Evaluating Detectors on Optimal Attack Vectors that Enable Electricity Theft and DER Fraud

Varun Badrinath Krishna, *Student Member, IEEE,* Carl A. Gunter, *Member, IEEE,*
and William H. Sanders, *Fellow, IEEE*

***Abstract*—Worldwide, utilities are losing billions of dollars annually because of electricity theft. The detection of electricity theft has been a topic of research for decades. In this paper, we extend our prior work in the context of advanced metering infrastructures, wherein smart meters are compromised and made to under-report consumption. To the best of our knowledge, this paper presents the first study of meter fraud in the context of distributed energy resources (DERs). With an increased penetration of DERs in modern power grids, and with the decline in electricity prices, we show that there is incentive for electricity generators to over-report generation. We quantify the economic impact of cyber-attacks (on meters) that are optimal in that they maximize fraud while circumventing detectors. In doing so, we use consumption data from Ireland, solar generation data from the U.S. and Australia, and wind generation data from France.**

***Index Terms*—Smart meter, smart grid, anomaly detection, distributed energy resource, attack, machine learning.**

I. Introduction

In modern power grids, electronic meters are used by utilities to measure both consumption and generation from customers for billing purposes. The meters use computer communication networks to communicate with utilities. Customers who have generation capabilities are paid for their net contribution of power to the grid, such that their consumption (if any) is subtracted from their total generation.

In 2010, the Cyber Intelligence Section of the FBI reported that smart meter consumptions were being under-reported in Puerto Rico, leading to annual losses for the utility estimated at $400 million [1]. In 2014, BBC News reported that smart meters in Spain were hacked to cut power bills [2]. While those hacking attempts involved under-reporting of consumption by consumers, customers with generation capabilities can similarly hack their meters to over-report their generation to make monetary gains. In addition to their operating costs, distributed energy resources (DERs) are associated with high capital costs for their owners. Although the cost of installation has been decreasing in recent years [3], it can take a long time to recover those costs through income from generation. In this paper, we show that there is a compelling motivation for DER generation fraud, because it can reduce the time it takes to recover the capital costs of DER installation by over 80%.

Manuscript received October 1, 2017.

The adoption of renewable energy generators has been increasing rapidly in recent years. From 2010 to 2015, photovoltaic adoption in the U.S. grew by 46%, 43%, and 101% for residential, commercial, and utility-scale installations, respectively [4]. Globally, solar capacity increased by 28% and wind capacity increased by 17% from 2014 to 2015 [3]. Most capacity additions in the U.S. came from wind power (41%) in that period [5], and by the end of 2016, wind surpassed hydro as the largest source of renewable energy in the U.S. [6]. The worrying trend is that demand for electricity has not grown with the generation, and as recently as April 2017, it was reported that wholesale electricity prices had dropped so low that at times they were even negative [7]. Since the operating costs of distributed energy resources (DERs) such as solar and wind (about $15/MWh [8]) often exceed the amount that DER owners are paid (ranging from $0 to $45/MWh [7]), there is a real motivation for DERs to compromise their reported generation in order to make fraudulent monetary gains and ensure that they profit.

We restrict our attention to solar and wind, which are the most prevalent DERs. We refer to customers who seek to make monetary gains through consumption or generation fraud as *attackers*. The attackers may be individuals or groups of individuals who own or operate electricity consumption facilities and/or DERs. In the context of consumption, the fraud involves under-reporting of consumption, which is equivalent to theft of electricity, so we refer to that as *electricity theft*. DER fraud involves over-reporting of generation, which also leads to monetary gains, but we refer to such fraud as *DER fraud* and not as electricity theft. We refer to both consumption and generation fraud together as *meter fraud*.

In this paper we extend prior work on the design of approaches to detect electricity theft and present the first study of approaches to detect DER fraud. The detection approaches, referred to as *detectors*, analyze smart meter readings collected at a central location, which is presumably at the utilities' data centers. The detectors then construct an unsupervised model of normal consumption and generation behaviors. The models of normal consumption/generation are unsupervised because there are infinitely many ways in which attackers could compromise their meter readings in order to make monetary gains, so labeling of attacks for supervised modeling cannot be done in a comprehensive manner. Our prior work in [9] presents the first formal framework to classify those attack vectors, which are false meter readings injected for the attackers' benefit. In this paper we present modeling approaches that are unique to DERs, and leverage correlations

between DERs and relevant weather data. The models are used to detect anomalies that are indicative of meter fraud.

We evaluate the detectors in two different ways. First, we derive the optimal attack against each detector, and thereby quantify the detector in terms of how much fraud could be realized while circumventing that detector (the worst-case scenario for the detector). By design, the true-positive rate for the detector, when evaluated against the optimal attack, is zero. Second, we compare detectors by using attacks that are not optimal against them. In doing so, we use receiver operating characteristics (ROCs) to compare true-positive and false-positive rates at different detection thresholds.

The paper is organized as follows. Related work is presented in Section II. The system model and threat model are presented in Section III. Prior work is summarized in Section IV, and the datasets used in this study are described in Section V. Prior work is extended with derivations of optimal attacks for electricity theft in Section VI, and comparative evaluations against suboptimal attacks are presented in Section VII. A framework for detecting DER fraud is presented in Section VIII. For DER fraud, optimal attacks are presented in Section IX, and suboptimal attacks are evaluated using ROCs in Section X. We present a profit analysis for DER fraud in Section XI and conclude in Section XII.

II. Related Work

In this section, we present related work on detection and mitigation of electricity theft. To the best of our knowledge, this paper is the first work on DER fraud, so there is no related work to present on that topic.

Electricity theft detection methods include those based on well-defined attack strategies [9], [10], [11], [12], [13], [14] and general consumption behavior anomalies [15]. In [12], the authors evaluate a few different attack detection algorithms for attacks in which the attackers do not change their consumption behavior, but report lower consumption readings by compromising their own smart meters. In [10], we evaluated a different attack strategy wherein the attacker steals electricity from a neighbor at no loss to the utility. The authors of [13] evaluate support vector machines in the context of attack strategies, such as random and constant scaling of readings, which we first presented in [11]. Those papers failed to capture other possible attack classes, because a comprehensive and fundamental approach to classification was not adopted. In [9], we filled in that gap with a framework that provides a comprehensive classification for better defense.

In [16], [17], and [18], the authors assume that smart meters have not been compromised, and use their readings to detect electricity theft. They do so by calculating the total power lost and estimating how much of the loss was due to electricity theft. Their methods fail under the realistic scenario in which smart meters have been compromised. Motivating factors for attackers who steal electricity are discussed in detail in [17].

In [19], the authors describe how they simulated consumption patterns of loads in households and detected changes in those patterns to report electricity theft achieved by tapping power lines. They reduced false-positive rates by fusing alerts reported by multiple sensors. Their approach is similar to ours in that they try to identify ways in which attacks can take place, and employ learning algorithms to detect attacks. The authors of [15], [19], and [20] all independently claim to have built comprehensive attack trees that span all possible electricity theft attacks. However, their attack trees all depend on existing technologies. In contrast, [9] analyzes the fundamental necessary conditions for the execution of a successful electricity theft attack, which are equally applicable to future, unknown technological approaches for such attacks.

It may be possible to use power grid state estimation methods to validate voltages and currents at different buses for detection of fraud. However, state estimation was shown in [21] to be ineffective for detection, as it can be spoofed. The authors of [22] discuss game-theoretic models of electricity theft detection.

Industry has also invested in mitigating electricity theft. Utilities such as BC Hydro and CenterPoint have implemented tamper detection features on smart meters [23]. Unfortunately, penetration testing on a variety of different smart meters has shown that such features are ineffective [24], and that despite decades of work on tamper detection schemes (the earliest patent on those schemes was awarded in 1980 [25]), better protections against electricity theft are needed. BC Hydro has worked with start-up Awesense to go one step further than tamper detection by placing distribution grid meters (which are different from consumer smart meters) at key nodes on BC Hydro's distribution grid [23]. Although these efforts have been tailored for line-tapping electricity theft, in [9] we showed that this investment in distribution grid meters can also be effective against cyber-intrusion-based theft attacks.

In this paper, we contribute to the state of the art and the state of the practice by evaluating detectors to mitigate cyber-based meter fraud. The attacks we consider are effective despite the presence of the security measures that are currently employed by industry. We are the first to employ and study the Kullback-Leibler (KL) divergence in the context of fraud detection, but this method has been used in other anomaly detection applications in [26] and [27].

III. Preliminaries

In this section, we describe the system model and threat model used in our study.

A. System Model

Both in this paper and in prior work, we assume that meter fraud is occurring in an electrical distribution network (as opposed to the high-voltage transmission network). We assume a radial topology, which can be represented as an unbalanced n-ary tree, where n represents the maximum number of consumers, or leaf nodes, connected to a single node. Another common topology is the loop system, which was designed to improve the reliability of power delivery. It is essentially radial, as the loop is closed only during a fault (see [28]). As a result, power to a consumer at any one time is supplied through a single path from the distribution substation, which we refer to as the *root node* of the n-ary tree. Through a series of

transformers and protective equipment (*internal nodes*), power is supplied from the root node to the leaf nodes, which are the consumers. This root node would typically lie in a substation that connects the transmission (high-voltage) electric grid with the distribution (low-voltage) electric grid.

B. Threat Model

The threat model in this paper (and in our prior work) is as follows. The attacker compromises smart meter readings in order to make fraudulent monetary gains. The attacker can do so by leveraging weak authentication firmware installed in hundreds of millions of smart meters deployed around the world. Physical access is required to compromise the meter, because the attacker gains entry to the firmware by means of an ANSI optical port, which typically uses the C.12.18 communication protocol [29]. Tools, such as Termineter [30], can be used to simplify the process of gaining access to the meter through the optical ports. Termineter leverages the fact that weak passwords are set for the optical interface, and those passwords can be cracked by brute force (trying out a list of commonly used weak passwords). A video demonstration of Termineter is included in [30]. Unfortunately, there is no secure mechanism in place to automatically roll out a security patch to the millions of meters that have this firmware vulnerability. As that proactive approach is infeasible, we propose a reactive approach using data-driven signal processing methods to detect attacks by analyzing the compromised readings.

The use of the optical port on the smart meter is the most common intrusion method, and it has been documented by the National Electric Sector Cybersecurity Organization Resource (NESCOR) in their publication [31], which is widely used by utilities in the U.S. It may be possible to compromise the meters through other approaches, but they may be more difficult, because they may require access to cryptographic keys that allow the decryption and injection of false readings into the communication channels. The Idaho National Laboratory has prepared a comprehensive list of vulnerabilities that an attacker could exploit, and they involve protocols, authentication, authorization, network access control, communication channels, and endpoints [32].

C. Simulation of Attacks

As described in Section I, there are infinitely many attack vectors that can achieve monetary gains for the attacker. In this paper (and in [10]), we simulate attacks that are optimal for the attacker given knowledge of a specific detector. Therefore, our results are conservative in that they represent the worst-case scenario for a given detector. Using this method, we are able to compare detectors based on the worst-case gains for an attacker for each detector. We do not have real data on how attacks are executed in reality, but even if we did, we believe that use of optimal attacks to quantify detectors is a more scientific approach than relying on attack data that may not be representative of worst-case scenarios.

IV. Summary of Prior Work

This paper extends our prior work [9], [10], [11], which discusses different methods for detecting anomalies in consumer meter readings. In this section, we present a summary of that prior work so that this paper is self-contained. Beyond this section, we present original work that extends the prior work.

