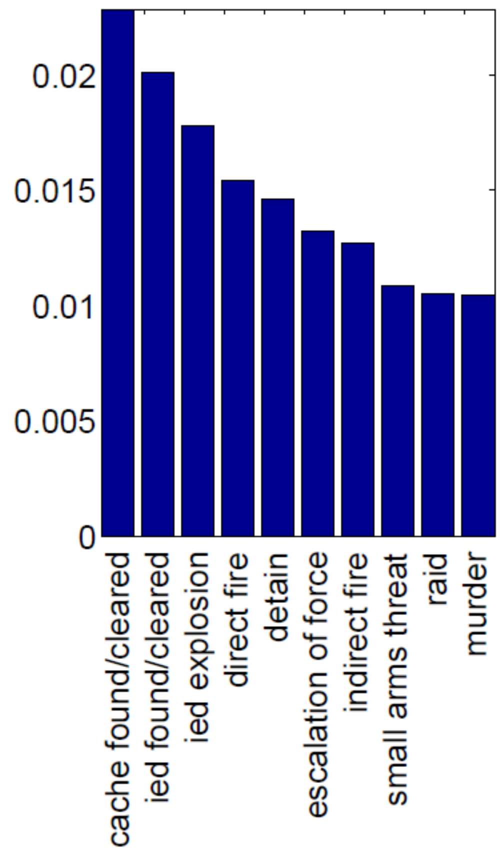
Held-out log-likelihood: Iraq



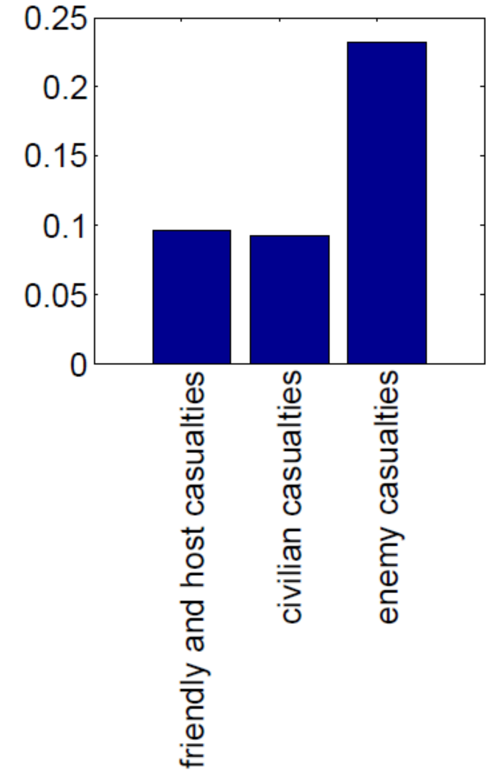
Visualization: Iraq, Laplace Mechanism
State 1: US military “doing well”



