Received: 24 January 2023 | Revised: 4 April 2023 | Accepted: 12 April 2023

IET Cyber-Physical Systems: Theory & Applications

DOI: 10.1049/cps2.12057

ORIGINAL RESEARCH

Detecting smart meter false data attacks using hierarchical feature clustering and incentive weighted anomaly detection

Martin Higgins[1] | Bruce Stephen[2] | David Wallom[1]

[1]Oxford e-Research Center, Department of Engineering Science, University of Oxford, Oxford, UK

[2]Advanced Electrical Systems Research Group, Institute of Energy and Environment, University of Strathclyde, Glasgow, UK

Correspondence

Martin Higgins.
Email: martin.higgins@eng.ox.ac.uk

Funding information

Engineering and Physical Sciences Research Council, Grant/Award Number: EP/S030131/1

Abstract

Spot pricing is often suggested as a method of increasing demand-side flexibility in electrical power load. However, few works have considered the vulnerability of spot pricing to financial fraud via false data injection (FDI) style attacks. The authors consider attacks which aim to alter the consumer load profile to exploit intraday price dips. The authors examine an anomaly detection protocol for cyber-attacks that seek to leverage spot prices for financial gain. In this way the authors outline a methodology for detecting attacks on industrial load smart meters. The authors first create a feature clustering model of the underlying business, segregated by business type. The authors then use these clusters to create an incentive-weighted anomaly detection protocol for false data attacks against load profiles. This clustering-based methodology incorporates both the load profile and spot pricing considerations for the detection of injected load profiles. To reduce false positives, the authors model incentive-based detection, which includes knowledge of spot prices, into the anomaly tracking, enabling the methodology to account for changes in the load profile which are unlikely to be attacks.

KEYWORDS

cyber-physical systems, data analysis, data privacy, embedded systems, security of data, smart cities, smart meters, smart power grids, telecommunication security

1 | INTRODUCTION

The contemporary power network is a cyber-physical system consisting of modern communication technologies working in conjunction with sophisticated power electronics. Up until recently, most of the power system innovations in real-time monitoring occurred at the transmission layer. However, the recent introduction of smart metering offers exciting opportunities for distribution level consumers and system operators to optimise their consumption of power. The increased granularity offered by load profile data offers new ways to reduce costs and encourage demand-side flexibility [1, 2]. Increasingly, variable tariffs are becoming popular which offer intraday variation in the electricity consumption price. These tariffs make utility level spot prices directly available to industrial consumers themselves [3]. Exposure to these spot curves offers some consumers an opportunity to save money. As consumers can now receive cheaper prices to consume power during non-peak hours, they can save cash if they act to adjust their demand curves. These advantages can also extend to the system operators in terms of potential benefits from lower intraday volatility in consumption which may increase network stability.

1.1 | Motivation

However, the introduction of spot prices also brings potential issues. Fraud is already a well-known issue in the modern distribution network. This is especially true in developing markets wherein upto 20% of the produced power might be stolen or consumed by customers committing fraud [4]. In the past, fraud has been limited to bypassing or stealing electricity. However, the advent of smart metering and variable tariffs will introduce new risks into the framework which are not currently considered. On one hand, smart meters can enable consumers to use spot pricing, which unlocks rewards for proactive consumers; on the other hand, the large intraday volatility of spot pricing means there is a direct cash incentive for malicious actors with the capabilities to bypass the relatively basic smart-metering cyber-infrastructure. These cash incentives are amplified for industrial users for whom electricity demand far outstrips the average consumer. In view of this, it is necessary to begin considering how users may try and exploit the system. In this work, we aim to provide a methodology for detecting smart meter attacks which take advantage of spot pricing. We are motivated by the risk of smart meter load profiles to financial fraud and seek to provide a detection methodology in order to protect them. The scope of this work is based around smart meter load profiles and the ability to detect changes in the smart meter profile via cyber-attack.

1.2 | Categorisation of load profiles

With the growing ubiquity of smart metering, researchers are increasingly investigating how to effectively utilise the data they capture. Load profile categorisation, either via clustering or using other techniques, has also become a popular subfield in the area of smart meter profile analysis. In the past, almost all studies involving smart meter load profiles have focussed on residential smart meter data. This is likely due to the relative availability of data compared with their industrial counterparts. Several works examine consumer data. For example, in Ref. [5] the authors explored a segmentation strategy for households using hourly data. A clustering approach for consumer smart meter data was examined in Ref. [6] and behavioural demand profiles were identified using smart meter data in Ref. [7]. In Ref. [8] a C-Vine Copula mixture model for clustering of residential data is examined. A non-Gaussian residual is used to model intraday forcasting at the feeder level in Ref. [9] while in Ref. [10] novel approaches for load profiling using smart meter data are explored. Again load profiles are used to cluster consumer profiles in Ref. [11]. In Ref. [12] Stephen et al presented several Linear Gaussian (LG) load profiling techniques. These were embedded within a mixture model framework, which allowed multiple behaviours to be considered with the most probable used for categorisation. The prior focus on consumer data is likely attributable to the relative availability of this data in comparison to industrial load flow profiles. However, some works have addressed non-residential flows. For example, in Ref. [13] self-organising maps are used to classify industrial loads. Previous works have also used customer-specific data to create use profiles [14] and analysed industrial electricity consumption with respective to behavioural dynamics [15]. The authors in Ref. [16] introduced a general scheme for analysing load patterns, while an overview of clustering techniques was presented in Ref. [17], which summarises and evaluates methods for load pattern classification. Often, these works stop short of finding a use case for the profiling. In Ref. [18], the authors applied a clustering-based framework for building energy-based benchmarks. Data extracted from smart meter load profiles were used to categorise buildings according to their operational characteristics. In Ref. [19], the load profiles of supermarket chains were predicted using machine learning. In Ref. [20] a novel probabilistic approach was proposed that utilises similar principal components. Hu et al. used interpretable feature extraction to categorise load profiles based on a combination of statistical and temporal features [21]. The authors in Ref. [22] examined load profiling and its applications in relation to demand response. An anomaly detection scheme for big industrial data sets is applied in Ref. [23]. A review of electric load classification in smart grids is available in Ref. [24].

1.3 | Attacks against metering infrastructure

Before the use of smart meter infrastructure, bypassing an electricity meter was a common method of defrauding utility operators. However, from the perspective of utility providers, direct bypass attacks are easy to identify using data driven methods as they are effectively a string of zeroes. In the case of smart meter infrastructure, while some commentators initially believed that smart meters would provide additional security, they have been proven to susceptible to hacking [25]. The available evidence suggests that in the future, smart meter attacks may aim to change the transmitted load curve completely, thereby reducing the cost of power consumption. In the past, these types of attacks have been called False Data Injection (FDI) attacks and have usually been suggested at the transmission layer.

FDI attacks began to gain consideration for power systems in Ref. [26]. Power system style (transmission layer) FDI attacks require the altering of system measurements in a very specific manner, dictated by network topology in order to corrupt a network operator's state estimation process [27]. This corruption can cause blackouts, line outages or indeed even hide outages on the network [28].

FDI attacks by their nature are ineffective if discovered. The principle aim of an FDI attack is to fool the SO about the current network state. As a result, stealthiness is key. In the power system, FDI attackers compete with the SO's state estimator which provides a bad data check. Initially, this meant that FDI attacks on power systems had a high knowledge component. However, later works invalidated this assumption [29–31]. Comprehensive reviews on FDI for power systems of this attack type can be found in Refs. [32, 33]. Indeed, the authors have explored these type of attacks previously in Refs. [34–36] in case studies where FDI attackers alter system measurements to spoof the transmission-level state estimation processes. However, while FDI style attacks on transmission level infrastructure have received significant research attention, limited research has examined the impact of these attacks on distribution-level systems such as smart meter load profiles.

A putative advantage of the FDI approach is that these attacks can utilise distributed and poorly protected measurements rather than attacking a well-defended central system operator. This is especially true for distribution-level attacks against metering infrastructure as these devices are usually decoupled from operational processes and not monitored in real-time by utility providers. Also, modern smart meter infrastructure has also been shown to suffer from several vulnerabilities, which can be exploited by motivated attackers [37]. Some works have explored detection of FDI attacks using predictive or analytic methodologies. In Ref. [38], the authors examine a prediction algorithm to enhance grid resilience with reference to wide area monitoring and control systems. In Ref. [39] a median regression based state estimation approach is taken to protect the system against data-driven cyber-attacks. While Inayat et al examine a learning-based method for cyber-attack detection in Ref. [40]. We also consider that at the distribution level, state estimation processes cannot be relied upon for identifying bad data. Therefore, in this work we consider new data driven approaches for detecting attacks against smart meter load profiles.

