Differential Privacy in Estimation and Control

George J. Pappas
Joseph Moore Professor
Department of Electrical and Systems Engineering
University of Pennsylvania
(joint work with Jerome Le Ny)
How did I get here?
The trillion sensor challenge

- 5 Billion people to be connected by 2015 (Source: NSN)
- 7 trillion wireless devices serving 7 billion people in 2017 (Source: WWRF)
  - 1000 wireless devices per person?

[Courtesy: Niko Kiukkonen, Nokia]
Industrial Internet: Pushing the Boundaries of Minds and Machines
General Electric's Industrial Internet

Wave 1
**Industrial Revolution**
Machines and factories that power economies of scale and scope

Wave 2
**Internet Revolution**
Computing power and rise of distributed information networks

Wave 3
**Industrial Internet**
Machine-based analytics: physics-based, deep domain expertise, automated, predictive
2014 will be a tipping point. Mobile, bandwidth and the internet of things are going to change people's lives fundamentally.

Marissa Mayer  
Chief Executive Officer  
Yahoo
Industrial buzz

- Internet of Everything
- Smarter Planet
- Machine to Machine (M2M)
- The Fog
- T Sensors (Trillion Sensors)
- The Industrial Internet
- Industry 4.0
- Internet of Things (IoT)
- The Swarm
TerraSwarms: Swarm at the edge of the cloud

Trillions of Distributed Connected Devices Opportunistically Collaborating to Present Unique Experiences or Fulfill Common Goals

[J. Rabaey, ASPDAC'08]
TerraSwarm Challenges
Privacy Concerns - Cyber-physical applications

Intelligent Transportation Systems

Location based services

Wireless Implantable Medical Devices

Deep Brain Neurostimulators
Cochlear Implants
Gastric Stimulators
Cardiac Defibrillators/Pacemakers
Foot Drop Implants
Insulin Pumps

Camera Networks

Building Automation
The time you jump into the shower in the morning, the time you finally flick off that TV at night — even the time you set your home security alarm.

Ontario’s privacy czar wants protection for these habits.

“[The] thing has to be protected like Fort Knox,” says Ontario’s privacy commissioner Ann Cavoukian.

Toronto Star, May 12, 2010
Tanya Talaga

 Hydro meter info a boon for thieves, marketers, and must be protected, privacy czar says

Can Smart Grid know too much?

Utilities work to prevent privacy backlash over smart grid

SHAWN MCCARTHY - GLOBAL ENERGY REPORTER,
OTTAWA — The Globe and Mail
Published Wednesday Last updated Wednesday

The Smart Grid and Privacy
Concerning Privacy and Smart Grid Technology

- California Protects the Privacy of Smart Meter Data: The California Public Utility Commission has established new rules to protect information about consumer use of "smart meter" electrical services. The California decision, the first in the country, establishes four information practice requirements, including a consumer right of access and control, data minimization obligations, use and disclosure limitations, and data quality and integrity requirements. Electric utilities and their contractors, as well as third parties who receive electricity usage data from utilities are subject to the new rules. EPIC submitted extensive comments to the Public Utility Commission regarding privacy safeguards for consumer energy usage data. For more, see EPIC Smart Grid Privacy. (Aug. 6, 2011)

- Consumer Groups Recommend Privacy Safeguards on “Smart Meter” Services: The Trans-Atlantic Consumer Dialogue (TACD), a coalition of consumer groups in Europe and North America, adopted a report on privacy and electrical services at the 12th Annual TACD meeting held recently in Brussels. The Smart Meter-
Privacy challenges

• What is privacy, formally?

• What is the tradeoff between privacy and utility?

• Privacy-aware estimation and control

• Systems and control tools for privacy
Privacy is not anonymity

- Privacy breaches generally due to existence of side information
  - Mass. GIC medical db w/ voter registration db (Sweeney, 1997)
  - Netflix prize w/ IMDB (Narayana & Shmatikov, 2008)
  - Individual online transactions w/ changes in public recommendation systems (Calandrino et al., 2011)
  - Anonymity in location based services
- Can’t know what the adversary knows, or might know in the future.
A Model-based Framework for Privacy and Security Analysis of Traffic Monitoring Systems

