Smart Meter Privacy
with an Alternative Energy Source

Deniz Gündüz
Imperial College London
joint work with
Giulio Giaconi (Imperial College), Onur Tan (CTTC)
Jesús Gómez Vilardebó (CTTC), H. Vincent Poor (Princeton)

24 November 2015
Lyon, France
Smart grid refers to the future energy grid that exploits information technologies to increase reliability, to increase efficiency and reduce carbon footprint, to incorporate renewable as well as traditional energy sources, to provide security, to introduce new services that cannot be foreseen today.
**Smart Meters**

Smart meters (SMs) are an essential component of smart grids; they enable many “smart” grid functionalities.

SMs introduce the ability to provide bi-directional communication between consumers and the energy providers/grid operator and to promote services that facilitate energy efficiency within the home\(^1\).

European Parliament’s directive\(^2\) requires 80\% SM adoption in all European households by 2020, and 100\% by 2022.

---

1. The European Commission’s Interpretative Note on Directive 2009/72/EC.
Benefits of Smart Meters

**Benefits for Consumers:**
- Ability to track energy consumption near real time,
- More accurate and timely billing services,
- Possibility to benefit from demand flexibility and time-of-use (ToU) pricing,
- Possibility to introduce safety solutions, detection of waste and unexpected activity or inactivity,
- Increase competition among energy providers.

**Benefits for the Energy Provider/ Grid Operator**
- Reduced cost of meter readings and back office rebilling processes,
- Misuse and fraud detection,
- Accurate sensing and monitoring capabilities to distribution network operators (DNOs)
- Possibility to introduce demand response management in order to reduce peak loads,
- Renewable integration and microgeneration management.
Netherlands: Senate voted against mandatory roll-out of SMs, found to be against European Convention on human rights

9000 consumers polled in 17 countries: 1/3 discouraged from using smart meters if it gave utilities greater access to data about their personal energy use
Non-intrusive load monitoring (NILM) techniques using high frequency SM data

Reveals information on user’s energy consumption behaviour: can track appliance usage patterns, home occupancy, even the TV channel user is watching ...

Smart Meter Privacy: Social Angle

- **Patterns** (behaviour profiling)
  - Watching too much TV?
  - Another microwave meal?

- **Real-time surveillance**
  - Were you home last night?
  - Did your friend move in?

- **Non-grid use of data**
  - Advertising and spam
  - Insurance
  - Appliance warranties

- **Information leakage**
  - Phishing, pharming, fraud

Security ≠ Privacy

Remote switching off capability of smart meters opens up new vulnerabilities (Stuxnet type cyber attacks)

Meters can be hacked by consumers or third parties to reduce/increase energy bill

- A utility in Puerto Rico lost $400 million in annual revenue after criminals hacked into smart meters to under-report electricity usage.

Highly connected metering infrastructure allows spread of malware

Wireless transmission of meter readings is prone to eavesdropping and data injection attacks
Focus of current SMs is on protection against manipulation by customers.

Grid operators/EPs can remotely update crucial meter parameters (e.g., cryptographic keys, sampling frequency), install new software, or disconnect energy.

Measurement data collected and stored in database of the operator.

Trust in grid operators: consumers are protected mainly by guidelines, audits, codes of behaviour.

To protect privacy we first need to measure it.
In this talk: Information theoretic privacy
   - A probabilistic approach
   - Privacy against any detection algorithm (independent of computation power)
   - Consider a storage device and/or an alternative energy source

Other measures are also possible:
   - Variation in consumption: keep consumption constant
   - Differential privacy
   - Cluster classification
   - Description length (interestingness)
Existing Approaches

Meter data is modified before being reported to the energy provider.

- Anonymization with/ without trusted third party (TTP), i.e., using pseudonyms instead of real identities.
  