1.4 | Contributions

While many papers have addressed the categorisation or clustering of load profile data, few demonstrate the utility that results from this categorisation. In this report, we propose both a methodology for grouping load profile data and also an application for this process within the realms of cyber-attacks. This work introduces an incentive-weighted detection model for cyber-attacks against smart meter infrastructure. Where previous works in the past, have looked at clustering of smart meter load flow data, few have analysed smart meter data sets under attack from an adversary.

The main contributions of this work are as follows:

- To start, the work introduces a new methodology for the clustering of load profile data. This methodology involves a two-step process that incorporates both clustering and silhouette scoring to establish a set of base models within each industry type. We use 20 features for this approach, which include a combination of global statistical, index and quartile statistical features.
- We use these average cluster groups to produce a scoring model for new inbound datasets. This scoring model incorporates both model departure and spot prices to present an incentive-weighted model of fraud detection in load profiles. We apply, for the first time, an ’incentive-weighted’ approach to anomaly detection. This incentive-weighted detection uses the weighted spot price to identify when attackers maybe trying to change profiles for financial gain as well as the core anomaly model itself.
- We then introduce several adversary models based around FDI attacks against smart meter infrastructure. We introduce these adversary models into our dataset and evaluate our anomaly detection approach on them.
- Finally, we develop our model, using real load profile data and real-life spot price data (not simulated data) to ensure our model is viable in real-world data-sets. This dataset is substantial (profiles for 12,055 businesses) and representative of actual network operation.

The next section presents the base model methodology used to build the average cluster models.

2 | BASE MODEL METHODOLOGY

2.1 | Input data

The input data were obtained as part of the Energy Demand Research Project (EDRP) and consists of industrial load flow profiles for 12,055 businesses operating over a 2-year period. The EDRP aims to understand and model how load user flexibility changes as consumers develop an awareness and understanding of their energy consumption. Within these businesses, we categorised business data, and took a subset of the businesses under the branch of consumer entertainment industrial parks. The reason we opted for this subset is that these businesses offer distinct business models that are easily interpretable, at a conceptual level, to the average user. The profiles consist of 48 consumption periods, with each period corresponding to 30-min power consumption windows within a given 24 h day. Within this, we focus on summer profile data sets (June through September) to maintain consistency in the underlying data. We believe it important to operate with a real rather than similar data as this helps validate the model in a realistic environment.

2.2 | Data pre-processing

In this subsection we outline the process of data cleaning, preparation and normalisation we have taken prior to implementing the anomaly detection. We deemed normalisation of the data set important as while some businesses may share similar relative statistical properties, the magnitude of energy consumption within business of the same type may vary considerably. When building our groupings, we intend to identify businesses based on the shape of operation and relative properties rather than straight magnitude. Therefore, for each individual business, we perform a max-min normalisation of the load profile data using the following equation:

$$z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \tag{1}$$

In this max-min normalisation equation, $x = (x_1, \ldots, x_n)$ represents an array of length equal to the number of load consumption measurements for a given business line. The normalisation enables us to capture departures from expected operation demand curves. Simple changes in consumption magnitude (such as a bypass attack) are typically easy to identify via conventional means, and so we focus on relative model departure. We also apply a ‘low touch’ data cleaning strategy, which aims to remove corrupted or incomplete data to the greatest possible extent, while minimising the discarded data. It is often tempting to be overzealous when cleaning data, but as we are working with a real-world data sample we wish for our model to incorporate as much actual data as possible.

2.3 | Feature-based clustering

We use feature-based clustering to establish the base models for our anomaly detection and incentive-weighted anomaly detection system. We use a set of 20 different features consisting of global statistical features, quartile statistical features, and index-based features. Table 1 summarises the features used in our clustering algorithm. The approach employed is similar to the one outlined in Ref. [21], with the exception that we also incorporate quartile statistical values.

2.4 | Hierarchical clustering

This subsection outlines the proposed combination of clustering and scoring used to establish the average profile groupings and cluster numbers. After data pre-processing, we perform agglomerative hierarchical clustering on the respective industrial load business types. Hierarchical clustering is also known as AGNES (agglomerative nesting) and refers to a bottom-up approach to clustering wherein each observation starts in a cluster on its own and clusters are slowly merged. The steps involved in AGNES are as follows:

TABLE 1 List of features used in clustering model.

Feature No.	Feature description	Feature type
G1	Mean	Global
G2	Standard deviation	Global
G3	Max	Global
G4	Min	Global
G5	Range	Global
G6	Sum	Global
G7	Skew	Global
G8	Kurtosis	Global
Q9	Sum 1–12	Quartile
Q10	Sum 12–24	Quartile
Q11	Sum 24–36	Quartile
Q12	Sum 36–48	Quartile
Q13	Standard deviation 1–12	Quartile
Q14	Standard deviation 12–24	Quartile
Q15	Standard deviation 24–36	Quartile
Q16	Standard deviation 36–48	Quartile
I17	Max time period	Index
I18	Min time period	Index
I19	Index > Mean	Index
I20	Index < Mean	Index

1. The proximity matrix for each point within the dataset is calculated.
2. The algorithm then considers each element as a cluster consisting of a single element cluster.
3. The two closest clusters are merged and the new proximity matrix is recalculated for the dataset.
4. Steps 1–3 are then repeated until the desired number of clusters is reached.

As we are using an unsupervised learning approach, we opted for this methodology to avoid manually inputting a cluster number for the algorithm to use and utilise an automated approach. Therefore, we incorporate an automatic cluster number selection feature, which utilises silhouette scoring.

2.5 | Silhouette coefficient

The silhouette coefficient is a method of quality checking and validating cluster consistency within groups. The coefficient measures the similarity of an object with respect to its given cluster. Each data point within a series is assigned a silhouette value. This silhouette value of an individual data point is given by the following:

$$s(\mathbf{z}) = \frac{b(z) - a(z)}{\max(a(z), b(z))}, \tag{2}$$

where $s(z)$ is the silhouette score for a given data point z, and $b(z)$ is the average minimum distance between z and the clusters that z is not located within and $a(z)$ is the average distance between z and all the other data points with the cluster z is located within.

The silhouette coefficient is then given by finding the maximum value of the mean silhouette score for a given number of clusters k such that

$$SC = \max_{k} \bar{s}(k), \tag{3}$$

where $\bar{s}(k)$ is the mean silhouette score across the entire dataset.

In this work, we employ a short loop. This compares the silhouette coefficient under different cluster numbers (up to 5) and selects for the maximum coefficient value. In turn, this is used to define the number of clusters involved in hierarchical clustering.

3 | INCENTIVE-BASED ANOMALY TRACKING

Attackers will usually have a reason for an attack dictated by an aim or goal. These goals are often financial. Historically, anomaly tracking in energy systems has been based on the analysis of simple departures from expected models. However, within the context of cyber-attacks, model departure is only an indication and not a guarantee of foul play. We consider that departure from a model is not merely an indication of a cyber-attack. Anomalous measurements do occur amid routine operation. We also consider that in a scenario considering financial or cost lowering attack, the attacker is unlikely to inject an attack vector which will increase his overall cost. Therefore, we can use considerations about the attack vector incentive as a method of reducing false positives. In this way we create an 'incentive-based' anomaly detection which considers the cash incentive of the attack as well as the direct anomaly.

3.1 | Detection model

Here we outline the detection model for the incentive-weighted anomaly detection. The steps involved are as follows:

1. The outlined hierarchical and silhouette scoring clustering model are leveraged to model expected behaviours in load profiles for respective industry types.
2. Unexpected departures from the underlying models in new inbound data are identified for the respective company groups. Also, a score is created based on how different these groups are, which is referred to as the violation scoring.
3. We then use the weighted average spot price to produce an incentive based scoring model which indicates whether a profile is financially preferable.
4. The scores are then combined to establish the incentive-weighted violation score to identify potential FDI attacks based on model difference and potential financial gain.

3.2 | Violation score

We consider a metric, dubbed a violation score, used to assess how different a new incoming dataset is compared to the previous model. The violation score is based on how often these inbound measurements violate a confidence interval of 2 standard deviations when compared to the average model for the business group. We start with the following equation:

$$\mathbf{VSD}^{n} = 2\mathbf{ASD} - \left\| \left(\mathbf{z}^{n}_{new} - \mathbf{AC} \right) \right\|, \tag{4}$$

where $\mathbf{VSD}^{p}$ is an array of length t that contains the respective violation decisions for a given consumption interval, $\mathbf{AC}$ is an array of length t representing the average cluster profile, $\mathbf{ASD}$ is an array of length t representing the standard deviations for the respective consumption periods, and $\mathbf{z}_{new}$ is an array of length t that represents the new measurement set which is being checked. Also, n refers to the number of days in the set. A violation is recorded if VS is a negative value such that

$$VSC^{n}_{t} = \begin{cases} 1 & \text{if } VSD^{n}_{t} < 0 \\ 0 & \text{if } VSD^{n}_{t} > 0 \end{cases} \tag{5}$$

In turn, this is presented as a percentage of the number of periods recorded:

$$VSP = \sum_{n=1}^{n} \sum_{t=1}^{t} \left(\frac{VSC^{n}_{t}}{nt} \right), \tag{6}$$

where VSP is the violation score. This score gives an initial indication as to whether there is a significant departure from the previously established cluster groups. A high violation score indicates that the model varies significantly from the average cluster model established by the clustering algorithm. We consider that a simple departure from the underlying model is not necessarily an indication of foul play and that we must also consider the impact of an attack.