Edward S. Canepa, Member, IEEE, and Christian G. Claudel, Member, IEEE

Abstract

Most large scale traffic information systems rely on fixed sensors (e.g. loop detectors, cameras) and user generated data, the latter in the form of GPS traces sent by smartphones or GPS devices onboard vehicles. While this type of data is relatively inexpensive to gather, it can pose multiple security and privacy risks, even if the location tracks are anonymous. In particular, creating bogus location tracks and sending them to the system is relatively easy. This bogus data could perturb traffic flow estimates, and disrupt the transportation system whenever these estimates are used for actuation. Another issue could be the possibility for an attacker to infer user location tracks from anonymous location data, which affects users privacy. In this article, we propose a new framework for solving a variety of privacy and cybersecurity problems arising in transportation systems. The state of traffic is modeled by the Lighthill-Whitham-Richards traffic flow model, which is a first order scalar conservation law with a concave flux function. Given a set of traffic flow data, we show that the constraints resulting from this partial differential equation are mixed integer linear inequalities for some decision variable. The resulting framework is very flexible, and can in particular be used to detect spoofing attacks in real time, or to carry out attacks on location tracks. Numerical implementations are performed on experimental data from the Mobile Century experiment.

I. Introduction

The convergence of mobile sensing, communication and computing has led to the rise of a

[Canepa and Claudel, HiCoNS’13]
In statistics
- Risk associated to small entries in table counts
- Techniques: cell suppression, clustering, perturbations...
  - [Duncan & Lambert, 1986], [Reiter, 2005]

In information theory
- Lower bound conditional entropy $H(\text{private info}|\text{public info})$ while still publishing useful information
  - [Sankar et al., 2010], [Venkitasubramaniam, 2013]

In computer science
- K-anonymity [Sweeney, 1998]
- Differential privacy [Dwork et al., 2006]

Privacy requires some form of obfuscation of the data published
Differential Privacy, informally

- **Set-up:**

  ![Diagram](image)

  - **Key Idea:**
    - A differentially private mechanism **randomly perturbs** the answers to a query so that the output distribution over answers does not vary much if any given individual participates or not.
    - Hard to infer if the data of any individual was used or not to answer the query.
Differential Privacy, informally
Differential Privacy, formally

• Set-up:

  - Formally
    - $\text{Adj}(d, d')$ a symmetric binary relation on the set $D$ of databases
    - Adjacent databases differ by the data of a single individual
    - A mechanism $M : D \times \Omega \rightarrow (R, \mathcal{M})$ is $(\varepsilon, \delta)$-DP if
      for all sets $\forall S \in \mathcal{M}$ and all databases $d, d'$ s.t. $\text{Adj}(d, d')$, we have

      $$\mathbb{P}(M(d) \in S) \leq e^\varepsilon \mathbb{P}(M(d') \in S) + \delta$$
Differential Privacy, formally

\[ \mathbb{P}(M(d) \in S) \leq e^\varepsilon \mathbb{P}(M(d') \in S) + \delta \]

- Constant \( \varepsilon \) is typically small (i.e. \( \sim 0.1 \)) - multiplicative error
- Constant \( \delta \) is very small (i.e. \( \sim 0.01 \)) - additive error
- If \( \delta=0 \) then we have \((\varepsilon,0)\)-DP or simply \(\varepsilon\)-DP
- Privacy definition depends on adjacency relation
Two Basic Differentially Private Mechanisms

- Database of MIT faculty salaries $d = [d_1, \ldots, d_n]$
- Adjacency: $\text{Adj}(d, d')$ iff for some $i$, $|d_i - d'_i| \leq \rho_i$
  \[ d_j = d'_j, j \neq i \]
  \[ R = \max \{\rho_i\} \]
- Analyst Query:
  \[ q(d) = \frac{1}{n} \sum_{i=1}^{n} d_i \]
- The mechanism $M(d) = q(d) + \text{Lap}(R/\epsilon)$ is $\epsilon$-DP
- The mechanism $M(d) = q(d) + N(0, (\kappa(\delta, \epsilon) R)^2)$ is $(\epsilon, \delta)$-DP

\[ \kappa(\delta, \epsilon) \leq 2\sqrt{2 \ln(2/\delta)}/\epsilon \]

\[ \left( \text{Lap}(b) \right. \text{ pdf: } \frac{1}{2b} e^{-|x|/b} \right) \]
More Advanced: The Exponential Mechanism

