Detection of Integrity Attacks to Smart Grids using Process Mining and Time-evolving Graphs

Short Paper

Simona Bernardi, Raquel Trillo-Lado, José Merseguer
Department of Computer Science and System Engineering
University of Zaragoza, Spain
Email: {simonab, raqueltl, jmerse}@unizar.es

***Abstract*—In this paper, we present a work-in-progress approach to detect integrity attacks to Smart Grids by analyzing the readings from smart meters. Our approach is based on process mining and time-evolving graphs. In particular, process mining is used to discover graphs, from the dataset collecting the readings over a time period, that represent the behaviour of a customer. The time-evolving graphs are then compared in order to detect anomalous behavior of a customer. To evaluate the feasibility of our approach, we have conducted preliminary experiments by using the dataset provided by the Ireland's Commission for Energy Regulation (CER).**

I. Introduction

Traditional power grids are networks of power lines, and their associated equipment, used to transmit and distribute a specific type of power or energy, over a geographic area. During the last decade, these grids have been incorporating Information and Communication Technologies (ICT) to enable bi-directional communication among their components to improve operations, maintenances, planning, coordination and control. This new type of architecture of power grids is called Smart Grids and it is characterized by an Advanced Metering Infrastructure (AMI) that enables the collection and distribution of information in real-time between smart meters, located at customer sites, and utilities [1].

Although Smart Grids provide a lot of benefits, they also pose several security challenges. Cyberattacks to Smart Grids take different forms, such as denial of service, gaining access to the power grids control system or stealing information [2], [3]. Our work focusses on data injection attacks on the meters of the customers [4], and on their detection. This kind of attacks causes mis-billing as well as quality of service depreciation, and may lead even to destabilize the energy market system [5]. As posted in [6], the FBI reported the first case of financial losses due to smart meter hacking suffered by a single electric utility, which were estimated of several hundreds of millions of dollars annually.

Most of the current approaches on detecting data integrity attacks are oriented to protect the infrastructure, the power suppliers and system operators, but disregard the integrity attacks targeted to the final customer [7], [8]. To the best of our knowledge, there are few approaches aimed to detect the corruption of data from a particular meter and they propose anomaly detection methods embedded in the smart meter [8], [9]. Our approach, instead, does not necessarily require additional devices or hardware mechanisms, and can be developed as a remote service. The anomaly detection method proposed in [5] combines two traditional data mining techniques, Principal Component Analysis (PCA) and clustering, to verify the smart meter measurements. However, our approach, that is based on [5], relies upon *process mining* and *time-evolving graphs*.

Process mining [10] is a relatively young discipline, whose main goal is to discover, monitor and improve business processes by extracting knowledge from *event logs*, produced by information systems in operation. In this work, we are interested in model discovery techniques based on *fuzzy mining* that are used to automatically generate graphs from event logs.

Time-evolving graphs [11] represent data at different time periods and are used to detect changes of behavior or temporal anomalous patterns. The dynamic graph-based techniques rely upon the concept of graph distance (or graph similarity), for anomaly detection: the main approach consists of comparing graphs related to consecutive time periods by using a distance or a similarity function [12]. When the distance is greater than a predefined threshold (or conversely, the similarity is smaller than a certain threshold), the corresponding graph is characterized as anomalous.

The rest of the paper is organized as follows. Section II describes the approach proposed. Conclusions and current work under development are given in Section III.

II. Approach overview

The rationale of this work is to develop an efficient automatic technique to detect integrity attacks to the smart-grid AMI. Therefore, we propose an approach to discover graphs that model the consume patterns of the customers from the readings of their smart meters over a time period, and detect outliers in a particular period. This approach has three main steps that are depicted in Figure 1: *Classification*, *Graph discovery*, and *Integrity attack detection*. In the following, each step is discussed in more detail.