A. Detection Framework

We presented *F-DETA*, a framework for detecting electricity theft attacks, in [9]. The framework allows utilities to discover meter fraud in the context of different settings, which include tariff schemes used by the customer (e.g., flat, time-of-use, or real-time pricing). An important constraint that was addressed by F-DETA is the *balance check*, which is an approach used in industry to mitigate theft. The check is performed by a redundant meter that has been installed upstream of consumer meters (at an internal node, such as a transformer or a bus). Consider an electric distribution network node I to which $M+1$ consumers are connected. The consumers are the attacker, denoted by A, and a set of M innocent neighbors $N = \{N_1, N_2, ..., N_M\}$. Then, in a discrete time period t,

$$D_I(t) = D_A(t) + \sum_{n \in N} D_n(t) + \sum_{l \in L} D_l(t), \tag{1}$$

where $D(t)$ refers to the true average demand during the time period t. The length of the period, denoted by Δt, is typically fixed at 5, 15, 30, or 60 minutes. L refers to the set of losses/line impedances that are downstream of I. We use D' to denote the reported measurements taken, which may or may not be compromised. Physically, that node may be a bus or a transformer. The utility uses the following balance check at node I at time t.

$$D'_I(t) = D'_A(t) + \sum_{n \in N} D'_n(t) + \sum_{l \in L} D_l(t). \tag{2}$$

Note that the losses and line impedances are static, so the corresponding demand can be easily precomputed and subtracted out from $D'_I(t)$ to get the sum of downstream meter readings $D'_A(t) + \sum_{n \in N} D'_n(t)$. The balance check assumes that the internal meter at node I is trusted, so $D'_I(t) = D_I(t)$. Then, from (1) and (2),

$$D_A(t) - D'_A(t) = \sum_{n \in N} [D'_n(t) - D_n(t)]. \tag{3}$$

One of the key propositions in F-DETA is that the attackers can circumvent the balance check given in (3) by under-reporting consumption (to lower their energy bills) as follows. The attackers compromise both their own smart meters and the smart meters of one or more neighbors in N. Then the attackers under-report demand ($D'_A(t) < D_A(t)$) while simultaneously over-reporting the demand of the neighbor(s) whose meter(s) they have compromised ($D'_n(t) > D_n(t)$ for some $n \in N$) so that (3) is satisfied. Since (3) is satisfied, the attack will go undetected by utilities who use only the balance check for detection. The attack assumes that the attackers can compromise their neighbors' smart meters, but not the meter at the internal node I. This assumption is easily justified given that most residential and commercial building meters are installed outside the buildings and are physically accessible by any passerby. Even if they are encased, the cases are usually transparent and accessible by optical probes, as discussed in

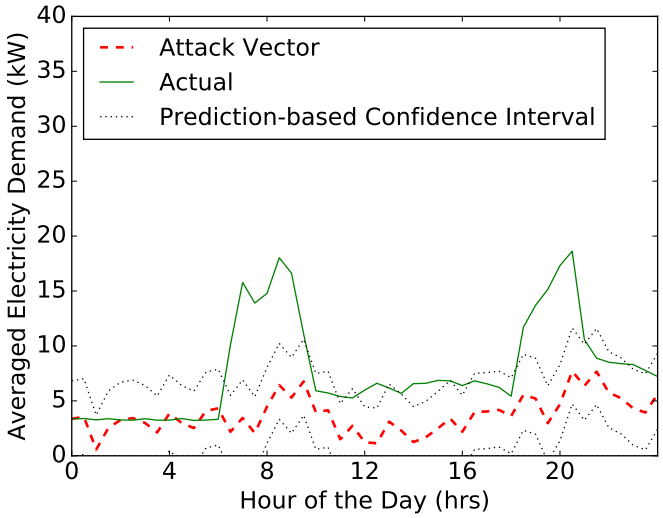


Fig. 1. Integrated ARIMA attack on a consumer. The attacker under-reports consumption by generating false readings per a truncated normal distribution whose maximum, minimum, average, and standard deviation are set in a manner that ensures maximum monetary gain while avoiding detection. The amount of electricity stolen is the difference between the actual consumption and the attack. This illustration was taken from [9].

Section III-B. Although it may be possible for an attacker to gain access to the meter at I, it is likely better protected and shielded, because pole-top meters need to be encased in weather-resilient cases. F-DETA is not dependent on the integrity of the meter at I; it simply shows that the balance check is an insufficient measure against theft detection.

Automated demand response to pricing signals is also included in F-DETA. In that scenario, the attacker can influence demand by compromising price signals. In considering automated demand response, pricing signals, and the presence of the balance check, F-DETA identifies seven attack classes, which compromise different types of smart meter measurements in different ways. The only attacks that we will consider in this paper are the ones in which the attackers do not change their actual behavior patterns, but report values that are lower than the typical consumption. Those attacks belong to classes that are labeled 2A and 2B in [9]. That labeling system will not be used in this paper, and we refer the interested reader to [9] for more details on the attack classes.

B. Attacks and Detectors in Prior Work

In the rest of this section, we describe the relevant attacks and detectors that we developed in prior work, and which we will further investigate in this paper.

1) Integrated ARIMA Attack: In [10], we presented a specific manner in which the meter readings can be manipulated, called the *integrated ARIMA attack*. Later in this paper, we will use that attack to evaluate our detectors. Now, we will explain how that attack was developed.

In [10], we used a prediction model to create a confidence interval for smart meter readings. The model used recent past readings in time to predict the range within which the next smart meter reading should lie. As such, it was suitable for real-time anomaly detection of smart meter measurements. We used the Auto-Regressive Integrated Moving Average (ARIMA) model to construct the confidence interval for the readings. First-order differencing was applied to remove the effects of seasonality in the data. The ARIMA model order was determined for each consumer in the dataset separately using the automated Hyndman-Khandakar algorithm [33]. Percentile points on the confidence interval were chosen to act as thresholds for anomaly detection. While attackers would be expected to set meter readings to abnormally low values, they would simultaneously set the readings for their neighbors to correspondingly high values (as discussed in Section IV-A). Therefore, a two-tailed test for anomalies was performed using the percentile points of the detector, which we called the *ARIMA detector*.

In order to circumvent the ARIMA detector, an attacker who has obtained knowledge about the detector could set the smart meter readings at the detection threshold. Since the compromised readings would not cross the detection threshold, the compromise would not be detected. The key point to be noted is that this attack maximizes the amount of electricity that can be stolen while avoiding detection. In that sense, the attack, which we called the *ARIMA attack*, is optimal for the attacker and the worst-case attack for the detector. By adding checks on the average and standard deviation in addition to the checks on the maximum and minimum values of the individual readings from the ARIMA detector, we obtained what we called the *integrated ARIMA detector*. That detector was effective in detecting and mitigating the ARIMA attack.

The optimal attack against the integrated ARIMA detector, which maximizes the amount of electricity that can be stolen while avoiding detection, is what we refer to as the *integrated ARIMA attack*, and it works as follows. The attacker generates an attack vector (a set of compromised meter readings) from a truncated normal distribution. This distribution has four parameters: the maximum value, the minimum value, the mean, and the standard deviation. The parameters are independent (unlike the uniform distribution, for example, for which the mean and variance are completely determined by the minimum and maximum values), and they can be set such that readings generated from the distribution fall just inside the integrated ARIMA detection boundary, thereby avoiding detection while maximizing theft. The integrated ARIMA attack is illustrated in Fig 1. The prediction-based confidence interval (ARIMA detector) is not effective because the confidence interval is based on the *reported* measurements as opposed to the actual measurements. The reason is that the ARIMA detector is run at the utility's server, which sees only the reported measurements.

2) KLD Detector: We proposed the *KLD detector* in [9] as a means to detect the integrated ARIMA attack. The detector is based on the Kullback-Leibler divergence (KLD), and it examines a full week of meter readings in the test set to determine whether the set was anomalous with respect to the expected patterns of sets of readings taken from weeks in the training set. That is different from the approach of the ARIMA detector, which examined one reading at a time. The integrated ARIMA detector, however, considered a mixture of individual readings (to produce the ARIMA max/min bounds) and sets of readings (to compute the mean and standard deviation).

The KLD detector was implemented *for each consumer, independently* as follows. For each consumer, we use D'_{Tr} to denote the training set. D'_{Tr} contains M weeks of consumption readings, and each week is denoted by D'_w, where $w \in \{1, 2, ..., M\}$. A histogram was computed on the readings in D'_{Tr} using $|B|$ bins, where B denotes the set of bins and $|.|$ denotes the cardinality operator. For a vector X indexed by time t, let $P_X(b)$ denote the probability $Prob(X(t) \in b)$ for $b \in B$. For each week, w in the training set, we calculate the KLD of D'_w from D'_{Tr}, denoted by $KLD(D'_w, D'_{Tr})$, as follows.

$$KLD(D'_w, D'_{Tr}) = \sum_{b \in B} P_{D'_w}(b) \log_2 \frac{P_{D'_w}(b)}{P_{D'_{Tr}}(b)}, \qquad (4)$$

where $w \in \{1, 2, ..., M\}$, giving us M KLD values. Those KLD values capture the *normal* or *non-malicious* deviations between the distribution of a week of consumption readings D'_w from the full training set D'_{Tr}. We set a threshold τ at a specified percentile point on the list of M KLD values. For example, it may be set on the 90^{th} percentile point.

A test vector of consumption readings D'_{Ts} is deemed to be an attack if it deviates too much from the training distribution D'_{Tr} or, more specifically, if $KLD(D'_{Ts}, D'_{Tr}) > \tau$.

3) PCA-DBSCAN Detector: We were the first to propose the PCA-DBSCAN detector in [11] to detect anomalies in consumption readings, but not solely for the purpose of electricity theft detection. The detector uses principal component analysis (PCA) as a dimensionality reduction method, leveraging the repetitive nature of electricity consumption across multiple weeks. Each week is a vector of readings in a high-dimensional space. For example, if the readings are taken every half-hour, then one week would contain $N = 336$ readings for each half-hour of the week.

Like the KLD detector, the PCA-DBSCAN detector examines anomalous sets of readings, each of which is collected over a full week. The detection procedure is similar to that of the KLD detector. For each consumer, a training matrix X was constructed with M rows and N columns, where M refers to the number of weeks in the training set, and N refers to time-aligned columns (for example, all Mondays are aligned across all weeks). This $M \times N$ matrix X was projected down to an $M \times R$ matrix Y by using PCA, which ensures that maximum variance is retained by the dimensionality reduction. Each week of N measurements in X could then be described by a point in the R-dimensional space of Y, where $R << N$.

After we reduced the dimensionality of X to obtain Y, the points that corresponded to weeks of consumption in Y were clustered using a density-based clustering approach called DBSCAN [34]. That approach was suitable for anomaly detection because it formed a single cluster around points that were densely grouped in the R-dimensional space, while points that were far from the cluster were marked as anomalies. The size of the cluster was determined by the DBSCAN parameter, ϵ. In [11], we chose $R = 2$ and $\epsilon = kS_n$, where S_n is the name of a statistical measure of inter-vector distances that is robust to outliers [35]. In this paper, we evaluate this detector on the integrated ARIMA attack for different values of R and k. In summary, the PCA-DBSCAN approach marked weeks of consumption readings as anomalous if they were located far from most other weeks in the R-dimensional space.