- Sometimes, adding noise to the query is not meaningful
  - Example: Who has the highest salary?
- The exponential mechanism (McSherry & Talwar, 2007)
  - A real-valued scoring function: \( u(q, d) \)
  - “Better” \( q \rightarrow \) larger \( u(q, d) \)
  - Example (highest salary): \( u(q = i, d) := \) “salary of person \( i \)”
  - Output \( q \) according to the probability distribution

\[
\frac{\exp(\varepsilon u(q, d)/2\Delta_u)}{\int_{q' \in \mathcal{M}} \exp(\varepsilon u(q', d)/2\Delta_u) \, dq'}
\]

\( \Delta_u := \max_q \max_{d, d' : \text{Adj}(d, d')} |u(q, d) - u(q, d')| \)

\( \varepsilon \)-differentially private

Global sensitivity of the scoring function
If mechanism $M(d)$ is $(\varepsilon, \delta)$-differentially private and $f$ is an arbitrary function, then $f(M(d))$ is also $(\varepsilon, \delta)$-differentially private.

- $f$ as the adversaries: Models arbitrary auxiliary or side information information the adversary may have. Privacy guarantee holds no matter what adversary does.
- $f$ as our algorithm: If we access the database in a differentially private way, we don’t have to worry about how our algorithm post-processes the result.
If mechanism

\[ M_1(d) \] is \( \epsilon_1 \)-differentially private and

\[ M_2(d) \] is \( \epsilon_2 \)-differentially private,

then the mechanism

\[ M(d) = (M_1(d), M_2(d)) \] is \( (\epsilon_1 + \epsilon_2) \)-differentially private
• Data mining with differential privacy A. Friedman 2010

• Combinatorial optimization with privacy A. Roth et al 2010

• Mechanism design via differential privacy M. Kearns et al 2012

• Privacy aware learning J. Duchi et al 2013

• Streaming continuous data C. Dwork et al 2010
Abstract—Emerging systems such as smart grids or intelligent transportation systems often require end-user applications to continuously send information to external data aggregators performing monitoring or control tasks. This can result in an undesirable loss of privacy for the users in exchange of the benefits provided by the application. Motivated by this trend, this paper introduces privacy concerns in a system theoretic context, and addresses the problem of releasing filtered signals that respect the privacy of the user data streams. Our approach relies on a formal notion of privacy from the database literature, called differential privacy, which provides strong privacy guarantees against adversaries with arbitrary side information. Methods are developed to approximate a given filter by a differentially private version, so that the distortion introduced by the privacy mechanism is minimized. Two specific scenarios are considered. First, the notion of differential privacy is extended to dynamic systems with many participants contributing independent input signals. Kalman filtering is also discussed in this context, when a released output signal must preserve differential privacy for the measured signals or state trajectories of the individual participants. Second, differentially private mechanisms are described to approximate stable filters when participants contribute to a single event stream, extending previous work on differential privacy under continual observation.

Index Terms—Estimation, filtering, Kalman filtering, privacy.

I. INTRODUCTION

A rapidly growing number of applications require users to release private data streams to third-party applications for signal processing and decision-making purposes. Examples include smart grids, population health monitoring, online recommendation systems, traffic monitoring, fuel consumption optimization, and cloud computing for industrial control systems. For privacy, confidentiality or security reasons, the participants benefiting from the services provided by these systems generally do not want to release more information than strictly necessary.

Manuscript received September 04, 2012; revised April 16, 2013; accepted August 22, 2013. Date of publication September 23, 2013; date of current version January 21, 2014. This work was supported in part by the Natural Sciences and Engineering Research Council under Grant RGPIN-419906-13 and in part by TerraSwarm, one of six centers of STARnet, a Semiconductor Research Corporation program sponsored by MARCO and DARPA. Preliminary versions of this paper appeared in [1] and [2]. Recommended by Associate Editor E. Scheuer.

J. Le Ny is with the Department of Electrical Engineering, École Polytechnique de Montréal, Montreal, QC, H3T 1J4, Canada (e-mail: jerome.le-ny@polymtl.ca).

