PHYSA: 17433 Model 3Gsc pp. 1–8 (col. fig: NIL)

ARTICLE IN PRESS

Physica A xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Physica A

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

Characterization of electric load with Information Theory quantifiers

Andre L.L. Aquino [a,*], Heitor S. Ramos [a], Alejandro C. Frery [a], Leonardo P. Viana [a], Tamer S.G. Cavalcante [b], Osvaldo A. Rosso [c,d,e]

[a] *LaCCAN/CPMAT – Instituto de Computação, Universidade Federal de Alagoas (UFAL), BR 104 Norte km 97, 57072-970 Maceió, Alagoas, Brazil*

[b] *Centro de Informática, Universidade Federal de Pernambuco (UFPE), Av. Professor Moraes Rego 1235, 50670–901 Recife, Pernambuco, Brazil*

[c] *Instituto de Física, Universidade Federal de Alagoas (UFAL), BR 104 Norte km 97, 57072-970 Maceió, Alagoas, Brazil*

[d] *Instituto Tecnológico de Buenos Aires (ITBA), CONICET, C1106ACD, Av. Eduardo Madero 399, Ciudad Autónoma de Buenos Aires, Argentina*

[e] *Complex Systems Group, Facultad de Ingeniería y Ciencias Aplicadas, Universidad de los Andes, Av. Mons. Álvaro del Portillo 12.455, Las Condes, Santiago, Chile*

HIGHLIGHTS

- We use Information Theory for the characterization of electric load.
- With our characterization was identified different regimes and behaviors in the electric load.
- We observe that types of appliances have characteristic footprints that form clusters in the CCE plane.

ARTICLE INFO

Article history:
Received 5 April 2016
Received in revised form 10 August 2016
Available online xxxx

Keywords:
Causality Complexity–Entropy Plane (CCEP)
Electric load
Information Theory
Signal processing

ABSTRACT

This paper presents a study of the electric load behavior based on the Causality Complexity–Entropy Plane. We use a public data set, namely REDD, which contains detailed power usage information from several domestic appliances. In our characterization, we use the available power data of the circuit/devices of all houses. The Bandt–Pompe methodology combined with the Causality Complexity–Entropy Plane was used to identify and characterize regimes and behaviors over these data. The results showed that this characterization provides a useful insight into the underlying dynamics that govern the electric load.

1. Introduction

Traditionally, power grids are used to carry power from a few central generators to a large number of users or customers. In contrast, Smart Grids use two-way flows of electricity and information to create an automated and distributed advanced energy delivery network [1]. Smart Grids enable the development of new applications related to advanced information metering, monitoring, and management, for instance, Non-Intrusive Load Monitoring (NILM) [2,3].

* Corresponding author.
E-mail address: alla.lins@gmail.com (A.L.L. Aquino).

http://dx.doi.org/10.1016/j.physa.2016.08.017
0378-4371/© 2016 Elsevier B.V. All rights reserved.

Zoha et al. [2] state that NILM requires disaggregating electrical loads by examining only appliance specific power consumption signatures within the aggregated load data. The data is acquired from the main electrical panel outside the building or the residence, hence it is considered to be non-intrusive as the method does not require any equipment installation inside the customer's property. The goal is to breakdown the whole-house building data into its major constituents.

A key aspect to explore in NILM applications is the adequate characterization of electrical energy consumption of domestic appliances. To tackle this issue we present a characterization of the behavior of electrical devices by using quantifiers stemming from Information Theory. The characterization is performed in two stages. Firstly, the original time series of the electrical consumption is transformed into a histogram with a nonparametric transformation that retains time causal information: the Bandt–Pompe methodology [4]. Secondly, this histogram is mapped onto the Causality Complexity–Entropy Plane (CCEP) [5], and its location is shown to serve as a characterization of a number of typical regimes. This plane is a compact manifold spanning values of the normalized Shannon entropy $\mathcal{H}$ and the statistical complexity $\mathcal{C}$.