3.3 | Incentive score

We consider incentive as a product of the relative gain that a change from the average profile gives to a customer. To do this, we incorporate the weighted average spot price WSP versus flat price to identify regions where there may be an incentive to change the input profile. The weighted spot price array is calculated as below:

$$\mathbf{WSP} = \frac{(\mathbf{CSP} - \mathbf{FP})}{FP}, \tag{7}$$

where $\mathbf{WSP}$ is an array of length t that represents the number of prices in the period (in this case, 48 half-hourly periods), and $\mathbf{FP}$ is an array of length t consisting of the flat price FP. The weighted spot price is then used to score the incentives given by the departure from the model, such that

$$ISC = \left\| \mathbf{z}^{n}_{new} - \mathbf{AC} \cdot \mathbf{WSP} \right\|, \tag{8}$$

where ISC is the incentive score, $\mathbf{z}_{new}$ is an array of normalised load profile measurements of length t. This gives an initial indication as to whether there is a significant departure from the underlying model. A high violation score indicates that the model varies significantly from the average cluster model established by the clustering algorithm.

3.4 | Incentive-weighted violation score

Finally, we consider the weighted incentive-based violation score *WIVS* as a simple product of the violation score and the incentive score such that

$$WIVS = ISC \cdot VSP \tag{9}$$

This yields a simple metric for each given business, with which it is possible to assess the likelihood of financial fraud. In the following section, these metrics are tested with existing business types and FDI profile sets to verify the effectiveness of the approach.

3.5 | Algorithm overview

The algorithm implementation is outlined in Figure 1. To surmise, the detection methodology takes 2 sets of inputs namely the consumption data and spot prices. One of these inputs is the consumer consumption data. The consumption data is used to create base models of which future consumption are compared. The spot pricing is used to create incentive models whereby it can be confirmed if a new inbound data set looks profitable for a consumer. Both of these are combined to create the incentive-weighted detection model as an output. This model allows us identify those profiles which are both anomalous and potentially valuable to an attacker.

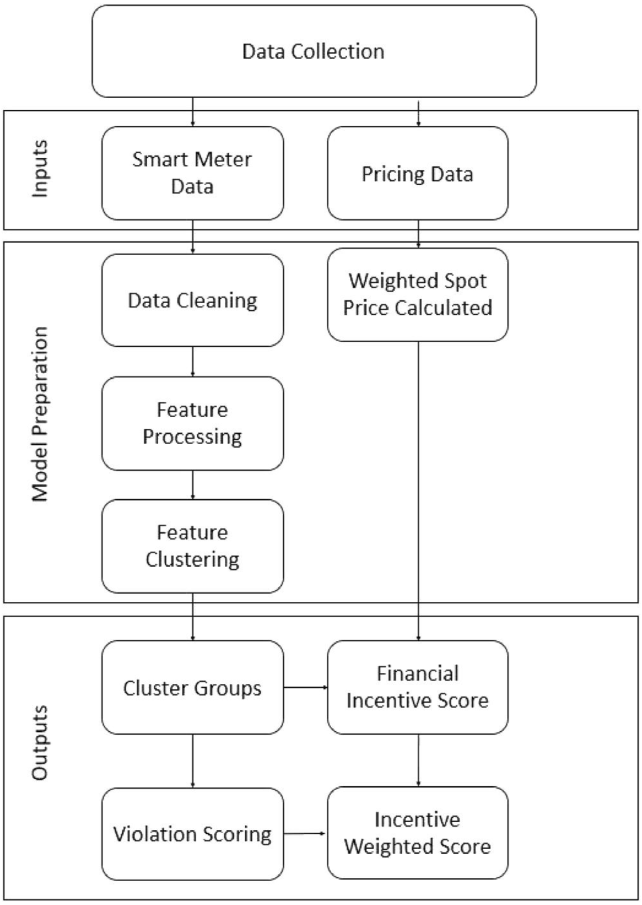

FIGURE 1 Overview of the incentive weighted detection algorithm.

3.6 | Approach rationale

This particular approach offers many advantages against other methodologies. One of the main advantages of hierarchical clustering is the ease of understanding. Hierarchical clustering provides significant explainability of the underlying groups. The ease of understanding, implementation and delivering data outputs were fundamental in our decision to use the algorithm. The relative data accuracy of hierarchical clustering was also a consideration. We also consider that the unsupervised nature is desirable for the application as the algorithm was required to run semi-autonomously without the need for arbitrary cluster selection from the user (which would be the case in other popular algorithms such as K-means). It is true that hierarchical clustering has drawbacks. The main one being relative time complexity. However, we consider that the time complexity while not ideal was acceptable for the purposes outlined in this report Ref. [41]. We also note, that this methodology could be later augmented to bring complexity down to $O\left(n^2\right)$ under certain circumstances. However we note that while running this method on an Intel Core i7-7820X CPU with 64 GB of ram running a Windows 10 system the entire data processing and results took less than a minute to process. We also note these time complexity issues are common with other alternative approaches and significant reductions are usually only viable in specific use cases. There are dimension reduction alternatives such as PCA clustering [42] which might allow for additional time savings by reducing the 48 point dataset into 2 or 3 dimensions. However, the flip-side of this is reduction in data accuracy as we reduce a data of 48 points into 2 or 3. This would reduce the overall effectiveness of the anomaly detection approach while offering only marginal benefit when dealing with a dataset of this size. The combination of global statistical, quartile statistical elements and index elements was crucial to developing an effective clustering model and this precluded many alternative models. Hence hierarchical clustering which allowed us to have a high numbers of dimensions for our feature selection which was essential.

4 | RESULTS

In this section we outline the results for the clustering technique, average profiles and the incentive-weighted detection algorithm. We initially show the results of the average profile model and then go onto examine the results of the anomaly detection algorithm on a combination of future load profile data and injected red team profiles.

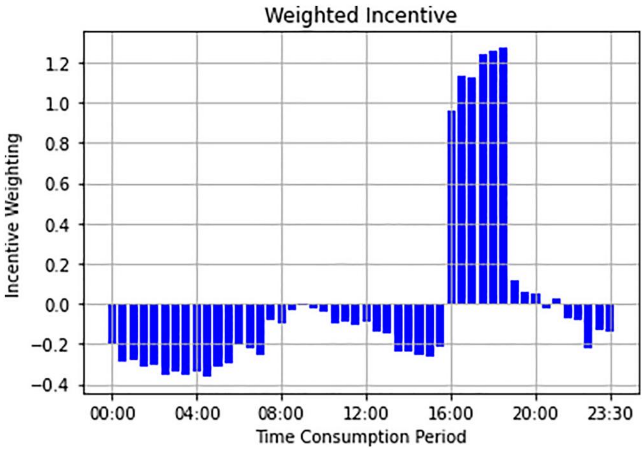

FIGURE 2 Incentive weighting for each consumption period.

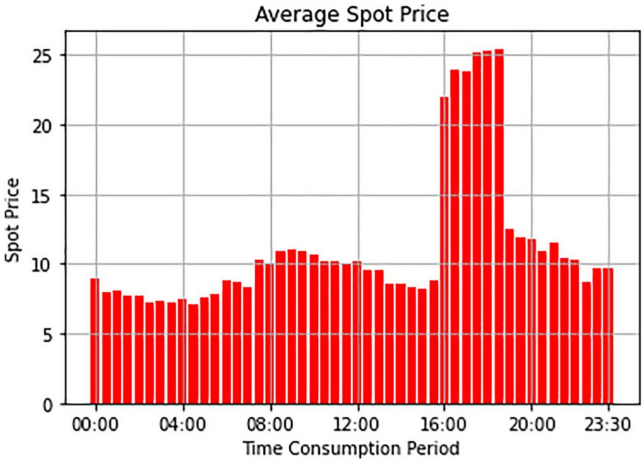

FIGURE 3 Average spot price for each consumption period.