Type

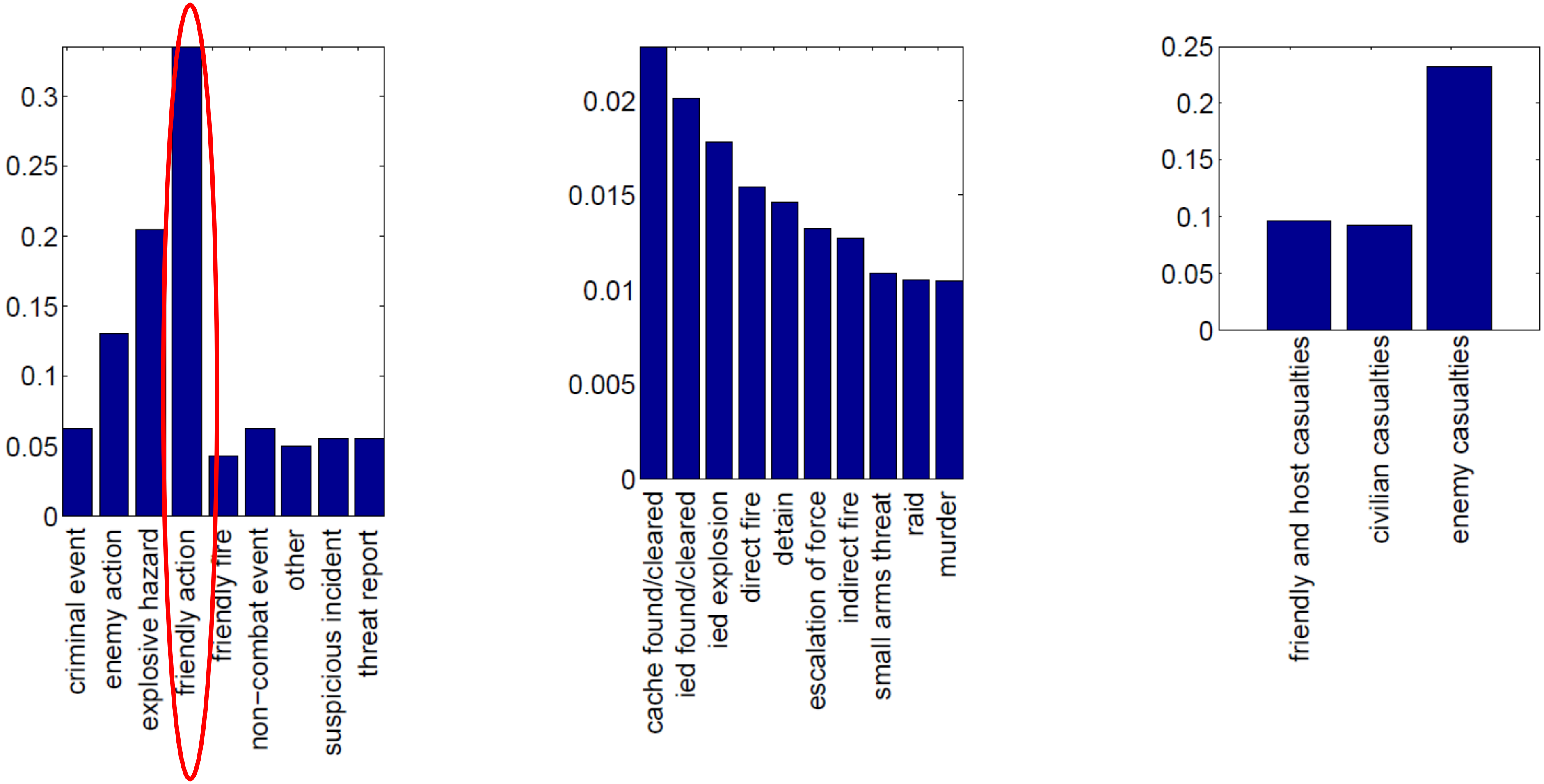


Category



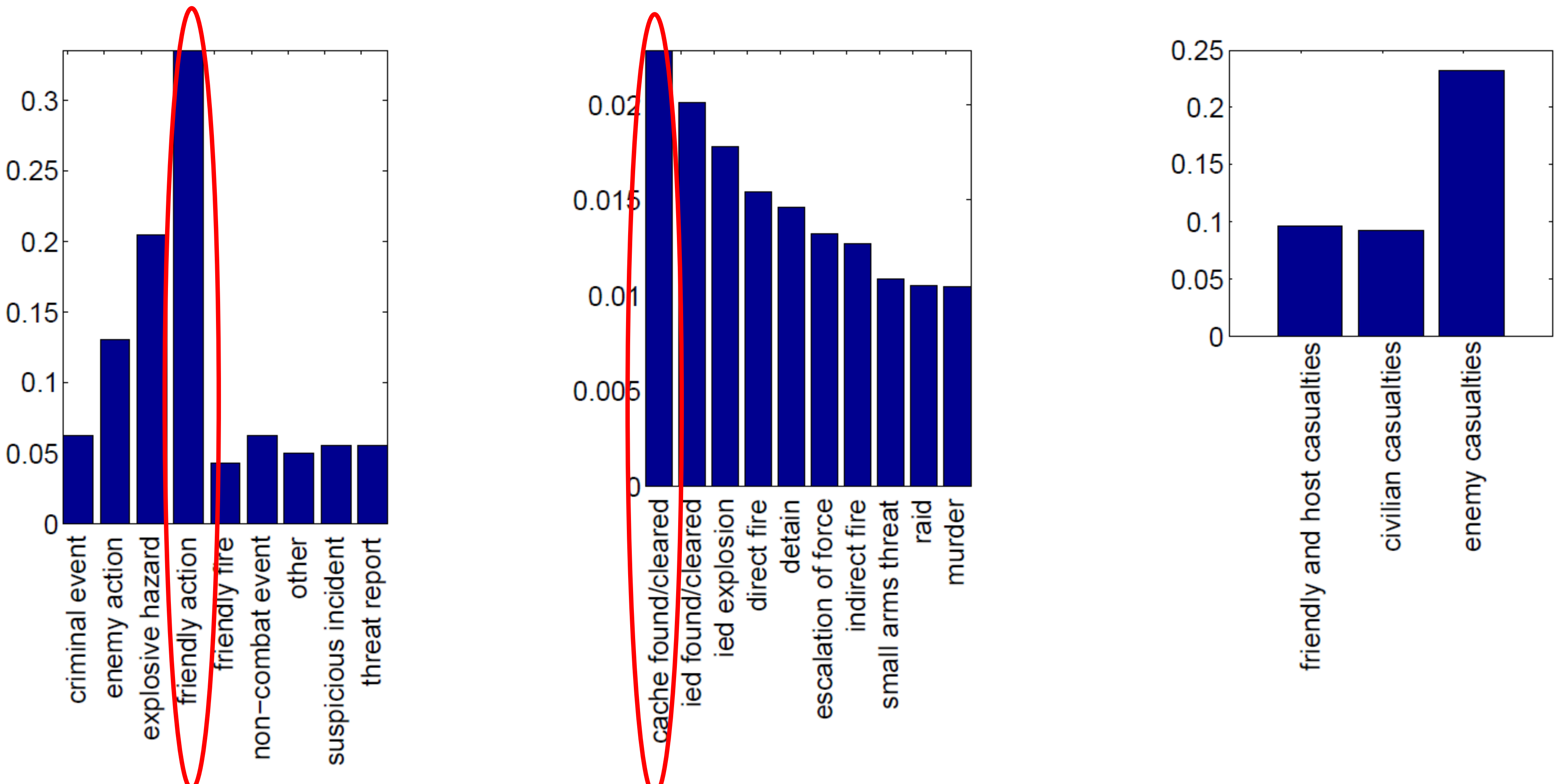
Casualties

Visualization: Iraq, Laplace Mechanism
State 1: US military "doing well"



