ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/320747284

A distributed intelligent framework for electricity theft detection using benford's law and stackelberg game

Conference Paper · September 2017

DOI: 10.1109/RWEEK.2017.8088640

CITATIONS
22

READS
834

5 authors, including:



Longfei Wei
Florida International University
17 PUBLICATIONS **247** CITATIONS
SEE PROFILE



Aditya Sundararajan
Oak Ridge National Laboratory
56 PUBLICATIONS **723** CITATIONS
SEE PROFILE



Arif Sarwat
Florida International University
193 PUBLICATIONS **2,742** CITATIONS
SEE PROFILE



Saroj Biswas
Temple University
110 PUBLICATIONS **731** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Analytics for Smart Grid Cybersecurity View project

Inductive Power Transfer Systems View project

All content following this page was uploaded by Aditya Sundararajan on 14 March 2018.

The user has requested enhancement of the downloaded file.

A Distributed Intelligent Framework for Electricity Theft Detection Using Benford's Law and Stackelberg Game

Longfei Wei, *Student Member, IEEE,* Aditya Sundararajan, *Student Member, IEEE,*
Arif I. Sarwat, *Senior Member, IEEE,* Saroj Biswas, *Member, IEEE,* and Erfan Ibrahim, *Member, IEEE*

***Abstract*—Electricity theft is a major contributor of non-technical losses in the distribution systems of the smart grid. However, owing to the resource-limitations of smart meters and the privacy requirement of electricity usage data, theft detection has become a challenging task for electric utilities. To address this problem, a Distributed Intelligent Framework for Electricity Theft Detection (DIFETD) is proposed and implemented in this paper. It is equipped with Benford's Analysis for initial but powerful diagnostics on smart meter big data. A Stackelberg game-theoretic model is formulated to analyze the strategic interactions between one utility and multiple electricity thieves, which is applied to data flagged suspicious by Benford's Analysis. The Stackelberg equilibrium provides sampling rate and threshold to conduct a Likelihood Ratio Test (LRT) to detect potentially fraudulent meters. The framework is validated on real interval electricity usage data from an electric utility in Florida to filter fraudulent meters in a community.**

***Index Terms*—Benford's Analysis, electricity theft, Likelihood Ratio Test, smart meter, Stackelberg game.**

I. Introduction

The transformation of the traditional power grid into a smart grid capable of advanced computing and communication functions, self-healing, and autonomy is expected to guarantee an improved efficiency, reliability, and security [1]–[3]. Advanced Metering Infrastructure (AMI) is a key component of the smart grid, which brings together technologies deployed at the customer to the utility level and supports smart meters capable of bidirectional communication [4]. However, proliferation of smart meters has rendered the distribution system vulnerable to cyber-attacks [5], [6]. Particularly, electricity theft is a major challenge for utilities, where malicious attackers alter usage measurements collected by smart meters. According to the U.S. Energy Information Administration in 2016, between 1.5 and 2% of electricity in the U.S. is lost due to theft, costing utilities as much as $6 billion annually [7], [8]. Although most traditional theft detection methods consider both technical as well as Non-Technical Losses (NTLs), this paper focuses on NTLs caused by deliberate acts to manipulate electricity usage data by actors called electricity thieves, referred to in this paper as *thieves*.

The main contribution of this paper is the development and implementation of a data-driven Distributed Intelligent Framework for Electricity Theft Detection (DIFETD) for smart meters. The proposed method first uses Benford's Analysis [9] to study the distribution of significant digits in a machine-generated data. The analysis is based on Benford's Law which states that the frequency of occurrence of the first significant digit in a machine generated and natural data is around 30.1%, which is much higher than the expected value of 11.1% [10]. The analysis compares the Probability Density Function (PDF) of the leading significant digits of the structured electricity usage data with the baseline curve to which it must ideally adhere. Any violation therein signifies the occurrence of suspected theft. The baseline curve is the PDF of the leading significant digits of Fibonacci series which strictly conforms to Benford's Law [11].

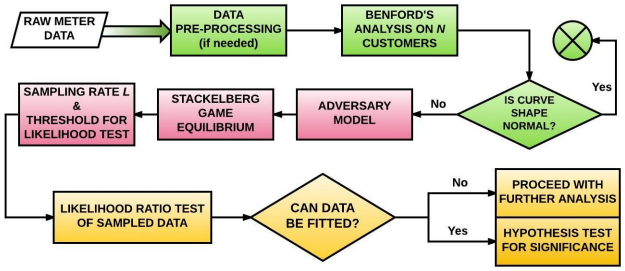


Fig. 1. Schematic of the proposed data-driven Distributed Intelligent Framework for Electricity Theft Detection (DIFETD).

Additionally, the proposed method provides a mathematical approach to maximize theft detection probability with minimum false positives. As shown in Fig. 1, for any AMI data that fails Benfords Analysis, the following steps are executed: **1)** Implementation of Maximum Likelihood Estimator (MLE) for training historical usage data of customers and development of the adversary model of multiple thieves; **2)** Formulation of a Stackelberg game where the utility aims to maximize theft detection probability while limiting false positives and the thieves engage in a non-cooperative strategic game by stealing different amounts of electricity to maximize the trade-off between rewards and probability of being detected; and **3)** Application of the proposed game to obtain a Stackelberg equilibrium from which the customer sampling rate l and a threshold value are derived. Likelihood Ratio Test (LRT) at the sampling rate l and the derived threshold value is implemented for classifying each customer into normal customers or thieves. Therefore, the specific meters being compromised are identified for final checking. It ensures faster resolution, minimal latency, and identification of suspicious meters from which associated customers could be discovered. The proposed method was implemented using real-world AMI interval data collected from an electric utility in Florida, USA, that contained multiple

This work was supported by the National Science Foundation under Grant CNS-1553494.