978-1-5386-8060-5/18/$31.00 ©2018 IEEE
DOI 10.1109/EDCC.2018.00032

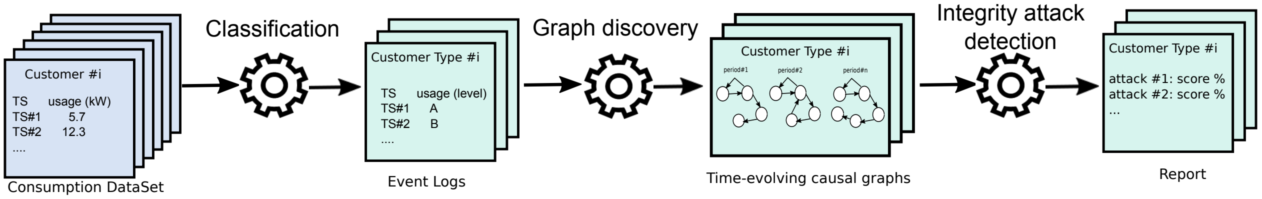

Figure 1. Overview of the approach

A. Classification

The first step corresponds to a pre-processing phase, where the *consumption dataset* is transformed to *event logs*. The consumption dataset collects the readings, from the smart meters, related to the energy/gas consumption of the customers over a time period. On the other hand, an event log is a set of execution traces, where each trace consists of a sequence of time-ordered events. Three main attributes characterize an event log: the *case identifier*, the *event timestamp* and the *event type*.

Since a smart meter is associated to a customer, the smart meter identifier can be trivially considered as a case identifier and the time of the reading can be taken as the event timestamp. Concerning the event types, they represent the consumption, whose types are real values (e.g., gas/energy consumed in kW), whereas the event types in logs need to be discrete values. Therefore, each value registered by a smart meter is associated to an ordered set of levels $\mathcal{L} = \{L_i\}$, where $L_i < L_{i+1}$ $(i = 1, \dots, N-1)$ by defining a monotonic mapping function $C : \mathbb{R} \rightarrow \mathcal{L}$ such that: $\forall x, y \in \mathbb{R} : x < y \Rightarrow C(x) \leq C(y)$.

Each level actually corresponds to an interval of gas/energy consumption (e.g., $[0, 5[$ Kw, $[5, 10[$ Kw, etc.).

In the following we unveil the problems that need to be addressed for this step.

Discussion. The main issue of this step is the definition of the mapping function to be considered, as that function should take into account several parameters such as, the type of customer. Each type of customer (e.g., residential customer, SME, etc.) has a different consumption behavior. For example, considering the dataset provided by the Ireland's Commission for Energy Regulation (CER) [13], Figure 2 shows the different range of the gas consumption, in kW, registered by three different smart meters during a week. In addition, the consumption behavior of a customer varies according on the epoch of the year (e.g., winter, summer, holidays, etc.). Therefore, the main statistical qualifiers (e.g., min, max, mean values, etc.) of the customer consumption, over a time period of reference, should be also considered when defining the set of the consumption levels.

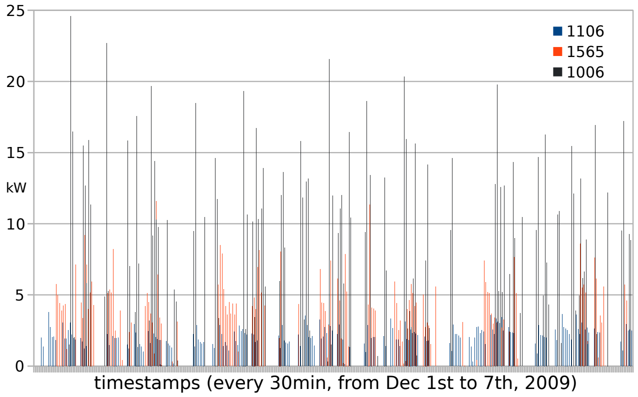

Figure 2. One week gas consumption by different smart meters.

B. Graph discovery

In this step, we resort to process mining and, in particular, model discovery techniques, to generate graphs from the event logs automatically. An event log, produced by the previous step and representing the consumption of a customer c over a time period T, can be defined as a set of ordered tuples:

$$\mathcal{EL}_c^T = \{\langle t_j, u_j \rangle\}_j,\ t_{j-1} \leq t_j,\ j = 1, \dots, M$$

where $u_j \in \mathcal{L}^T \subseteq \mathcal{L}$ is the consumption level registered by the smart meter during the time interval $[t_{j-1}, t_j]$.

A graph should capture at least the consumption levels and the change of consumption levels observed during T.