V. Datasets

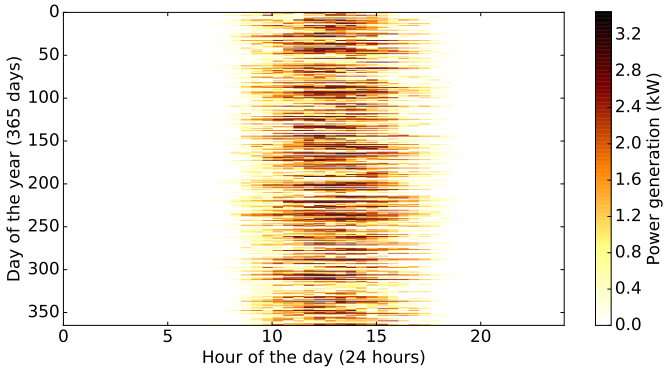
We use four datasets to evaluate our meter fraud detectors. *The fact that the datasets are all freely available makes it possible for researchers to replicate and extend our results*. We assume that the datasets have not been compromised by an attacker, and use the data to model normal behavior from which attack behavior can be distinguished. Note that the datasets may have anomalous behavior that can lead to false positives. The data come from meters installed at consumers and generators in Australia, France, Ireland and the U.S.

1) CER Consumption Dataset: This is a freely available dataset of smart meter readings collected by Ireland's Commission for Energy Regulation (CER). The dataset was collected at a half-hour time resolution, over a period of up to 74 weeks. We extracted from the dataset a set of 500 consumers, all of whom had continuous measurements reported over the 74 weeks. The consumers include 404 residential consumers, 36 small and medium enterprises (SMEs), and 60 unclassified by CER. Only this dataset was used in our prior work. We split the dataset into 60 weeks for training and 14 weeks for testing. As there are 52 weeks in a year, having 52 weeks would be the minimum training set size to ensure that all seasons of the year are covered in the training set. In conjunction with this dataset, we obtained time-of-use (TOU) pricing rates from the Nightsaver plan offered by Electricity Ireland [36].

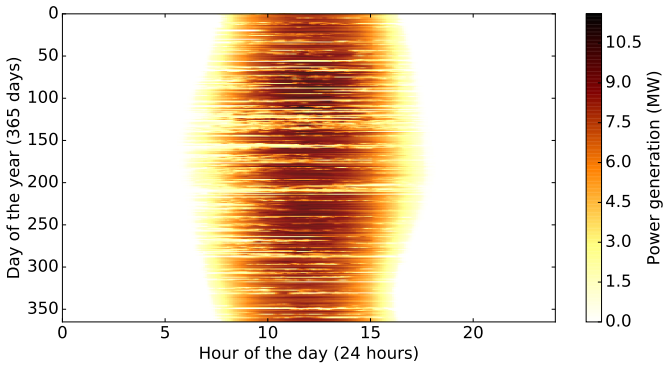
2) Ausgrid Solar Dataset: This is an openly available dataset of electricity consumption and generation measurements taken from a real deployment of 300 customers in the Sydney area [37] who have rooftop solar panels on their homes. Readings were taken at a half-hour time granularity for one year. In *net metering*, the primary purpose of the solar panel is to meet the customer's own consumption needs. If consumption exceeds the solar generation, the deficit power is supplied by the utility grid at the prevailing retail price. If generation exceeds consumption, the excess generation is sold back to the utility, because the customers do not have storage on their premises. The net generation is illustrated in Fig. 2(a).

3) NREL Solar Dataset: This dataset was created by the National Renewable Energy Laboratory (NREL) to be representative of solar output characteristics across the U.S. [38]. We examine the metered generation from 238 distributed photovoltaics in California in the dataset. Those photovoltaics have ratings ranging from 4 MW to 121 MW, and the data for one photovoltaic are plotted in Fig. 2(b). The dataset was produced at a 5-minute granularity for a period of one year.

4) Engie Wind Dataset: This is an openly available dataset of wind power generation from four 2 MW turbines in Meuse, France. The data was provided by Engie, a French utility company. Readings were taken at a 10-minute granularity, and we extracted a period of one year in which all four turbines were continuously operational. One sample turbine group rated at 16 MW is illustrated in Fig. 3.



(a) Ausgrid: Net solar generation



(b) NREL: Gross solar generation

Fig. 2. Solar generation datasets: Heatmap illustrations of daily repeating patterns for (a) one photovoltaic in the Ausgrid dataset (rated at 9 kW) and (b) one photovoltaic in the NREL dataset (rated at 13 MW).

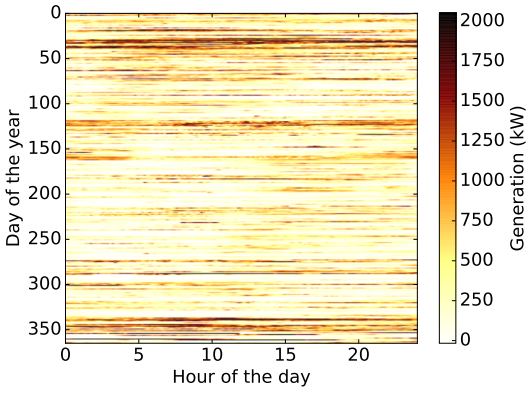


Fig. 3. Engie wind dataset: Sample utility-scale turbine rated at 2 MW.

VI. Optimal Attack Vectors for Electricity Theft

An optimal attack would maximize the amount of electricity stolen while going undetected. In [10], we obtained the ARIMA attack and the integrated ARIMA attack as the optimal attacks against the ARIMA detector and the integrated ARIMA detector, respectively. In this section, we derive the optimal attacks against the KLD detector and the PCA-DBSCAN detector.

A. Optimal Attack Against the KLD Detector

In order to construct the optimal attack against the KLD detector, the attacker (denoted by A) would need to have access to three pieces of information: 1) the set of bins being used, B; 2) the distribution of the training data, D'_{Tr}, over those bins, and 3) the percentile threshold calculated on the training data, τ. The optimal attack vector for a week of readings, D'^{*}_{A}, maximizes the attacker's profit as follows.

$$D'^{*}_{A} = \arg\max_{D'_A} \sum_{t=1}^{T} \lambda(t)\Delta t[D_A(t) - D'_A(t)] \tag{5}$$

$$\text{subject to } \text{KLD}(D'_A, D'_{Tr}) \leq \tau, \tag{6}$$

where $\lambda(t)$ is the electricity price at time t and Δt is the time period between the collections of readings. The objective function is the profit given in dollars ($), Δt is in hours (h), $D_A(t)$ is in kilowatts (kW), and $\lambda(t)$ is in $/kWh. The space of readings is very large (but countably finite because readings are rounded off to the nearest watt and bounded below by zero). Therefore the search space is exponentially large in T. However, we show that the problem can be reformulated as a convex optimization problem and efficiently solved using free solvers, like SCS [39], which comes packaged with CVXPY [40] for Python. The trick is as follows. Instead of solving for D'^{*}_{A}, we solve for the optimal distribution of D'^{*}_{A} such that the KLD from the training data remains within the threshold τ. Once we have the optimal distribution, we can generate D'^{*}_{A} as per that optimal distribution.

We can restate the objective function in Equation (6) as follows by removing all the constants.

$$D'^{*}_{A} = \arg\max_{D'_A} \sum_{t=1}^{T} \lambda(t)[D_A(t) - D'_A(t)] \tag{7}$$

$$= \arg\min_{D'_A} \sum_{t=1}^{T} \lambda(t) D'_A(t) \tag{8}$$

$$= \arg\min_{D'_A} \frac{1}{T} \sum_{t=1}^{T} D'_A(t), \tag{9}$$

where the last equality assumes that $\lambda(t)$ is constant for all t (we will later relax that assumption). Therefore, by minimizing the mean of the reported readings, we can maximize the profits for the attacker. Let the probability distribution of $D'_{Tr}(t)$ be discretized into $|B|$ bins with bin centers $X_{Tr}(b)$ for $b \in B$. Let $P_A(b)$ denote the probability $Prob(D'_A(t) \in b)$ for $b \in B$. Then the mean given in Equation (9) can be expressed in terms of the expectation of the probability distribution as follows: $\frac{1}{T}\sum_{t=1}^{T} D'_A(t) = \sum_{b\in B} X_{Tr}(b)P_A(b)$. With that formulation, we can use $P_A(b)$ as the optimization variable, and that is convenient, because the KLD constraint can be expressed directly in terms of $P_A(b)$ and the corresponding training probabilities $P_{Tr}(b)$. The equivalent convex optimization formulation of the optimal attack against the KLD detector

is given as follows.

$$P_A^* = \arg\min_{P_A} \sum_{b\in B} X_{Tr}(b) P_A(b) \tag{10}$$

$$\text{subject to } \sum_{b\in B} P_A(b) \log \frac{P_A(b)}{P_{Tr}(b)} \leq \tau, \tag{11}$$

$$P_A(b) \geq 0, \tag{12}$$

$$\text{and } \sum_{b\in B} P_A(b) = 1. \tag{13}$$

The first constraint ensures that the KLD value is less than the threshold τ, and the other two constraints ensure that the probability values are valid. X_{Tr}, P_{Tr}, and τ are constants. The objective function is a linear sum, and the probability validity constraints are also linear. The KLD can be expressed as the sum of the negative entropy function of $P_A(b)$ (which is convex) and a linear function of $P_A(b)$, because $\log P_{Tr}(b)$ is a constant with respect to the optimization parameter.

$$\sum_{b\in B} P_A(b) \log \frac{P_A(b)}{P_{Tr}(b)} = \tag{14}$$

$$\sum_{b\in B} P_A(b) \log P_A(b) - \sum_{b\in B} P_A(b) \log P_{Tr}(b) \tag{15}$$

Since the KLD is a sum of convex and linear functions, it is convex. Therefore, the optimization problem is a convex optimization problem that solves for the $P_A(b)$ values from which the attack vector can be generated. An example of an optimal attack for one particular consumer in the CER dataset is illustrated in Fig. 4 with $|B| = 10$ and τ set at the 90^{th} percentile of the KLD values in the training set. Notice that the attacker is trying to under-report consumption, as evidenced by the fact that the bin corresponding to the lowest consumption readings has a probability associated with it in the attack distribution that is higher than in the training distribution. Also, the bins corresponding to larger consumption readings have lower probabilities associated with them. The smallest consumption value in each bin is, by design, the value at the left edge of the bin; $D_A'^*$ can be generated by making copies of the smallest value in each bin in a manner that adheres to the optimal distribution.

The time ordering of readings $D_A'^*$ generated from the optimal distribution P_A^* can be chosen to maximize the monetary gains from fraud. Earlier, we had assumed that the electricity price $\lambda(t)$ was constant for all t, and in that scenario the time ordering of readings did not matter. But if $\lambda(t)$ were time-varying, then $D_A'^*$ could be ordered such that the smallest values in $D_A'^*$ would correspond to the largest values of $\lambda(t)$, and vice versa, so that $\sum_{t=1}^{T} \lambda(t) D_A'^*(t)$ is minimized. That is illustrated in Fig. 5, in which TOU pricing corresponding to the CER dataset was applied, and larger values in the optimal attack vector were injected during off-peak periods, when the price was low. As a result, the attacker was charged less for larger consumption values.