L. Wei, A. Sundararajan, and A. I. Sarwat are with the Department of Electrical and Computer Engineering, Florida International University, Miami, FL, USA (e-mail: lwei004@fiu.edu, asund005@fiu.edu, and asarwat@fiu.edu). S. Biswas is with the Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA, USA (e-mail: saroj.biswas@temple.edu). E. Ibrahim is the Center Director of Cyber-Physical Systems Security & Resilience R&D at National Renewable Energy Lab (NREL), Golden, CO, USA (e-mail: erfan.ibrahim@nrel.gov).

978-1-5090-6055-9/17/$31.00 ©2017 IEEE

suspected electricity thefts. It is noteworthy that the terms *attack* and *theft* are used interchangeably.

The rest of this paper is organized as follows. A summary of related work is provided in Section II. Benford's Analysis for preliminary theft detection is explained in Section III. A Stackelberg game between utility and thieves is proposed in Section IV. LRT for theft detection is described in Section V. While Section VI discusses data cleansing and the results, Section VII provides conclusion and future work.

II. Related Work

The significance of cybersecurity for smart meters has been a well-researched topic in the literature, where the focus is on ensuring power availability at all times [12]–[15].Traditional research on electricity theft detection has focused on employing specific devices, like wireless sensors and balance meters, to provide a high electricity theft detection accuracy [16], [17]. In [16], an AMI intrusion detection system was proposed to accurately detect electricity theft, where anti-tampering sensors were embedded into smart meters. A set of trusted balanced meters were implemented in the distribution network of smart grid to detect electricity theft [17]. Although these research works reduce the risks due to unmeasured and non-billed usage of electricity, they do not identify specific meters being compromised. Further, these methods significantly increase the cost of deploying and operating millions of smart meters.

Despite cyber-physical vulnerabilities of smart meters, the high-resolution data is a promising tool to complement traditional tools for theft detection [18], and is also considered one of the preliminary tools for auditors [19]–[22]. A Fast NTL Fraud Detection (FNFD) method using Recursive Least Squares (RLS) was proposed by [23] to detect frauds in real-time, and its performance was shown to be better than Intrusion Detection Systems, Binary Coded Grouping-based Inspection (BCGI) [24], NTL Fraud Detector (NFD) and Difference-comparison-based detection [25]. However, it's performance for real data was not accounted for, nor was its convergence speed. Considering not all datasets conform to Benford's Law, a model-based Digits Analysis was proposed using log-Pearson Type IV model [26].

Additionally, statistics and machine learning have been used to train a classifier based on detailed electricity usage measurements [27]–[30]. In [27], average historical electricity usage under the same conditions was used for constructing an electricity theft detector, and an alarm was raised if the average usage was below a predefined detection threshold. Principal Component Analysis (PCA) based anomaly detection was proposed in [28], where anomalies were deviations from the normal usage behavior. In [29], usage data was proved to be non-stationary, and Auto-Regressive Integrated Moving Average (ARIMA) forecasting methods were proposed to validate readings. A Consumption Pattern-Based Energy Theft Detector (CPBETD) that employed a multi-class Support Vector Machine (SVM) for each customer was formulated in [30]. However, these works ignored the attack model of potential thieves, and the effectiveness of anomaly detector was only evaluated based on a dataset of attack examples.

The problem of electricity theft detection was formulated as a game between the utility and the thieves in [31], in which the utility intended to maximize the detection probability and minimize the investment in monitoring fraud, while each electricity thief was to steal a certain amount of electricity and minimize the probability being detected. However, these works assume all electricity thieves as a player, and the competition between thieves was ignored. If thieves add high loads to the distribution networks and steal electricity at the same time, the resulting power surges and electrical system failures can cause power outages, causing thefts to be detected.

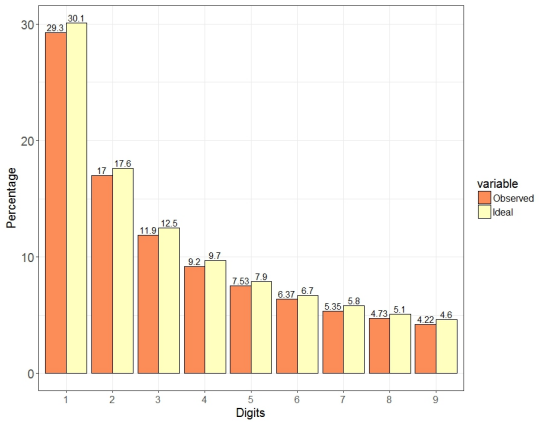


Fig. 2. Ideal vs. observed Benford's distribution for normal meter dataset.

III. Benford's Analysis

The first step in the proposed framework is Benford's Analysis, as depicted by Fig. 1 [32]. Consider an electric utility serving a set of customers, denoted by $\mathcal{N} := \{1, ..., N\}$. Assuming these customers have a similar preference of electricity usage however some customers have the ability to tamper with the smart meter data, they can be separated into two classes, such as normal customers and thieves. The goal of this analysis is to determine whether the electricity usage data belonging to these $\mathcal{N}$ customers is potentially tampered with or not.