We have developed an algorithm that generates a weighted graph $G_c^T = \langle N, E, W \rangle$, where each node $n \in N$ corresponds to a consumption level and an arc $e \equiv (n_k, n_l) \in E$ represents the change of consumption from level n_k to level n_l. In addition, both nodes and arcs are characterized by a weight $(W : N \cup E \rightarrow \mathbb{N})$ that represents the frequency of the consumption level, $W(n)$, or the frequency of the change of consumption level from n_k to n_l, $W(e)$, in the log.

The tool Disco [14], among other functionalities, can support this algorithm. Figure 3 shows two graphs, generated by Disco[1] from two event logs $\mathcal{EL}_{1565}^{W0}$ and $\mathcal{EL}_{1565}^{W1}$ without filtering option, that represent the consumption behavior of a customer during two subsequent weeks. Each node represents a consumption level that corresponds to an

[1] Our algorithm produces the same graphs, though only in textual format.

interval of gas consumption (i.e., $5[i-1, i[$ kW, $i = 1, 2, 3$). The numbers associated to nodes and arcs correspond to the frequencies.

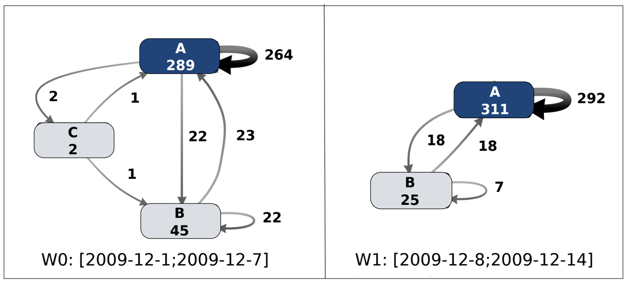

Figure 3. Consumption graphs of customer #1565 of two consecutive weeks.

Discussion. The number of consumption levels defined by the mapping function (see Subsection II-A) is an upper bound for the number of nodes of a graph. Indeed, the mapping function is defined over a time period of reference (e.g., a year), whereas an event log used to produce a graph represents the consumption over a shorter period (e.g., a week or a month). Moreover, reasonable sizes for the set of consumption levels $\mathcal{L}$ is of order of tens. Therefore, we do not expect to deal with *large* graphs.

On the other hand, a graph generated by the aforementioned algorithm is characterized by one type of feature, that is the node/arc frequency. Similarly to fuzzy mining [10], which relies upon an interesting set of significance and correlation metrics to generate fuzzy nets, we could consider other graph features to be exploited in the graphs comparison (subsection II-C).

C. Integrity attack detection

In this step, we use the time-evolving graphs, generated in the previous step, to detect integrity attacks to the AMI.

The detection problem can be stated as follows: Given an ordered sequence of graphs $\{G_{T_i}\}_{i=1,\dots,K}$, that models the consumption behavior of a customer during consecutive time periods T_i, find those time periods that correspond to *anomalous* behaviors. An anomalous behavior corresponds to integrity attacks that produce counterfeit readings.

In [15], random scale and average attacks, already studied in [5], were considered, using the Ireland's CER dataset [13]. We generated two synthetic datasets from the original one, that represent the anomalous behavior of the considered customer according to two types of integrity attacks and, then, we obtained the corresponding time-evolving graphs, $\{G^{rs}_{T_i}\}_{i=1,\dots,K}$ and $\{G^{avg}_{T_i}\}_{i=1,\dots,K}$, by applying the previous two steps of the approach (subsections II-A and II-B).

In time-evolving graph techniques [11], the anomalous behavior is detected by comparing consecutive graphs using a distance (or similarity) function and verifying whether the distance is greater (or smaller) than a predefined threshold. In our approach, we chose one distance and one similarity measure: the Hamming distance, that is purely structural, and the cosine similarity measure, that takes into account frequencies associated to the nodes and arcs of the graphs.

To evaluate the effectiveness of the two measures in the detection of the two types of attacks, we applied the *paired-t* approach, that is based on the computation of the confidence interval for the mean difference [16]. The approach allowed us to determine the statistical and practical significance of the difference between the normal behavior (i.e., no attacks) and the anomalous behavior due to an integrity attack.

