Electricity Theft Analysis using Benford's Law and Stackelberg Game Theory: Point of View on Trialling on GB Energy Theft

1. Definitions

Benford's Law:

Benford's Law is a statistical principle that predicts the frequency distribution of digits in many real-life datasets. In sets that obey the law, lower numbers appear as the leading digits more frequently than higher numbers. Since it has been shown that this law applies to a wide variety of data sets, including electricity consumption/bills, sudden deviations from the expected distribution might indicate potential issues like meter tampering and theft.

Stackelberg Game Theory

The Stackelberg game is a strategic interaction between two players where one player, called the leader, makes the first move and the other player, the follower, observes this move before making their own decision. This setup creates a sequential decision-making process where the leader's actions influence the follower's decisions. In the context of electricity theft detection, the Stackelberg game is can be used to model the interactions between the utility company (acting as the leader) aiming to detect theft by strategically observing and setting parameters for detection, and the potential thieves (acting as the followers) who respond to these parameters by deciding how much electricity to steal while minimising the risk of being caught.

2. Method Overview

The paper defines a data-driven **Distributed Intelligent Framework for Electricity Theft Detection (DIFETD)** which aims to maximise theft detection probability with minimum false positives The components of this framework are summarised below:

1. Benford's Law
   - Checks if the first digit distribution in a customer's electricity usage dataset conforms to the expected distribution according to Benford's Law.
   - If there are anomalies or irregularities, then then the data is flagged as potential theft and subject to further analysis.

2. Maximum Likelihood Estimator (MLE) and Adversary Model Development:
   - Collect historical data on customer usage.
   - Use MLE method (a statistical model) to develop a model based on historical usage data to understand normal usage patterns
   - Create an adversary model to detect anomalies or potential theft.

3. Formulation of a Stackelberg Game:
   - Create a strategic game where the utility company aims to maximise the probability of detecting theft while keeping false positives low.
   - Thieves engage in a non-cooperative game by stealing various amounts of electricity, trying to balance their rewards against the risk of being caught.

4. Application of the Game and Obtaining Equilibrium:

- Apply the formulated game to reach a Stackelberg equilibrium, a stable point where both the utility company and thieves have optimised their strategies.
- From this equilibrium, determine:
    a. Customer sampling rate (how frequently to check customers for potential theft).
    b. Threshold values for identifying potential thieves.
- Implement the Likelihood Ratio Test (LRT) using the determined sampling rate and threshold values to classify customers as normal or potential thieves.
- Identify compromised meters for further inspection, ensuring quicker resolution, minimal delay, and the discovery of suspicious meters and associated customers.

3. Required Data Sources

The table below is an indicative list of data sources that would be required to implement the DIFETD framework.

Data Item	Description	Possible Source
Minimum Required Data		
Historical Customer Usage Data	The following data items for a minimum of 5,000 domestic MPANs per fuel type (gas/electricity) over a two-year time period, including a minimum within this pot of 10% where theft was found during the two year period: • MPXN number • HH consumption per settlement period/or with timestamp • HH meter readings per settlement period/or with timestamp This data will be used to apply the methodology, with it applied to both the HH consumption and the HH meter readings to test if both follow Benford's Law.	Utility Companies
Theft Instances or Known Anomalies	The complete set of Theft Detection Incentive Scheme (TDIS) files held by RECCo. This data will be used to establish the efficacy of the methodology, as well as enable investigation into irregularities. Note, we recommend the full set of TDIS files are utilised, as these incidents can serve as a basis for identifying irregularities for example: • Historic theft on the meter pre the two-year trial period could indicate a higher risk of undetected theft occurring during the trial period. • For confirmed thefts, the type of theft may impact results e.g., fiscal theft will not impact meter reading values, so may not be identifiable with this methodology.	RECCo
Geographical and Demographic Data	• Geographical data to understand regional variations in electricity usage patterns. • Demographic information that might influence usage patterns, such as household size, type of residence, or economic factors.	Open-source datasets available

Additional Value-Add Data		
Meter and Customer Information	• Information on the meters themselves, their specifications, locations, and associated customer details. This data can help correlate irregularities or theft-related patterns with specific meters or customer profiles. It is proposed that this data, where available, be obtained from the Electricity Enquiry Service and Gas Enquiry Service.	EES and GES
Utility Infrastructure Data	• Data related to the overall electricity distribution network, including the topology, connections, and load distribution across regions. This can provide contextual information about the broader electricity distribution system, helping to identify potential areas or patterns of theft.	Utility Companies

4. Additional Analysis

In obtaining data to trial Benford's Law, it is noted that there is opportunity to also trial other potential methods of identifying energy theft. For example, analysing whether there is a correlation with sudden and sustained drops in consumption.

5. Summary and Recommendations

This is a novel methodology that takes a bottom-up approach to detecting theft (at a customer level). With that comes some risks:

- The approach would rely heavily on historic customer electricity usage data obtained from smart meters. If the data quality collected is incomplete or unreliable, it would impact the effectiveness of the method.
- There is a risk that customers who engage in theft would not use smart meters and employ alternative payment methods. This can create limitations and challenges for the theft detection approach based solely on smart meter data analysis.
- The method would require inputs from data scientists, domain experts and stakeholders to ensure a comprehensive understanding of the problem and the effectiveness of the proposed method, including its drawbacks and limitations.
- There will be data privacy considerations since customer-level data would be required to perform the modelling and statistical analysis. The granularity of data may pose GDPR issues.