G. J. Pappas is with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: pappas@seas.upenn.edu).

Digital Object Identifier 10.1109/TAC.2013.2280096

In a smart grid for example, a customer could receive better rates in exchange of continuously sending to the utility company her instantaneous power consumption, thereby helping to improve the demand forecast mechanism. In doing so however, she is also informing the utility or a potential eavesdropper about the type of appliances she owns as well as her daily activities [3]. Similarly, individual private signals can be recovered from published outputs aggregated from many users, and anonymizing a dataset is not enough to guarantee privacy, due to the existence of public side information. This is demonstrated in [4], [5] for example, where private ratings and transactions from individuals on commercial websites are successfully inferred with the help of information from public recommendation systems. Emerging traffic monitoring systems using position measurements from smartphones [6] is another application area where individual position traces can be re-identified by correlating them with public information such as a person’s location of residence or work [6], [7]. Hence, the development of rigorous privacy preserving mechanisms is crucial to address the justified concerns of potential users and thus encourage an increasing level of participation, which can in turn greatly improve the efficiency of these large-scale systems. Precisely defining what constitutes a breach of privacy is a delicate task. A particularly successful recent definition of privacy used in the database literature is that of differential privacy [8], which is motivated by the fact that any useful information provided by a dataset about a group of people can compromise the privacy of specific individuals due to the existence of side information. Differentially private mechanisms randomize their responses to dataset analysis requests and guarantee that whether or not an individual chooses to contribute her data only marginally changes the distribution over the published outputs. As a result, even an adversary cross-correlating these outputs with other sources of information cannot infer much more about specific individuals after publication than before [9].

Most work related to privacy is concerned with the analysis of static databases [8], [10]–[12], whereas cyber-physical systems clearly emphasize the need for mechanisms working with dynamic, time-varying data streams. Recently, the problem of releasing differentially private statistics when the input data takes the form of a binary stream describing event occurrences aggregated from many participants has been considered in [13]–[15]. This work forms the basis for the scenario studied in Section VI, and is discussed in more details in Section VI-C. However, most of this paper is devoted to a different situation where participants individually provide real-valued signals. A differentially private version of the iterated averaging algorithm for consensus is considered in [16]. In this case, the input data to protect consists of the initial values of the participants and is...
Private Filtering: Event Streams Counters

- Processing binary input signal (0-1 events) – DP linear filter approximation
  - Event-level privacy: each user contributes a unique event: two inputs differing by a single event must be hard to distinguish
  - Ex: Counter [Dwork et al., 2010], [Chan et al., 2011] (unstable stable filter)
  - Ex: Certain stable filters with slowly decreasing impulse response [Bolot et al., 2011]
  - Complicated algorithms (non-recursive, not finite memory), hard to generalize, poor performance of the approach for the approximation of stable filters

![Diagram of private filtering with moving averages (MA) for event streams counters](image)

\[ y_t = \sum_{i=1}^{n} \frac{1}{l} \sum_{k=t-l+1}^{t} u_{i,k} \]
• Approximate filter \( y = \sum_{i=1}^{n} G_i u_i \) by a differentially private version

\[
\begin{align*}
 & u_1 \quad \rightarrow \quad G_1 \\
 & u_2 \quad \rightarrow \quad G_2 \\
 & \vdots \\
 & u_n \quad \rightarrow \quad G_n \\
 & + \quad \rightarrow \quad y
\end{align*}
\]

• Adjacency relation

\[ \text{Adj}^b(u, u') \text{ iff for some } i, \|u_i - u'_i\|_2 \leq B, \text{ and } u_j = u'_j \text{ for all } j \neq i. \]
**Differentially Private Filtering**

- Approximate filter \( y = \sum_{i=1}^{n} G_i u_i \) by a differentially private version
  - Adjacency relation
    \[ \text{Adj}^b(u, u') \text{ iff for some } i, \|u_i - u'_i\|_2 \leq B, \text{ and } u_j = u'_j \text{ for all } j \neq i. \]

**Theorem**: For \((\varepsilon, \delta)\)-differential privacy, can add white Gaussian noise proportional to the maximum incremental gain with respect to the input channels

- Incremental gain: \( \| (Gu)_{0:T} - (Gu')_{0:T} \|_2 \leq \gamma \| u_{0:T} - u'_{0:T} \|_2, \forall u, u', \forall T \)
- Generalizes mechanism of [Dwork et al., 2006] to dynamic setting with continuous streams of real-time data
- System and control theoretic tools can be used to design differentially private mechanisms for continuous data streams
Approximate filter $y = \sum_{i=1}^{n} G_i u_i$ by a differentially private version