It has been empirically proven by statisticians and mathematicians that PDF of the leading significant digits of a data which is either randomly distributed or is a result of mathematical operations on multiple randomly distributed data that is most likely to follow a trend as illustrated in Fig. 2 [33]–[35]. This is counter-intuitive to the presumption that the PDF of such digits should be uniformly distributed with a percentage of occurrence around 11% each. Electricity usage falls under this category because it is derived by the product of voltage and current, both of which are randomly distributed. Thus, by virtue of this property, this data is expected to conform to Benford's Law. A mathematical representation of Benford's Law for first significant digits in decimal (base 10) system can be written as [36]:

$$\mathcal{P}(d) = log_{10}(d+1) - log_{10}(d) = log_{10}(1 + \frac{1}{d}), \qquad (1)$$

where d is the first (leading) significant digit of the data and $\mathcal{P}(\cdot)$ denotes the PDF of the leading digit. PDF of the leading digits of Fibonacci series is used as the baseline since real-world data does not conform with zero error. The frequency of occurrence of the leading significant digits is expressed in percent for better intuition in Fig. 2. Also shown in the same

figure is the PDF of the leading significant digit frequency for preprocessed meter data. It can be seen that this data is close to the ideal trend with a small, acceptable deviation. A smart meter data that is tampered by humans is expected to have a PDF that violates Benford's Law.

IV. Game Model for Strategic Interactions between Utility and Thieves

Let $\mathcal{M} := \{1, ..., M\} \subseteq \mathcal{N}$ be the set of thieves among the N customers of an electric utility. Then, a single leader, multi-follower Stackelberg game can be formulated between the utility and M thieves to characterize and analyze strategic interactions between the two. In this game, a subset from N customers is chosen for an anomaly detection test $\mathcal{D}$, aiming to reduce NTLs. Based on the limited sampling rate l, the utility intends to maximize detection probability and minimize false positives, while the thieves interact with one another using a non-cooperative game to identify optimal quantities of electricity to steal in response to the utility's detection strategy. In the following subsection, the thieves' game is first analyzed before finding the Stackelberg game solution.

A. Thieves' Non-cooperative Game Formulation

A non-cooperative game is formulated in this section to analyze optimal decision making of M thieves in response to any arbitrary utility's action. This game is formulated in normal form, $\Xi := \langle \mathcal{M}, (\mathcal{A}_i)_{i\in\mathcal{M}}, (R_i)_{i\in\mathcal{M}} \rangle$, where $\mathcal{M}$ is the set of M thieves. $\mathcal{A}_i$ is the set of actions available to thief $i \in \mathcal{M}$, where $\boldsymbol{a}_i \in \mathcal{A}_i$ is represented by the expected amount of consumed and stolen electricity by thief i, denoted by q_i and q_i^S, respectively. Additionally, $R_i(\boldsymbol{a}_i)$ is the reward function of thief i under action $\boldsymbol{a}_i$. Thus, thief i selects an action $\boldsymbol{a}_i := \{q_i, q_i^S\}$ that maximizes its reward R_i, which can be defined as below:

$$R_i(q_i, q_i^S) := B(q_i^S) - p_i^D(l, q_i^S)P(q_i^S), \tag{2}$$

where $B(\cdot)$ represents the utility's electricity billing function; $B(q_i^S)$ gives the amount of electricity bill that is not paid by thief i; $p_i^D(l, q_i^S)$ denotes the probability of thief i being detected when the sampling rate is l and the amount of stolen electricity is q_i^S; and $P(q_i^S)$ indicates the penalty function activated upon the successful detection of thief i for stealing a power of q_i^S.

The reward function of each thief reflects the financial benefit obtained by stealing electricity. Thus, for a given sampling rate l and theft detection mechanism $\mathcal{D}$, the goal of thief i is to optimize the following problem (**Problem 1**):

$$\begin{aligned} \max_{\boldsymbol{a}_i \in \mathcal{A}_i} \quad & R_i(\boldsymbol{a}_i, \boldsymbol{a}_{-i}) \\ \text{s.t.} \quad & 0 \le p_i^D(l, q_i^S) \le p_{i\max}^D, \\ & 0 \le q_i^S \le q_i, \\ & 0 \le q_i \le q_{\max}, \\ & 0 \le \sum_{i=1}^{M} q_i \le q_{\mathrm{Tmax}}, \end{aligned} \tag{3}$$

where $\boldsymbol{a}_{-i}$ denotes the actions of all thieves except thief i. In order to control the risk, $p_{i\max}^D$ represents the upper bound for the probability of being detected for thief i. $q_{\max}$ indicates the electricity usage limitation for thief i. Similarly, q_{Tmax} gives the total electricity usage constraint for the community. Following this formulation, one popular solution of the thieves' game is the Generalized Nash Equilibrium (GNE) [37]:

Definition 1: Consider the proposed noncooperative game $\Xi := \langle \mathcal{M}, (\mathcal{A}_i)_{i\in\mathcal{M}}, (R_i)_{i\in\mathcal{M}} \rangle$. GNE is a state of the game in which each electricity thief aims at maximizing Problem 1. As a response to optimal chosen actions of other thieves, a thief aims at choosing the actions, in the restricting subset dictated by the choice of other thieves that maximizes their own reward.

To find a GNE that can be reached by the thieves, the distributed learning algorithm using the framework of learning automata [37] is implemented, in which each thief only knows its own action space and its own reward after choosing an action.

B. Utility's Side Analysis

Under the derived GNE of M thieves, the utility needs to selects a defense action $\mathbf{a}_0$ that maximizes its reward R_0. It is assumed that the utility's defense action is determined by two variables including: (1) Detection mechanism, $\mathcal{D}$, used to identify electricity theft, and (2) Customer sampling rate, l.

