Differentially Private State Estimation in Distribution Networks with Smart Meters

Henrik Sandberg, György Dán, and Ragnar Thobaben
ACCESS Linnaeus Centre
KTH Royal Institute of Technology, Stockholm, Sweden
Motivation

- **The promise:** Smart meters help in demand response, billing, etc.

- Few real-time measurements in today’s distribution networks → Enabler for state estimation?

- **The threat:** Customers’ privacy concerns (among others)

- **The opportunity:** Privacy-preserving monitoring and control techniques
Distribution Network Model

- Conservation of currents:
  \[ I_j = \sum_{k>j} L_k, \quad j = 0, 1, \ldots, N - 1 \]

- Large resistance in distribution grids → use currents
- Little dynamics in (current) distribution grids → study steady-state

- **Operator**: Desires to estimate load \( L_j \)
- **Customer** \( C \): Desires to keep his/her real-time load private
The Base Scenario – Total current with “physical meter noise”:

\[ Z_0 = I_0 + W_0 \quad \text{with} \quad W_0 \sim \mathcal{N}(0, R_0) \]

The Smart Meter Scenario – Load current with “privacy noise”:

\[ Z_j = L_j + W_j \quad \text{with} \quad W_j \sim \text{Lap}(b_j) \]

\[ p_{W_j}(w) = \frac{1}{2b_j} e^{-|w|/b_j}, \quad R_j = 2b_j^2 \]
Problem Formulation:
Characterize Estimation vs. Privacy Trade-Offs

![Graph showing estimation improvement vs. loss of privacy for different customers]
Related Work

• **Differential privacy:**
  • Dwork, McSherry, Nissim, Smith, 2006

• **Differential privacy in control:**
  • Le Ny, Pappas, 2014
  • Huang, Wang, Mitra, Dullerud, 2014

• **Privacy for Smart Meters:**
  • Ács, Castelluccia, 2011
  • Tan, Gunduz, Poor, 2013
Differential Privacy [Dwork et al., 2006]

- Two adjacent data vectors:
  \[ l = (l_1 \ l_2 \ \ldots \ l_i \ \ldots \ l_{m-1} \ l_m)^T \]
  \[ l' = (l_1 \ l_2 \ \ldots \ l_i \pm \Delta \ \ldots \ l_{m-1} \ l_m)^T \]

- Measurement policy (\( q \) deterministic, \( W \) stoch. noise)
  \[ Z(l, W) = q(l) + W \]
  (We will use \( q(l) = \sum_k l_k \))

**Definition:** Measurement \( Z \) is \((\epsilon, \delta)\)-differentially private if for all events \( E \):
\[
\Pr[Z(l, W) \in E] \leq e^\epsilon \Pr[Z(l', W) \in E] + \delta
\]
Example: $\epsilon$-Differential Privacy with Laplacian Noise

Measurements of adjacent data vectors virtually indistinguishable for small $\epsilon$
The Base Scenario \((Z_0)\): Customer \(C\) has \((\epsilon_0, \delta_0)\)-differential privacy where

\[
\epsilon_0 = \frac{\Delta K}{\sigma_0} + \frac{\Delta^2}{2\sigma_0^2}, \quad K = K(\delta_0) = Q^{-1}(\delta_0)
\]
Optimal Estimate: Load Model

Suppose loads have a known normal distribution:

\[ L \sim \mathcal{N}(m, P) \]

\[
m = \begin{pmatrix} m_1 \\ \vdots \\ m_N \end{pmatrix} \quad P = \begin{pmatrix} P_{11} & \cdots & P_{1N} \\ \vdots & \ddots & \vdots \\ P_{N1} & \cdots & P_{NN} \end{pmatrix} \quad \begin{pmatrix} P_1 \\ \vdots \\ P_N \end{pmatrix} = \begin{pmatrix} P_{11} + \cdots + P_{1N} \\ \vdots \\ P_{N1} + \cdots + P_{NN} \end{pmatrix} = P_1
\]
Optimal Estimate: Base Scenario

\[ W_0 \sim \mathcal{N}(0, \sigma_0^2) \]

\[ Z_0 \]

\[ \text{Substation} \]

\[ I_0 \quad I_1 \quad I_j \quad I_{N-1} \]

\[ L_1 \quad \ldots \quad L_j \quad \ldots \quad L_N \]

**MMSE estimate:**

\[ \hat{L}_j^0 := \mathbf{E}[L_j | Z_0] = m_j + \frac{P_j}{P_0 + R_0} (Z_0 - m_0) \]

**MMSE error:**

\[ Q_{j,j}^0 := \mathbf{E}[(\hat{L}_j^0 - L_j)^2] = P_{j,j} - \frac{P_j^2}{P_0 + R_0} \]
Optimal Estimate: Smart Meter Scenario

$L_0 \sim \mathcal{N}(0, \sigma_0^2)$ \quad $Z_0$

$L_j \sim \text{Lap}(b_j)$

LMMSE estimate:

$$\hat{L}_j^{0,j} := \mathbb{E}^{\text{lin}}[L_j | Z_0, Z_j]$$

$$= \hat{L}_j + K_j \left[ (Z_j - m_j) - \frac{P_j}{R_0 + P_0} (Z_0 - m_0) \right]$$

LMMSE error:

$$K_j = \frac{(R_0 + P_0)P_{jj} - P_j^2}{R_0 + P_0} \in [0, 1]$$

$$Q_j^{0,j} := \mathbb{E}[(L_j - \hat{L}_j^{0,j})^2] = Q_j^0 (1 - K_j) \leq Q_j^0$$
Trade-Off: Estimation Quality vs. Privacy

Dimensionless quantities:

- Customers’ relative importance at site $j$: 
  $$\eta_j := \frac{\Delta^2}{P_{jj}}$$

- Site $j$’s relative importance on the line: 
  $$\zeta_j := \frac{P_{jj}}{P_0 + R_0}$$

Substation

$$C \in [l_i - \Delta, l_i + \Delta]$$

$$L_j \sim \mathcal{N}(m_j, P_{jj})$$
Trade-Off: Estimation Quality vs. Privacy

Baseline privacy:

\[ \epsilon_0^2 \approx \frac{\Delta^2 K^2}{R_0} \]

\[ = \eta_j \zeta_j K^2 \left(1 + \frac{P_0}{R_0}\right) \]

Est. improvement:

\[ K_j = \frac{1}{1 + \frac{2\eta_j}{\epsilon^2 (1 - \zeta_j)}} \]

\[ \approx \frac{\epsilon^2 (1 - \zeta_j)}{2\eta_j} \]
Summary

Simple analytical treatment of trade-off between state estimation quality and customers’ privacy loss $\epsilon$

Estimation gain $\sim \left(\frac{\epsilon}{\epsilon_0}\right)^2$ → Customers with high baseline privacy can make a large difference!

Possible extensions: Dynamics, general topologies, active/reactive power flows

Est. improvement with Smart Meters:

$$\geq \frac{1}{1 + \frac{2n_j}{\epsilon^2(1-\zeta_j)}} \approx \frac{K^2}{2} \left(1 + \frac{P_0}{R_0}\right) \zeta_j(1 - \zeta_j) \left(\frac{\epsilon}{\epsilon_0}\right)^2$$

- SNR at substation
- Site size
- Norm. privacy loss