**Theorem:** The mechanism $M(u) = G(u) + w$ where $w$ is white noise with

$$w_t \sim \text{Lap}(B/\varepsilon)^m$$

is $\varepsilon$-differentially private.
Approximate filter $y = \sum_{i=1}^{n} G_i u_i$ by a differentially private version

$$\text{Theorem: The mechanism } M(u) = G(u) + w \text{ where } w \text{ is white noise with }$$

$$w_t \sim N(0, \sigma^2 I_m)$$

is $(\varepsilon, \delta)$-differentially private.
Two basic architectures for \((\epsilon, \delta)\)-differential privacy

\[
\begin{align*}
  w & \sim \mathcal{N}(0, \sigma^2) \\
  \sigma & = \kappa(\epsilon, \delta) B \\
  \text{MSE} & = \sigma^2 \sum_{i=1}^{n} \| G_i \|_2^2
\end{align*}
\]

\[
\begin{align*}
  w & \sim \mathcal{N}(0, \sigma^2) \\
  \sigma & = \kappa(\epsilon, \delta) B \max_{1 \leq i \leq n} \{ \| G_i \|_\infty \} \\
  \text{MSE} & = \sigma^2
\end{align*}
\]
Event Streams Counters

\[ y_t = \sum_{i=1}^{n} \frac{1}{l} \sum_{k=t-l+1}^{t} u_{i,k} \]

\[ \| G_i \|_2^2 = \frac{1}{l}, \| G_i \|_\infty = 1 \]

→ output perturbation better than input perturbation iff \( n > l \)
• Two binary-valued $u$, $u'$ signals are adjacent iff $u-u' = \pm \delta$

• $l_2$-Sensitivity of a linear filter $G$ is its $H_2$ norm:

$$\|Gu - Gu'\|_2 = \|g \ast \delta\|_2 = \|g\|_2 = \|G\|_2$$

• Privacy architecture: add noise proportional to $\|G_1\|_2$

---

Post-processing with $G_2$ does not alter privacy guarantee

Measure quality by MSE of error signal $e$
• $G_1$ is fixed $\rightarrow$ noise level $w$ is fixed $\rightarrow$ design of $G_2$ is a linear estimation (equalization) problem

• Zero-forcing equalizer: $G_1$ stable, minimum-phase, $G_2 = GG_1^{-1}$

• **Theorem** (Optimization over $G_1$):

$$MSE = \kappa(\delta, \epsilon)^2 \|G_1\|_2^2 \|GG_1^{-1}\|_2^2 \geq \kappa(\delta, \epsilon)^2 \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |G(e^{j\omega})| d\omega \right)^2$$

  – Bound achieved for: $|G_1(e^{j\omega})|^2 = |G(e^{j\omega})|$
Event Streams Counters

\[ y_t = \sum_{i=1}^{n} \frac{1}{l} \sum_{k=t-l+1}^{t} u_{i,k} \]
A Traffic Monitoring Example

• Traffic velocity estimation using individual location traces from smartphones (ex: Mobile Millenium, Bayen et al.)

\[ x_{i,t+1} = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix} x_{i,t} + \sigma_{i1} \begin{bmatrix} T_s^2/2 \\ T_s \end{bmatrix} w_{i,t}, \]

\[ y_{i,t} = \begin{bmatrix} 1 & 0 \end{bmatrix} x_{i,t} + \sigma_{i2} \begin{bmatrix} 0 & 1 \end{bmatrix} w_{i,t}. \]

Estimate \( \frac{1}{n} \sum_{i=1}^{n} x_{i,2,t} \)
• For Kalman Filtering, we have additional public information about the dynamics generating the user signals

\[
x_{i,t+1} = A_i x_{i,t} + B_i w_{i,t} \\
y_{i,t} = C_i x_{i,t} + D_i w_{i,t}
\]

- Estimation objective:

\[
z_t = \sum_{i=1}^{n} L_i x_{i,t} \\
\min_{\hat{z}} \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E} \left[ \| z_t - \hat{z}_t \|_2^2 \right]
\]