According to the findings obtained by Rosso et al. [5], chaotic maps have intermediate $\mathcal{H}$ values, while $\mathcal{C}$ stays close to the maximum possible complexity value [6]. For regular processes, entropy and complexity have small values, close to zero. Finally, totally uncorrelated stochastic processes are located close to the $(1, 0)$ point. It has also been found that $1/f^k$ correlated stochastic processes with $1 \leq k \leq 3$ are characterized by intermediate permutation entropy and intermediate statistical complexity values [5]. Additionally, they found that fractional Gaussian noise (fGn) fingerprints lie in the same region than $0 \leq k < 1$ noise. Fractional Brownian motion (fBm) sweeps the region equivalent to $1 \leq k \leq 3$ noise, with two important particular cases: antipersistent fBm ($1 \leq k < 2$), and persistent fBm ($2 \leq k \leq 3$). Note that these processes lie in different positions in the CCEP and, thus, can be characterized by two watersheds: those below and above what looks like a division line. Our assumption is that using a similar methodology we would be able to characterize the dynamic behavior of appliances.

We conduct an exploratory study to test our methodology with time series from the REDD (Reference Energy Disaggregation Data Set) data set, which describes the power usage information of several domestic appliances [7]. The characterization herein performed uses the available power data of circuits/devices collected from five houses. We analyze devices that can be classified into two different modes of operation: (i) devices that are continuously switched on, such as refrigerators, and (ii) devices that can be switched on or off due to human intervention or any kind of automation, such as oven, lamp and washing machines. Although the information about the intervals that devices are switched on or off can be used to characterize the device, in this work we are mostly interested in the dynamics of the power consumption behavior only when a device is switched on. Thus, we pre-processed the data to rule out all zero readings to capture this situation.

This work is organized as following: Section 2 presents the related work. Section 3 explains the Information Theory quantifiers. Section 4 discusses the energy information data set. Section 5 presents the obtained results. Finally, Section 6 concludes the manuscript.

2. Related work

The characterization of electrical energy consumption of domestic appliances in NILM applications is largely underexplored. Among the problems tackled in the literature, Refs. [8–10] discuss techniques for event detection. Rather than focusing on such problem, other approaches treat the issue of energy disaggregation, i.e., they estimate the energy consumed by devices operating in different ranges of power consumption [11–13]. These works are based on finding correspondences between known and observed patterns.

A number of tools have been used to deal with the NILM problem, event detection or complete disaggregation, among them: hidden Markov models [7,12]; fuzzy systems [13,14]; k-nearest neighbors [15]; evolutionary algorithms [16,17]; k-means [11]; and support vector machines [18]. These works analyze the data for specific applications in mind, whereas in this work we propose an application-free approach that uses a tool based on information-theoretic descriptors that have not been employed previously for this problem.

Several data sets are available, among them: Reference Energy Disaggregation Data Set (REDD) [7]; Building-Level fUlly-labeled data set for Electricity Disaggregation (BLUED) [19]; UK recording Domestic Appliance-Level Electricity (UK-DALE) [20]; Smart* [20]; and the Berkeley campus energy portal (Openbms) [21]. Our testbed is the REDD data base because its low sampling frequency turns the characterization problem more challenging and realistic. This kind of data is the one produced by off-shelf equipment and, thus, contributes for reproducibility of the study both experimentally and theoretically.

Finally, several works investigate the usage of information theory quantifiers to characterize the dynamics underlying time series, for instance, the effects of streamflow dynamics [22], unsupervised edge map scoring [23], and stochastic resonance in a bistable system [24]. Additionally, theoretical advances in statistical complexity measure are discussed in Refs. [25–27].

3. Time series and information theory quantifiers

Bandt and Pompe [4] introduced a method to associate a probability distribution from a time series taking into account the time causality of the process. Given a time series $\boldsymbol{X}(t) = \{x_t : t = 1, \ldots, M\}$, an embedding dimension $D \geq 2$ ($D \in \mathbb{N}$),

Please cite this article in press as: A.L.L. Aquino, et al., Characterization of electric load with Information Theory quantifiers, Physica A (2016), http://dx.doi.org/10.1016/j.physa.2016.08.017