The objective of the utility is to maximize the following problem (**Problem 2**):

$$\begin{aligned} \max_{\mathcal{D}, l} \quad & R_0 := \sum_{i\in\mathcal{M}} p_i^D(l, q_i^S)P(q_i^S), \\ \text{s.t.} \quad & 0 \le p_i^E(l, q_i^S) \le p_{i\max}^E, \\ & 0 \le l \le l_{\max}, \end{aligned} \tag{4}$$

where q_i^S is derived from the GNE of the proposed noncooperative game Ξ; $p_i^E(l, q_i^S)$ represents the false alarm probability for thief i; $p_{i\max}^D$ gives the constraint of the false alarm probability for thief i; and $l_{\max}$ indicates sampling rate limitation for the utility. The Stackelberg solution concept is adequate for games with hierarchy in which the leader enforces its strategy and the followers respond, rationally (i.e. optimally), to the leaders strategy. The optimal response of M thieves to an action $\mathbf{a}_0$ by the utility is written as $\mathcal{A}^{\text{theft}}(\mathbf{a}_0) = \{\mathbf{a}_1^*, ..., \mathbf{a}_M^*\}$. This optimal strategy denotes the equilibrium strategy profile of the attackers as a response to the defenders strategy. In this regard, $\mathbf{a}_0^* \in \mathcal{A}_0$ is a Stackelberg equilibrium [38] if it minimizes the utility's reward function R_0. In other words,

$$R_0(\mathbf{a}_0^*, \mathcal{A}^{\text{theft}}) \le R_0(\mathbf{a}_0, \mathcal{A}^{\text{theft}}), \forall \mathbf{a}_0 \in \mathcal{A}_0. \tag{5}$$

In the Stackelberg equilibrium, the optimal customer sampling rate l and the threshold selected for the the detection mechanism $\mathcal{D}$ can be derived; they determine the detection probability p_i^D and false alarm probability p_i^E, $i \in \mathcal{N}$. In the thieves' game, M thieves make their decisions simultaneously at each step of the evolutionary process, playing a GNE between themselves. A multimodal Genetic Algorithm [38] is implemented for computing the Stackelberg equilibrium for the utility.

V. Likelihood Ratio Test

Following the scenario that the smart meter data from a community fails Benford's Analysis, the proposed DIEFTED implements the Likelihood Ratio Test (LRT) for further scrutiny in order to identify potentially fraudulent meters within this community using results from the Stackelberg game. For each customer $i \in \mathcal{N}$, the utility collects a time series of electricity

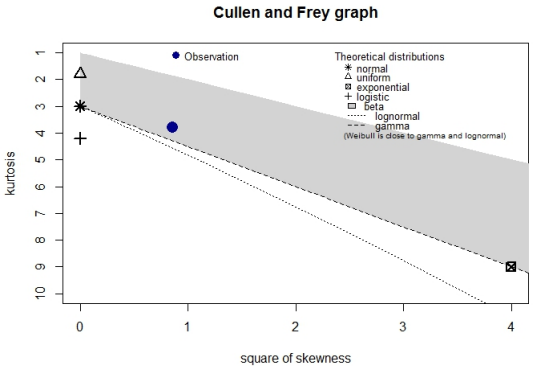


Fig. 3. Skewness vs. Kurtosis for the data of one meter.

usage measurements, denoted by x_t^i, from time 1 to time T. It is assumed that the collected meter measurements x_t^i, $t = 1, .., T$, are independently drawn from identically distributed random variables following the PDF of f_i^0. Then, the expected value of x_t^i for customer i is q_i. In other words, it is the expected amount of electricity consumed by customer i at one time step. However, if customer i is an electricity thief, they may have comprised their smart meter and thus can propagate a falsified time series, $\hat{x}_t^i$, to lower their billable energy. Let the fraudulent meter measurements, $\hat{x}_t^i$, follow the PDF of f_i^1 and thus the expected value of $\hat{x}_t^i$ is $q_i - q_i^S$, which is the expected amount of electricity billed by customer i. Therefore, a binary hypothesis testing problem can be formulated as follows:

$$\begin{aligned} \mathcal{H}_0 &: x_t^i \overset{iid}{\sim} f_i^0,\ \mathbb{E}[x_t^i] = q_i, \\ \mathcal{H}_1 &: x_t^i \overset{iid}{\sim} f_i^1,\ \mathbb{E}[x_t^i] = q_i - q_i^S, \end{aligned} \tag{6}$$

where $t = 1, .., T$. The null hypothesis $\mathcal{H}_0$ indicates that customer i's time series follows the PDF of normal customers, while the alternative hypothesis $\mathcal{H}_1$ states that customer i's time series follows the PDF of thieves.

For customer i, let the time series for a normal customer follow lognormal distribution with scale parameter μ_i^0 and shape parameter δ_i^0. Also let the time series for a thief follow lognormal distribution with scale parameter μ_i^1 and shape parameter δ_i^1. Therefore, the LRT can be expressed as:

$$\begin{aligned} \Lambda &= \ln \frac{\prod_{t=1}^{K} f_i^n(x_t^i)}{\prod_{t=1}^{K} f_i^f(x_t^i)} \\ &= \sum_{t=1}^{T} \ln(|\frac{\delta_i^1}{\delta_i^o}| \times e^{\frac{(\ln(x_t^i)-\mu_i^1)^2}{2{\delta_i^1}^2} - \frac{(\ln(x_t^i)-\mu_i^0)^2}{2{\delta_i^0}^2}}) \gtrless \gamma, \end{aligned} \tag{7}$$

where γ represents the threshold to classify customer i as "Normal" or "Thief". If $\Lambda > \gamma$, customer i is considered Normal, and a Thief otherwise. If $\delta_i^0 = \delta_i^1 = \delta$, then:

$$\Lambda = \sum_{t=1}^{T} \frac{(2 \ln x_t^i - \mu_i^1 - \mu_i^0)(\mu_i^0 - \mu_i^1)}{2\delta^2} \gtrless \gamma. \tag{8}$$