- Adjacency relation:

\[
\text{Adj}^\rho (x, x') \text{ iff for some } i, \| x_i - x'_i \|_2 \leq \rho_i, \text{ and } x_j = x'_j \text{ for all } j \neq i.
\]

- Cannot distinguish between two sufficiently close state trajectories of a user
Differentially Private Kalman Filtering

\[ x_{i,t+1} = A_i x_{i,t} + B_i w_{i,t} \]
\[ y_{i,t} = C_i x_{i,t} + D_i w_{i,t} \]

\[ \hat{z}_t = \sum_{i=1}^{n} L_i x_{i,t} \]

Adj \( S(x,x') \) iff for some \( i \), \( \| S_i x_i - S_i x'_i \|_2 \leq \rho_i \), \( (I-S_i)x_i = (I-S_i)x'_i \), and \( x_j = x'_j \) for all \( j \neq i \).

**Theorem 5:** Let \( \epsilon, \delta > 0 \). A mechanism releasing \( (\sum_{i=1}^{n} L_i K_i y_i) + \gamma \kappa(\delta, \epsilon) \nu \), where \( \nu \) is a standard white Gaussian noise independent of \( \{w_i\}_{1 \leq i \leq n}, \{x_i,0\}_{1 \leq i \leq n} \), and \( \gamma = \max_{1 \leq i \leq n} \left\{ \gamma_i \rho_i \right\} \), with \( \gamma_i \) the \( H_\infty \) norm of \( L_i K_i C_i S_i \), is \( (\epsilon, \delta) \)-differentially private for the adjacency relation (11).
• Can add the previous input and output perturbation schemes to the standard Kalman filter

\[ x_{i,t+1} = A_i x_{i,t} + B_i w_{i,t} \]
\[ y_{i,t} = C_i x_{i,t} + D_i w_{i,t} \]

• For the input perturbation scheme, can take into account the additional privacy-preserving noise in the redesign of the KF
Filter Redesign for Output Perturbation Scheme

\[ x_{i,t+1} = A_i x_{i,t} + B_i w_{i,t} \]
\[ y_{i,t} = C_i x_{i,t} + D_i w_{i,t} \]

• For the output perturbation scheme, can redesign the filter to trade-off the estimation error and the \( H_\infty \) norm of the filter
  – Overall MSE is
  \[
  \left( \sum_{i=1}^{n} \| TF(w_i \rightarrow e_i) \|_2^2 \right) + \kappa(\delta, \epsilon)^2 \max_{1 \leq i \leq n} \{ \rho_i^2 \| TF(x_i \rightarrow \hat{z}_i) \|_\infty^2 \}
  \]

• Multi-objective \( H_2/H_\infty \) optimization problem
  – Lyapunov shaping using Linear Matrix Inequalities
  – Distinguish between stable and unstable dynamics
• Traffic velocity estimation using individual location traces from smartphones (ex: Mobile Millenium, Bayen et al.)

\[
x_{i,t+1} = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix} x_{i,t} + \sigma_{i1} \begin{bmatrix} T_s^2/2 \\ T_s \end{bmatrix} w_{i,t},
\]

\[
y_{i,t} = \begin{bmatrix} 1 & 0 \end{bmatrix} x_{i,t} + \sigma_{i2} \begin{bmatrix} 0 & 1 \end{bmatrix} w_{i,t}
\]

Estimate \( \frac{1}{n} \sum_{i=1}^{n} x_{i,2,t} \)

\[
\rho = 100m, \\
\epsilon = \ln 3, \\
\delta = 0.05, \\
T_s = 1s, \\
\sigma_{i1} = \sigma_{i2} = 1, \\
n = 200
\]
Traffic Monitoring – input versus output architectures

Diagram showing traffic monitoring with input and output architectures.
Traffic Monitoring – utility/privacy tradeoff

The graph shows the RMSE (km/h) on the y-axis and ε on the x-axis. There are three lines in the graph:
- Blue dashed line: Input pert. + original KF
- Green dotted line: Output pert. + original KF
- Red solid line: Input pert. + compensating KF

The RMSE decreases as ε increases for all three lines, indicating a tradeoff between utility and privacy.
Traffic Monitoring – convergence/privacy tradeoff

The diagram illustrates the convergence time (s) as a function of ε. Two lines are plotted:

- **Input pert. + compensating KF**: A blue line that shows a significant decrease in convergence time as ε increases.
- **Output pert. + original KF**: A green dotted line that remains relatively flat and lower compared to the input perturbation line.