4.1 | Weighted spot price calculation

For the weighted spot price calculation, we used data provided by Octopus energy prices. As model building relies on using summer data, we employ equivalent summer data to build our weighted spot price. The Octopus data is split into 14 sub-regions accounting for regions within the UK, such as London, East England, and Midlands. We note that despite these regions being split into groups, the level of inter-regional price volatility is low. For simplicity, we take a simple mean average of all these regions combined, which is used as the basis for our weighted spot price. This average spot price per consumption period is shown in Figures 2 and 3.

4.2 | False data injection profiles

We introduce two FDI style profiles into the load datasets. One profile is a simple meter bypass, which is typical of the current state of play in physical attacks; in meter bypass, the profile is simply replaced with a set of zeroes representing no load. We also introduce a more sophisticated reduced-cost spot attack (RCSA) profile. This RCSA profile represents an attacker attempting to create a non-zero load profile to attain significant reductions via the spot price. The RCSA profile is shown in Figure 4.

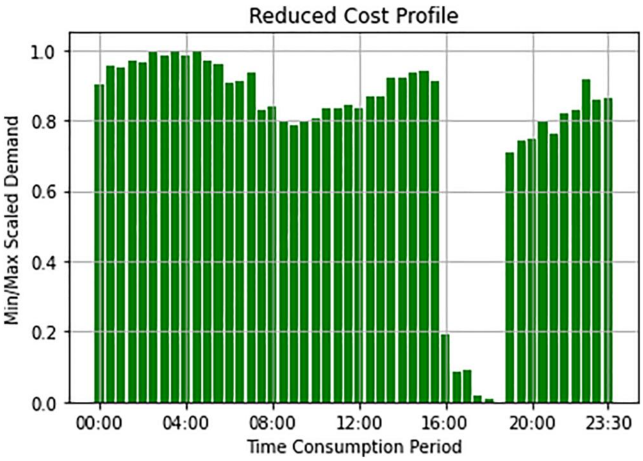

FIGURE 4 Reduced cost red-team profile for each consumption period.

4.3 | Industrial cluster groups

Figure 5 shows the individual measurement sets for the cluster groupings, while Figure 6 presents the average corresponding model. These models are built using the 2009 summer data set. We note that although there are various industrial business models, we often see a trend towards a limited number of common models not dissimilar to the typical consumer load. We also illustrate this in 4 which shows the time dynamics and relative density of the respective profiles in 3D (Figure 7).

In Cinema cluster 1, we observe a consumption peak at approximately 12:00, which falls off after around 16:00. This shape is somewhat unusual as we might expect typical a cinema business to continue into the late evening. However, it may be necessary to distinguish between 'mom-and-pop' style cinemas, which might shut relatively early, and large cinema corporations that run into the night. Cinema 2 resembles a more typical consumer load flow profile, with a broad consumption peak running from 08:00–20:00. These relationships can also be seen via the histograms in Figure 7.

The most atypical business grouping in the dataset is Club cluster 1. Where the other businesses have the expected peaks around the common spot peak consumption times, Club cluster 1 exhibits a later peak at around the 20:00 mark, which gradually increases into the night. This is consistent with the nature of business, given that the main hours of operation for nightclubs are during the night.

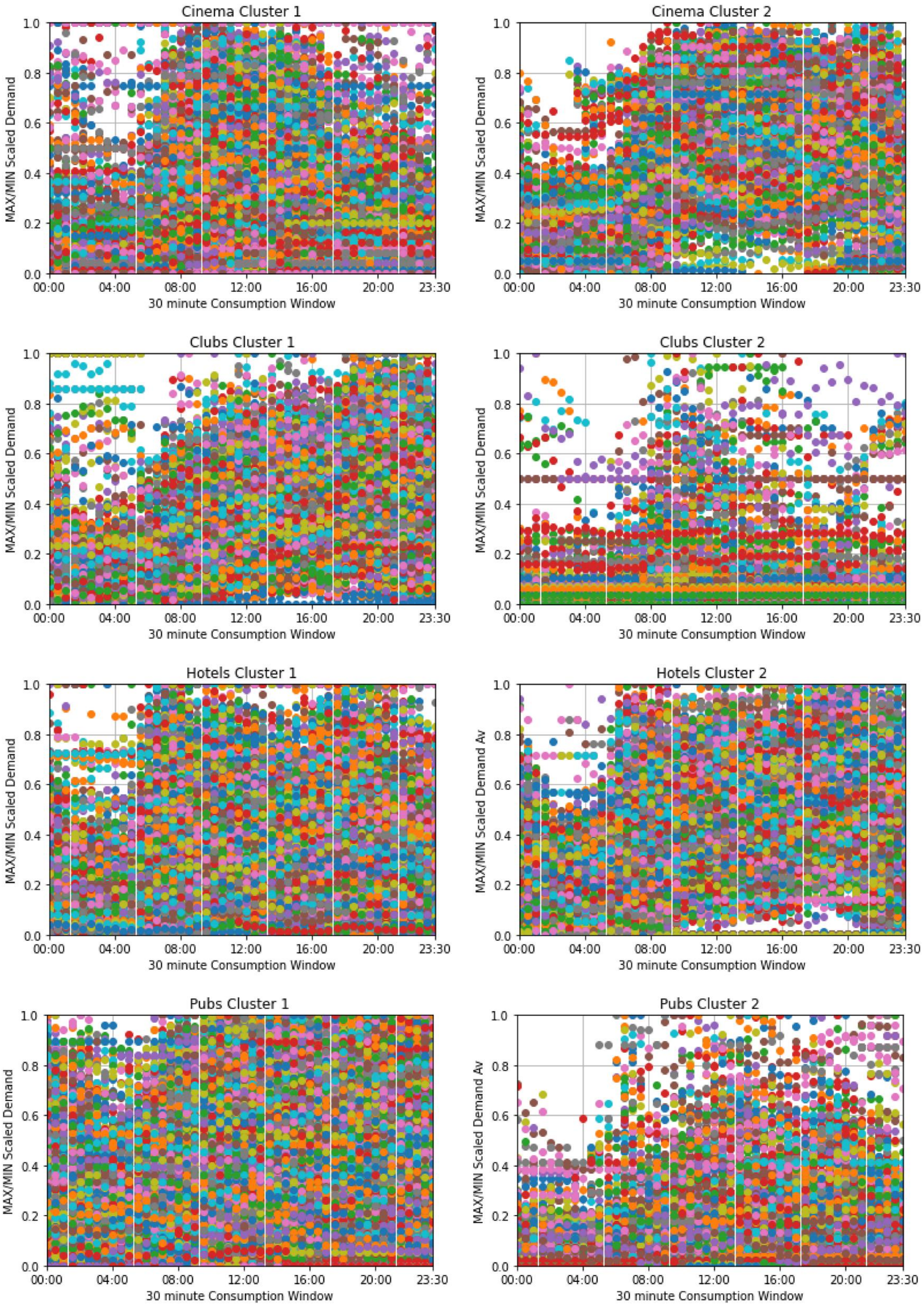

FIGURE 5 All data points for respective industry clusters. Based on weekend profile data during summer 2009.

Hotel cluster 1 exhibits a notable peak at 06:00, which is potentially attributable to breakfast preparations. Curiously, a similar peak does not occur for lunch, but we do see it for the dinner menu at approximately 16:00. Clubs cluster 2 is also atypical in that it has a low range between the midday peaks and the overnight operation. Similar to other groups Hotels