B. Optimal Attack Against the PCA-DBSCAN Detector

In principle, the PCA-DBSCAN detector is very similar to the KLD detector. M weeks in the training set, each

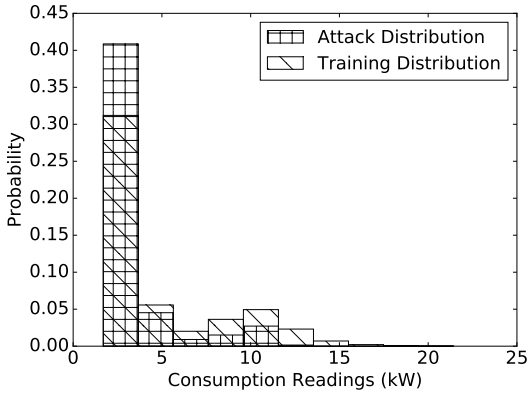


Fig. 4. Distribution of the optimal attack against the KLD detector in comparison to the training distribution.

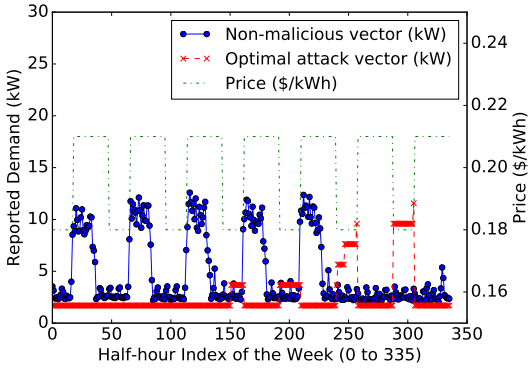


Fig. 5. Illustration of optimal attack against KLD with TOU pricing.

containing N consumption readings, are transformed into an R-dimensional space, where $R << N$. In the CER dataset, $N = 336$. Let V be an $N \times R$ matrix that performs the PCA transformation $Y = XV$ of the training set X. Note that X is $MC \times N$ and contains the training data for all C consumers, and V is calculated on that full dataset. For the sake of obtaining the detection boundary, the training data for each consumer, denoted by D'_{Tr}, are independently projected into the R-dimensional space by using V. V, however, is calculated on X, which combines the data for all consumers.

As per the approach in [11], the DBSCAN algorithm finds a subset of the M training weeks that are not anomalous in the vector space of Y. First, it determines *core weeks*, which are points in Y for each consumer that contain $M/2$ neighbors in Y that are within an ϵ radius (as measured by the Euclidean $L2$ norm). Any points in Y that lie outside of the ϵ radius of all core weeks are deemed anomalous, and indicative of an attack. An example for a consumer in the CER dataset is illustrated in Fig. 6, in which the core weeks are projected from a 336-dimensional space to a 2-dimensional space. The *safe* region is shaded in yellow and contains overlapping circles centered at core weeks with radius ϵ. All points outside that region are marked as attacks. For example, the *zero attack* vector, which sets all consumption readings to zero, is marked as an attack because its projection lies far outside the safe region.

Let η denote the set of core weeks, a subset of the M

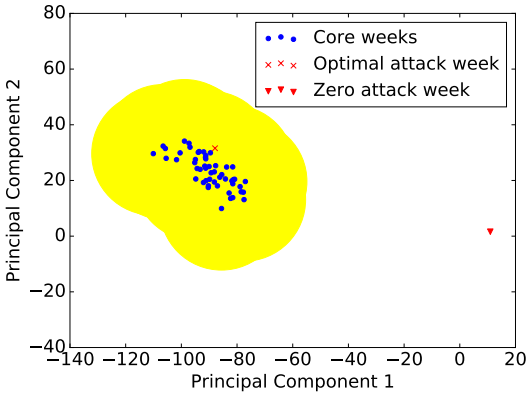


Fig. 6. Core weeks and attacks projected using PCA, onto a two-dimensional space. Points outside the yellow region, which was formed by overlapping circles centered on core weeks, are marked as attacks. The optimal attack circumvents detection and lies within the detection boundary.

training weeks for each consumer, and it is determined by DBSCAN. In the example illustrated in Fig. 6, there are $M = 60$ weeks in the training set, of which $|\eta| = 52$ are core weeks. The remaining 8 weeks are possibly anomalous, and are not used to model normal consumption behavior. The optimal attack vector contains $M = 336$ readings in one week and must project into the safe region. The objective is taken from Equation (9), and the problem is formulated as follows.

$$D'^{*}_{A} = \arg\min_{D'_A} \sum_{t=1}^{N} \lambda(t) D'_A(t) \tag{16}$$

$$\text{subject to } D'_A \geq 0, \tag{17}$$

$$D'_A \leq max(D'_{Tr}), \tag{18}$$

$$||D'_A V - D'_n V|| \leq \epsilon, \forall n \in \eta, \tag{19}$$

where $\lambda(t)$ would be known beforehand in flat-pricing or time-of-use schemes. $D'_A \leq max(D'_{Tr})$ is an upper-bound constraint imposed based on the maximum of the historic data in the training set. The distance between the projected attack, $D'_A V$, and the projected core week, $D'_n V$, is measured by the $L2$ norm, and that constraint is repeated $|\eta|$ times over all the different core weeks. Since the $L2$ norm is a convex function and the objective function is linear, the optimization problem is convex and can be solved efficiently using solvers like SCS [41]. An example of the optimal attack in comparison to core weeks is illustrated in the lower-dimensional space in Fig. 6 and in the higher-dimensional space in Fig. 7. In Fig. 7, notice that the optimal attack has mostly zeroed values, but the few non-zero values are large and ensure that the projection lies in the safe region. In addition, the nonzero values in the optimal attack coincide with peaks in the consumption readings of the core weeks. That ensures that the optimal attack vector preserves the trend in the training data. In preserving that trend, the optimal attack vector does not adversely affect the low-rank approximation and lie far from the core weeks in the low-dimensional space.

Note that the optimal attack vector against the PCA-DBSCAN detector incorporates the timing of reported meter readings with the TOU electricity prices. It is different from

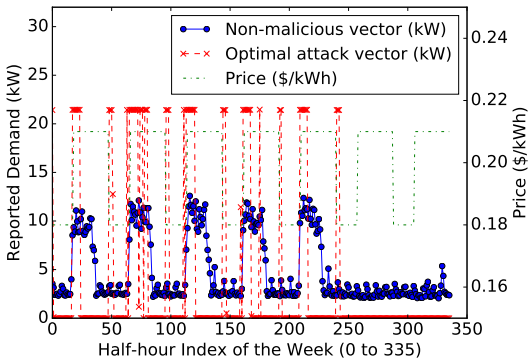


Fig. 7. Illustration of optimal attack against PCA in the original dimension.

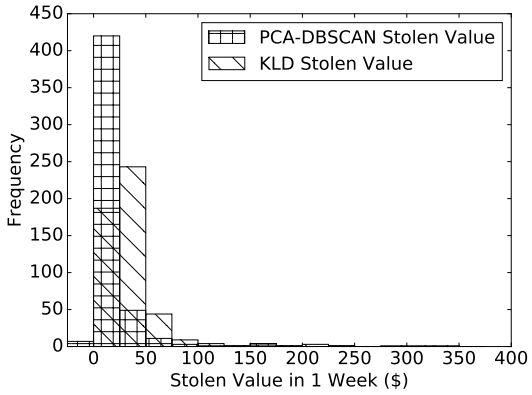
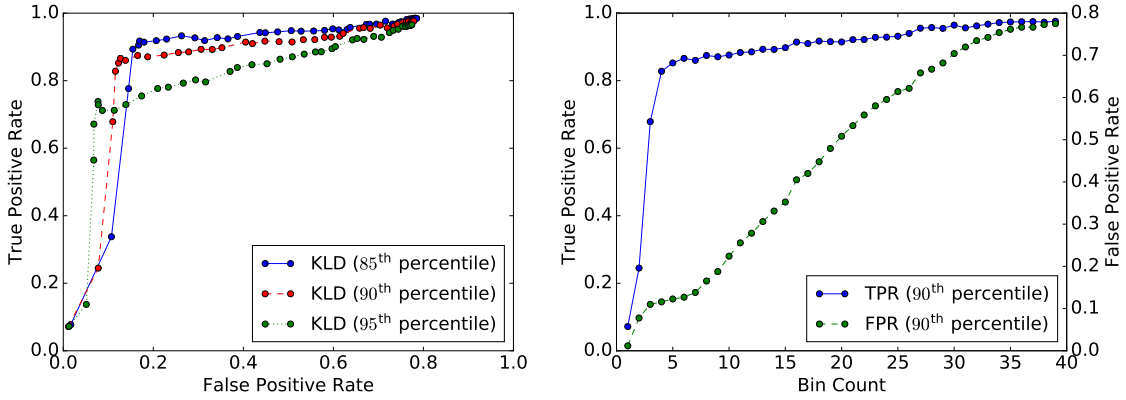


Fig. 8. Distribution of how much electricity can be stolen in one week through the use of optimal attacks against the PCA-DBSCAN detector ($18 on average) and the KLD detector ($37 on average).

the optimal attack vector against the KLD detector, which is agnostic to the time-ordering and allows the freedom to inject larger consumption values when the price is low. Also notice that both optimal attack vectors require consumption to be over-reported at certain times, in order to avoid detection.

C. Comparative Evaluation of the KLD Detector and the PCA-DBSCAN Detector using Optimal Attack Vectors

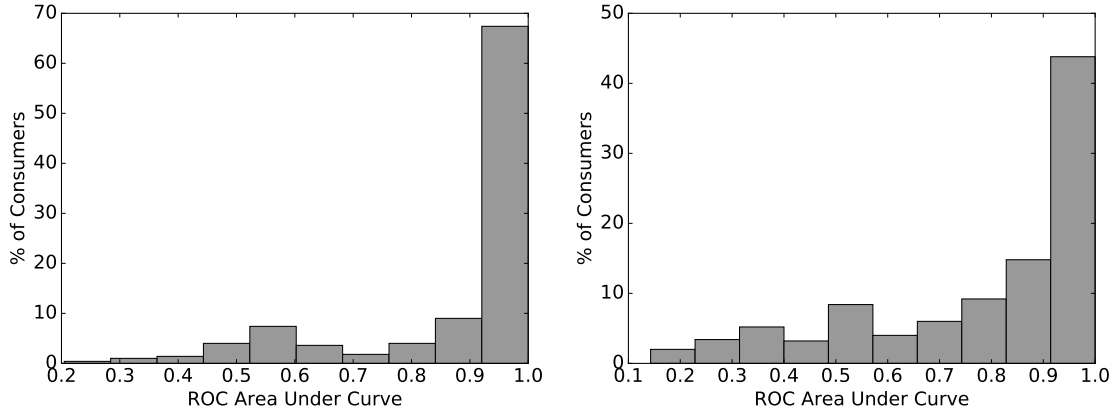
We compare the KLD and PCA-DBSCAN detectors using the CER dataset in terms of how much electricity can be stolen by using the optimal attack vector against each detector. 60 weeks of training data were used to construct the optimal attacks, as described previously in this section. Note that the optimal attacks depend on the detection thresholds. Therefore, a very tight threshold would not only mitigate the amount of electricity that could be stolen through the optimal attack, but also lead to a large false-positive rate. In this evaluation, we chose thresholds to ensure false-positive rates of less than 7.5%, as measured on 14 weeks of test data. The results were computed for the consumers in the CER dataset, considered independently as potentially malicious, and histograms of electricity values stolen by those consumers are plotted Fig. 8. The stolen values represent the difference between the actual consumption in the test set (D_A) and the reported consumption through the optimal attack vector (D'^{*}_{A}). The optimal attack vector was replayed during the 14-week-period of the test data,



(a) KLD ROC curve

(b) KLD TPR & FPR vs. $|B|$

Fig. 9. ROC for KLD detector on the Integrated ARIMA attack. (a) ROC curves for three different thresholds on the KLD distribution. (b) TPRs and FPRs across different bin counts ($|B|$) at a threshold set at the 90$^{\text{th}}$ percentile.