The x-axis represents ε values ranging from 0 to 1.2, while the y-axis represents the convergence time (s) ranging from 0 to 160.
Differential Privacy in Decision Making

Available at: http://arxiv.org/abs/1411.4105
Motivating Example: Charging of Electric Vehicles

Each user has his own charging specification (max charge rate, total energy).

Central mediator tries to minimize a certain objective that depends on the aggregate load.

Optimization problem:

\[
\begin{align*}
\min_{\{r_i\}_{i=1}^n} & \quad U \left( \sum_{i=1}^n r_i \right) \\
\text{s.t.} & \quad 0 \leq r_i \leq \bar{r}_i, \quad 1^T r_i = E_i, \quad i = 1, 2, \ldots, n.
\end{align*}
\]
• The EV charging problem has the following form:

\[
\begin{align*}
\min_{\{r_i\}_{i=1}^n} & \quad U \left( \sum_{i=1}^n r_i \right) \\
\text{s.t.} & \quad r_i \in C_i, \quad i = 1, 2, \ldots, n.
\end{align*}
\]

• Objective function: Depends on the “aggregate action”
• Constraints: Decoupled among users
• Why distributed: For scalability
• Can be viewed as *Distributed Projected Gradient Descent*

**Algorithm 1** Distributed projected gradient descent (with a fixed number of iterations).

<table>
<thead>
<tr>
<th>Input:</th>
<th>$U$, ${C_i}<em>{i=1}^n$, $K$, and step sizes ${\alpha_k}</em>{k=1}^K$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output:</td>
<td>${r_i^{(K+1)}}_{i=1}^n$.</td>
</tr>
<tr>
<td>Initialize</td>
<td>${r_i^{(1)}}_{i=1}^n$ arbitrarily. For $k = 1, 2, \ldots, K$, repeat:</td>
</tr>
</tbody>
</table>

1. Compute $p^{(k)} := \nabla U \left( \sum_{i=1}^n r_i^{(k)} \right)$. \hspace{1cm} Gradient
2. For $i \in [n]$, update $r_i^{(k+1)}$ according to

$$r_i^{(k+1)} := \Pi_{C_i} (r_i^{(k)} - \alpha_k p^{(k)}).$$ \hspace{1cm} Projection
• The constraint set $C_i$ may contain sensitive information

In electric vehicle charging:

• Both $r_i$ and $E_i$ contain sensitive information about the user

• Example on $r_i$: “Unable to charge my car between 8-10pm” → User may not be at home from 8-10pm.

• Example on $E_i$: “Need to charge my car by a large amount at night” → This user may have used the car heavily during the day.
The Distributed Projected Grad. Descent is Non-Private

- The gradient $p^{(k)}$ is available to all users
- $p^{(k)} = \nabla U \left( \sum_{i=1}^{n} r_{i}^{(k)} \right)$ depends on the aggregate load
- With enough side information, an adversary (who pretends to be a participating user) can learn from $p^{(k)}$ about the load profile of another user
- The load profile can reveal $\bar{r}_{i}$ and $E_{i}$ (which are sensitive)

- The goal of a privacy-preserving algorithm: Solve the constraint optimization problem, while ensuring diff. privacy on the gradients.
Putting Things in the Differential Privacy Framework

• Database: \( D = \{C_i\}_{i=1}^n \)
  – Two databases are adjacent if at most one pair \((C_i, C'_i)\) is different

• Mechanism: \( M_p := (\hat{p}^{(1)}, \hat{p}^{(2)}, \ldots, \hat{p}^{(K)}) \) (i.e., approx. gradients)

• **Problem statement**: Find a distributed optimization algorithm in which the published gradients \( M_p \) satisfy diff. privacy, i.e.:

\[
\mathbb{P}(M_p(D) \in \mathcal{R}) \leq e^\epsilon \mathbb{P}(M_p(D') \in \mathcal{R})
\]

for all adjacent \( D, D' \) and all \( \mathcal{R} \subseteq \text{range}(M_p) \)
Theorem: Algorithm 2 preserves $\epsilon$-differential privacy of the gradients $M_P := (\hat{p}^{(1)}, \hat{p}^{(2)}, \ldots, \hat{p}^{(K)})$.

