Applied Energy 141 (2015) 190–199

Contents lists available at ScienceDirect

Applied Energy

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

A clustering approach to domestic electricity load profile characterisation using smart metering data

Fintan McLoughlin [a,*], Aidan Duffy [a], Michael Conlon [b]

[a] School of Civil Engineering and Dublin Energy Lab, Dublin Institute of Technology, Bolton St, Dublin 1, Ireland
[b] School of Electrical & Electronic Engineering and Dublin Energy Lab, Dublin Institute of Technology, Kevin St, Dublin 4, Ireland

HIGHLIGHTS

- We characterise diurnal, intra-daily, seasonal and between customer electricity use.
- A series of profile classes reflective of home electricity use are constructed.
- We examine the influence of household characteristics on patterns of electricity use.

ARTICLE INFO

Article history:
Received 1 August 2014
Received in revised form 4 December 2014
Accepted 21 December 2014
Available online 6 January 2015

Keywords:
Domestic electricity load profile
Segmentation
Clustering

ABSTRACT

The availability of increasing amounts of data to electricity utilities through the implementation of domestic smart metering campaigns has meant that traditional ways of analysing meter reading information such as descriptive statistics has become increasingly difficult. Key characteristic information to the data is often lost, particularly when averaging or aggregation processes are applied. Therefore, other methods of analysing data need to be used so that this information is not lost. One such method which lends itself to analysing large amounts of information is data mining. This allows for the data to be segmented before such aggregation processes are applied. Moreover, segmentation allows for dimension reduction thus enabling easier manipulation of the data.

Clustering methods have been used in the electricity industry for some time. However, their use at a domestic level has been somewhat limited to date. This paper investigates three of the most widely used unsupervised clustering methods: k-means, k-medoid and Self Organising Maps (SOM). The best performing technique is then evaluated in order to segment individual households into clusters based on their pattern of electricity use across the day. The process is repeated for each day over a six month period in order to characterise the diurnal, intra-daily and seasonal variations of domestic electricity demand. Based on these results a series of Profile Classes (PC's) are presented that represent common patterns of electricity use within the home. Finally, each PC is linked to household characteristics by applying a multi-nominal logistic regression to the data. As a result, households and the manner with which they use electricity in the home can be characterised based on individual customer attributes.

1. Introduction

Throughout the European Union, there has been a move towards smarter electricity networks, where increased visibility over electricity generation and consumption has been achieved with the installation of Advanced Metering Infrastructure (AMI). Smart metering is part of this and is seen as a necessary component to achieve EU 20-20-20 energy policy goals by the year 2020: to cut greenhouse gas emissions by 20%, to improve energy efficiency by 20% and for 20% of EU energy demand to come from renewable energy resources [1].

In recent years, smart meter installations have increased worldwide in a bid to modernise aging electricity networks [2]. Furthermore, improvements in the regulatory environment, particularly within the residential sector in Europe has resulted in a number of smart metering pilot programmes [3]. As a consequence, a wealth of new data exists for utilities, giving detailed electricity consumption at increased granularity for a large number of customers within the residential sector [4]. The availability of this source of data can potentially be used by utilities to create customised electricity load Profile Classes (PC) and can assist in areas such

* Corresponding author. Tel.: +353 (0)1 4027912; fax: +353 (0)1 4024035.
E-mail address: fintan.mcloughlin@dit.ie (F. McLoughlin).

http://dx.doi.org/10.1016/j.apenergy.2014.12.039
0306-2619/

as: improved load planning and forecasting; Time of Use (ToU) tariff design; electricity settlement; and Demand Side Management (DSM) strategies [5].

This paper presents a new methodology for electricity load profile characterisation. In doing so, a series of domestic electricity PC's are constructed that are reflective of the varied manner with which electricity is used within the home. Currently, PC's are derived based on aggregating many dissimilar patterns of electricity use together [6]. The application of this type of approach, where individual households which may use electricity in very different ways get lumped together, results in the formation of highly aggregated load profiles. However, in reality this is not a true reflection of how electricity is actually consumed and which can change considerably between different customers [7]. The paper proposes an alternative method which uses clustering to identify similar patterns of electricity use before any aggregation processes are applied. In this way, information pertaining to the electricity load profile shape is not lost. In addition, the paper also presents a method of linking PC's to individual customers so that a household and the manner with which they use electricity within the home can be characterised based on their individual customer attributes.

The paper is structured as follows. Section 2 illustrates existing methods used for electricity load profile characterisation and their limitations in dealing with smart metering data. Section 3 presents the structure of the data on which the analysis was carried out. Section 4 provides the methodological approach for the paper which is divided into three distinct sections: clustering; electricity load profile characterisation; and customer profile classification. Section 5 presents and discusses results with Section 6 containing concluding remarks.