In particular, we considered the following sets of pairs, where each pair models the behavior of the customer (smart meter $id = 1565$), in two consecutive weeks:

- $\{(G_{T_i}, G_{T_{i+1}})\}$, the set represents the *normal* behavior assuming the integrity of the smart meter;
- $\{(G_{T_i}, G^{rs}_{T_{i+1}})\}$, the set represents the case of a random scale attack in the period T_{i+1}, and
- $\{(G_{T_i}, G^{avg}_{T_{i+1}})\}$, the set represents the case of an average attack in the period T_{i+1}.

Each set, consists of 77 pairs, covering an overall period of 78 weeks. For each pair of graphs, we computed the distance (similarity) according to the selected measure f, i.e., $f_i = f(G_{T_i}, G_{T_{i+1}})$, $f^{rs}_i = f(G_{T_i}, G^{rs}_{T_{i+1}})$ and $f^{avg}_i = f(G_{T_i}, G^{avg}_{T_{i+1}})$, and the corresponding differences between the normal and anomalous behavior due to the two types of attacks, i.e., $\zeta^{rs}_i = f_i - f^{rs}_i$ and $\zeta^{rs}_i = f_i - f^{avg}_i$, respectively. Finally, each set of differences $\{\zeta^{rs}_i\}$ and $\{\zeta^{avg}_i\}$ was used to compute a confidence interval for the mean of the difference ζ random variable, considering the Student t-distribution with N-1=76 degrees of freedom.

Random scale attacks for the Smart Meter Id 1565					
	$\bar{\zeta}$	σ_ζ	$lb_{0.10}$	$ub_{0.10}$	R.E.
ζ_H	0.004149	0.330856	-0.071147	0.079444	10.12%
ζ_{Cos}	-4.8227E-05	0.002	-0.0005034	0.000407	0.05%
Average attacks for the Smart Meter Id 1565					
	$\bar{\zeta}$	σ_ζ	$lb_{0.10}$	$ub_{0.10}$	R.E.
ζ_H	0.393579	0.387513	0.305389	0.481769	43,43%
ζ_{Cos}	0.023530	0.041048	0.014189	0.032872	1.42%

The previous table summarizes the results for the two types of attacks and each type of measure, i.e., the Hamming distance ζ_H, and the cosine similarity ζ_{Cos}. The columns of the Table (from 2^{nd} to 6^{th}) show respectively: the sample mean $\bar{\zeta}$, the sample standard deviation σ_ζ, the lower $lb_{0.10}$ and upper $ub_{0.10}$ bounds of the 95% confidence interval, and the minimum relative error (R.E.). In particular, R.E. $= min(\frac{lb_{0.10}}{\bar{f}}, \frac{ub_{0.10}}{\bar{f}})$, where $\bar{f}$ is the sample mean of the set of distances (similarities) $\{f_i\}$.

From the experiments, we learnt that the cosine similarity measure seems to be not a good choice since we were not able to discern between normal changes in the behavior and anomalous behavior. Indeed, for random scale attacks, the differences are not statistically significant, i.e., $0 \in [lb_{0.10}, ub_{0.10}]$, and for average attacks, the minimum relative error is negligible. Concerning random scale attacks, the results are not statistically significant also in the case the Hamming distance is used. On the other hand, the results obtained by using the Hamming distance are promising in the detection of average attacks (i.e., the differences are statistically significant and the R.E. is above the 40%).

Discussion. We have identified two main open issues in this step. First, the choice of a *good* distance measure that effectively enables to detect the anomalous behavior. In particular, we cannot draw general conclusions about the suitability of the Hamming and cosine measures initially considered, since the results concern the behavior of just one customer. Second, the definition of a threshold that enables to discern between the change of behavior, due to the epoch of the year, and the anomalous behavior due to an integrity attack. Both issues can be addressed with a validation of the approach, considering the consumption data collected by different smart meters and tuning the parameters (e.g., observation periods, attack model parameters) to minimize false positives and false negatives.

III. Conclusions

Most current techniques to detect integrity data attacks on Smart Grid are focussed on protecting markets, power suppliers and system operators, but not the AMI as our approach does. Moreover, the approach does not necessarily requires additional devices or hardware mechanisms and can be developed as a remote service. At this respect privacy maybe a concern, in particular, when the service is provided by third-parties [17] and will be considered as future work.