Therefore, LRT can be simplified as $\sum_{t=1}^{T} \ln x_t^i \gtrless \tau$, where $\tau = \frac{2\gamma\delta^2}{\mu_i^0 - \mu_i^1} + \frac{T(\mu_i^0 + \mu_i^1)}{2}$. Under this assumption, successful

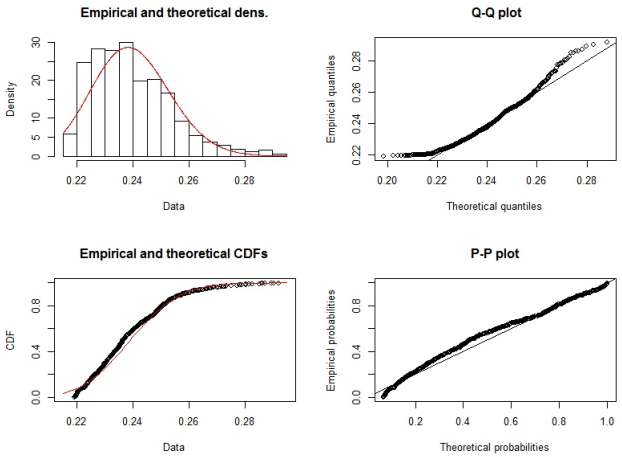


Fig. 4. Fitting meter data to a lognormal distribution using MLE.

detection probability p_i^D and the false alarm probability p_i^E are defined as follows:

$$\begin{aligned} p_i^D &= \Pr\{x_t^i < e^\tau | H_1\} = \int_0^{e^\tau} f_i^1(x)dx, \\ p_i^E &= \Pr\{x_t^i < e^\tau | H_0\} = \int_0^{e^\tau} f_i^0(x)dx, \end{aligned} \tag{9}$$

where $f_i^0(x)$ is the lognormal PDF of the time series consumed by the normal customer with scale parameter μ_i^0 and shape parameter δ_i, and $f_i^1(x)$ is the lognormal PDF of the time series consumed by the thief with scale parameter μ_i^1 and shape parameter δ_i. Based on historical time-series on normal customers, the parameters μ_i^0 and δ_i can be determined by the Maximum Likelihood Estimator (MLE), where $\mathbb{E}[X] = e^{\mu_i^0 + \delta_i^2/2} = q_i$, and $Var[X] = e^{2\mu_i^0 + \delta_i^2}(e^{\delta_i^2} - 1)$. For a thief, $\mathbb{E}[\hat{X}] = e^{\mu_i^1 + \delta_i^2/2} = q_i - q_i^S$, and $Var[\hat{X}] = e^{2\mu_i^1 + \delta_i^2}(e^{\delta_i^2} - 1)$.

The descriptive parameters for one meter's interval usage data is illustrated in Fig. 3 using a Cullen and Frey Graph that plots kurtosis against skewness [39]. While Kurtosis is a measure of how heavy-tailed the distribution of a given data is, Skewness is a measure of its symmetry. For example, the PDF of an ideal normally distributed data has a Skewness of 0 and a Kurtosis of 3. It can be seen from the figure that the PDF of the considered dataset, denoted by a blue colored filled circle, is likely to show a fit for lognormal distribution denoted by the dotted line. A similar Kurtosis-Skewness plot was obtained for data from majority of the other meters. It was further deduced that data from all meters were within the shaded region, implying they were all likely to show a fit for beta distribution. These observations are again evident in the PDF of the data shown in the top-left of Fig. 4.

The Quantile-Quantile (Q-Q) and Probability-Probability (P-P) plots shown in the top and bottom right of the same figure illustrate the agreement between the theoretical and empirical quantiles and probabilities, respectively. The empirical data is drawn from the sample observations while the theoretical data belongs to the known lognormal distribution. While Q-Q plot compares the quantiles, P-P plot compares the Cumulative Distribution Functions (CDFs), which is affirmed by the agreement between the theoretical and empirical CDFs shown in the

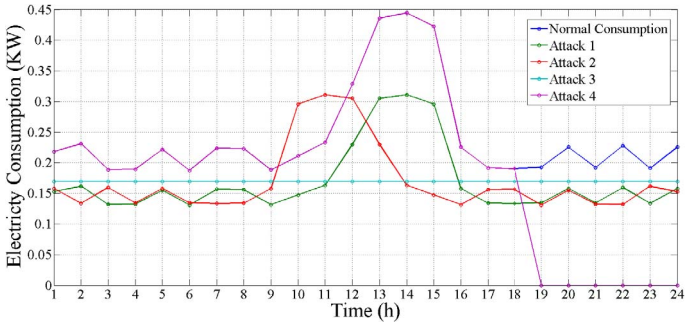


Fig. 5. An example of the normal daily electricity usage and four theft scenarios.

bottom-left of Fig. 4. It can be concluded both mathematically as well statistically that lognormal distribution can be viewed as the most likely fit for the smart meter interval data.

VI. Results and Discussion

In order to implement the proposed DIFETD, real-world data from 103 AMI smart meters was collected from an electric utility in Florida, for the month of May, 2013. The interval data was recorded every one hour, annotated with timestamp. It is to be noted that the sample of 103 meters was chosen to demonstrate the feasibility of the proposed framework alone, and that in the future, the number of meters would be scaled to validate the framework against real-world test cases.

1) Data Preprocessing: Raw data was available as DAT Header files containing values separated by semicolon delimiters (;). Appropriate preprocessing steps were implemented since the data was in a non-numeric format and contained non-intuitive header format. The resulting dataset was structured, time-series and multivariate in nature. In order to maintain privacy of the consumers, original meter IDs were renamed $Meter1$ through $Meter103$.