- Aggregation with/ without trusted third party (TTP), i.e. summing measurements over a group of users,


- Obfuscation, i.e., adding noise to data (Differential privacy, quantization)


Energy consumption is modified

- Through storage devices, i.e., filtering energy consumption
  


- Through alternative energy sources (renewables, uninterruptable power supplies),

- Sampling approach, i.e., reducing sampling rate of measurements.
Discrete time model:
- Energy demand (input load): $X_t$,
- Energy from grid (output load): $Y_t$
- Remainder from alternative energy source (AES): $X_t - Y_t$
- SM reads and reports $Y_t$
Energy management policy: \( f_t : \mathcal{X}^t \times \mathcal{Y}^{t-1} \rightarrow \mathcal{Y} \)

- Peak power constraint:

\[
0 \leq X_t - Y_t \leq \bar{P}
\]

- Average power from AES:

\[
P_n = \mathbb{E} \left[ \frac{1}{n} \sum_{t=1}^{n} (X_t - Y_t) \right]
\]

- Privacy: Information leakage rate

\[
I_n \triangleq \frac{1}{n} I(X^n; Y^n) = \frac{1}{n} [H(X^n) - H(X^n|Y^n)]
\]

For given \( \bar{P} \), pair \((I, \hat{P})\) is achievable if there exist energy management policies with \( \lim_{n \rightarrow \infty} I_n \leq I \) and \( \lim_{n \rightarrow \infty} P_n \leq \hat{P} \).

Privacy-power function, \( \mathcal{I}(\bar{P}, \hat{P}) \): For given \( \bar{P} \) and \( \hat{P} \), minimum information leakage rate such that \((I, \hat{P})\) is achievable.
Energy Management Policy

- Energy management policy: $f_t : \mathcal{X}^t \times \mathcal{Y}^{t-1} \to \mathcal{Y}$
- Peak power constraint:

$$0 \leq X_t - Y_t \leq \bar{P}$$

- Average power from AES:

$$P_n = \mathbb{E}\left[\frac{1}{n} \sum_{t=1}^{n} (X_t - Y_t)\right]$$

- Privacy: Information leakage rate

$$I_n \triangleq \frac{1}{n} I(X^n; Y^n) = \frac{1}{n} [H(X^n) - H(X^n|Y^n)]$$

- For given $\bar{P}$, pair $(I, \hat{P})$ is achievable if there exist energy management policies with $\lim_{n\to\infty} I_n \leq I$ and $\lim_{n\to\infty} P_n \leq \hat{P}$.
- Privacy-power function, $\mathcal{I}(\bar{P}, \hat{P})$: For given $\bar{P}$ and $\hat{P}$, minimum information leakage rate such that $(I, \hat{P})$ is achievable.
Energy management policy: \( f_t : \mathcal{X}_t \times \mathcal{Y}_{t-1} \to \mathcal{Y} \)

Peak power constraint:

\[
0 \leq X_t - Y_t \leq \bar{P}
\]

Average power from AES:

\[
P_n = \mathbb{E} \left[ \frac{1}{n} \sum_{t=1}^{n} (X_t - Y_t) \right]
\]

Privacy: Information leakage rate

\[
I_n \triangleq \frac{1}{n} I(X^n; Y^n) = \frac{1}{n} \left[ H(X^n) - H(X^n | Y^n) \right]
\]

For given \( \bar{P} \), pair \((I, \hat{P})\) is achievable if there exist energy management policies with \( \lim_{n \to \infty} I_n \leq I \) and \( \lim_{n \to \infty} P_n \leq \hat{P} \).

Privacy-power function, \( \mathcal{T}(\bar{P}, \hat{P}) \): For given \( \bar{P} \) and \( \hat{P} \), minimum information leakage rate such that \((I, \hat{P})\) is achievable.
Energy Management Policy

- Energy management policy: \( f_t : \mathcal{X}^t \times \mathcal{Y}^{t-1} \rightarrow \mathcal{Y} \)
- Peak power constraint:

\[
0 \leq X_t - Y_t \leq \bar{P}
\]

- Average power from AES:

\[
P_n = \mathbb{E} \left[ \frac{1}{n} \sum_{t=1}^{n} (X_t - Y_t) \right]
\]

- Privacy: Information leakage rate

\[
I_n \triangleq \frac{1}{n} I(X^n; Y^n) = \frac{1}{n} [H(X^n) - H(X^n | Y^n)]
\]

For given \( \bar{P} \), pair \((I, \hat{P})\) is achievable if there exist energy management policies with \( \lim_{n \rightarrow \infty} I_n \leq I \) and \( \lim_{n \rightarrow \infty} P_n \leq \hat{P} \).