Our current work, instead, is focussed on overcoming the issues reported in the *Discussion* sections and on carrying out an in-depth evaluation of the approach by considering different datasets and more distance metrics.

Acknowledgment

We thank the Lt. Etienne Chotard for the preliminary analysis done under the supervision of the authors. The used smart meter dataset is accessed via the Irish Social Science Data Archive (www.ucd.ie/issda). This work has been funded by the projects CyCriSec [TIN2014-58457-R] and Aragon Government Ref. T27-DISCO research group.

References

[1] E. Ancillotti, R. Bruno, and M. Conti, "The role of communication systems in smart grids: Architectures, technical solutions and research challenges," *Computer Communications*, 2013.

[2] Y. Mo, T. H. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli, "Cyber-physical security of a smart grid infrastructure," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 195–209, 2012.

[3] W. Wang and Z. Lu, "Cyber security in the smart grid: Survey and challenges," *Computer Networks*, vol. 57, no. 5, pp. 1344 – 1371, 2013.

[4] A. Giani, E. Bitar, M. J. Garcia, M. McQueen, P. P. Khargonekar, and K. Poolla, "Smart grid data integrity attacks," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1244–1253, 2013.

[5] V. Badrinath-Krishna, G. Weaver, and W. Sanders, "PCA-Based Method for Detecting Integrity Attacks on Advanced Metering Infrastructure," in *Proc. of QEST 2015 - Volume 9259*. New York, NY, USA: Springer-Verlag New York, Inc., 2015, pp. 70–85.

[6] R. Former, "Fbi: Smart meter hacks likely to spread," 2012, Krebs on Security. In depth security news and investigation, Available: https://krebsonsecurity.com/2012/04/fbi-smart-meter-hacks-likely-to-spread/.

[7] R. Tan, V. B. Krishna, D. K. Y. Yau, and Z. Kalbarczyk, "Integrity attacks on real-time pricing in electric power grids," *ACM Trans. Inf. Syst. Secur.*, vol. 18, no. 2, pp. 5:1–5:33, Jul. 2015.

[8] W. Yu, D. Griffith, L. Ge, S. Bhattarai, and N. Golmie, "An integrated detection system against false data injection attacks in the smart grid," *Security and Communication Networks*, vol. 8, no. 2, pp. 91–109, 2014.

[9] M. Raciti and S. Nadjm-Tehrani, "Embedded Cyber-Physical Anomaly Detection in Smart Meters," in *Critical Information Infrastructures Security*, B. Hämmerli, N. Kalstad, and J. Lopez, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 34–45.

[10] W. M. P. van der Aalst, *Process Mining - Data Science in Action, 2^{nd} Edition*. Springer, 2016.

[11] L. Akoglu, H. Tong, and D. Koutra, "Graph based anomaly detection and description: a survey," *Data Mining and Knowledge Discovery*, vol. 29, no. 3, pp. 626–688, May 2015.

[12] S. Cha, "Comprehensive survey on distance/similarity measures between probability density functions," *Int. Journ. of Mathematical Models and Methods in Applied Sciences*, vol. 1, no. 4, pp. 300–307, 2007.

[13] "Commission for Energy Regulation," Irish Social Science Data Archive. URL: https://www.ucd.ie/issda/.

[14] C. W. Günther and A. Rozinat, "Disco: Discover Your Processes." *BPM (Demos)*, vol. 940, pp. 40–44, 2012.

[15] E. Chotard, "Use of process mining techniques for the detection of integrity attacks to a SmartGrid," Master's thesis, Ecole Speciále Militaire de Saint Cyr, 2017.

[16] M. L. Averill, *Simulation Modeling and Analysis*. McGraw-Hill, 2015.

[17] O. Stan, M. Zayani, R. Sirdey, A. Ben-Hamida, A. Ferreira, and M. Mziou-Sallami, "A new crypto-classifier service for energy efficiency in smart cities," in *Proc. of the 7th International Conference on Smart Cities and Green ICT Systems, Funchal, Madeira, Portugal*. SciTePress, 2018, pp. 78–88.