2) Theft Scenario Construction: This subsection considers the different electricity theft scenarios that will be used to demonstrate the feasibility of the proposed framework. It is to be noted that, while the attacks were synthetically generated, their distributions are randomized to ensure the exact percolation of the attacks into the meter data is still unknown. This ensures a minimization of any inherent bias the proposed framework might have and keeps the overall data structure intact (as would be the case in reality). Consider the electricity bill function as $B(q) = Bq$, and the fine penalty as $B(q) = P$, with $p^D_{i\max} = 0.5$ and $p^E_{i\max} = 0.3$, $\forall i \in \mathcal{N}$, and $l_{\max} = 0.7$. Therefore, the Stackelberg equilibrium can be derived. Based on the derived q^S_i in the game equilibrium, for each meter measurement x_t, $t = 1, .., 744$, in the dataset, four electricity theft scenarios can be formulated:

1) $A^1_t = (q_i - q^S_i)/q_i \times (x_t)$;
2) $A^2_t = (q_i - q^S_i)/q_i \times (x_{744-t})$;
3) $A^3_t = (q_i - q^S_i)/q_i \times \mathbb{E}[x_t]$;
4) $A^4_t = [x_1, ..., x_{t_1}, 0, ..., 0]$, where t_1 is the earliest time satisfies $\sum_1^{t_1} x_t \geq q_i - q^S_i$.

One customer's daily electricity usage and the four suspected scenarios of electricity theft are shown in Fig. 5.

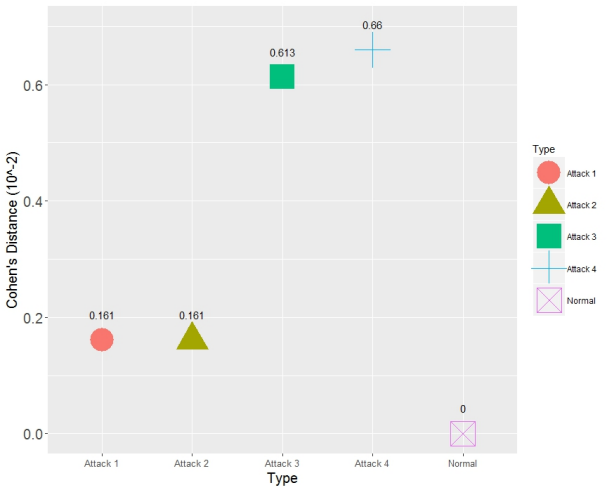


Fig. 6. Cohen's distance showing effect of size measure for normal and tampered data against ideal dataset.

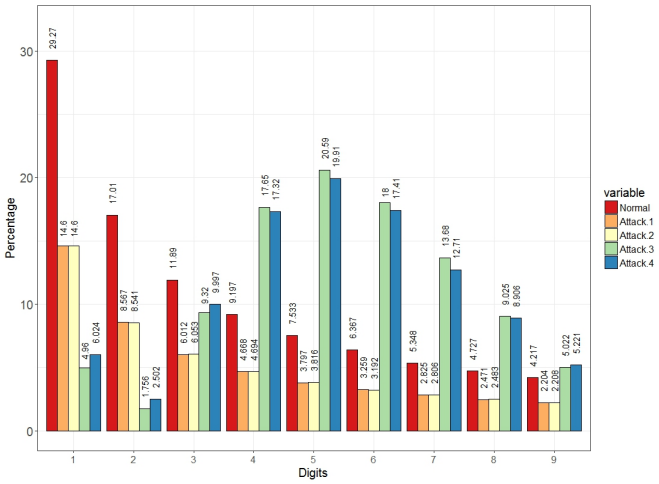


Fig. 7. Benford analysis for normal and tampered data against ideal dataset.

3) Benford's Analysis: To validate the data's adherence to Benford's Law, several Goodness of Fit (GoF) tests have been proposed in the literature. Of them, Pearson's Chi-Square test and Euclidian distance are two of the widely considered methods to test for significance. These tests assume a null hypothesis ($\mathcal{H}_0$) of the data conforming to the Law, and an alternate hypothesis ($\mathcal{H}_1$) representing the conclusion otherwise. As the sample size grows larger, as is the case in this study, the significance level represented by p-value decreases substantially, approaching zero. Furthermore, threshold of significance, denoted by α is arbitrarily set to a value of 0.05 in most studies. However, even a slight difference between the mean of the two groups of data becomes statistically significant for such a large sample size. Hence, it is necessary to justify hypothesis testing with a measure of how big the effect of size is. Cohen's distance is used as an effect of size measure to evaluate the difference between the mean of two groups. Here, group 1 comprises the Fibonacci series that precisely conforms to the Benford's Law curve. It is used as the baseline, while group 2 comprises different types of meter data: normal, and those subject to Attacks 1 through 4 as described in Section VI-1. Illustrated by Fig. 6, it can be seen that the difference in mean

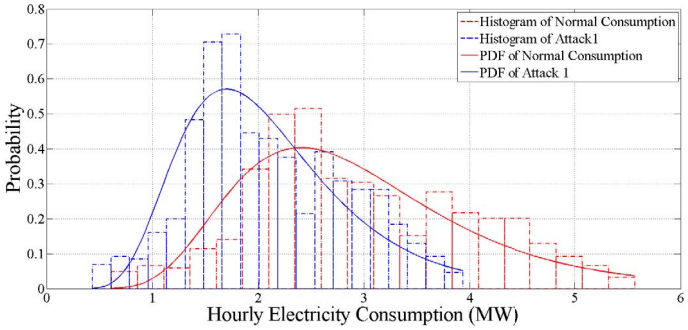


Fig. 8. The LRT between normal daily electricity usage and Attack 1.