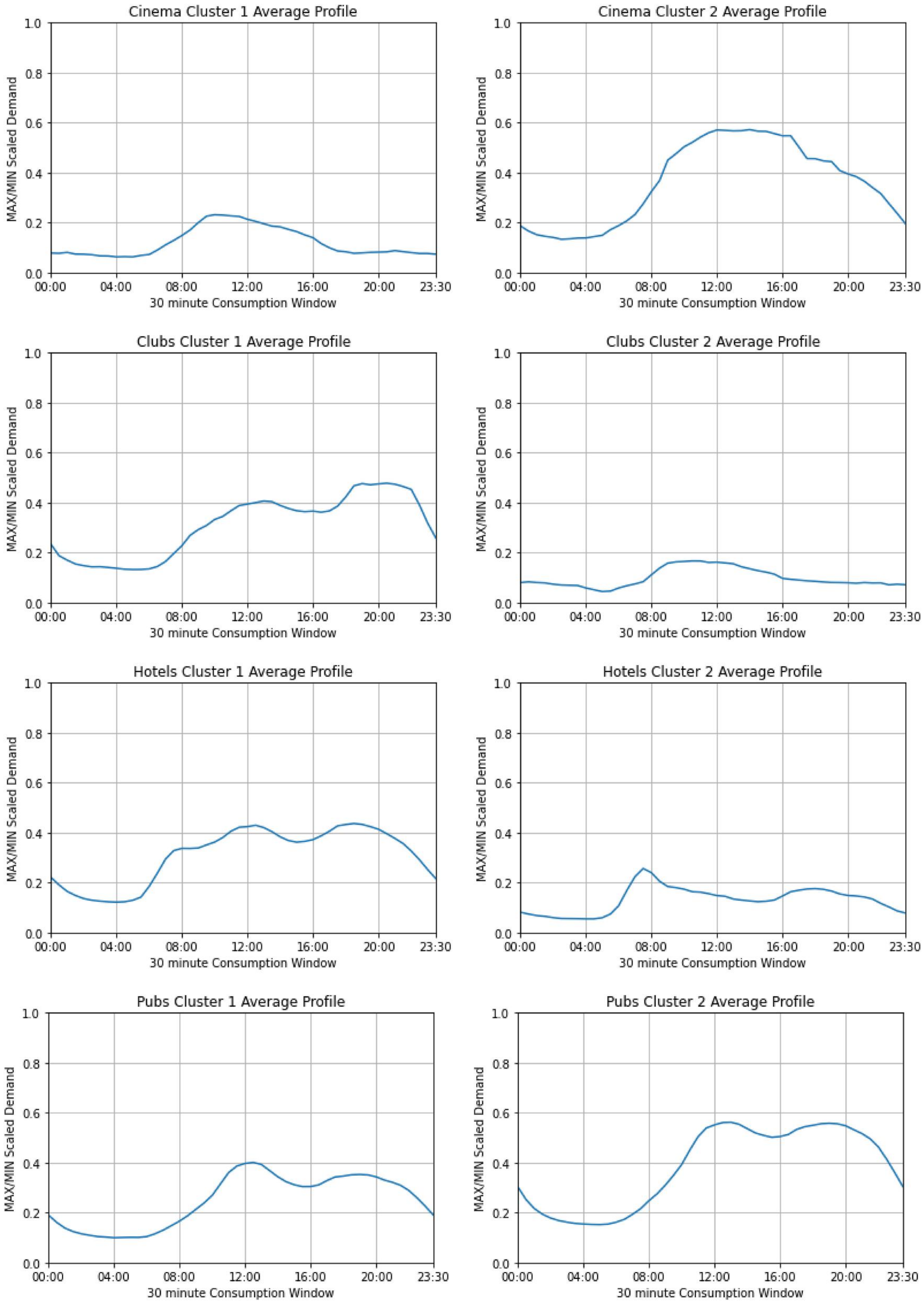

FIGURE 6 Average cluster based on weekend profile data during summer 2009.

cluster 2 and Pubs cluster 1 exhibit similar patterns to a residential consumer load profile. However, we note that, generally, industrial profiles have broader operation periods, which is reflected in the peak and intraday consumption windows.

4.4 | False profile detection

The performance of the detection model is shown in Figures 8,9,10-12, indicating the violation score, incentive weight,

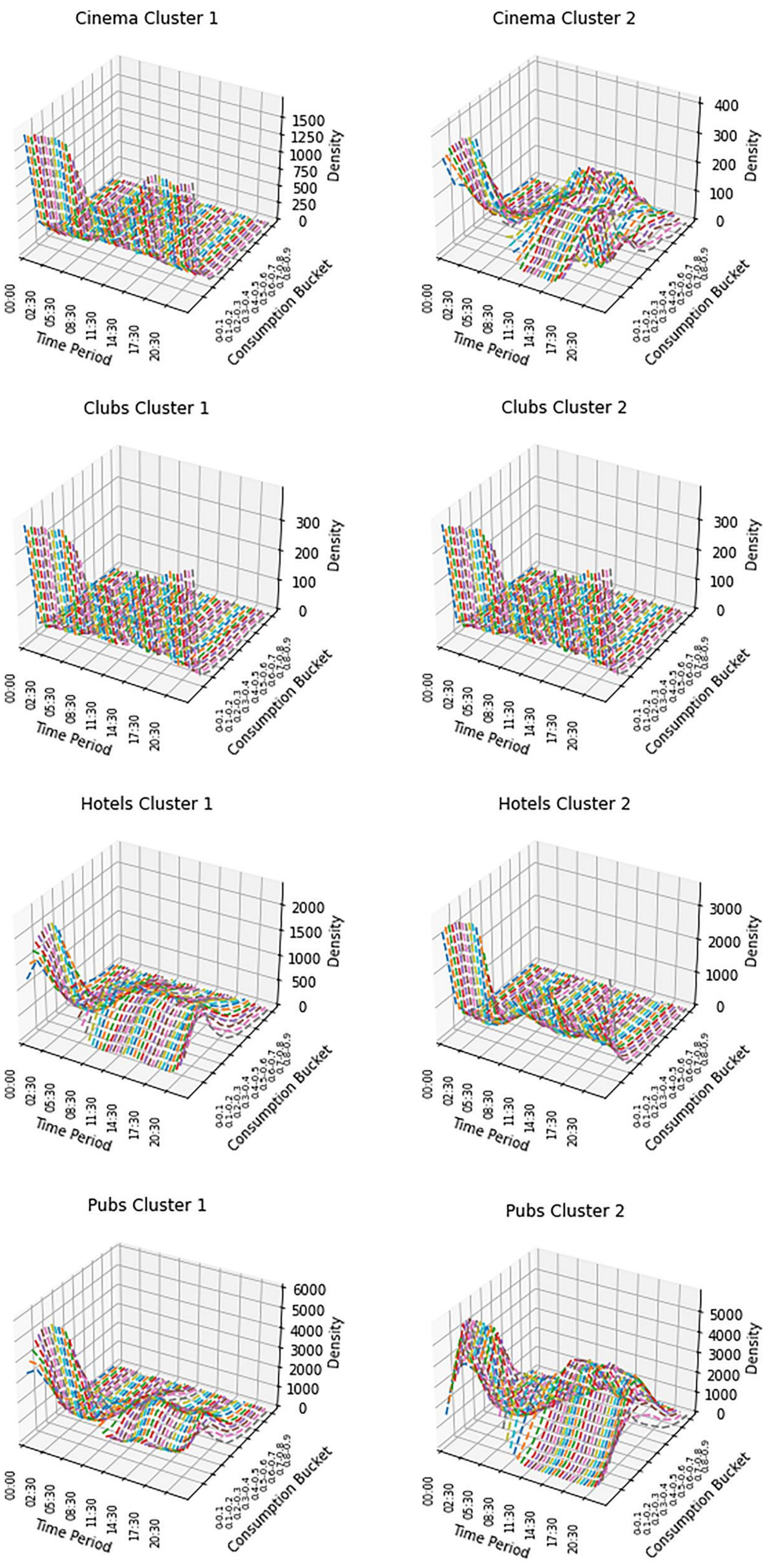

FIGURE 7 3d histograms of all data points for respective industry clusters. Based on weekend profile data during summer 2009.

and incentive-weighted violation score for both identified hotel cluster types. The initial cluster models for each respective business were built using the 2009 summer weekend dataset through a combined clustering and scoring approach. They are then cross-compared with the 2010 summer weekend dataset using the anomaly detection technique. In each of these figures, the last two data points are the bypass vector and RCSA attack, respectively. Consistent detection levels were observed for the RCSA attack for all three types of detection. Incentive scoring adds a heavy weighting to detection for the RCSA attack. For

