Electrical Power and Energy Systems 101 (2018) 301–310

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Clustering-based novelty detection for identification of non-technical losses

Joaquim L. Viegas[a,*], Paulo R. Esteves[b], Susana M. Vieira[a]

[a] IDMEC, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1, 1049-001 Lisbon, Portugal
[b] PowerData, Portugal

ARTICLE INFO

Keywords:
Clustering
Data mining
Detection of non-technical losses
Electricity theft
Novelty detection
Smart metering

ABSTRACT

The reduction of non-technical losses is a significant part of the total potential benefits resulting from implementations of the smart grid concept. This paper proposes a data-based method to detect sources of theft and other commercial losses. Prototypes of typical consumption behavior are extracted through clustering of data collected from smart meters. A distance-based novelty detection framework classifies new data samples as malign if their distance to the typical consumption prototypes is significant. The proposed method works on the space of four different indicators of irregular consumption, enabling the easy interpretation of results. A use case based on real data is presented to evaluate the method. The threat model considers sixteen different possible types of changes in consumption pattern that result from non-technical losses, including attacks and defects present since the first day of metering. The proposed clustering-based novelty detection method for identification of non-technical losses, using the Gustafson-Kessel fuzzy clustering algorithm, achieves a true positive rate of 63.6% and false positive rate of 24.3%, outperforming other state-of-the-art unsupervised learning methods.

1. Introduction

In electrical grids, non-technical losses (NTLs) are equal to the difference between electricity supplied and electricity paid for, subtracting the energy lost through heat in lines, transformers and other equipments. NTLs are the result of electricity theft, fraud or deficient metering assets and have significant financial impact to utilities and economies. Theft is widespread in many developing economies, such as India, where theft has been estimated to amount to more than 1% of the country gross domestic product (GDP) [1]. The impact of NTLs is also significant in developed countries, in the UK electricity theft is estimated at £173 million every year [2], in the US it may be worth up to $6 billion [3].

The growing proliferation of the smart grid concept and implementation of advanced metering infrastructure (AMI) systems results in grids with many digitally interconnected assets, enabling complete remote control and monitoring. Two-way communications between assets and utility systems have the potential to enable better grid management. Meanwhile, this wide use of cyber-physical systems opens the door for hacking and cyber-attacks.

Usually, the reported sources of NTLs are fraud through meter manipulation, tapping distributions lines and non-payment [4–6]. Deficient meters and utility systems that compromise measurements and collusion with utility employees can also result in losses. The use of smart meters (SMs) for remote control and consumption data collection widens the attack surface for electricity theft [7,8]. Through meter hacking, manipulation and spoofing of communication individuals can enact false data and bad data injection (BDI) attacks [9].

Multiple data-based classification and estimation techniques have been tested to detect NTLs, such as state estimation [10], clustering [11], neural-networks [12], support-vector machines (SVM) [13] and decision trees [14,15]. Some of the studies only deal with electricity theft while other studies deal with aggregated NTLs, not being able to pin-point the exact location of their source [16,17]. Multiple authors are starting to deal with the resulting potential threats that come from the extended attack-surface due to smart meters [7,18,19]. Recent research focuses on challenges in dealing with large and imbalanced datasets collected through smart grid assets, usually using artificial intelligence techniques [15,20,21].

Taking into account the potential of sophisticated fraudsters and cyber-attacks, [22,23] propose game-theoretic frameworks to deal with electricity theft. In [18], multiple classifiers are evaluated in an adversarial environment, analyzing the worst case scenario assuming attackers have knowledge of the detection technique used. In [7], supervised and non-supervised classification techniques are tested to detect a synthetic consumption pattern that result from theft, achieving best results with SVM classification.

This paper proposes a method to detect sources of NTLs in smart grids. We focus on all types of losses that can result in changes in the consumption data that is collected by a SM and communicated to the

* Corresponding author.
E-mail address: joaquim.viegas@tecnico.ulisboa.pt (J.L. Viegas).

https://doi.org/10.1016/j.ijepes.2018.03.031
Received 12 October 2017; Received in revised form 21 February 2018; Accepted 21 March 2018

utility. Firstly, indicators of irregular consumption are computed from the collected consumption data, representing changes in behavior or irregularities in comparison to similar consumers. Secondly, the data of a set of benign consumers is clustered to uncover the prototypes of legitimate behavior, representing the different patterns of indicators that result from normal consumption. Thirdly, the prototypes are used in a distance-based novelty detection method. The farther away data from an analyzed consumer is from the normal prototypes, the higher their NTLs score is, indicating they may be stealing electricity or metering equipment is malfunctioning.

We propose the use of fuzzy Gustafson-Kessel clustering (GK) to detect consumption patterns resulting from the presence of NTLs, which we show is well suited for the application and has not been used in the current literature on novelty detection. A novelty detection framework has not been used before in the field of detection of NTLs and electricity theft. The method is tested on a use case that extends the threat models proposed in [7,21,24]. A complete set of possible changes of consumption, including sources of NTLs active from day of connection, are considered. Results of the use case show the potential of the method, achieving good results, out-performing other tested techniques proposed in the literature to deal with equivalent data. The proposed indicators enable an easy interpretation of the scores given by the detection method, which contrasts to the non-transparent nature of most techniques used in the literature.