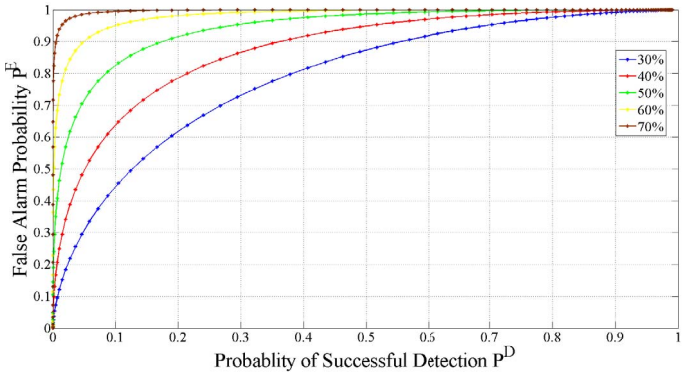


Fig. 9. The ROC Curve (probability of successful detection versus false alarm probability) of the LRT for different levels of stealing by thieves.

between Normal and the baseline is 0, implying that there is no effect of size between the two. However, the Cohen's d-value increases relatively significantly when the data is subject to the Attacks 1 through 4. Greater the d-value, more obvious is the difference between the normal and abnormal datasets. The small d-values implicitly suggest that the four attacks considered are quite intelligent, probably perpetrated by an attacker who is aware of typical electricity usage profiles.

Benford's Analysis for normal data against that compromised by the four attacks is depicted by Fig. 7. It can be seen that the PDFs of the leading significant digits of data manipulated by Attacks 1 through 4 (colored orange, yellow, green and blue, respectively) show significant deviations from the expected PDF represented by the normal data (colored red). This deviation signifies the comparison step shown in Fig. 1. This, hence, illustrates the power of Benford's Analysis in determining whether the data is potentially tampered with in some form. However, this analysis fails to pin-point specific meters that are manipulated. Thus, this analysis provides a crucial preliminary high-level picture for the operators and analysts to study the meter data of $\mathcal{N}$ consumers and decipher anomalies in them for the next stage. If flagged by this analysis, the data is sent to the next stage of the proposed DIFETD.

4) Likelihood Ratio Test: The lognormal PDF of a normal customer and a thief was trained using normal electricity usage and attack patterns, respectively. Figure 8 gives one normal customer's lognormal PDF and corresponding malicious lognormal PDF. Figure 9 shows the Receiver Operating Characteristic (ROC) curve of the LRT for different levels of stealing by thieves, where $q^S/q = 30\%, ..., 70\%$. In this figure, it can be seen that, with increase of the percent of electricity being stole, the LRT test exhibits better performance. For instance, when $q^S/q = 30\%$, the probability of successful detection probability p^D is close to the false alarm probability p^E. However, when $q^S/q = 70\%$, the probability of successful detection probability p^D is close to 1, which is much larger than p^E.

VII. Conclusion and Future Work

Electricity theft is a major concern for utilities all over the world, since it accounts for billions of dollars in losses every year. Considering crucial power system applications such as demand response and state estimation utilize smart meter data, undetected thefts such as FDI could pose serious threats to the reliable operation of smart grid distribution network, raising resiliency concerns. Thus, electricity theft can be considered a very key precursor to the issue of resiliency in smart grid. To address this growing concern, DIFETD is proposed in this paper to detect electricity theft in smart meters using Benford's Analysis for preliminary diagnostics and Stackelberg game for strategic interactions between one utility and multiple thieves. The game equilibrium provides optimal sampling rate and threshold value for LRT. The capability of the proposed framework was validated against four intelligent theft scenarios with real usage data. For each smart meter, the successful detection rate is achieved more than 95% and the false alarm is controlled beyond 10%, if the electricity is stolen in 50%. In the future, more sophisticated types of theft scenarios on a community of more number of smart meters will be used to evaluate the performance of the proposed work. Here, realistic attack data will be used to validate the framework, and the number of meter samples drawn will be further increased to demonstrate the real-world complexity and how the proposed framework handles it. Further, the uncertainty in theft detection will be considered in the adversary model formulation, in order to detect specific fraudulent meters.

References

[1] W. Saad, Z. Han, H. V. Poor, and T. Basar, "Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 86–105, Sep. 2012.

[2] L. Wei, A. Sarwat, W. Saad, and S. Biswas, "Stochastic games for power grid protection against coordinated cyber-physical attacks," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–1, 2017.

[3] I. Parvez, M. Jamei, A. Sundararajan, and A. I. Sarwat, "Rss based loop-free compass routing protocol for data communication in advanced metering infrastructure (ami) of smart grid," in *2014 IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG)*, Dec 2014, pp. 1–6.

[4] I. Parvez, A. I. Sarwat, L. Wei, and A. Sundararajan, "Securing metering infrastructure of smart grid: A machine learning and localization based key management approach," *Energies*, vol. 9, no. 9, 2016.

[5] R. Jiang, R. Lu, Y. Wang, J. Luo, C. Shen, and X. S. Shen, "Energy-theft detection issues for advanced metering infrastructure in smart grid," *Tsinghua Science and Technology*, vol. 19, no. 2, pp. 105–120, April 2014.

[6] L. Wei, A. I. Sarwat, and W. Saad, "Risk assessment of coordinated cyber-physical attacks against power grids: A stochastic game approach," in *2016 IEEE Industry Applications Society Annual Meeting*, 2016, pp. 1–7.

[7] U. S. EIA, "Internatioanl energy outlook 2016," May 2016.

[8] L. Wei, A. H. Moghadasi, A. Sundararajan, and A. Sarwat, "Defending mechanisms for protecting power systems against intelligent attacks," in *Proc. IEEE 10th SoSE Conf.*, San Antonio, the United States, May 2015.