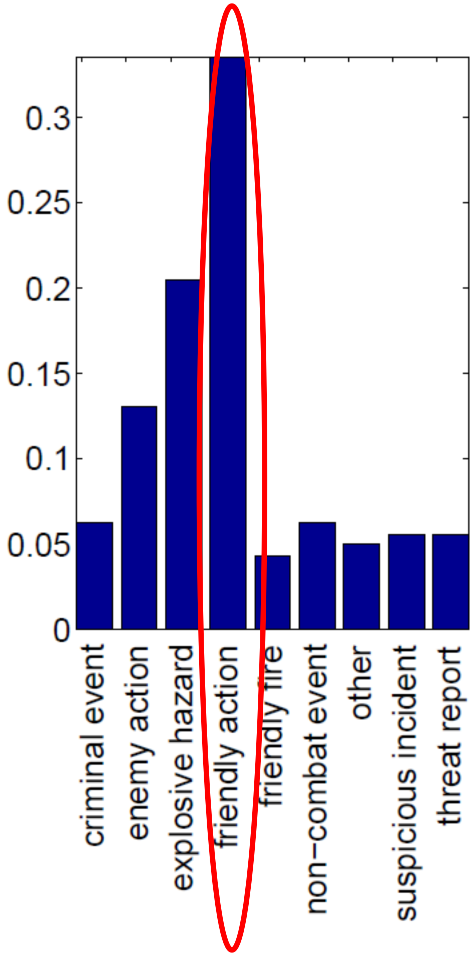
Type **Category** **Casualties**

Visualization: Iraq, Laplace Mechanism
State 1: US military "doing well"

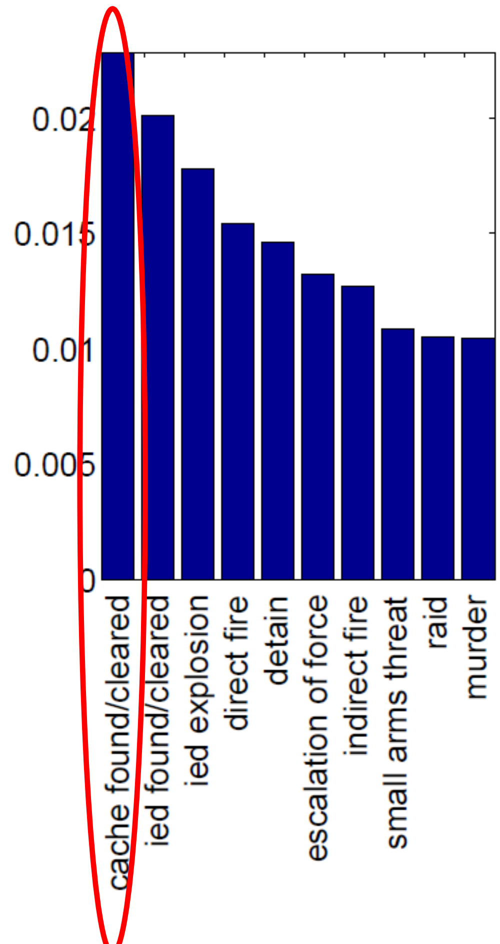


Type **Category** **Casualties**

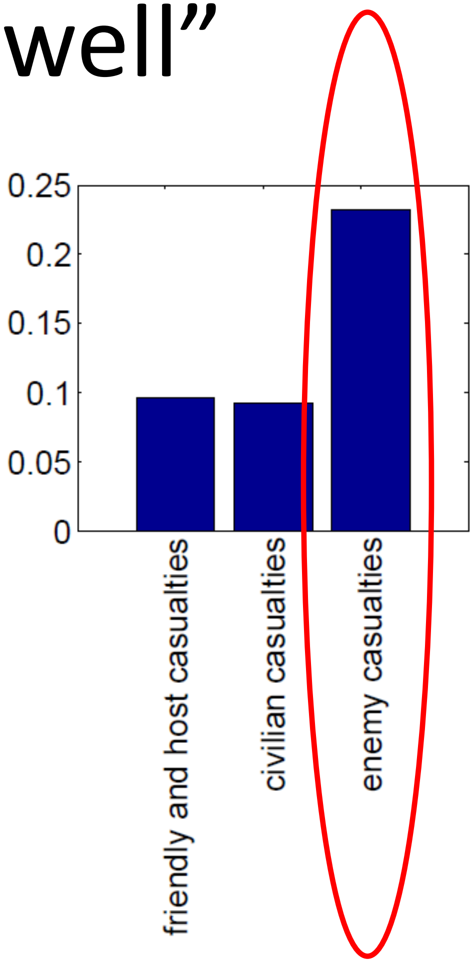
Visualization: Iraq, Laplace Mechanism
State 1: US military "doing well"