Algorithm 2 Differentially private distributed projected gradient descent.

Input: $U$, $L$, $\{C_i\}_{i=1}^n$, $K$, $\{\alpha_k\}_{k=1}^K$, $\eta \geq 1$, $\Delta$, and $\epsilon$.

Output: $\{\hat{r}_i^{(K+1)}\}_{i=1}^n$.

Initialize $\{r_i^{(1)}\}_{i=1}^n$ arbitrarily. Let $\hat{r}_i^{(1)} = r_i^{(1)}$ for all $i \in [n]$ and $\theta_k = (\eta + 1)/(\eta + k)$ for $k \in [K]$.

For $k = 1, 2, \ldots, K$, repeat:

1) If $k = 1$, then set $w_k = 0$; else draw a random vector $w_k \in \mathbb{R}^T$ from the distribution (proportional to)
   $\exp \left( -\frac{2\epsilon \|w_k\|}{K(K-1)L\Delta} \right)$

2) Compute $\hat{\nabla} := \nabla U \left( \sum_{i=1}^n r_i^{(k)} \right) + w_k$.

3) For $i \in [n]$, compute:
   \[
   r_i^{(k+1)} := \Pi_{C_i} (r_i^{(k)} - \alpha_k \hat{\nabla}) ,
   \hat{r}_i^{(k+1)} := (1 - \theta_k) \hat{r}_i^{(k)} + \theta_k r_i^{(k+1)}.
   \]

min. $U(\sum_{i=1}^n r_i)$

s.t. $r_i \in C_i$, $i = 1, 2, \ldots, n.$

sensitivity of projection (defined later)

additive noise to ensure privacy

noise magnitude $\propto \frac{K(K-1)\Delta}{\epsilon}$
Sensitivity of Projection (onto the Constraint Set)

\[ \Delta := \max_{i \in [n]} \max \left\{ \| \Pi_{C_i}(r) - \Pi_{C'_i}(r) \| : r \in \mathbb{R}^T \right\}, \]

\[ C_i \text{ and } C'_i \text{ are adjacent} \}

Looks complicated, but can be computed in certain cases!

**Example:** Computation of \( \Delta \) for the EV charging problem

- Adjacency relation: Recall that \( C_i \) is defined through \( \bar{r}_i \) and \( E_i \)
- \( C_i \) and \( C'_i \) are adjacent if (\( \delta r \) and \( \delta E \) are design choices)

\[ \| \bar{r}_i - \bar{r}_i' \|_1 \leq \delta r, \quad |E_i - E_i'| \leq \delta E \]

**Theorem:** The sensitivity for the EV charging problem is given by

\[ \Delta = 2\delta r + \delta E \]

Proof: Use results from sensitivity analysis in optimization theory.
Simulation Results: EV Charging

- Number of vehicles: 100,000
- Privacy level: \( \text{epsilon} = 0.1 \)
- Objective: Minimize load variance

DP solution is very close to the optimal solution!
Choosing the Number of Iterations

- # of iterations can not be too large nor too small
- Too few: Affects convergence of algorithm
- Too many: More gradients to preserve privacy → more noise
**Theorem:** The suboptimality of the differentially private distributed projected gradient descent algorithm is given by

\[
\mathbb{E} \left[ U \left( \sum_{i=1}^{n} \hat{r}_i^{(K+1)} \right) - U^* \right] \leq O \left( \eta T^{1/4} (\Delta/n\epsilon)^{1/4} \right)
\]

- Suboptimality decreases as epsilon grows (less privacy)
- Suboptimality decreases as n grows (more users)
• Many cyber-physical applications raise privacy concerns that need to be addressed to encourage user participation

• Need privacy-preserving mechanisms for various types of dynamic systems and data

• Characterizing privacy-utility tradeoff requires a quantitative definition of privacy

• System and control theoretic tools (optimal estimators, system gains) can be used to design differentially private mechanisms
Next steps

- Fundamental limits between privacy and performance
- Optimal mechanism design
- Scaling laws
- Apply to SmartMeter and transportation models
- Protect controllers and strategy not data
- Privacy pricing