Pattern Recognition 74 (2018) 317–325

Contents lists available at ScienceDirect

Pattern Recognition

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

On the distinctiveness of the electricity load profile

M. Bicego[a,*], A. Farinelli[a], E. Grosso[b], D. Paolini[b], S.D. Ramchurn[c]

[a] University of Verona, Ca' Vignal 2, Strada Le Grazie 15, Verona, 37134, Italy
[b] University of Sassari, Italy
[c] University of Southampton, UK

ARTICLE INFO

Article history:
Received 27 June 2017
Revised 8 September 2017
Accepted 24 September 2017
Available online 25 September 2017

Keywords:
Energy market
Load profile
Biometrics
Classification
Pre-processing

ABSTRACT

The recent increasing availability of fine-grained electrical consumption data allows the exploitation of Pattern Recognition techniques to characterize and analyse the behaviour of energy customers. The Pattern Recognition analysis is typically performed at group level, i.e. with the aim of discovering, via clustering techniques, *groups of users* with a coherent behaviour – this being useful, for example, for targeted pricing or collective energy purchasing. In this paper we took a step forward along this direction, investigating the possibility of discriminating the behaviours of *single users* – i.e., in a biometrics sense. This aspect has not been properly addressed and would pave the way to crucial operations, such as the derivation of alternative advertising schemes based on behavioural targeting. To investigate the uniqueness of the load profiles (i.e. the daily consumption of electrical energy), in our study we used the raw data (the original energy consumption time series) as well as different types of features such as frequency coefficients and normalized load shape indexes, together with various classification schemes. Results obtained on two real world datasets suggest that the load profile does contain significant distinctive information about the single user.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Over the past two decades, power and energy systems have been experiencing a huge transformation, due to the increase of importance of renewable energy sources such as solar, hydroelectric and wind power. In this perspective, the balancing of power sources and consumer demand becomes a serious challenge that cannot totally rely on local production and energy storage systems, but rather requires non isolated grids and intelligent reversal of the load flows, following customer needs. This drastic change opens new and challenging problems for intelligent control systems which must face a number of new interesting issues: in this sense Pattern Recognition tools [1] may be of paramount importance, being able to provide solutions to problems such as forecasting of energy prices, optimal dispatching, consumer segmentation, and energy demand allocation [2–5]. In particular, the availability of fine-grained electrical consumption data (due to the recent large scale deployment of intelligent metering infrastructures), coupled with an increasing and worldwide energy market liberalization, results in a growing interest in discovering and categorizing *groups of users* which share similar behaviours. This is usually done via clustering of the so called user *load profiles*, i.e., the users' consumption of electrical energy over a given period, as measured by the so called Advanced Metering Systems (AMS). Interestingly, this has to do with models taking into account the dynamics and the magnitude of the consumption and the ability to capture in the models exogenous factors (type of appliances, insulation) or context factors (occupancy, weather, seasons, holidays). With reference to Fig. 1, a typical processing scheme includes the following steps [6]:

- temporal aggregation;
- context filtering;
- metadata generation;
- data analysis.

In this scheme, temporal aggregation is used to define the temporal granularity of the data consumption collection (hourly, daily, etc.) while the context filtering stage takes into account specific factors such as holidays, seasons and temperatures. Metadata generation is probably the most critical step in the proposed processing scheme. In fact, starting from a coherent set of temporal measures (load profiles), a number of quantitative descriptors can be derived; in the literature, these descriptors are often denoted as feature functions [6] because they act on the load profiles transforming the original time representation into a more compact or more discriminative representation. Examples of quantitative descriptors are the load factor and the night/lunch impact [7,8]; the

* Corresponding author.
E-mail address: manuele.bicego@univr.it (M. Bicego).

https://doi.org/10.1016/j.patcog.2017.09.039
0031-3203/© 2017 Elsevier Ltd. All rights reserved.

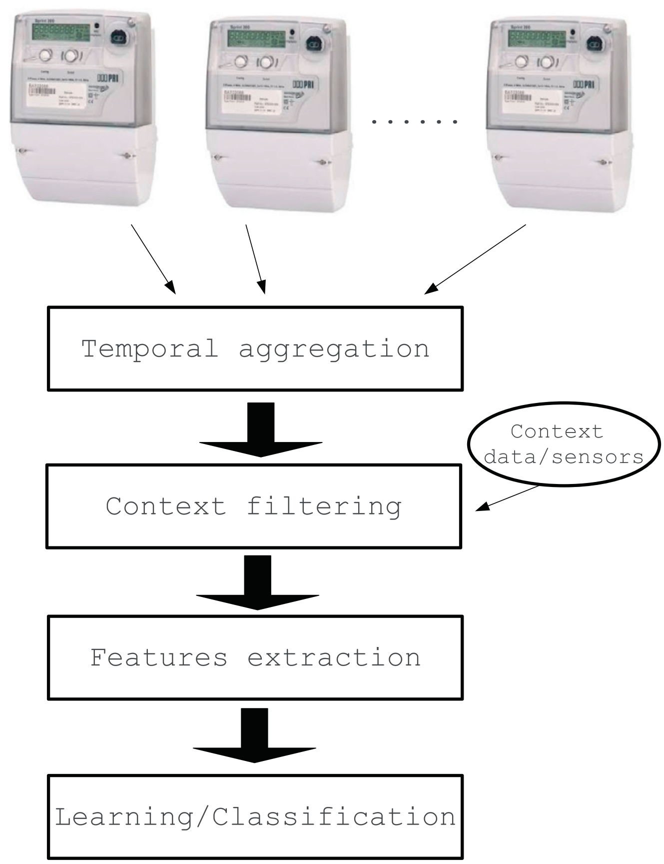

Fig. 1. Typical processing scheme of data collected from AMS.