Type

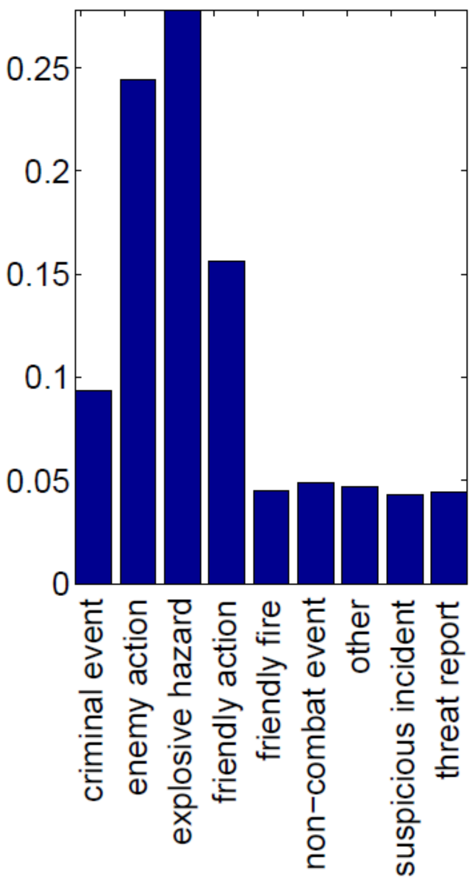


Category

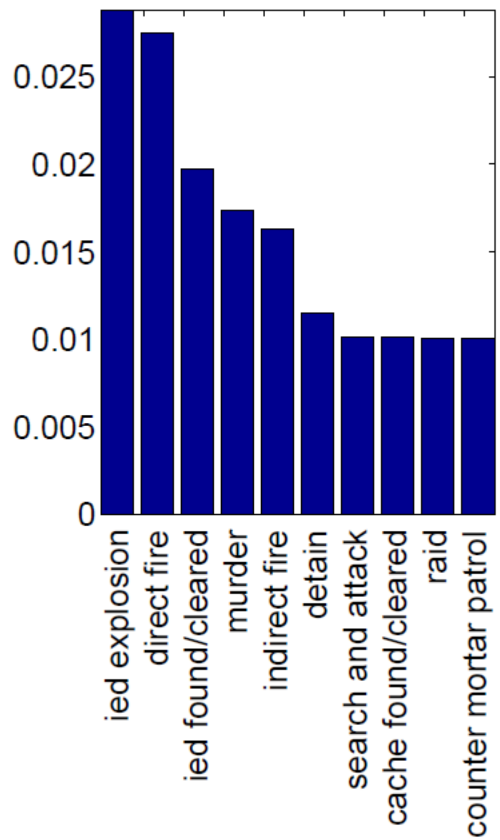


Casualties

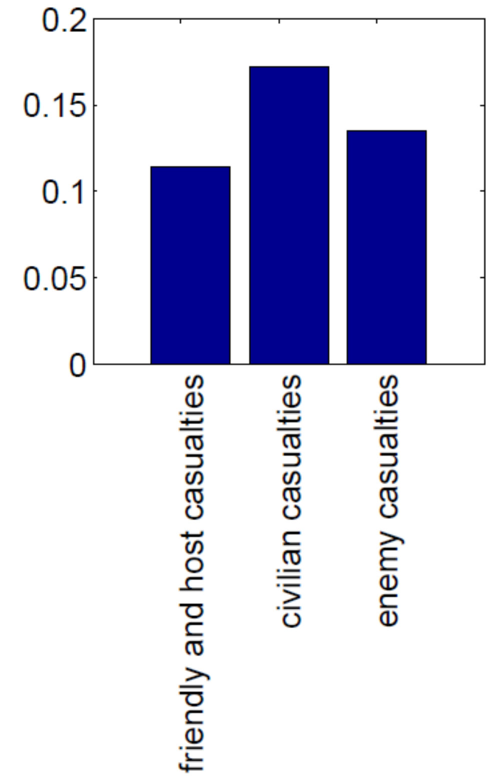
Visualization: Iraq, Laplace Mechanism State 2: US military "doing not so well"