Privacy-power function, \( \mathcal{I}(\bar{P}, \hat{P}) \): For given \( \bar{P} \) and \( \hat{P} \), minimum information leakage rate such that \((I, \hat{P})\) is achievable.
Energy Management Policy

- Energy management policy: $f_t : \mathcal{X}^t \times \mathcal{Y}^{t-1} \to \mathcal{Y}$
- Peak power constraint:

\[
0 \leq X_t - Y_t \leq \bar{P}
\]

- Average power from AES:

\[
P_n = \mathbb{E} \left[ \frac{1}{n} \sum_{t=1}^{n} (X_t - Y_t) \right]
\]

- Privacy: Information leakage rate

\[
I_n \triangleq \frac{1}{n} I(X^n; Y^n) = \frac{1}{n} [H(X^n) - H(X^n | Y^n)]
\]

For given $\bar{P}$, pair $(I, \hat{P})$ is achievable if there exist energy management policies with $\lim_{n \to \infty} I_n \leq I$ and $\lim_{n \to \infty} P_n \leq \hat{P}$.

Privacy-power function, $\mathcal{I}(\bar{P}, \hat{P})$: For given $\bar{P}$ and $\hat{P}$, minimum information leakage rate such that $(I, \hat{P})$ is achievable.
The privacy - power function $\mathcal{I}(\bar{P}, \hat{P})$ for an i.i.d. input load $X$ with distribution $p_X(x)$ is given by

$$\mathcal{I}(\bar{P}, \hat{P}) = \inf_{p_Y \mid X (y \mid x) : E[X - Y] \leq \hat{P}, \quad 0 \leq X - Y \leq \bar{P}} I(X; Y)$$

Lemma

The privacy - power function is a non-increasing convex function of $\hat{P}$ for a given $\bar{P}$.

Optimal energy management policy is memoryless and stochastic: randomly generate output load based on instantaneous input load.

The privacy - power function $\mathcal{I}(\bar{P}, \hat{P})$ for an i.i.d. input load $X$ with distribution $p_X(x)$ is given by

$$
\mathcal{I}(\bar{P}, \hat{P}) = \inf_{p_Y|X(y|x); \mathbb{E}[X-Y] \leq \hat{P}, \ 0 \leq X-Y \leq \bar{P}} I(X; Y)
$$

Lemma

The privacy - power function is a non-increasing convex function of $\hat{P}$ for a given $\bar{P}$.

Optimal energy management policy is memoryless and stochastic: randomly generate output load based on instantaneous input load.

Theorem

The privacy - power function $I(\bar{P}, \hat{P})$ for an i.i.d. input load $X$ with distribution $p_X(x)$ is given by

$$I(\bar{P}, \hat{P}) = \inf_{\substack{p_Y|X(y|x) \cdot E[X-Y] \leq \hat{P}, \\ 0 \leq X-Y \leq \bar{P}}} I(X; Y)$$

Lemma

The privacy - power function is a non-increasing convex function of $\hat{P}$ for a given $\bar{P}$.

Optimal energy management policy is memoryless and stochastic: randomly generate output load based on instantaneous input load

Privacy-power function is a rate-distortion function with difference distortion measure:

\[ d(x, y) = \begin{cases} 
  x - y & \text{if } 0 \leq x - y \leq \bar{P}_z, \\
  \infty & \text{otherwise}. 
\end{cases} \]

- No digital interface: \( Y^n \) is direct output of “encoder”, rather than the reconstruction of the decoder based on the transmitted index
- EMU does not operate over blocks: \( Y_t \) decided instantaneously based on previous input/output loads
- If all future energy demands were known, same privacy could be achieved by deterministic block-based energy management policy
Continuous output alphabet: Infinitely many variables

**Theorem**

*Without loss of optimality output load alphabet \( \mathcal{Y} \) can be constrained to the input load support set, i.e., \( \mathcal{Y} = \mathcal{X} \).*