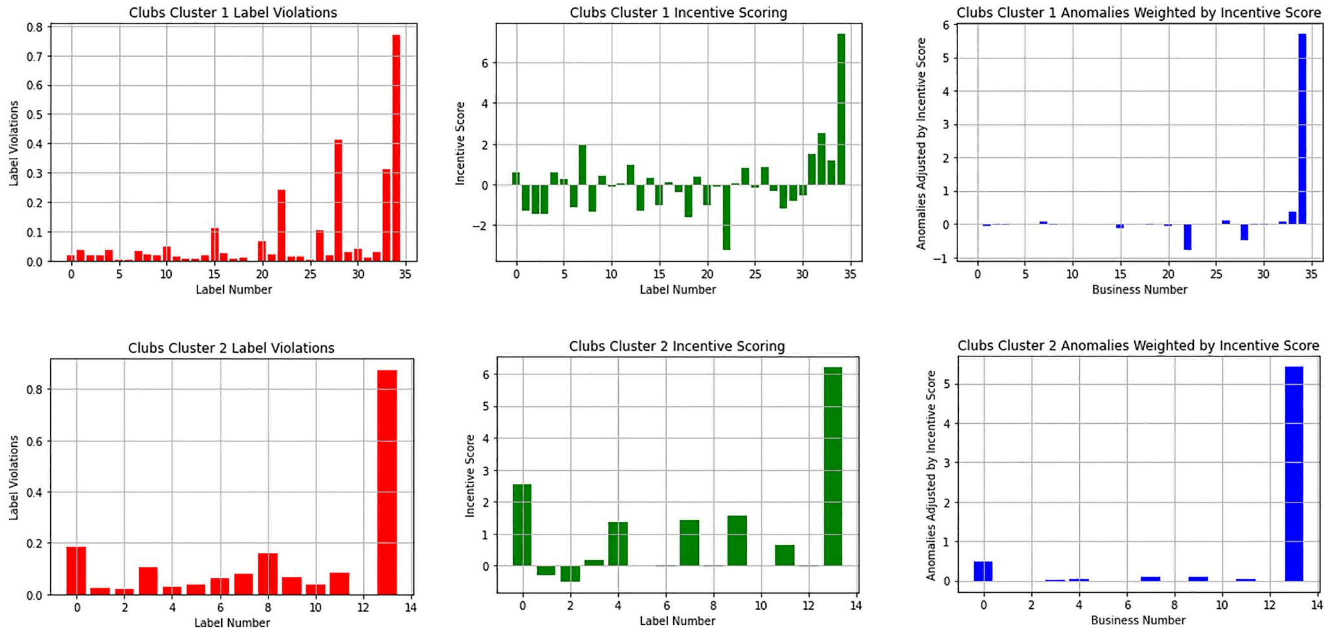

FIGURE 8 Clubs label violations, incentive score, and weighted incentive score.

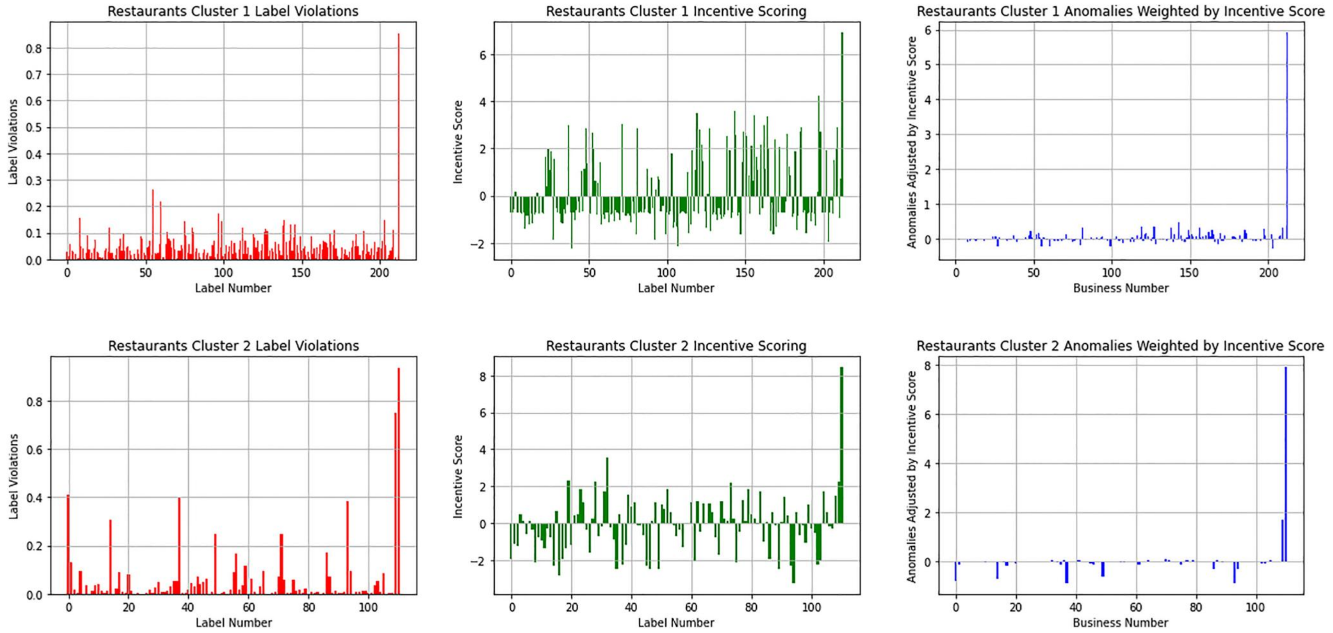

FIGURE 9 Restaurants label violations, incentive score, and weighted incentive score.

both clusters 1 & 2, incentive-weighted detection for the RCSA attack was clear and present. We note a similar result in Figures 9,11,12, with consistent incentive-weighted detection for the RCSA style of attack. However, we note that the incentive-weighted detection is not as effective for the bypass attack. The bypassed data is not easily identified using the incentive-weighted approach. However, there are already several methods for identifying a bypass style string of zeroes. Indeed, we do see generally higher than average scoring for the bypass vector in some cluster groups (e.g. clubs cluster 1). Generally, however, this method of scoring for this type of attack is undermined due to incentive weighting. As the bypass has no clear incentive weighting, this reduces the impact.

5 | CONCLUSION & FUTURE WORK

Smart meters are a weakly defended, distributed infrastructure that represent an easy attacking opportunity for a cyber-attacker. Load profile altering attacks in spot price

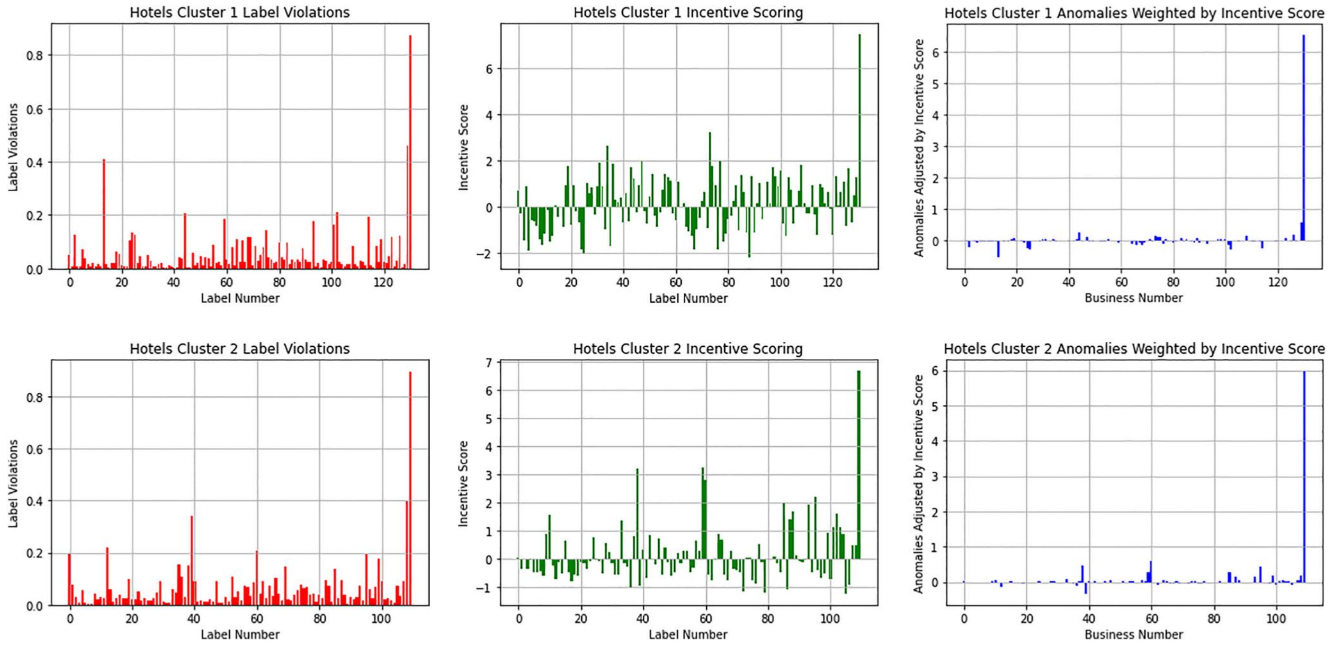

FIGURE 10 Hotel label violations, incentive score, and weighted incentive score.