2. Domestic electricity load profile characterisation

Based on the literature, existing methods used to characterise domestic electricity use can generally be divided into four categories: statistical; engineering; time series and clustering. Statistical methods have been widely used in de-regulated electricity markets to form standard load PC's [6]. Standard load PC's are used for the purposes of settlement and provide an estimate as to the quantity and Time of Use (ToU) of electricity being used. A series of PC's are produced for different segments of the market (e.g. residential, commercial, industrial) and are derived based on the average for all customers contained within a single customer class [8]. The UK electricity market has two domestic PC's; Unrestricted and Economy 7. In Ireland, four PC's exist for the domestic sector; 24 h and Night Saver which are split by urban and rural divide [9]. Although PC's are suitable for the purposes of settlement, in reality they are not reflective of how electricity is actually consumed within the home on a daily basis and merely represent the average for all customers contained within the same class. Other statistical techniques consist of using descriptive statistics and probability [10–16] as well as regression [17–22] to describe electricity use within the home. Similar to that stated above, these methods produce highly diversified load profile shapes, a result of combining many dissimilar patterns of electricity use together [10].

Engineering approaches to domestic load profile characterisation are varied but generally characterise electricity use as a function of parameters such as occupancy or appliance ownership [23–28]. These methods are considered to be a bottom up approach where multiple profiles are constructed for different households and therefore do not suffer from the same problem highlighted above for statistical approaches. However, engineering methods are difficult to generalise and require detailed knowledge of household occupant and appliance Time Use (TU) [29]. In contrast time series approaches have been limited in their application to domestic households, but this is most likely due to a historical lack of available data for the sector [7]. The methods have been used extensively to describe electricity use at a Transmission System Operator (TSO) level [30–34]. However, these approaches suffer from a similar problem to that highlighted above for statistical techniques when many dissimilar profiles are aggregated together resulting in diversified electricity load profile shapes [35].

Finally data mining techniques such as cluster analysis have been used to group customers which exhibit similar electrical behaviour through ToU smart meter data, but have mostly been applied at an aggregated level [36–38]. Furthermore, customers have also been clustered based on aggregated parameter values such as annual electricity use or features relating to the electricity load profile shape (e.g. load factor) [39,40]. Similarly, load profiles have been constructed for commercial, industrial and mostly aggregated residential customers based on clustering methods: Self Organising Maps (SOM), k-means and Follow the Leader [41–43]. In particular, one large study of approximately 3000 residential customers was monitored over a period of a single year and used methods: SOM; k-means; and hierarchical to cluster and construct load profiles [44]. However, the analysis was restricted to only a small portion of the time series (5%) due to computational demands. Clustering methods do not suffer from many of the problems highlighted above particularly when it is applied prior to carrying out any statistical analysis. Furthermore with improvements in computer hardware tasks such as clustering, which can be computationally intensive have become easier to implement.

This paper fills a gap in the literature by clustering based on ToU for a large sample of residential customers over a period of six months. This allows for load PC's to be derived based on individual patterns of electricity use within the home over this period and does not suffer from some of the same aggregation problems highlighted above. Furthermore, as the entire dataset is clustered, diurnal, intra-daily and seasonal patterns to electricity use can be characterised, as well as between customer variations. Moreover, as dwelling, occupant and appliance characteristics are correlated with each PC's it also provides a method of assigning patterns of electricity use to individual customers. Finally, as the sample size is relatively large the PC's can be considered to be representative of the wider population in Ireland. A similar method could also be used in other electricity markets outside of Ireland.

3. Data structure

The smart metering trial carried out by Commission for Energy Regulation (CER) provided the necessary information to segment the domestic electricity market in Ireland based on ToU [45]. The trial was conducted between 2009 and 2010 and consisted of installing smart meters in over 4000 residential dwellings in Ireland. Electricity demand at half hourly intervals as well as detailed information on dwelling, occupant and appliance characteristics for a representative sample of dwellings in Ireland was recorded [46,47]. The data provided was in anonymised format in order to protect personnel and confidential information relating to the customer.

The data used in the analysis was taken over the period 1st July to 31st December 2009. The sample size was trimmed to 3941 customers in total on account of missing information due to technology communication problems. Matlab and its respective statistical (ver. 7.3) and neural network toolboxes (ver. 6.0.4) were used to carry out manipulation and analysis of the data [48]. SPSS was used to analyse dwelling, occupant and appliance characterises with a unique service ID providing the link between the two software programs [49].