(a) KLD area under ROC curve (90$^{\text{th}}$ percentile)

(b) PCA-DBSCAN area under ROC curve (3 dimensions)

Fig. 10. Area under the ROC curve (AUC) for KLD and PCA-DBSCAN detectors on the integrated ARIMA attack. The larger the area, the better the detection performance. For a large fraction of consumers, the detector had near-perfect performance (close to 1).

and the stolen values illustrated in Fig. 8 are the averages of the values taken across the 14 weeks.

The PCA-DBSCAN detector outperformed the KLD detector because less electricity could be stolen through the optimal attack against the PCA-DBSCAN detector. That is evident from Fig. 8. Both detectors performed well in the sense that less than $50 could be stolen per week for over 80% of the consumers. Surprisingly, there were 7 consumers for whom the optimal attack against the PCA-DBSCAN detector produced losses for the attacker (average stolen value in 1 week was negative). That is because the optimal attack vector was computed without knowledge of the test data and, for those consumers, the consumption in the 14 weeks of the test data happened to be less than would be expected from the 60 weeks of the training data.

VII. ROCs for Electricity Theft Detectors

In this section, we extend prior work [9] in evaluating the KLD detector on the integrated ARIMA attack. We also evaluate the PCA-DBSCAN detector on the same attack and compare its results with those of the KLD detector. In doing so, we use the CER consumption dataset. We find that although the PCA-DBSCAN detector outperformed the KLD detector in terms of the impact of optimal attack vectors, the KLD detector outperformed the PCA-DBSCAN detector against the integrated ARIMA attack.

The two main limitations of our prior work were that 1) we did not sufficiently explore the parameter space of the proposed detectors, and 2) we did not consider the false-positive rates in evaluating the detectors. As false positives are expensive to investigate, a utility would like to minimize the false-positive rate. A well-known caveat in detection theory is that true-positive rates (TPRs) trade-off with false-positive rates (FPRs). That trade-off is illustrated by ROC curves. The curves are generated by plotting the TPR against the FPR for various detection thresholds. Note that TPR is nonzero in this section because the integrated ARIMA attack is not designed to evade the KLD detector or the PCA-DBSCAN detector.

A. ROC for the KLD Detector

For the KLD detector, two parameters need to be chosen by a practitioner, and they determine the effectiveness of the

detector. The first parameter is $|B|$, which is the number of bins that we would like to use to describe the nonparametric distribution of meter readings in the training set. In general, fewer bins produce fewer true and false positives. In the extreme case, a single bin produces 0% true and false-positive rates because all readings lie in the same bin and one cannot set an appropriate detection threshold. Similarly, a very large number of bins would produce higher true and false-positive rates, because the distribution becomes too fine-grained and over-fits the data. The second parameter to be chosen is the detection threshold on the distribution of KLD values. We had used percentiles to set the threshold, and had evaluated two choices (90^{th} and 95^{th} percentiles) in [9].

Fig. 9(a) illustrates the ROC curves for the KLD detector. The TPRs and FPRs were averaged over all consumers. The ROC curves are not monotonically increasing, because B is based on a nonparametric distribution that is not smooth. As a result, there might be data points in the test set that exist in an empty bin, increasing the KLD metric because those points were not expected from the training set. The relationship between the TPRs and FPRs with the bin size is more explicitly shown in Fig. 9(b). It can be seen that a practitioner would do well to choose $|B| = 5$ to achieve a good trade-off between TPRs and FPRs. With $|B| = 4$ and $|B| = 6$ there would be no major gains or losses over the choice of $|B| = 5$. In other words, the system is stable to small perturbations around $|B| = 5$.

Similarly, the operator would be well-advised to pick a detection threshold at the 90^{th} percentile, so that the detector can operate at 87% TPR and 12% FPR. We believe that that may be an acceptable FPR, but the utility could pick a different setting that achieves their preferred trade-off. The choice of percentile threshold is not immediately clear because the ROC curves for the different thresholds cross in Fig. 9(a). One commonly used approach compares ROC curves based on the *areas under the curves* (AUCs). A perfect detector has an AUC of 1. We do not see such perfect performance in the ROC curves shown in Fig. 9, because the TPRs and FPRs were averaged over all consumers. To illustrate the performance of the detector for each consumer considered separately, we provide a histogram of AUCs in Fig. 10(a). The AUCs were computed using the composite trapezoidal rule of integration.

Note that the FPR was dramatically decreased (in many cases by up to 20 percentage points) when the detector was "turned off" during the two weeks spanning Christmas and New Year's day. That period produced the maximum number of false positives across all consumers, likely because consumption patterns were affected by holiday schedules. The ROC curves in this section were calculated *after* the detectors were turned off during Christmas and New Year's day.

B. ROC for the PCA-DBSCAN Detector

The PCA-DBSCAN detector first projects the data into a lower-dimensional space. The dimensionality of that lower-dimensional space is a parameter that can be chosen by the operators. A smaller value will retain less of the variance from the original dataset. A larger value will include more noise

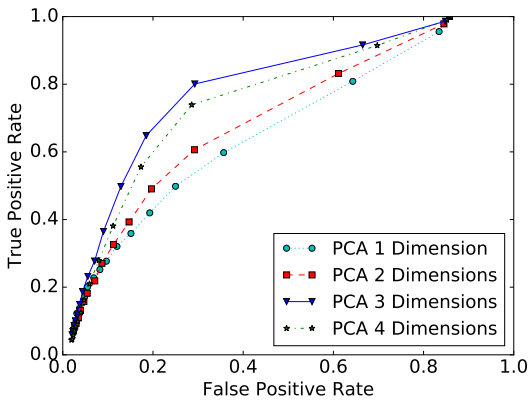


Fig. 11. ROC for PCA-DBSCAN detector on the Integrated ARIMA attack.

from the dataset. Thus there is an optimal dimension for a given dataset, and we found it to be equal to 3 dimensions for the CER dataset, as illustrated in Fig. 11. For each dimension, we generated the ROC curves by varying the ϵ parameter in DBSCAN. It is clear from Fig. 11 that 3 dimensions work best, because the corresponding curve lies entirely above the other curves. A histogram of AUCs across all consumers, considered separately as potentially malicious, is plotted in Fig. 10(b).

While it was clear from the ROC curves that the best KLD detector setting outperformed the best PCA-DBSCAN detector setting, we were surprised to find that both detectors had perfect performance on a large fraction of consumers, as seen in Fig. 10. Upon investigating the consumers that performed poorly, we found that some of them had near-zero consumption throughout the 74-week period, while others had zero consumption during the period of the test set alone. Therefore, those consumers could not be distinguished from malicious consumers even to the naked eye, and they would need to be investigated by a utility. We spoke with a representative from the Pacific Gas & Electric company in California, and he told us that they use knowledge of move-in and move-out dates of residents and check that low consumption readings were seen after move-outs. If we had had access to those dates for the consumers in the CER dataset, we might have been able to further reduce the FPRs.

VIII. Framework for Detecting DER Fraud

In this paper, we present the first analysis of how much attackers, who own or operate DERs, would stand to gain by fraudulently reporting that they generate more than they actually do. The detection of DER fraud, in the case of solar and wind, is different from that of consumer fraud in that those DERs generate electricity based on weather conditions. The dependence of those DERs on weather creates correlations among nearby DERs that can be leveraged for detection.

The framework for detecting DER fraud is an extension of F-DETA, which was described in Section IV-A. The meter readings of a generator C represent the *average net generation* $G_C(t) \in \mathbb{R}$ during each time period t, and are measured in kW/MW. $G'_C(t)$ is the reported value corresponding to $G_C(t)$. If $G'_C(t) \neq G_C(t)$, the meters are not reporting their actual values and must be investigated.

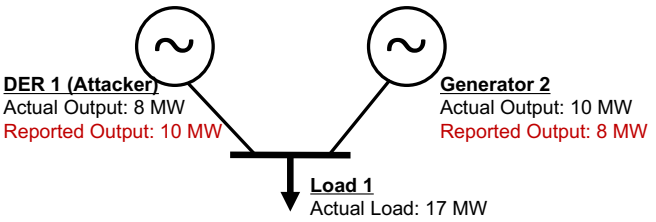


Fig. 12. How attackers can circumvent the balance check by over-reporting their own generation and simultaneously under-reporting another generator's output (or by over-reporting the load). Power losses amount to 1 MW.

Let $\lambda(t)$ denote the electricity price during the time period t, where $\lambda(t) \in \mathbb{R}$. Note that the price does not necessarily change between smart meter polling periods, and that price updates are usually less frequent than polling reports. We assume that in any time period t the electricity price $\lambda(t)$ is common to all customers.

The attackers' monetary advantage through fraud, α, is given by the difference between what the utility should pay them based on the actual generation, B_{Utility}, and what the utility actually pays them based on the reported generation, B'_{Utility}. If the billing cycle contains T time periods, then the attacker, A, can make a monetary gain through fraud if and only if the following condition holds:

$$\begin{aligned}\alpha &\triangleq B'_{\text{Utility}} - B_{\text{Utility}} \\ &= \sum_{t=1}^{T} \lambda(t) G'_A(t) \Delta t - \sum_{t=1}^{T} \lambda(t) G_A(t) \Delta t \\ &> 0,\end{aligned} \tag{20}$$

where the units are given as follows: λ is in \$/kWh, G is in kW, Δt is in hours, and α is in \$ (dollars). The attackers' objective is to maximize α subject to the constraint that the attack must go undetected. Since $\Delta t > 0$, (20) holds only if $\text{sgn}(\lambda(t))[G'_A(t) - G_A(t)] > 0$ for some t, where sgn is the sign function. The statement is evident and the proof follows from the proof of Proposition 1 in [9]. Therefore, the attackers must over-report their generation in order to make a monetary gain when $\lambda(t) > 0$ and under-report their generation when $\lambda(t) < 0$. We design attacks for the far more common case in which $\lambda(t) > 0$.

A naive way to detect such an attack would be to use a variant of the balance check, described in Section IV-A, that would use redundant meters to ensure that the total amount generated is the total amount consumed. The balance check is naive in that it can easily be circumvented, as follows. Consider the schematic diagram in Fig. 12. The two generators illustrated are logically separated. Each generator may be composed of multiple individual generators whose values sum up to the values shown in the figure; the same applies to loads. DER 1 represents a group of DERs whose reported output exceeds the actual output, while Generator 2 represents a group of generators whose reported output is less than the actual output. The load may also be misreported, but that is not illustrated in the figure for simplicity. Energy is conserved because the total actual generation is equal to the total reported generation. That value (18 MW) is the sum of the load (17 MW) and losses due to electric line impedance and transformer cores (1 MW). DER 1 effectively steals from Generation 2.

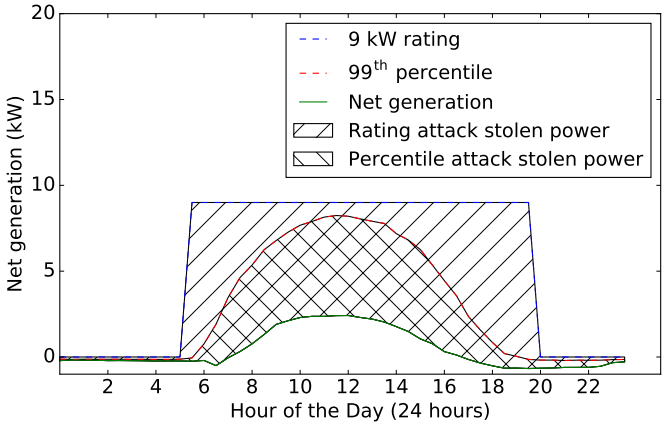


Fig. 13. Rating and percentile attacks illustrated for one customer in the Ausgrid solar dataset. The shaded regions represent the stolen electricity.