- Discrete input/output alphabets: Convex optimization problem
- Blahut-Arimoto algorithm
Continuous output alphabet: Infinitely many variables

**Theorem**

*Without loss of optimality output load alphabet $\mathcal{Y}$ can be constrained to the input load support set, i.e., $\mathcal{Y} = \mathcal{X}$.***

- Discrete input/output alphabets: Convex optimization problem
- Blahut-Arimoto algorithm
Continuous output alphabet: Infinitely many variables

**Theorem**

*Without loss of optimality output load alphabet $\mathcal{Y}$ can be constrained to the input load support set, i.e., $\mathcal{Y} = \mathcal{X}$.***

- Discrete input/output alphabets: Convex optimization problem
- Blahut-Arimoto algorithm
Uniform demand over \( \{0, c, 2c \ldots, 20c\} \), such that \( E[X] = 1 \)

- Time division: Either from AES or grid
- Limit max output load: \( Y(t) \leq C \)
Continuous Input Loads

- Continuous input and output alphabets
- No efficient numerical computation method (infinite dimensional optimization problem)
- Shannon Lower Bound (SLB):
  \[
  \mathcal{I}(P) \geq (h(X) - \ln(P))^+ \text{ nats}
  \]
- Not tight in general
- Exponential input load, \(X \sim \exp(\lambda)\): SLB is tight
  \[
  \mathcal{I}(P) = \left(\ln\left(\frac{\lambda}{P}\right)\right)^+ \text{ nats.}
  \]

Total information leakage: $I(X^n; Y^n)$

Total power constraint: $P_n = \mathbb{E}\left[\sum_{i=1}^{N} \frac{1}{n} \sum_{t=1}^{n} (X_i(t) - Y_i(t))\right]$  

Exponential input load with mean $\lambda_i$, i.e., $X_i \sim \text{Exp}(\lambda_i)$

Minimum leakage rate:

$$\mathcal{I}_{E_i}(P_i) = \begin{cases} 
\ln \left( \frac{\lambda_i}{P_i} \right), & \text{if } P_i \leq \lambda_i, \\
0, & \text{otherwise.}
\end{cases}$$

Optimal power allocation: reverse waterfilling

$$P_i^* = \begin{cases} 
\lambda, & \text{if } \lambda < \lambda_i, \\
\lambda_i, & \text{if } \lambda \geq \lambda_i.
\end{cases}$$

where $\sum_{i=1}^{N} P_i^* = P$. 
Total information leakage: $I(X^n; Y^n)$

Total power constraint: $P_n = \mathbb{E} \left[ \sum_{i=1}^{N} \frac{1}{n} \sum_{t=1}^{n} (X_i(t) - Y_i(t)) \right]$

Exponential input load with mean $\lambda_i$, i.e., $X_i \sim \text{Exp}(\lambda_i)$

Minimum leakage rate:

$$I_{E_i}(P_i) = \begin{cases} \ln \left( \frac{\lambda_i}{P_i} \right), & \text{if } P_i \leq \lambda_i, \\ 0, & \text{otherwise}. \end{cases}$$

Optimal power allocation: reverse waterfilling

$$P_i^* = \begin{cases} \lambda, & \text{if } \lambda < \lambda_i, \\ \lambda_i, & \text{if } \lambda \geq \lambda_i, \end{cases}$$

where $\sum_{i=1}^{N} P_i^* = P$. 
Instantaneous Battery Constraints

- Energy is harvested over time
- Harvested energy at time $t$ is a random variable $E_t$
- Stored in a battery of capacity $B$
- Cumulative energy constraints:

\[ \sum_{t=1}^{n} (X_t - Y_t) \leq \sum_{t=1}^{n} E_t, \quad \forall n. \]

**Theorem**

If $B = \infty$, minimum information leakage rate $I_\infty$ for an i.i.d. input load $X$, and an energy harvesting process with an average power $\bar{P}_E$, is given by

\[ I_\infty \triangleq I(\infty, \mathbb{E}[E]) \]

This is a lower bound on information leakage under battery constraints. Equivalent to average-power-constrained case.

