Robust Electricity Theft Detection Under Data-Poisoning Attacks in Smart Grids

SETU South East Technological university · Carlow

Surekha Suresh | C00317211 | Supervisor: Milu Philip | MSc in Data Science (SE61H) | Department of Computing, SETU Carlow

INTRODUCTION

Smart grids use Advanced Metering Infrastructure (AMI) to collect high-resolution electricity consumption data. ML models classify usage patterns as benign or malicious to detect electricity theft in near real-time.

Electricity theft causes losses of up to $6 billion annually in the US & Canada. While smart meters reduce physical theft, they introduce new cyber vulnerabilities — particularly data-poisoning attacks where mislabelled samples shift decision boundaries and degrade detection.

"Most detectors implicitly assume correctly labelled training data — a critical weakness exploited by adversarial data-poisoning."

RESEARCH QUESTIONS

RQ1. How does data poisoning quantitatively affect shallow and deep theft detection models?

RQ2. Do deep learning architectures show greater inherent robustness vs. classical ML?

RQ3. Can ensemble-based mitigation significantly reduce robustness degradation?

RQ4. What trade-offs exist between robustness and computational complexity?

Hypotheses: H1: DR degrades | H2: Deep>Shallow | H3: Ensemble Δ

3 LITERATURE REVIEW

- Jokar et al. [1] — ML outperforms rule-based detection in AMI networks
- Jindal et al. [2] — SVM & Decision Trees effective but feature-sensitive
- Ismail et al. [5] — GRU-RNNs capture temporal correlation; improved accuracy
- Takiddin et al. [4] — first systematic poisoning study; DR drop up to 17%
- Zhu et al. [7] — hybrid-order features; adversarial resilience not analysed
- Dietterich [9] — ensemble methods reduce variance, improve generalisation

Research Gap: No study provides structured Δ benchmarking across poisoning levels comparing shallow ML and deep learning simultaneously.

4 THE DATA

Irish Smart Energy Trial Dataset (Sustainable Energy Authority of Ireland)

3,000 Residential units	18 mo Timespan
25,000 Records/customer	30 min Read interval

5 METHODOLOGY

1. Data prep — Irish Smart Energy Trial; ADASYN balancing; zero-mean unit-variance scaling
2. Poisoning sim — label-flipping at 0%, 10%, 20%, 30% controlled penetration
3. Baselines — Random Forest, AdaBoost, SVM (shallow) + GRU, AEA (deep)
4. Ensemble — Averaging vs. Sequential (AEA → GRU → FC Layer)
5. Optimisation — Sequential grid search: layers, neurons, optimisers, dropout
6. Evaluation — DR, FA, Precision, F1, AUC-ROC + Δ = (Clean − Poisoned)

ATTACK PIPELINE

Fig. 1 — Poisoned readings enter AMI and flow into the ensemble detector

SEQUENTIAL ENSEMBLE ARCHITECTURE

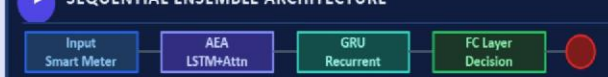

Fig. 2 — AEA extracts anomalies → GRU captures temporal patterns → FC classifies

DETECTION RATE VS. POISONING LEVEL

Fig. 3 — Seq. ensemble Δ 1–3% vs. up to 17% for benchmarks at 30% poisoning

6 KEY FINDINGS [4]

95.2%	Sequential ensemble DR — stable under all poisoning levels
17%	Max DR degradation in benchmarks at 30% poisoning
1–3%	Sequential ensemble Δ vs. 13–17% for all benchmarks
+12%	Deep learning advantage over shallow ML in clean conditions

This research extends [4] by formal Δ robustness metric, structured benchmarking at 0/10/20/30% poisoning, and computational complexity trade-off analysis for AMI.

7 TECHNOLOGIES

Python	TensorFlow	Keras	scikit-learn
NumPy	pandas	Jupyter	GitHub

8 EXPECTED CONTRIBUTIONS

- Reproducible Δ robustness benchmarking framework for smart grid ML
- Systematic evaluation of shallow ML vs. deep learning at 0–30% poisoning
- Ensemble mitigation strategy with quantified robustness improvement
- Computational complexity trade-off analysis for real AMI deployment

R REFERENCES

→ Jokar et al. IEEE Trans. Smart Grid, vol.7, no.1, 2016
→ Jindal et al. IEEE Trans. Ind. Inform., vol.12, no.3, 2016
→ Jagielski et al. IEEE Symp. Security & Privacy, 2018
→ Takiddin et al. IEEE Trans. Smart Grid, vol.12, no.3, 2021
→ Ismail et al. IEEE Trans. Smart Grid, vol.11, no.4, 2020

CONNECT

LinkedIn / QR

in LinkedIn

Ollscoil Teicneolaíochta an Oirdheiscirt
South East Technological University