We believe the method is well suited to be used in areas of a smart grid where significant aggregated NTLs are detected through calculation of the difference between supplied and billed electricity. In this case, the method can pin-point the thieving individual or faulty equipment.

2. Threat model

The considered threat model identifies the possible attack vectors and main system vulnerabilities related to electricity theft in smart grids. The term *attack vectors* refers to the ways an individual can maliciously affect the electricity network or the utilities systems to pay less than the full amount they owe for the electricity they consume. Other kinds of NTLs can also be detected using this framework, as they also result in changes or irregularities in the consumption data sent to the utility by the SM.

We propose a model that extends the ones presented in [7,21,24]. This paper presents an extended analysis of the attack surface, considers false data attacks with higher complexity such as proposed in [21], and includes cases where the losses start on the first day of consumption data acquisition (we refer to these cases as first-day attacks).

This paper considers a smart grid environment with an AMI system, characterized by presence of SMs at all the consumption endpoints. SMs have advanced communication capabilities and automatically send consumption data to the utility. The attack surface is said to be increased with the use of SMs. New cyber and data attack/vulnerabilities, such as the possibility of sending false readings, appear with the use of these equipments [7,25,26]. BDI can be used to steal electricity and breakdown grid assets, possibly having catastrophic consequences [9]. Current literature on detection of theft and NTLs and electricity is giving an increased importance to this issue [27–31].

The different NTLs sources and attack/vulnerability vectors are pictured in Fig. 1. The encircled points indicate the different possible attack vectors.

NTLs can be detected through the analysis of metering data. The proposed method deals with the types of NTLs that result in a change or irregular consumption pattern (e.g. if a consumer connects an equipment to a distribution line their consumption is lowered). Table 1 lists the different attack/vulnerability vectors, scenarios and expected changes in metered consumption data. Column *Point* indicates the related point in Fig. 1. The first-day attack scenario is considered. Note most scenarios are expected to result in a variation or irregularity of the metered consumption data.

Scenarios relating to billing were not listed because they result in changes done after the processing consumption data collected by SMs. They include non-payment (point 4), collusion with utility employees (points 5 and 6), cyber attacks to commercial systems and erroneous billing (point 7).

Cases resulting in a constant reduction of consumption, such as the disconnection of a meter or use of a strong magnet to interfere with it, can be detected through straightforward methods such as slope analysis and rule-based systems [14,32]. If the attackers are highly resourceful, they may send false consumption data (e.g. BDI) which is seemingly legitimate [7,18]. In an adversarial environment the attacker evolves trough time and information on past attacks may not be useful to prevent future ones [18]. Also, if the attack is made from the day of connection to the grid (first-day), no reduction or change in consumption can be detected, only the comparison to similar consumers is effective [33].

As past examples of theft may not be suitable, different types of attacks are generated to test the proposed detection technique. Also, according to [7,8,10], real data samples of electricity fraud are not easily available as the smart grid is not fully implemented yet. Six of the attacks are the ones presented in [7]. The two other complex attacks are proposed by [24]. One deals with the manipulation of data to shift a significant amount of consumption from peak hours to lower valley hours, taking advantage of two-part and three-part tariffs that are higher at peak time. The other considers a especially resourceful thief, manipulating their consumption data to look completely legitimate while lowering their total bill. Each one of the 8 types of attacks is considered in two versions, starting in a day posterior to the first day of consumption data and starting in the first day, resulting in a final set of 16 attack types. h_1 to h_8 are attacks that start later than the day the metering starts and h_{10} to h_{80} represent the first-day versions of the attacks.

1. Random constant reduction of consumption (h_1 and h_{10} for the zero day scenario);
2. Drop of consumption to zero during a random period of the day (h_2 and h_{20});
3. Random hourly reduction of consumption (h_3 and h_{30});
4. Random hourly consumption pattern with reduced average consumption (h_4 and h_{40});
5. Constant hourly consumption equal to the average (h_5 and h_{50});
6. Reversed hourly consumption: switch consumption of hour 1 with hour 24, etc. (h_6 and h_{60});
7. Shift of consumption from peak hours to the rest of the day (h_7 and h_{70});
8. Shift the consumption data to the one of a legitimate consumer with lower electricity needs (h_8 and h_{80})

The following notation is adopted: we work with a smart metering dataset M with N consumers. $\mathbf{m}_i$ are the meter consumption readings from consumer i. The dimension of $\mathbf{m}_i$ is $n = r \times n_d$ where n_d is the number of days and r is the number of consumption readings per day. In this work 24 readings per day are used (one per hour), the simplified notation is: $m_i^{d,t}$ is the consumption in day d for hour t. $\mathbf{m}_i^d = (m_i^{d,1}, m_i^{d,2}, \ldots, m_i^{d,24})$ is the 24 h vector of metered data of consumer i in day d.

To compare similar consumers, a dataset of consumer characteristics S is used. $\mathbf{s}_i$ are the characteristics of consumer i with dimension p equal to the number of characteristics. The following equations describe the way an attack starting on day d by consumer i affects their consumption data. These are used to generate the synthetic attacks used to test the proposed method. μ represents the average function.

- $h_1(m_i^{d,t}) = \alpha m_i^{d,t}, \quad \alpha = \text{random}(0.1, 0.8)$
- $h_2(m_i^{d,t}) = \beta^h m_i^{d,t}$