Fast Fourier Transform is another example of data manipulation giving evidence to the content, in the frequency domain, of the original time representation.

The final analysis step is typically devoted to the clustering of the various load profiles, in order to detect coherent groups of users. This task can be very difficult due the number of groups which is generally unknown and the number of users that can be very high in real applications.

Different approaches have been proposed to face this problem: for example, in [7] authors propose a framework to characterize groups of users based on simple load descriptors. They also prove the robustness of the proposed method with respect to missing data and outliers. Carpaneto and colleagues [9] propose a scheme based on the frequency domain. In contrast, other approaches [10,11] consider each load profile as time sequence of load measurements and apply various unsupervised learning techniques for clustering (Self Organizing Maps, K-means and Hidden Markov Models among others). A good overview of applicable pattern recognition tools is given in [12]; this paper also includes a detailed description of most interesting clustering methods and proposes several consistency measures adequate to evaluate the performance of these methods. Another review is given in [13], discussing in particular how the number of categories can vary depending on locations and type of loads (public, industrial, residential). A deep investigation of clustering methods applied to the domestic sector in Ireland has been recently presented in [14]; authors consider few profile categories and a customer is essentially defined by a vector of likelihood coefficients, showing the statistical association of the customer to each of the profile groups defined.

Even though the above work represent an impressive progress in this area, there is an urgent need for advanced analysis of energy usage data. For example energy companies are becoming more and more interested in targeted advertisement, personal tarification [15] or even in detecting frauds [16] and changes in the composition of the group of people living in a given house. Analysts have begun to use these data for different goals, such as for example the optimal allocation of the energy flows and the reduction of purchase prices, or to help retailers designing new pricing models for implementing more accurate demand and supply profiles [8,10,11]. For all these applications, methods which work *at group level* are not enough, since the characterization and discrimination should be done *at the user level*: in other words, there is a need for automatic systems able to characterize and discriminate *every user* related to a single metering system: this crucial aspect has never been investigated in the literature, and represents the main goal of this paper. In particular, starting from some preliminary and encouraging results [17], this paper investigates different types of metadata and classification schemes to understand if an answer exists to the following key question: *does every single user have a unique behaviour when consuming electrical energy?* or, in different terms, *can the electrical energy consumption related to a single AMS be considered as a distinctive behavioural trait?* As better explained in the following, an answer to this question may open the possibility of devising novel targeting strategies and at the same time it would spur an important discussion on important privacy issues.

To answer the above question, in this paper we develop a *classification system* to identify a specific user (or, more precisely, a specific AMS) among several users, on the base of the electrical consumptions over a given period of time (i.e., a load profile). We investigate different metadata characterizing load profiles, including raw measurements, frequency characterizations and typical load shape indexes. We also investigate two classification schemes: the former is based on the classical Nearest Neighbour rule (i.e., it assigns an unknown object to the class of its Nearest Neighbour); the second scheme follows the classical Bayesian classification [1], based on Hidden Markov Models (HMM – [18]). This probabilistic approach has been widely used to characterize sequential data, and has been recently applied to the problem of clustering load profiles (in particular to characterise relationships between consumers' preferences or behaviours and electricity consumption [11]). The empirical evaluation is based on two databases composed of real load profiles, collected in the UK and in Portugal from several hundreds of metering systems. The system is trained on a known set of profiles, and tested on an hold out set. Our classification results suggest that the energy load profile does indeed contain user-specific discriminative information.

The rest of the paper is organized as follows: Section 2 details the problem of personal tarification and behavioural targeting, unfolding the complexity of the problem and the potential benefits of a behavioural analysis, also from an economic point of view. Section 3 presents the proposed approach, also in relation to related works, detailing both the choice of metadata and the proposed classification scheme. The empirical evaluation of the proposed approach is given in Section 4, while Section 5 concludes the paper.

2. Personal targeting for the energy market

In this section we provide some considerations on the impact that the distinctiveness of the user load profiles may have on the energy market. In particular, we are convinced that the behavioural peculiarities of the load profile may lead to the so-called behavioural targeting (BT) [19] in the energy market. BT is a kind of advertising that is based on the analysis of the peculiar and dis-

Download English Version:

https://daneshyari.com/en/article/4969504

Download Persian Version:

https://daneshyari.com/article/4969504

Daneshyari.com