Human-Factor-Aware Privacy-Preserving Aggregation in Smart Grid

Asmaa Abdallah
Outlines

• Introduction
• human-factor-aware differential aggregation (HDA) attack
• Privacy-preserving Metering Aggregation Scheme
• Conclusion
Introduction

• Smart grid integrates power and communication network to provide some features:
  • two-way transmission
  • demand response (consumer involved)
  • improves reliability (self-healing, automated outage management)
  • Efficiency (cost-effective power generation, transmission, and distribution)
  • sustainability (accommodation of renewable power sources)
  • Security and privacy.
Introduction

• Smart meters are devices deployed in each home and customer place to aggregate the electricity consumption of these places and send periodic reports to the utility (via local aggregator) to compute the electricity bill.

• the smart grid uses this information to balance the consumption between peak and off-peak periods.

• However, this information exposes the user’s personal behavior to the utility and may be adversaries.
human-factor-aware differential aggregation (HDA) attack

- The attacker can infer the metering information of a user by comparing the aggregated results when the user is in and out of the home.
- The adversary may collude with a set of corrupted meters in the region.
- The corrupted meters reveal their readings to the adversary to help him to know the honest meters’ measurements.
- If the adversary knows the exact aggregated results before the user comes back home (by the help of the malicious meters).
- Then adversary infers the user activities in the house by comparing the new aggregated output with the results when user is out of the house.
Privacy-preserving Metering Aggregation Scheme

• The basic idea of the proposed scheme is that each meter randomly divide its reading to m shares.
• meter i’s readings \( R_i \) are divided into \((r_{i1}, r_{i2}, \ldots, r_{im})\) such that \( R_i = r_{i1} + r_{i2} + \cdots + r_{im} \).
• All the n meters’ reading shares can be represented in the following matrix
Privacy-preserving Metering Aggregation Scheme

- Each row represents the data shares for one meter.
- Ri is represented by the ith row (ri1, ri2, . . . , rim)
Privacy-preserving Metering Aggregation Scheme

• The proposed scheme consists of three phases:
  
  – *System Initialization*
  – *Meter’s Reading Report*
  – *Privacy-Preserving Aggregation*
System Initialization

• Let H be a hash function $H : \{0,1\} \rightarrow \mathbb{Z}_p$
  $\alpha_1, \alpha_2, \ldots, \alpha_m$ are the seeds for H.
  Each meter’s readings are in the range $\{0, 1, \ldots, \Delta\}$.
  Then, the sum of n meter readings is in the range $\{0, 1, \ldots, n\Delta\}$.
  $p > n\Delta$.

• The public parameters in the system are
  $= (H, p, \alpha_1, \ldots, \alpha_m)$.

• The trusted server generates $m \cdot (n+1)$ random numbers
  $(s_01, s_02, \ldots, s_0m; s_11, s_12, \ldots, s_1m; \ldots; s_n1, s_n2, \ldots, s_nm) \subseteq \mathbb{Z}_p$
  such that $\sum_{i=0}^{n} s_{ij} = 0$ for $j = 1, 2, \ldots, m$, $m$ is the number of shares.

• The aggregator obtains $sk_0 = (s_01, s_02, \ldots, s_0m)$, and smart meter $i$ has
  the secret key $ski = (si_1, si_2, \ldots, sim)$. 
In order to conceal the meter readings $R_i$ in the time slot $t$, smart meter $i$ performs the following steps.

**Step 1):**
- meter $i$ divides its readings $R_i$ into $m$ shares
  $$R_i = r_{i1} + r_{i2} + r_{i3} + \cdots + r_{im} \mod p.$$ 
- meter $i$ computes the time-series information for each time slot $t$ by performing the following:
  $$(x_1, x_2, \ldots, x_m) = (H(\alpha_1 \parallel t), H(\alpha_2 \parallel t), \ldots, H(\alpha_m \parallel t))$$
Meter’s Reading Report

Step 2):

• using the private key ski, meter i encrypts each shares into \((y_{i1}, y_{i2}, \ldots, y_{im})\)

\[
\begin{align*}
    y_{i1} &= r_{i1} + r_{i2}x_1 + \cdots + r_{im}x_1^{m-1} + s_{i1}x_1^m \mod p \\
    y_{i2} &= r_{i1} + r_{i2}x_2 + \cdots + r_{im}x_2^{m-1} + s_{i2}x_2^m \mod p \\
    \cdots & \quad \cdots \quad \cdots \\
    y_{im} &= r_{i1} + r_{i2}x_m + \cdots + r_{im}x_m^{m-1} + s_{im}x_m^m \mod p
\end{align*}
\]
Privacy-Preserving Aggregation

Step 1): to aggregate the time slot t’s readings, the aggregator also computes the time-series information \((x_1, x_2, \ldots, x_m)\) as the meters.

Step 2): using the private key \(sk_0\), \((x_1, x_2, \ldots, x_m)\), and the collected data \((y_{i1}, y_{i2}, \ldots, y_{im})\) for all meter \(i (i = 1, 2, \ldots, n)\), the aggregator constructs the linear equations with the \((A_1, A_2, \ldots, A_m)\) unknown variables
Privacy-Preserving Aggregation

\[
\begin{align*}
\sum_{i=1}^{n} y_{i1} + s_{01}x_{1}^m &= A_1 + A_2x_1 + \cdots + A_mx_1^{m-1} \mod p \\
\sum_{i=1}^{n} y_{i2} + s_{02}x_{2}^m &= A_1 + A_2x_2 + \cdots + A_mx_2^{m-1} \mod p \\
\vdots & \quad \vdots \quad \vdots \\
\sum_{i=1}^{n} y_{im} + s_{0m}x_{m}^m &= A_1 + A_2x_m + \cdots + A_mx_m^{m-1} \mod p
\end{align*}
\]

- **Step 3):** the aggregator computes the solutions of the linear equation set and outputs the aggregated readings by adding the solutions together

\[
\text{Result} = A_1 + A_2 + \cdots + A_m.
\]
Conclusion

- The proposed privacy-preserving data aggregation scheme support efficient data aggregation without leaking the individual private information.

- It is suitable for the limited bandwidth smart meters because of its light computation and communication overhead.