[9] A. Jamain, "Benford's law," *Imperial College of London*, 2001.

[10] S. Newcomb, "Note on the frequency of use of the different digits in natural numbers," *American Journal of Mathematics*, vol. 4, no. 1, pp. 39–40, 1881.

[11] O. Kafri, "Entropy principle in direct derivation of benford's law," *Varicom Communication*, 2009.

[12] A. Sanjab, W. Saad, I. Guvenc, A. Sarwat, and S. Biswas, "Smart grid security: Threats, challenges, and solutions," in *arXiv:1606.06992*, Jun. 2016.
[13] A. Sundararajan, A. Pons, and A. Sarwat, "A generic framework for eeg-based biometric authentication," in *12th International Conference on Information Technology-New Generations (ITNG)*, Las Vegas, Apr. 2015.
[14] A. Anzalchi and A. Sarwat, "A survey on security assessment of metering infrastructure in smart grid systems," in *IEEE Southeast Conference*, Fort Lauderdale, 2015.
[15] I. Parvez, A. Sundararajan, and A. Sarwat, "Frequency band for han and nan communication in smart grid," in *IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG)*, Orlando, Dec. 2014.
[16] S. McLaughlin, B. Holbert, A. Fawaz, R. Berthier, and S. Zonouz, "A multi-sensor energy theft detection framework for advanced metering infrastructures," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1319–1330, July 2013.
[17] E. de Buda, "System for accurately detecting electricity theft," Patent US 20 100 007 336 A1, January, 2010.
[18] M. Jamei, A. Sarwat, S. Iyengar, and F. Kaleem, "Security breach possibility with rss-based localization of smart meters incorporating maximum likelihood estimator," in *International Conference on Systems Engineering*, Las Vegas, 2015.
[19] C. Durtschi, W. Hillison, and C. Pacini, "The effective use of benford's law to assist in detecting fraud in accounting data," *Journal of Forensic Accounting*, vol. V, pp. 17–34, 2004.
[20] R. Joannes-Boyau, T. Bodin, A. Scheffers, M. Sambridge, and S. May, "Using benford's law to investigate natural hazard dataset homogeneity," *Nature Scientific Reports*, 2015.
[21] G. Bella and F. Grigoli, "Power it up: Strengthening the electricity sector to improve efficiency and support economic activity," *IMF Working Paper Technical Report*, 2016.
[22] T. Mir, "Citations to articles citing benfords law: a benford analysis," *arXiv*, Feb. 2016.
[23] W. Han and Y. Xiao, "Fnfd: A fast scheme to detect and verify non-technical loss fraud in smart grid," *Proc. 2016 ACM International on Workshop on Traffic Measurements for Cybersecurity*, pp. 24–34, 2016.
[24] X. Xia, W. Liang, Y. Xiao, and M. Zheng, "Bgci: A fast approach to detect malicious meters in neighborhood area smart grid," in *IEEE International Conference on Communications*, London, UK, Jun. 2015.
[25] X. Xia, W. Liang, Y. Xiao, M. Zheng, and Z. Xiao, "Difference-comparison-based approach for malicious meter inspection in neighborhood area smart grids," in *Proc. 50th International Conference on Communications*, London, UK, Jun. 2015.
[26] C. Winter, M. Schneider, and Y. Yannikos, "Model-based digit analysis for fraud detection overcomes limitations of benford analysis," *2012 Seventh International Conference on Availability, Reliability and Security*, 2012.
[27] E. Telecom., "Fighting electricity theft with advanced metering infrastructure," March 2011.
[28] V. Badrinath Krishna, G. A. Weaver, and W. H. Sanders, *PCA-Based Method for Detecting Integrity Attacks on Advanced Metering Infrastructure*. Cham: Springer International Publishing, 2015, pp. 70–85.
[29] V. Badrinath Krishna, R. K. Iyer, and W. H. Sanders, *ARIMA-Based Modeling and Validation of Consumption Readings in Power Grids*. Cham: Springer International Publishing, 2016, pp. 199–210.
[30] P. Jokar, N. Arianpoo, and V. C. M. Leung, "Electricity theft detection in ami using customers' consumption patterns," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 216–226, Jan 2016.
[31] S. Amin, G. A. Schwartz, A. A. Cardenas, and S. S. Sastry, "Game-theoretic models of electricity theft detection in smart utility networks: Providing new capabilities with advanced metering infrastructure," *IEEE Control Systems*, vol. 35, no. 1, pp. 66–81, Feb 2015.
[32] W. K. T. Cho and B. J. Gaines, "Breaking the (benford) law: Statistical fraud detection in campaign finance," *The American Statistician*, vol. 61, no. 3, pp. 218–223, Aug. 2007.
[33] P. D. Scott and M. Fasli, "Benford's law: An empirical investigation and a novel explanation," *CSM Technical Report 349*, 2001.
[34] S. Miller, *Benford's Law: Theory and Applications*. Princeton Press, 2015.
[35] T. Sugiarto, "Application of first digits benford law: A case study of an indonesian company," *International Journal of Management & Organizational Studies*, vol. 5, no. 2, Jun. 2016.
[36] F. Benford, "The law of anomalous numbers," in *Proc. American Philosophical Society*, vol. 78, no. 4, 1938.
[37] E. Alpaydin, *Introduction to Machine Learning*. MIT Press, Aug. 2012.
[38] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. Philadelphia, PA: SIAM Series in Classics in Applied Mathematics, 1999.
[39] A. Cullen and H. Frey, *The Use of Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. New York, NY: Plenum Press, Plenum Publishing Corporation, 1999.

View publication stats