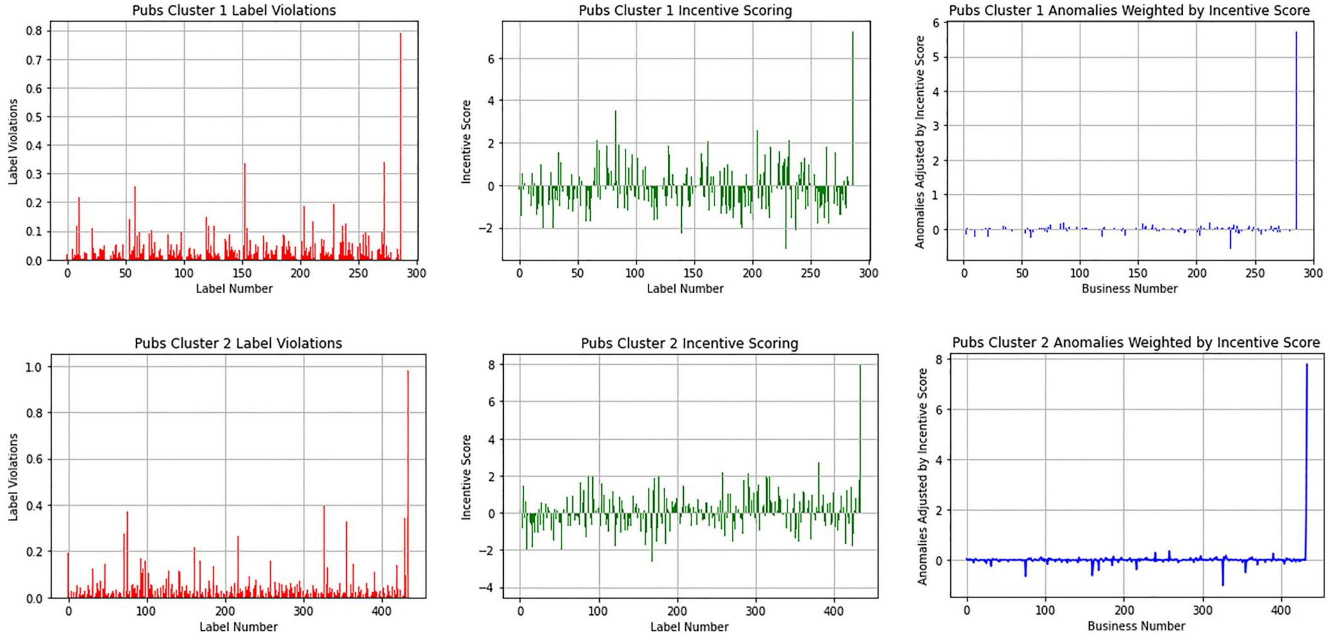

FIGURE 11 Pubs label violations, incentive score, and weighted incentive score.

markets can provide a highly lucrative financial incentive for attackers with currently few methodologies to detect this threat.

In this paper, we have examined an incentive weighted detection model for FDI style attacks against load-profile datasets. Through feature-based clustering, we examined different groupings within industrial load profiles and created an incentive-weighted detection methodology to examine potential fraud. This incentive-weighted methodology incorporates spot prices to identify when an attacker might be taking advantage of the spot curve. In short, this work has investigated how to improve corporate fraud detection in smart data through clustering and an incentive-weighted detection approach.

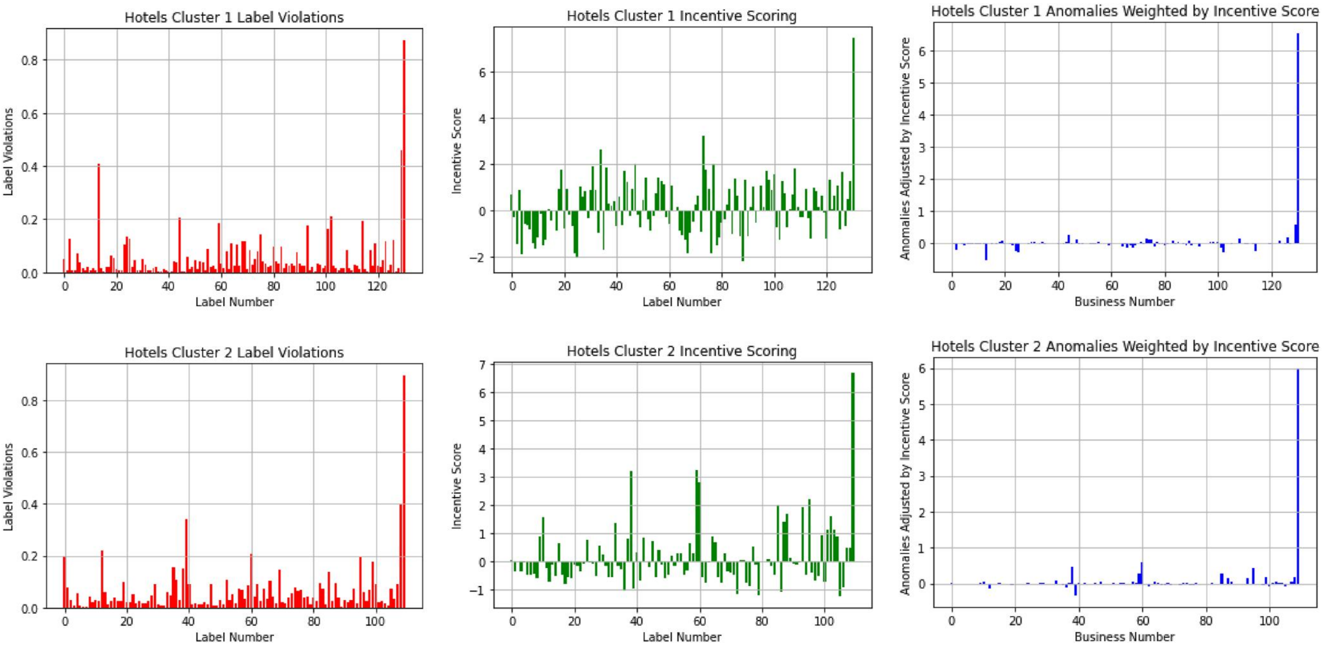

FIGURE 12 Cinema label violations, incentive score, and weighted incentive score.

In the first contribution of this paper, we examined how to establish a combination of hierarchical clustering and silhouette scoring base models for industrial load profiles. We incorporated real-life datasets and, for illustrative purposes, analysed businesses from entertainment-style industrial parks, due to the general familiarity of the nature of these businesses.

We enhance this anomaly detection model using an incentive-weighted violation approach. This approach incorporates both spot pricing and the level of departure from the expected model to detect attacks which try to gain financial advantage from the spot curve.

To analyse the effectiveness of our model under attack, we injected fraudulent profiles into the datasets. These were based on two likely red-team cases: first, a bypass profile representing a common 'direct bypass' of the metering infrastructure; and second, an RCSA profile, which attempts to exploit variations in price resulting from spot pricing. We found our model had consistently good detection rates.

In future work, dis-aggregation of the modern additions to the distribution network which influence load profiles which will be useful 'future proofing' of the model for contemporary power systems. These could include solar panels, heat pumps, and storage devices. It would be worthwhile to understand how these might impact the detection of attacks against load profiles.

NOMENCLATURE

a	average minimum distance between z and other data points
AC	average cluster profile
ASD	array of consumption period standard deviations
b	average minimum distance between z and clusters
CSP	array of average spot prices
FP	array of average flat price
ISC	incentive score
s_z	silhouette score for given point
SC	silhouette coefficient
VSC	the violation score for a given point
VSD	the violation decision value for a given consumption point
VSP	the violation score
VSP	array of violation decisions
WIVS	incentive weighted violation score
WSP	array of spot prices incentive weighted by flat price
x_i	load consumption measurement
z	array of length t normalised load consumption point
z_i	normalised load consumption point

AUTHOR CONTRIBUTIONS

Martin Higgins: Conceptualization; Data curation; Software; Writing – original draft. **Bruce Stephen**: Funding acquisition; Investigation; Project administration; Supervision; Writing – review & editing. **David Wallom**: Conceptualization; Funding acquisition; Investigation; Project administration; Supervision; Writing – review & editing.

ACKNOWLEDGEMENTS

This work was performed as part of the Analytical Middleware for Informed Distribution Networks (AMIDiNe) project under EPSRC grant EP/S030131/1.