Type

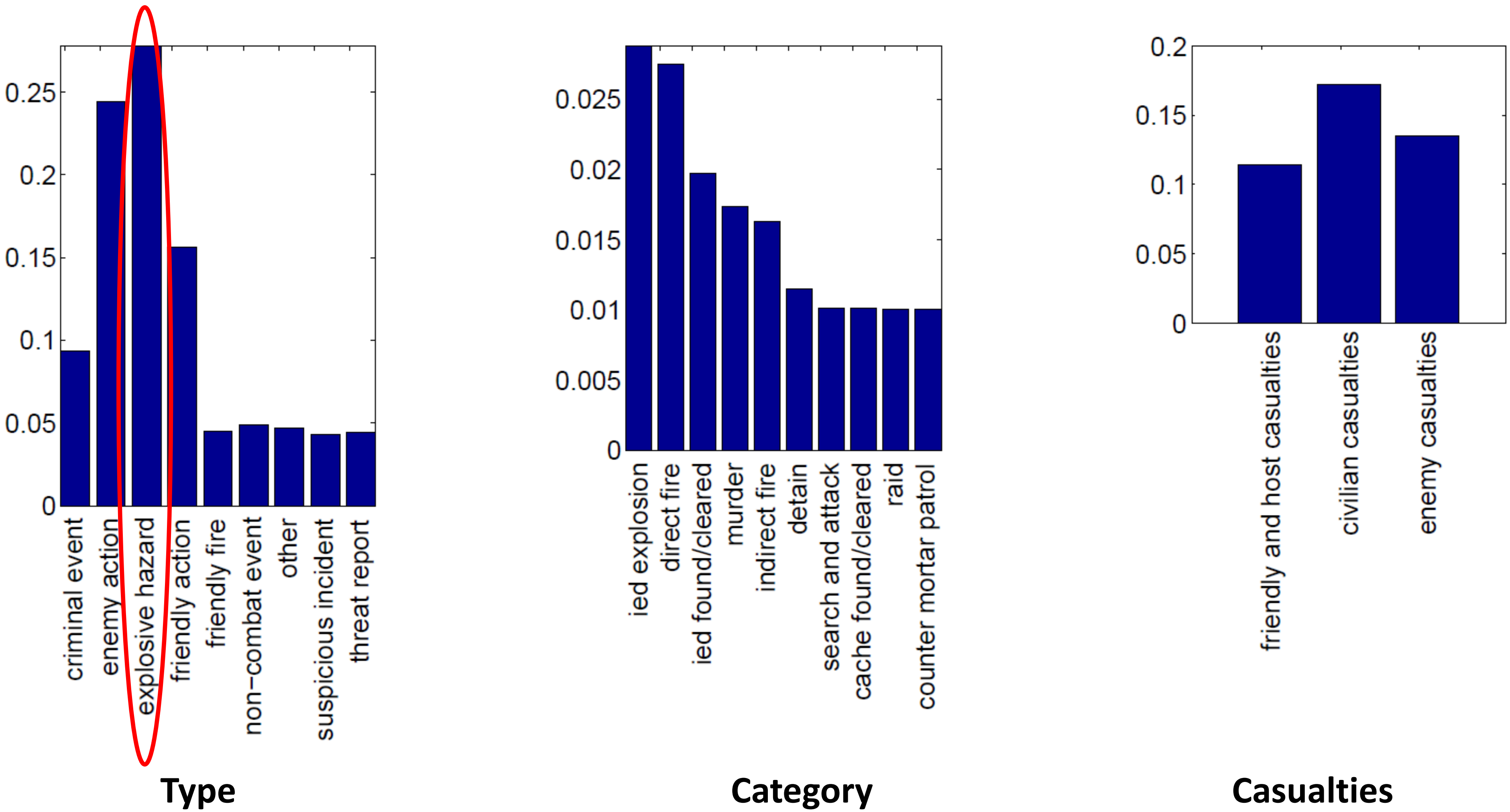


Category

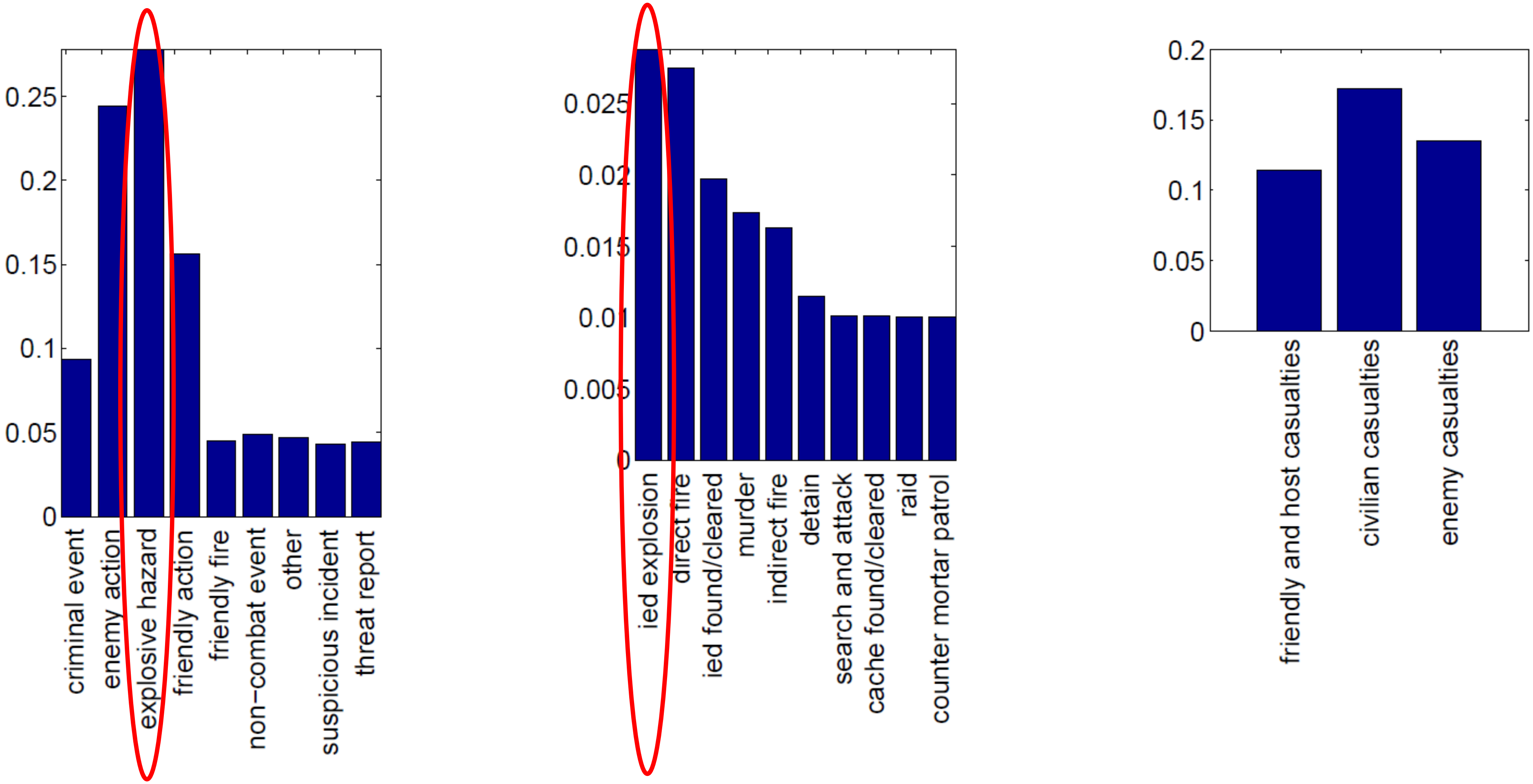


Casualties

Visualization: Iraq, Laplace Mechanism State 