IX. Optimal Attack Vectors for DER Fraud

As described in the attack model in Section VIII, an attacker may commit extensive fraud by over-reporting generation by an arbitrarily large value. In this section, we present detectors that can mitigate such fraud. For each detector, we identify the optimal attack vector, which maximizes the fraud while avoiding detection. In Section XI, we will evaluate each detector based on how much an attacker can gain using the optimal attack that circumvents that detector.

1) Rating Attack: A rating-based detector to limit the over-reporting of attacker generation would ensure that the reported generation does not exceed the DER's rating. The optimal attack for this detector sets the generation readings at the rating threshold. As a result, the attack vector does not exceed the threshold. Simultaneously, the attackers maximize how much they can steal by over-reporting their generation. This is the optimal attack for a rating-based threshold, and we refer to it in this paper as the *rating attack*.

In the case of solar, the generation is zero before sunrise and after sunset. Therefore, in designing the rating-based detector for solar generation, we ensured that the detection threshold is set such that the generation can never exceed zero before sunrise and after sunset. From sunrise to sunset, the upper threshold is the solar panel's rating. The amount of electricity stolen is the difference between the rating and the actual generation, as illustrated in Fig. 13.

In a net metering system, if a DER owner were to claim that the net generation was equal to the rating of the panel, they would effectively be claiming that their consumption was zero. In doing so, they not only over-report their generation, but also under-report their consumption, which is theft.

Unlike solar generation, wind generation of a turbine can reach its rated capacity at any time of the day or night. Therefore, the detection threshold would be set at the rated capacity of the turbine throughout the 24-hour day.

2) Percentile Attack: The rating attack, particularly for solar power, is naive in that it does not capture diurnal variations in generation throughout the day. For solar, for example, the output steadily increases until midday and then steadily decreases in the evening, according to the solar irradiance. In order to determine whether each solar output reading at a particular time is anomalous, one approach may leverage the diurnal patterns, and compare that reading with readings taken at the same time on previous days. Our percentile threshold accomplishes that by setting a threshold at the 99^{th} percentile of data points seen at the same time on previous days. For example, to determine whether a reading at 10:00 a.m. on a given day is anomalous, we check whether that reading is greater than the 99^{th} percentile of generation values taken at 10:00 a.m. on previous days. Our choice of percentile point is based on achieving an acceptable trade-off between true positives and false positives. For the percentile-based threshold, the optimal attack, which we refer to as the *percentile attack*, requires that the attacker know where the threshold has been set. In doing so, they would not exceed the threshold, and could maximize how much they can steal by over-reporting their generation while going undetected. The fraudulent gain is the difference between the percentile threshold and the actual generation, as illustrated in Fig. 13. That fraudulent gain is always less than the gain that can be made through the use of the rating attack. Therefore, the percentile-based threshold mitigates the extent of possible fraud, relative to the rating-based threshold.

3) Correlation Attack: One way to detect attacks in the context of DERs, like solar and wind, is to leverage their dependence on the availability of sunlight and wind, respectively. Therefore, one would expect the generations of different DERs to be correlated. We verified that with all three DER datasets; the results are illustrated in Fig. 14. The heatmaps in the figure were obtained by computing the pairwise Pearson correlation coefficients between the DERs in the datasets.

Let A be the attacker and C be a DER used by the utility to check whether A is anomalous. Correlation implies a linear relationship, which is modeled as follows.

$$G'_A(t) = mG'_C(t) + c + \epsilon, \tag{21}$$

where m and c are the slope and intercept obtained from linear regression, and $\epsilon \sim N(0, \sigma^2)$ is zero-mean Gaussian noise. It is Gaussian because linear regression inherently minimizes the squared L2 norm of the fitting error, which in turn maximizes the likelihood that the errors were Gaussian. All three parameters were obtained from the training set.

We claim that $G'_A(t)$ in the test set is anomalous if the following condition holds:

$$|G'_A(t) - (mG'_C(t) + c)| > k\sigma, \tag{22}$$

where m, c, and σ were obtained from the training set and G'_C was obtained at the same time as G'_A. k is the threshold parameter that determines the ROC for this detector. The optimal attack, which we refer to as the *correlation attack*, is achieved when A sets their generation reading as follows:

$$G'^{*}_A(t) = \min(mG'_C(t) + c + k\sigma, R_A(t)), \tag{23}$$

where $R_A(t)$ is A's rating, which should not be exceeded. In order to accomplish this attack, A would need to know that the utility is using C's readings for anomaly detection, and A would then need to monitor C's readings. In addition, A would need to know k. The difficulty of obtaining all that information may make this attack much less likely than the previous attacks, which looked only at A's own past readings.

4) Weather-Based Detectors: A limitation of using historical data in the previously described methods is that we need to assume that the training data have not been tampered with. If the data have been tampered with, then statistical learning methods trained on that data could become biased in a way that the attacker could escape detection.

We now discuss the use of weather data to perform detection. The assumption is that weather data can be obtained from a completely different data source, which the attacker has not compromised. For example, a utility could use IBM's Deep Thunder [42], which provides wind speed and wind direction at turbine altitudes with a spatial resolution of 1 to 2 km. It also provides solar irradiance data at that spatial granularity.

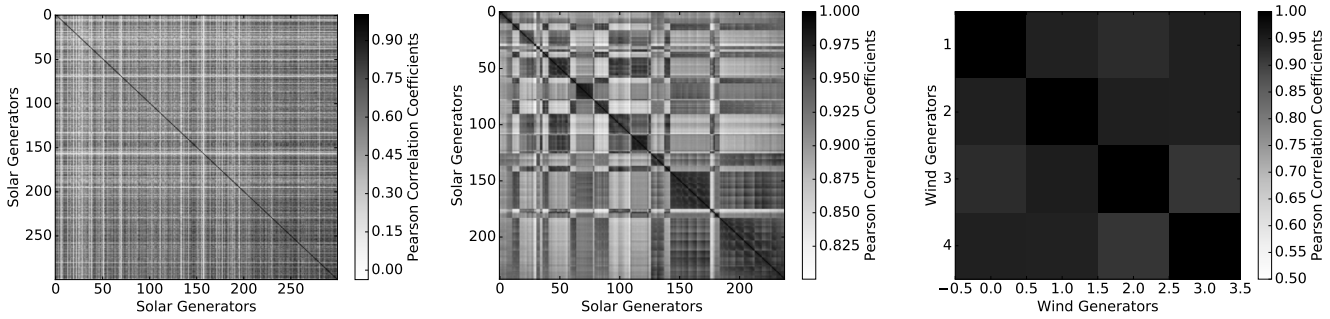
For a fixed wind turbine configuration, the power produced can be obtained from wind speed measurements by using a well-known physical relationship called the *power curve*. Power curves for over 200 manufacturers' turbines are provided in [43] as look-up tables that map wind speed to expected power for each of those turbines. The operator could use those data to detect anomalous deviations between every wind turbine's expected and reported output.

Similarly, solar generation can be predicted by using irradiance data along with solar panel configuration details such as the tilt angle (with respect to the axis perpendicular to the surface of the earth) and azimuth angle (with respect to true north). The NREL PVWatts calculator ([44]) estimates the expected solar output from the photovoltaic array. If the reported solar output significantly deviates from the expected output, the operator must investigate the cause of the deviation, as it may be indicative of an attack.

Since the Engie dataset contains wind power and wind speed measurements, we created an empirical model of the power curve, and used it to design what we call the *power curve detector*. The model is simple; we calculate the probability of wind power measurements, G'_A, given wind speed measurements, S'_A. We then extract distribution parameters for $P(G'_A|S'_A)$, such as the median and the median absolute deviation (MAD), and use those for anomaly detection as follows.

$$|G'_A(t) - \text{median}(G'_A|S'_A)| > k\text{MAD}(G'_A|S'_A), \tag{24}$$

where $P(G'_A|S'_A)$ is obtained from the training data. We used the medians and MADs because they are robust statistics, unlike the means and standard deviations. Fig. 15 illustrates the power curve from the training set; the anomalies are not malicious, but are present in the dataset. If we had used means and standard deviations, those anomalies would have skewed the model and made it less effective for detection of anomalies in the test set. Once again, the optimal attack for this detector sets the generation at the detection threshold, thereby ensuring maximum gains.



(a) Ausgrid residential-scale solar (b) NREL utility-scale solar (c) Engie wind turbines

Fig. 14. Cross-correlations. These heatmaps plot the Pearson correlation coefficient between all pairs of DERs in the dataset. Note that the minimum values on the scales are not zero, so the cross-correlations are all high.

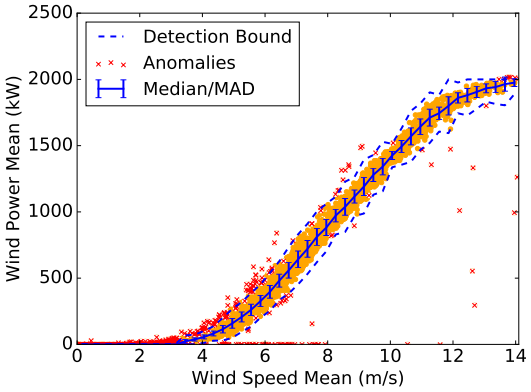


Fig. 15. Statistical estimation of wind power from wind speed. The data points are classified into normal (yellow circles) and anomalies (red crosses).

5) KLD Attack: Although the KLD detector works exactly the same way for DER fraud as it does for consumption fraud, the optimal attack against the detector is slightly different. The optimal attack described in Section VI-A involves minimizing the average consumption subject to the KLD threshold constraints. In the DER fraud case, the optimal attack, which we refer to as the *KLD attack*, involves maximizing the average reported generation subject to the same KLD threshold constraint. That is formulated as follows.

$$P_A^* = \arg\max_{P_A} \sum_{b\in B} X_{Tr}(b) P_A(b) \tag{25}$$

$$\text{subject to } \sum_{b\in B} P_A(b) \log \frac{P_A(b)}{P_{Tr}(b)} \leq \tau, \tag{26}$$

$$P_A(b) \geq 0, \tag{27}$$

$$\text{and } \sum_{b\in B} P_A(b) = 1, \tag{28}$$

where $P_A(b)$ denotes the probability $Prob(G'_A(t) \in b)$ for $b \in B$. Note that the objective function changes, but the constraints do not change from the consumption case described in Section VI-A. The problem remains convex, and the attack vector G'^*_A is generated from the distribution given by P^*_A. The ordering of values in G'^*_A is chosen such that $\sum_{t=1}^{T} \lambda(t) G'^*_A(t)$ is maximized.

6) PCA-DBSCAN Detector: We had proposed this detector in the context of consumption readings, but we found that it was not suitable for DER fraud detection. In order for this detector to be successful, the generation patterns projected onto a lower-dimensional space would have to be tightly clustered so that density-based clustering could be used for anomaly detection. We found that that was not true of our solar and wind datasets, because the data were not tightly clustered in the lower-dimensional space.