CONFLICT OF INTEREST STATEMENT

None.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Martin Higgins https://orcid.org/0000-0002-1816-8333

REFERENCES

1. Gelazanskas, L., Gamage, K.A.A.: Demand Side Management in Smart Grid: A Review and Proposals for Future Direction (2014).
2. Aduda, K.O., et al.: Demand side flexibility: potentials and building performance implications. Sustain. Cities Soc. 22, 146–163 (2016). https://doi.org/10.1016/j.scs.2016.02.011
3. Garcia, E.V., Runnels, J.E.: The utility perspective of spot pricing. In: IEEE Transactions on Power Apparatus and Systems, 1985, PAS-104, pp. 1391–1393.
4. De.Faria, R.A., et al.: Collusion and fraud detection on electronic energy meters - a use case of forensics investigation procedures. In: Proceedings - IEEE Symposium on Security and Privacy, pp. 65–68. Institute of Electrical and Electronics Engineers Inc. (2014).
5. Kwac, J., Flora, J., Rajagopal, R.: Household energy consumption segmentation using hourly data. IEEE Trans. Smart Grid 5(1), 420–430 (2014). https://doi.org/10.1109/tsg.2013.2278477
6. McLoughlin, F., Duffy, A., Conlon, M.: A clustering approach to domestic electricity load profile characterisation using smart metering data. Appl. Energy 141, 190–199 (2015). https://doi.org/10.1016/j.apenergy.2014.12.039
7. Haben, S., Singleton, C., Grindrod, P.: Analysis and clustering of residential customers energy behavioral demand using smart meter data. IEEE Trans. Smart Grid 7(1), 136–144 (2016). https://doi.org/10.1109/tsg.2015.2409786
8. Sun, M., Konstantelos, I., Strbac, G.: C-vine Copula mixture model for clustering of residential electrical load pattern data. IEEE Trans. Power Syst. 32(3), 2382–2393 (2017). https://doi.org/10.1109/tpwrs.2016.2614366
9. Bruce, S., Telford, R., Galloway, S.: Non-Gaussian residual based short term load forecast adjustment for distribution feeders. IEEE Access 8, 10731–10741 (2020). https://doi.org/10.1109/access.2020.2965320
10. Khan, Z.A., Jayaweera, D., AlvarezAlvarado, M.S.: A novel approach for load profiling in smart power grids using smart meter data. Elec. Power Syst. Res. 165, 191–198 (2018). https://doi.org/10.1016/j.epsr.2018.09.013
11. Tureczek, A., Nielsen, P.S., Madsen, H.: Electricity consumption clustering using smart meter data. Energies 11(4), 859 (2018). https://doi.org/10.3390/en11040859
12. Stephen, B., et al.: Enhanced load profiling for residential network customers. IEEE Trans. Power Deliv. 29(1), 88–96 (2014). https://doi.org/10.1109/tpwrd.2013.2287032
13. Verdú, S.V., et al.: Classification, filtering, and identification of electrical customer load patterns through the use of self-organizing maps. IEEE Trans. Power Syst. 21(4), 1672–1682 (2006). https://doi.org/10.1109/tpwrs.2006.881133
14. Räsänen, T., et al.: Data-based method for creating electricity use load profiles using large amount of customer-specific hourly measured electricity use data. Appl. Energy 87(11), 3538–3545 (2010). https://doi.org/10.1016/j.apenergy.2010.05.015
15. Wang, Y., et al.: Clustering of electricity consumption behavior dynamics toward big data applications. IEEE Trans. Smart Grid 7(5), 2437–2447 (2016). https://doi.org/10.1109/tsg.2016.2548565
16. Chicco, G.: Overview and performance assessment of the clustering methods for electrical load pattern grouping. Energy 42(1), 68–80 (2012). https://doi.org/10.1016/j.energy.2011.12.031
17. Chicco, G., et al.: Load pattern-based classification of electricity customers. IEEE Trans. Power Syst. 19(2), 1232–1239 (2004). https://doi.org/10.1109/tpwrs.2004.826810
18. Zhan, S., et al.: Building categorization revisited: a clustering-based approach to using smart meter data for building energy benchmarking. Appl. Energy 269, 269 (2020). https://doi.org/10.1016/j.apenergy.2020.114920
19. Granell, R., et al.: Predicting electricity demand profiles of new supermarkets using machine learning. Energy Build. 234, 234 (2021). https://doi.org/10.1016/j.enbuild.2020.110635
20. Elnozahy, M.S., Salama, M.M.A., Seethapathy, R.: A probabilistic load modelling approach using clustering algorithms. IEEE Power Energy Society General Meeting, 19 (2013)
21. Hu, M., et al.: Classification and characterization of intra-day load curves of PV and non-PV households using interpretable feature extraction and feature-based clustering. Sustain. Cities Soc. 75, 75 (2021). https://doi.org/10.1016/j.scs.2021.103380
22. Wang, Y., et al.: Load profiling and its application to demand response: a review. Tsinghua Sci. Technol. 2, 117–129 (2015). https://doi.org/10.1109/tst.2015.7085625
23. Caithness, N., Wallom, D.: Anomaly detection for industrial big data. In: 7th International Conference on Data Science, Technology and Applications - DATA, vol. 1644, pp. 97–104. American Institute of Physics Inc. (2018).
24. Zhou, K.L., Yang, S.L., Shen, C.: A review of electric load classification in smart grid environment. Renew. Sustain. Energy Rev. 11(24), 103–110 (2013). https://doi.org/10.1016/j.rser.2013.03.023
25. Tangsunantham, N., et al.: Experimental performance analysis of current bypass anti-tampering in smart energy meters. In: 2013 Australasian Telecommunication Networks and Applications Conference, ATNAC 2013, pp. 124–129. IEEE Computer Society (2013).
26. Liu, Y., Ning, P., Reiter, M.K.: False data injection attacks against state estimation in electric power grids. In: ACM Transactions on Information and System Security, vol. 14 (2011)
27. Xu, W., et al.: Blending Data and Physics against False Data Injection Attack: An Event-Triggered Moving Target Defence Approach (2022). http://arxiv.org/abs/2204.12970
28. Liu, X., et al.: Masking transmission line outages via false data injection attacks. IEEE Trans. Inf. Forensics Secur. 11(7), 1592–1602 (2016). https://doi.org/10.1109/tifs.2016.2542061
29. Rahman, M.A., MohsenianRad, H.: False data injection attacks with incomplete information against smart power grids. In: GLOBECOM - IEEE Global Telecommunications Conference, pp. 3153–3158 (2012).
30. Esmalifalak, M., et al.: Stealth false data injection using independent component analysis in smart grid. In: 2011 IEEE International Conference on Smart Grid Communications, SmartGridComm 2011, pp. 244–248 (2011).
31. Deng, R., Liang, H.: False data injection attacks with limited susceptance information and new countermeasures in smart grid. IEEE Trans. Ind. Inf. 15(3), 1619–1628 (2019). https://doi.org/10.1109/tii.2018.2863256
32. Deng, R., et al.: False data injection on state estimation in power systems-attacks, impacts, and defense: a survey. IEEE Trans. Ind. Inf. 13(2), 411–423 (2017). https://doi.org/10.1109/tii.2016.2614396
33. Wang, Q., et al.: Review of the false data injection attack against the cyber-physical power system. IET Cyber-Physical Systems: Theory and Applications 4(2), 101–107 (2018). https://doi.org/10.1049/iet-cps.2018.5022
34. Higgins, M., Mayes, K., Teng, F.: Enhanced cyber-physical security using attack-resistant cyber nodes and event-triggered moving target defence. IET Cyber-Physical Systems: Theory and Applications 6(1), 12–26 (2021). https://doi.org/10.1049/cps2.12002
35. Higgins, M., et al.: Cyber-Physical Risk Assessment for False Data Injection Attacks Considering Moving Target Defences (2022). http://arxiv.org/abs/2202.10841
36. Higgins, M., et al.: Topology learning aided false data injection attack without prior topology information. In: IEEE Power and Energy Society General Meeting, IEEE Computer Society (2021).
37. Rehman, Ur., et al.: Security issues in smart metering systems. In: International Conference on Smart Energy Grid Engineering, SEGE 2015. Institute of Electrical and Electronics Engineers Inc. (2015).
38. Musleh, A.S., et al.: A prediction algorithm to enhance grid resilience toward cyber attacks in WAMCS applications. IEEE Syst. J. 13(1), 710–719 (2019). https://doi.org/10.1109/jsyst.2017.2741483

39. Khalid, H.M., et al.: Wide area monitoring system operations in modern power grids: a median regression function-based state estimation approach towards cyber attacks. Sustain. Energy Grids Netw. 34, 101009 (2023). https://doi.org/10.1016/j.segan.2023.101009

40. Inayat, U., et al.: Learning-Based Methods for Cyber Attacks Detection in IoT Systems: Methods, Analysis, and Future Prospects. MDPI (2022).

41. Patel, S., Sihmar, S., Jatain, A.: A study of hierarchical clustering algorithms. In: INDIAcom (ed.) International Conference on Computing for Sustainable Global Development (INDIACom)., pp. 1–2. INDIAcom, New Delhi (2015).

42. Ramachandran, R., Ravichandran, G., Raveendran, A.: Evaluation of dimensionality reduction techniques for big data. In: Proceedings of the 4th International Conference on Computing Methodologies and Communication, ICCMC 2020, pp. 226–231. Institute of Electrical and Electronics Engineers Inc. (2020).

How to cite this article: Higgins, M., Stephen, B., Wallom, D.: Detecting smart meter false data attacks using hierarchical feature clustering and incentive weighted anomaly detection. IET Cyber-Phys. Syst., Theory Appl. 1–15 (2023). https://doi.org/10.1049/cps2.12057