2: US military “doing not so well”

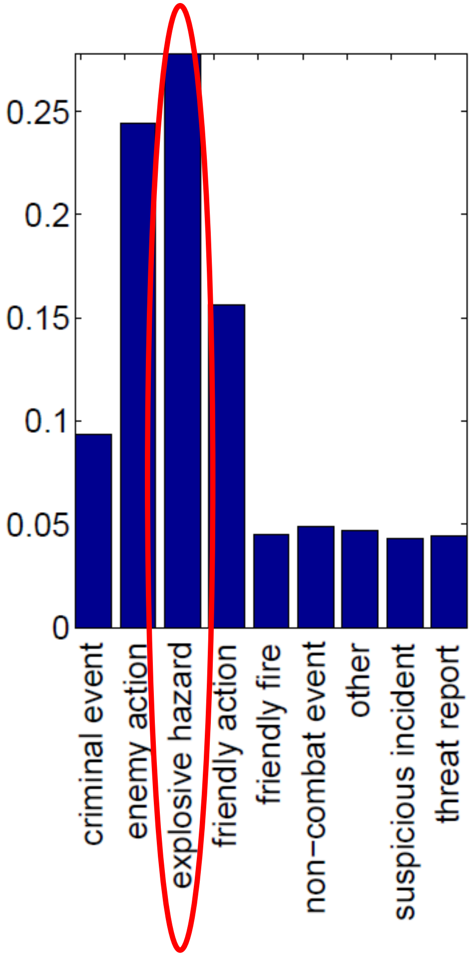


Visualization: Iraq, Laplace Mechanism State 2: US military "doing not so well"