Battery-dependent memoryless policy: Use conditional distribution $p_i(y|x)$ if battery state $B_t = B_i$.

No privacy for low battery states: if $i > B_{th}$, where $B_{th} \approx E[X] \cdot \log n$.

Perfect privacy for high battery states: When $B_t > B_{th}$ use $p^*(y|x)$.
$B = 0$. Harvested energy $E_t$ acts as state information, known causally to the EMU. It represents a peak power constraint on energy requested from AES. Energy constraints:

$$0 \leq X_t - Y_t \leq E_t, \quad t = 1, \ldots, n.$$  

**Remark**: Past has no influence

- Assume AES harvests a constant and fixed amount of energy in every time slot, i.e., $E_t = e, \forall t$.

**Corollary**

If $B = 0$ and $E_t = e, \forall t$, the privacy-power function $\mathcal{I}(e)$ for an i.i.d. input load $X$ is given by $\mathcal{I}(e, e)$. 
$B = 0$. Harvested energy $E_t$ acts as state information, known causally to the EMU. It represents a peak power constraint on energy requested from AES. Energy constraints:

$$0 \leq X_t - Y_t \leq E_t, \quad t = 1, \ldots, n.$$  

**Remark**: Past has no influence

Assume AES harvests a constant and fixed amount of energy in every time slot, i.e., $E_t = e$, $\forall t$.

**Corollary**  
If $B = 0$ and $E_t = e$, $\forall t$, the privacy-power function $\mathcal{I}(e)$ for an i.i.d. input load $X$ is given by $\mathcal{I}(e, e)$. 
$B = 0$. Harvested energy $E_t$ acts as state information, known causally to the EMU. It represents a peak power constraint on energy requested from AES. Energy constraints:

$$0 \leq X_t - Y_t \leq E_t, \quad t = 1, \ldots, n.$$  

**Remark**: Past has no influence

- Assume AES harvests a constant and fixed amount of energy in every time slot, i.e., $E_t = e, \forall t$.

**Corollary**  
If $B = 0$ and $E_t = e, \forall t$, the privacy-power function $I(e)$ for an i.i.d. input load $X$ is given by $I(e, e)$. 
AES without a Battery

**AES State Known Only by EMU**

If $B = 0$, and state of AES is i.i.d. with $p_E$, minimum information leakage rate is

$$I_0 \triangleq \inf_{p_{Y|x, E}(y|x, e): 0 \leq x - y \leq E} I(X; Y).$$

**AES State Known by UP**

If $B = 0$, and state of AES is available at UP, minimum information leakage rate $\bar{I}_0$ is

$$\bar{I}_0 \triangleq \inf_{p_{Y|x, E}(y|x, e): 0 \leq x - y \leq E} I(X; Y|E) = \mathbb{E}_E[I(E, E)].$$
AES without a Battery

**AES State Known Only by EMU**

If $B = 0$, and state of AES is i.i.d. with $p_E$, minimum information leakage rate is

$$\mathcal{I}_0 \triangleq \inf_{p_{Y|X,E}(y|x,e):0 \leq X-Y \leq E} I(X;Y).$$

**AES State Known by UP**

If $B = 0$, and state of AES is available at UP, minimum information leakage rate $\bar{I}_0$ is

$$\bar{I}_0 \triangleq \inf_{p_{Y|X,E}(y|x,e):0 \leq X-Y \leq E} I(X;Y|E) = E[\mathcal{I}(E,E)].$$
\( X \) is Bernoulli with \( q_x \)

\( E \) is Bernoulli with \( p_e \).
The greater the battery capacity, the better the privacy level achieved.
Conclusions

- Smart meter privacy is an important and non-trivial problem
- Significantly different from typical data privacy (no data in the strict sense)
- Information theoretic privacy on the energy management level may be the only viable solution to provide theoretical privacy guarantees
- Tools from rate-distortion theory and other information theoretic applications are instrumental
- Concepts may generalize to other cyber physical systems (e-health, electric cars, IoT, etc.) as well as social networks
THANKS FOR YOUR ATTENTION!