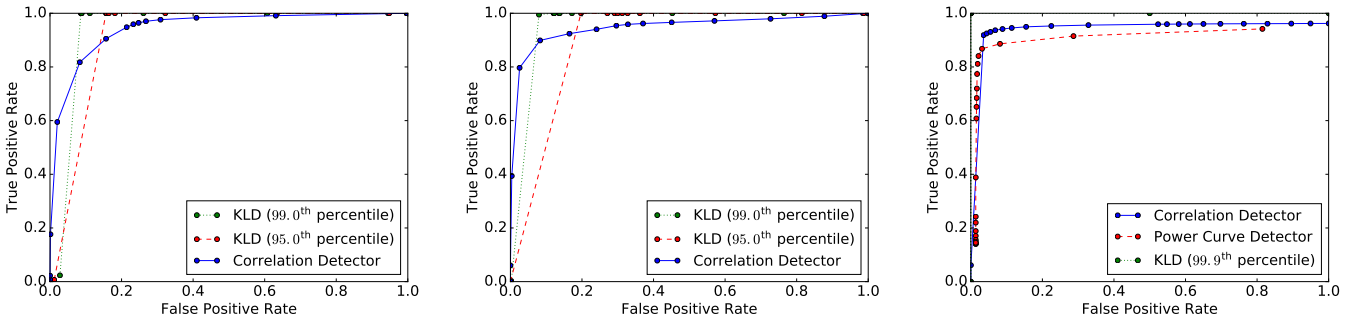
X. ROCs for DER Fraud Detectors

We evaluate the KLD detector against the correlation detector in terms of how well they detect the *percentile attack*. The ROC curves for the results are presented in Fig. 16. It can be seen that the KLD detector narrowly outperforms the correlation detector for both the solar and wind datasets. In certain settings, as seen in Fig. 16(b), the correlation detector may achieve a more desirable trade-off with a lower FPR than the KLD detector.

The KLD detector waits until a week of readings has been obtained and then determines that the readings have been over-reported consistently over the week. As a result, the percentile attack produces a probability distribution of readings that differs greatly from the probability distribution of the training set. As seen in Fig. 16(c) for wind, the KLD detector achieves perfect detection performance, with an AUC of 1, at a 99.9^{th} percentile detection threshold. In the case of solar, its performance (in terms of AUC) is comparable with that of the correlation detector.

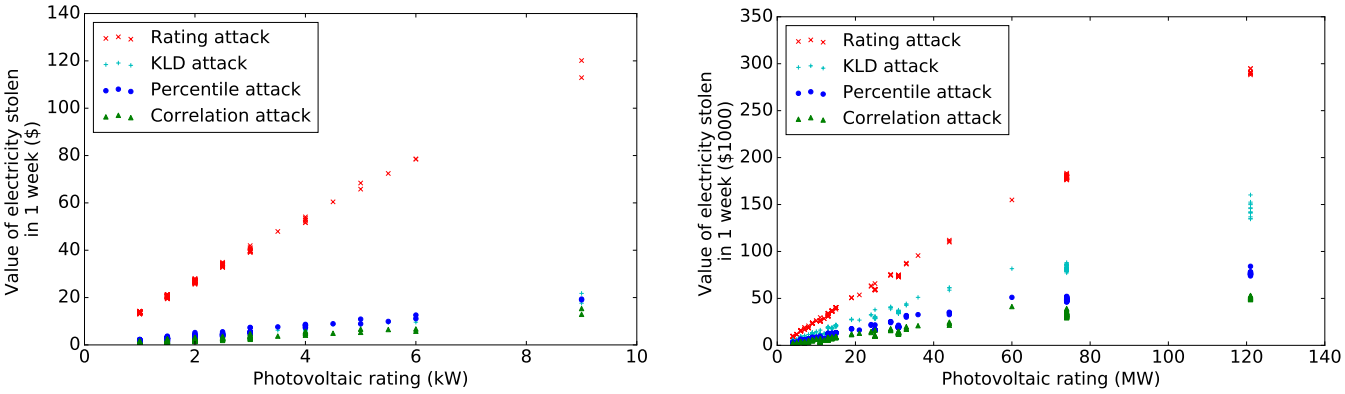
The fact that the correlation detector can work in real-time is an advantage over the KLD detector. However, we believe that an operator could use both detectors together, one in real-time and one at a periodicity of one week.

The ROC curve for the power curve detector is illustrated in Fig. 16(c) and is produced by varying k. As is evident from the illustration, the power curve detector is inferior to the correlation detector. For $k = 0$ it does not have a high TPR or a high FPR because it turned out that the value of the percentile attack vector was lower than the median power for the given speeds.



(a) Ausgrid residential-scale solar (b) NREL utility-scale solar (c) Engie wind turbines

Fig. 16. ROC curves for DER mitigation methods. The KLD detector performs the best.



(a) Residential-scale solar (Ausgrid) (b) Utility-scale solar (NREL)

Fig. 17. Values of monetary gains obtained through the use of attack vectors that are optimal against detectors of solar generation fraud.

XI. Profit Analysis of DER Fraud

The values of monetary gains obtained through the use of attack vectors that are optimal against detectors of solar generation fraud are illustrated in Fig. 17. Those values were calculated by multiplying the over-reported generation with the electricity price (the Ausgrid feed-in tariff of \$0.07/kWh [45] was applied to the Ausgrid dataset; the wholesale electricity price of \$0.03/kWh was applied to the NREL dataset [46]). As seen in Fig. 17, the value of the electricity stolen in an average week varies linearly with the rating of the solar panel for all the optimal attacks.

Monetary gain is the attackers' primary incentive in the model described in Section IV-A. However, before a gain is made, they must recover their DER capital costs. In this section, we quantify how long it takes for them to recover those capital costs through the optimal attacks discussed in Section IX. Since capital costs are often large, the ability to speedup the return on investment through fraud may serve as an incentive for attackers to commit that fraud.

For the Ausgrid dataset, we use the solar installation costs given in [47]; the costs vary with the size of the installation. For the NREL solar dataset, we obtained the cost per watt from the commercial solar power model given in [4]. We assume 1 MW panel array increments, and use the corresponding cost per watt of \$2.03. For the Engie wind dataset, we assume a cost of \$4 million for the installation of each 2 MW turbine at each site. That is a conservative estimate taken from the range of \$3–4 million given in [48].

TABLE I
Years to recover DER capital costs

	Small-Scale Solar (Ausgrid)	Utility-Scale Solar (NREL)	Utility-Scale Wind (NREL)
No Attack	6.8 | 11.8 | 23.7	47.9 | 54.4 | 90.1	37.2 | 40.8 | 48.7
Rating Attack	1.5 | 2.4 | 3.0	12.4 | 12.4 | 12.4	7.6 | 7.6 | 7.6
KLD Attack	4.7 | 8.4 | 14.9	17.7 | 20.7 | 23.3	8.5 | 9.2 | 9.3
Percentile Attack	4.7 | 8.0 | 13.8	25.2 | 28.1 | 30.2	9.0 | 9.3 | 10.2
Correlati -on Attack	5.3 | 9.1 | 15.4	30.0 | 33.5 | 40.3	17.7 | 20.4 | 23.6

Values are formatted as 'min | median | max' across all DERS in each type.

The years needed to recover DER capital costs, aggregated across all DERs in each dataset, are given in Table I. "Small-scale solar" refers to residential-scale solar. For all three datasets, it is evident that by committing fraud, attackers can recover their capital costs much faster than they could have if they had not committed fraud. That provides them with an additional incentive to commit fraud. Across all types of DERs, the rating attack reduced the time it took to recover the capital costs by around 80% on average. Less gains were obtained through the KLD attack in comparison to the rating attack, but the KLD attack was more advantageous for the attacker than the other attacks. The percentile attack benefited solar installations less than wind installations, because solar

generation was less erratic and approached the percentile-based threshold more often than wind generation approached that threshold. Therefore, the opportunity for fraud with the percentile attack was less with solar than it was with wind. As the correlation attack was the least beneficial to the attacker, the correlation detector was the most beneficial to the defender. In that sense, the correlation detector mitigates the other attacks by forcing the attackers to wait much longer to recover their capital costs. The hope is that the additional wait time will disincentivize the attacker from committing fraud.

Similar to the case of consumption fraud, the KLD detector performed the best in terms of ROCs against other attacks (in this case the percentile attack). However, the optimal attack against the KLD detector allowed greater gains for the attacker in comparison to the optimal attack against the other detectors. In that sense, the KLD detector performed the worst in the worst-case scenario. The correlation detector performs almost as well as the KLD detector in terms of ROCs against the percentile attack, and it performs much better against its optimal attack. In that sense, the correlation detector performs the best overall in detecting and mitigating DER fraud.

XII. Conclusion

In this paper, we presented signal processing methods to detect electricity theft and DER fraud. We derived optimal attack vectors against those methods. We used ROC curves to compare the TPRs and FPRs of those methods, allowing utilities to choose a threshold that produces a suitable trade-off. We used examples from wind and solar generation to illustrate how much an attacker would stand to gain monetarily from DER fraud. We showed that that gain can enable attackers to decrease the time it would take them to recover the capital cost of their solar or wind installations. We presented various detection mechanisms that could be used to detect and mitigate such fraud. The detectors were evaluated based on how much an attacker could possibly gain by evading them. The evaluation was driven by freely available data from Australia (provided by Ausgrid), France (provided by Engie), Ireland (provided by CER), and the U.S. (provided by NREL). We hope that our work will motivate research and development efforts to secure advanced metering infrastructure from electricity theft and DER fraud. Our code is freely available on GitHub [49].

Acknowledgment

We thank Prof. Ravishankar Iyer, Dr. Kiryung Lee, and Dr. Gabriel Weaver for their contributions to prior work [9], [10], [11]. We thank Jenny Applequist for editorial comments. This material is based upon work supported by the Department of Energy under Award Number DE-OE0000097 and the Siebel Energy Institute. The CER smart meter dataset used in this paper was accessed via the Irish Social Science Data Archive at www.ucd.ie/issda.

References

[1] Cyber Intelligence Section, *Smart Grid Electric Meters Altered to Steal Electricity (Intelligence Bulletin)*, Federal Bureau of Investigation, May 2010. [Online]. Available: http://krebsonsecurity.com/wp-content/uploads/2012/04/FBI-SmartMeterHack-285x305.png

[2] M. Ward, "Smart meters can be hacked to cut power bills," *BBC News*, Oct. 2014. [Online]. Available: http://www.bbc.com/news/technology-29643276

[3] Renewable Energy Policy Network for the 21st Century (REN21), *Renewables 2016: Global Status Report*, 2016. [Online]. Available: http://www.ren21.net/wp-content/uploads/2016/06/GSR_2016_Full_Report.pdf

[4] R. Fu, D. Chung, T. Lowder, D. Feldman, K. Ardani, and R. Margolis, *U.S. Solar Photovoltaic System Cost Benchmark: Q1 2016*, Nat. Renewable Energy Laboratory (NREL), Sep. 2016. [Online]. Available: http://www.nrel.gov/docs/fy16osti/66532.pdf

[5] R. Wiser and M. Bolinger, *2015 Wind Technologies Market Report*, U.S. Dept. of Energy, Aug. 2016, http://www.nrel.gov/docs/fy16osti/66532.pdf.