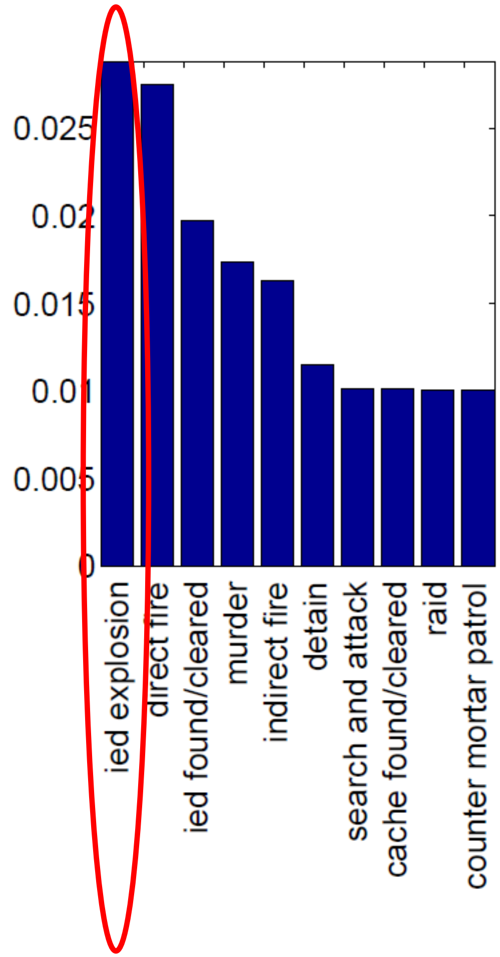


Type **Category** **Casualties**

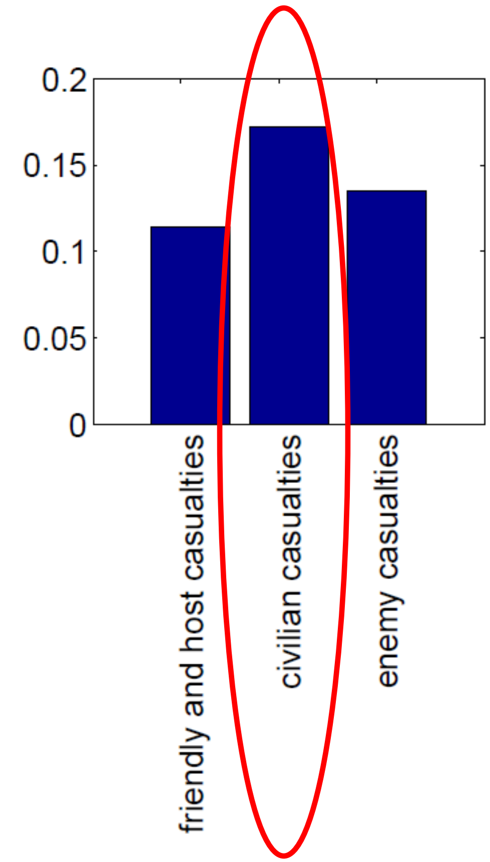
Visualization: Iraq, Laplace Mechanism State 2: US military "doing not so well"



Type

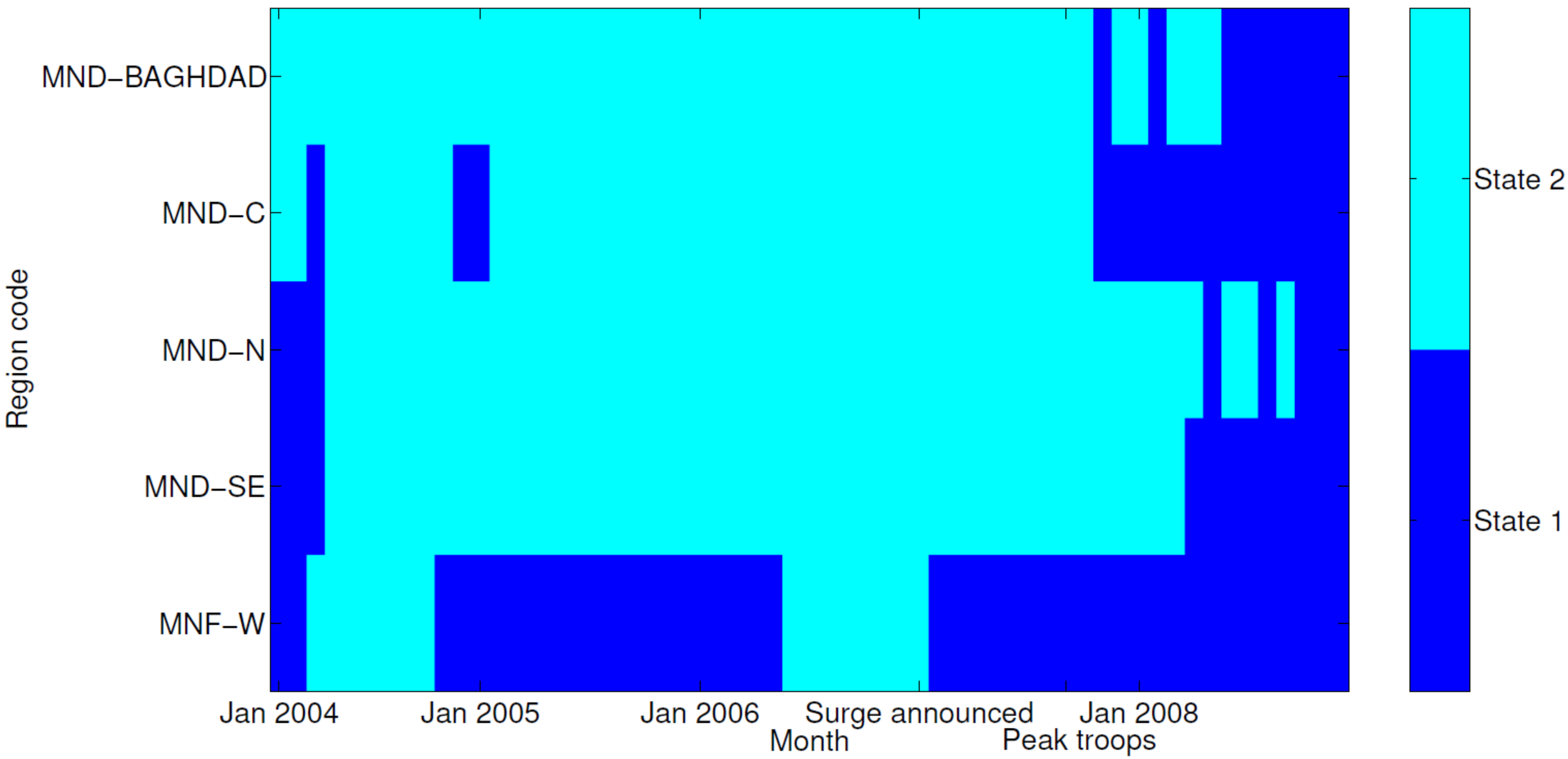


Category



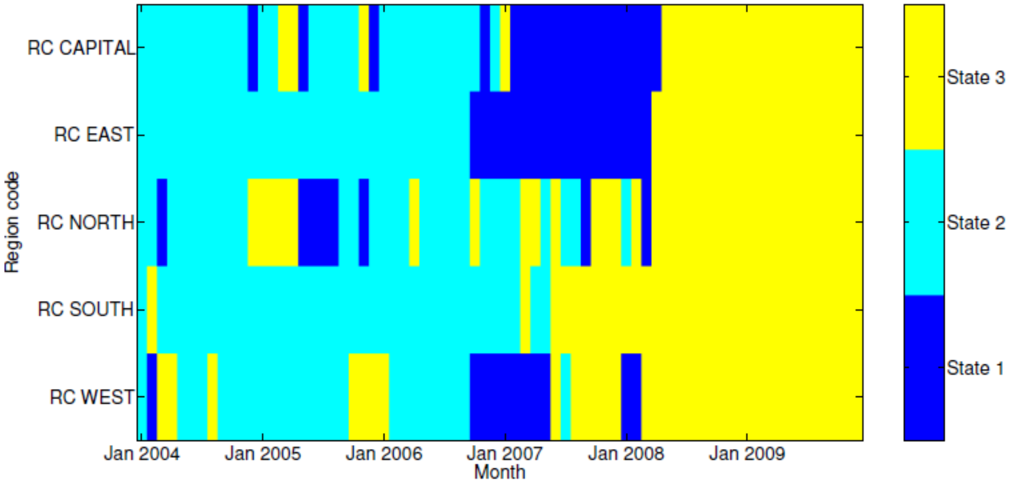
Casualties

Visualization: Iraq, Laplace Mechanism

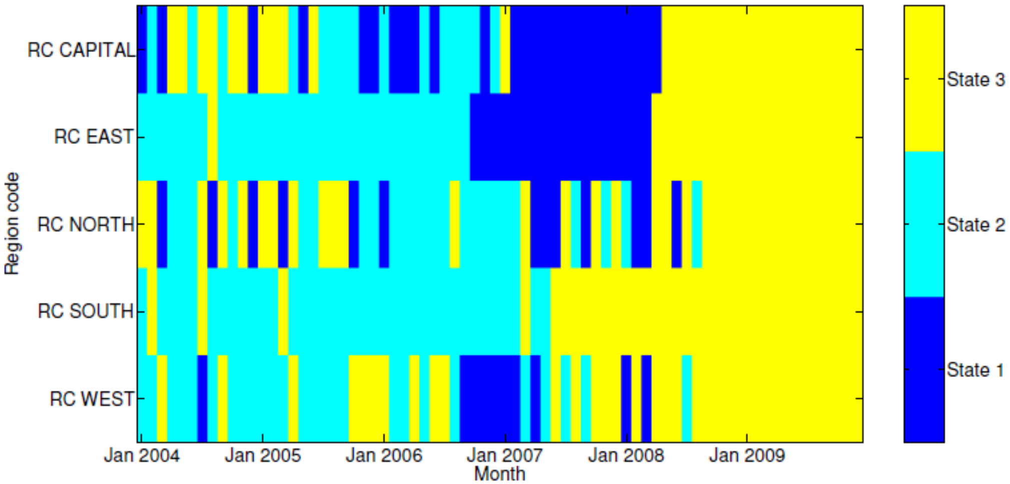


Visualization: Afghanistan, Exponential Mechanism

Last 100 samples:



Last 1 samples:



Conclusions

- We have proposed a Laplace mechanism approach for privacy-preserving Bayesian inference, as an alternative to the exponential mechanism (OPS) approach
- Asymptotic relative efficiency theorem shows data efficiency advantages vs exponential mechanism
- Privacy-preserving Gibbs sampling via exponential and Laplace mechanisms
- We demonstrated the benefits of our approach in a case study on an HMM time-series analysis of sensitive military records disclosed by Wikileaks

Future work

- Other approximate inference algorithms
  - In appendix, we analyze privacy of Metropolis-Hastings and annealed importance sampling.
  - Open problem to make better use of privacy budget to make these practical
  - New preprint on privacy-preserving EM!
    - **M. Park, J. R. Foulds, K. Chaudhuri, M. Welling. Practical Privacy for Expectation Maximization. ArXiv preprint arXiv:1605:06995 [cs.LG]**

- Practical applications to other sensitive real-world datasets: MOOCS, email data, genetic data…

- We have argued that asymptotic efficiency is important in a privacy context.
  - Open problem: How large is the class of privacy preserving algorithms that are asymptotically efficient?

Acknowledgements

- Collaborators:

Joseph Guemlek Max Welling Kamalika Chaudhuri

Conclusions

- We have proposed a Laplace mechanism approach for privacy-preserving Bayesian inference, as an alternative to the exponential mechanism (OPS) approach
- Asymptotic relative efficiency theorem shows data efficiency advantages vs exponential mechanism
- Privacy-preserving Gibbs sampling via exponential and Laplace mechanisms
- We demonstrated the benefits of our approach in a case study on an HMM time-series analysis of sensitive military records disclosed by Wikileaks

Thanks for your attention!