[6] A. Lee, *U.S. Wind Generating Capacity Surpasses Hydro Capacity at the End of 2016*, U.S. Energy Inform. Admin., accessed May 2017. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=30212

[7] U.S. Energy Inform. Admin., "Rising solar generation in California coincides with negative wholesale electricity prices," Apr. 2017, accessed May 2017. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=30692

[8] Int. Renewable Energy Agency, *Wind Power (Renewable Energy Technologies: Cost Analysis Series: Vol. 1: Power Sector, Issue 5/5)*, Jun. 2012, accessed May 2017. [Online]. Available: https://www.irena.org/DocumentDownloads/Publications/RE_Technologies_Cost_Analysis-WIND_POWER.pdf

[9] V. Badrinath Krishna, K. Lee, G. A. Weaver, R. K. Iyer, and W. H. Sanders, "F-DETA: A framework for detecting electricity theft attacks in smart grids," in *Proc. 2016 46th Annu. IEEE/IFIP Int. Conf. Dependable Systems and Networks*, pp. 407–418.

[10] V. Badrinath Krishna, R. K. Iyer, and W. H. Sanders, "ARIMA-based modeling and validation of consumption readings in power grids," in *Proc. Int. Conf. Critical Inform. Infrastructures Security*, ser. LNCS, vol. 9578. Springer, 2015, pp. 199–210.

[11] V. Badrinath Krishna, G. A. Weaver, and W. H. Sanders, "PCA-based method for detecting integrity attacks on advanced metering infrastructure," in *Proc. Int. Conf. Quantitative Evaluation of Syst.*, ser. LNCS, vol. 9259. Springer, 2015, pp. 70–85.

[12] D. Mashima and A. A. Cárdenas, "Evaluating electricity theft detectors in smart grid networks," in *Proc. Int. Workshop Recent Advances in Intrusion Detection*, ser. LNCS, vol. 7462. Springer, 2012, pp. 210–229.

[13] P. Jokar, N. Arianpoo, and V. C. M. Leung, "Electricity theft detection in AMI using customers' consumption patterns," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 216–226, Jan. 2016.

[14] J. L. Viegas, P. R. Esteves, and S. M. Vieira, "Clustering-based novelty detection for identification of non-technical losses," *International Journal of Electrical Power & Energy Systems*, vol. 101, pp. 301 – 310, 2018.

[15] R. Jiang, R. Lu, L. Wang, J. Luo, S. Changxiang, and S. Xuemin, "Energy-theft detection issues for advanced metering infrastructure in smart grid," *Tsinghua Sci. and Technol.*, vol. 19, no. 2, pp. 105–120, Apr. 2014.

[16] S. Depuru, L. Wang, and V. Devabhaktuni, "Support vector machine based data classification for detection of electricity theft," in *Proc. IEEE/PES Power Systems Conf. and Expo.*, Mar. 2011, pp. 1–8.

[17] S. Depuru, L. Wang, V. Devabhaktuni, and N. Gudi, "Measures and setbacks for controlling electricity theft," in *Proc. North Amer. Power Symp.*, Sep. 2010, pp. 1–8.

[18] D. N. Nikovski, Z. Wang, A. Esenther, H. Sun, K. Sugiura, T. Muso, and K. Tsuru, "Smart meter data analysis for power theft detection," in *Proc. Int. Workshop Machine Learning and Data Mining in Pattern Recognition*, ser. LNCS, vol. 7988. Springer, 2013, pp. 379–389.

[19] S. McLaughlin, B. Holbert, S. Zonouz, and R. Berthier, "AMIDS: A multi-sensor energy theft detection framework for advanced metering infrastructures," in *Proc. SmartGridComm'12*, Nov. 2012, pp. 354–359.

[20] S. McLaughlin, D. Podkuiko, and P. McDaniel, "Energy theft in the advanced metering infrastructure," in *Proc. Int. Workshop Critical Inform. Infrastructures Security*, ser. LNCS, vol. 6027. Springer, 2010, pp. 176–187.

[21] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Comput. and Commun. Security*, 2009, pp. 21–32.

[22] 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.

[23] P. Kelly-Detwiler, "Electricity theft: A bigger issue than you think," *Forbes*, Apr. 2013. [Online]. Available: http://www.forbes.com/sites/peterdetwiler/2013/04/23/electricity-theft-a-bigger-issue-than-you-think/
[24] S. McLaughlin, D. Podkuiko, S. Miadzvezhanka, A. Delozier, and P. McDaniel, "Multi-vendor penetration testing in the advanced metering infrastructure," in *Proc. 26th Ann. Comput. Security Appl. Conf.*, 2010, pp. 107–116.
[25] S. McWilliams, "Tamper protection for an automatic remote meter reading unit," U.S. Patent 4,357,601, 1982, Bell Telephone Laboratories.
[26] M. Afgani, S. Sinanovic, and H. Haas, "Anomaly detection using the Kullback-Leibler divergence metric," in *Proc. 1st Int. Symp. Appl. Sciences on Biomedical and Communication Technologies*, 2008, pp. 1–5.
[27] J. Harmouche, C. Delpha, and D. Diallo, "Faults diagnosis and detection using principal component analysis and Kullback-Leibler divergence," in *Proc. 38th Annu. Conf. IEEE Industrial Electronics Society*, 2012, pp. 3907–3912.
[28] J. D. Glover, M. Sarma, and T. Overbye, *Power System Analysis & Design, SI Version*. Cengage Learning, 2011.
[29] National Electrical Manufacturers Association, *Protocol Specification for ANSI Type 2 Optical Port.*, Amer. Nat. Standards Inst., Inc. Std., 2006.
[30] S. McIntyre, *Termineter*, Jan. 2018, https://github.com/securestate/termineter/blob/master/README.md.
[31] *Electric Sector Failure Scenarios and Impact Analyses Version 3.0*, National Electric Sector Cybersecurity Organization Resource (NESCOR), Dec. 2015. [Online]. Available: http://smartgrid.epri.com/doc/NESCOR%20Failure%20Scenarios%20v3%2012-11-15.pdf
[32] Idaho National Laboratory, *Vulnerability Analysis of Energy Delivery Control Systems*, Sep. 2011, accessed May 2017. [Online]. Available: https://www.energy.gov/sites/prod/files/Vulnerability%20Analysis%20of%20Energy%20Delivery%20Control%20Systems%202011.pdf
[33] R. J. Hyndman and Y. Khandakar, "Automatic time series forecasting: The forecast package for R," *J. Statistical Software*, vol. 27, no. 3, pp. 1–22, 2008.
[34] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise." in *Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining*, vol. 96, 1996, pp. 226–231.
[35] P. J. Rousseeuw and C. Croux, "Alternatives to the median absolute deviation," *J. Amer. Statistical Assoc.*, vol. 88, no. 424, pp. 1273–1283, 1993.
[36] E. Ireland. (2015, Nov.) Valuesaver nightsaver electricity price plan. Electric Ireland. [Online]. Available: https://www.electricireland.ie/switchchange/detailsValueSaverNightSaver.htm
[37] Ausgrid, *Solar Home Electricity Data*, Jul. 2010. [Online]. Available: http://www.ausgrid.com.au/Common/About-us/Corporate-information/Data-to-share/Solar-home-electricity-data.aspx
[38] National Renewable Energy Laboratory, *Solar Power Data*, Dec. 2016. [Online]. Available: http://www.nrel.gov/grid/solar-power-data.html
[39] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd, "SCS: Splitting conic solver, version 2.0.2," Nov. 2017. [Online]. Available: https://github.com/cvxgrp/scs
[40] S. Diamond and S. Boyd, "CVXPY: A Python-embedded modeling language for convex optimization," *J. Machine Learning Res.*, vol. 17, no. 83, pp. 1–5, 2016.
[41] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd, "Conic optimization via operator splitting and homogeneous self-dual embedding," *Journal of Optimization Theory and Applications*, vol. 169, no. 3, pp. 1042–1068, Jun. 2016. [Online]. Available: http://stanford.edu/~boyd/papers/scs.html
[42] IBM, "Deep Thunder," Aug. 2003. [Online]. Available: http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepthunder/
[43] *The Wind Turbine Database*, WindPower Program, http://www.wind-power-program.com/download.htm#database. Accessed Jun. 2017.
[44] National Renewable Energy Laboratory, *PVWatts Calculator*, accessed Apr. 2017. [Online]. Available: http://pvwatts.nrel.gov/pvwatts.php
[45] Independent Pricing and Regulatory Tribunal, *Solar Feed-in Tariffs: Setting a Fair and Reasonable Value for Electricity Generated by Small-Scale Solar PV Units in NSW: Energy – Final Report*, Mar. 2012. [Online]. Available: https://www.ipart.nsw.gov.au/files/sharedassets/website/trimholdingbay/final_report_-_solar_feed-in_tariffs_-_march_2012.pdf
[46] K. Shallenberger, "CAISO: Wholesale power prices dropped 9% in 2016 to $34/MWh average," *Utility Dive*, May 2017, accessed Oct. 2017. [Online]. Available: http://www.utilitydive.com/news/caiso-wholesale-power-prices-dropped-9-in-2016-to-34mwh-average/442626/
[47] Solar Choice, "1.5kW solar PV systems: Pricing, outputs and payback," Aug. 2016. [Online]. Available: http://www.solarchoice.net.au/blog/1-5kw-solar-pv-systems-price-output-payback
[48] Windustry, "How much do wind turbines cost?" accessed Dec. 2016. [Online]. Available: http://www.windustry.org/how_much_do_wind_turbines_cost
[49] V. Badrinath Krishna, *Algorithms and Code for Detecting Electricity Theft and DER Fraud*, University of Illinois at Urbana-Champaign. [Online]. Available: https://github.com/varunbk/etheft/

Varun Badrinath Krishna is a Ph.D. candidate in the Department of Electrical and Computer Engineering and a research assistant in the Information Trust Institute at the University of Illinois at Urbana-Champaign (UIUC). He is advised by Prof. William H. Sanders, and is researching data-driven methods to improve resilience, trust, and efficiency in power grids. His work on detecting electricity theft won best paper awards at QEST 2015 and CRITIS 2015, and seed funding from the Siebel Energy Institute. His work with IBM Research, on improving the prediction accuracy for wind power, led to two U.S. patent applications with him as the primary inventor. Varun was inducted into the Siebel Scholar Class of 2018 for academic, research, and leadership excellence in energy sciences.

Carl A. Gunter is a professor in the Department of Computer Science and the College of Medicine at the University of Illinois at Urbana-Champaign. He serves as the director of the Illinois Security Lab and the Health Information Technology Center (HITC). Professor Gunters contributions to the field of formal methods include the Packet Language for Active Networks (PLAN), the WRSPM reference model for requirements and specifications, the first formal analyses of Internet and ad hoc routing protocols, the Verisim system for analyzing network simulations, and the exploitation of bandwidth contention as a DoS countermeasure. His work on security and privacy included the first research on certificate retrieval for trust management and the formal analysis of regulatory privacy rules. His recent research focuses on security and privacy issues for the electric power grid and healthcare information technologies.

William H. Sanders is a Donald Biggar Willett Professor of Engineering and the Head of the Department of Electrical and Computer Engineering (ECE) at the University of Illinois at Urbana-Champaign. He is a professor in the Department of ECE and in the Department of Computer Science. He is a Fellow of the IEEE, the ACM, and the AAAS; a past Chair of the IEEE Technical Committee on Fault-Tolerant Computing; and past Vice-Chair of the IFIP Working Group 10.4 on Dependable Computing. He served as the Director and PI of the DOE/DHS Trustworthy Cyber Infrastructure for the Power Grid (TCIPG) Center, which did research at the forefront of national efforts to make the U.S. power grid smart and resilient. He was the 2016 recipient of the IEEE Technical Field Award, Innovation in Societal Infrastructure, for "assessment-driven design of trustworthy cyber infrastructures for electric grid systems."