Featuring of Electricity Consumption Behavior towards Big-Data Applications

T. Kavitha, K. Thejaswi, C. Sushama
Dept. of Computer Science and Systems Engineering,
SreeVidyanikethan Engineering College, A. Rangampet, Chittoor (Dt), A.P, INDIA

Abstract-There is growing interest in discerning behaviors of electricity users in both the residential and commercial sectors. With the advent of high-resolution time-series power demand data through advanced metering. Large volumes of smart meter data gives opportunity for load serving entities to improve their knowledge on customers electricity consumption behavior via load profiling. This paper implements a novel approach for clustering of electricity consumption behavior dynamics.first for each individual customer symbolic aggregate approximation(SAX) to reduce the scale of the data set,and time based Markov model is applied to model the dynamics of electricity consumption, transforming the large set of load curves to several state transition matrixes. A density-based clustering technique, CFSFDP, is performed to discover the typical dynamics of electricity consumption and segment customers into different groups.

Index Terms: *Load Profiling, big data, Markov model, electricity consumption, distributed clustering, demand response.*

*****

1. Introduction

In the modern society the deregulated electricity industryhaving a new set of opportunities which are continually appearing for the participators. In many countries with liberalized markets the focus of interest is gathered in the retail side. The viability and profitability of the retailers and aggregators are dependent on the information of their consumers` demand patterns. The nations around the world have a set of goals to reduce the power of monopolic towards liberated markets on demand side. The load serving entities(LSEs) are developed in a great number[1].in order to tell about the electricity consumption patterns and realizing the personalized power managements are effective ways to enhance the competitiveness of LSEs[2]. Meanwhile smart grids have been developed in a great manner in order to build a more robust and efficient power grid has been revolutining toward two-way flow of information and power. As from the demand side the advance metering infrastructure have grown significantly useful in which gives the detail description about the demand patterns in detail used by the consumers which are related to residential and commercial sectors. AMIs allows LSEs to obtain electricity consumption data at high frequency, e.g., minutes to hours[3]. The Large volumes of electricity consumption data reveal information of customers that can potentially be used by LSEs to manage their generation anddemand resources efficiently and provide personalized service.

Load profiling refers to the formulation of representative load curves for single consumers and groups of consumers and also tells about electricity consumption behaviors of customers over a specific period, e.g., one day, can help LSEs understand how electricity is actually used for different customers and obtain the customers' load profiles or load patterns. Based on certain criteria, the consumers are grouped together in a number of classes. Each class has a representative daily load curve which is the weighted average of the curves that belong to the cluster. According to this approach, the consumers are not only distinguished in macro-categories like residential, industrial, etc., but subcategories are formatted within the macro-categories. Such a detailed analysis entails advantages for the most market entities. Load profiling enables the retailers to make lower risk settlement in spot market. Load profiles can be the basis for the design of flexible tariffs [4],nodal or customer scale load forecasting[5], demand response and energy efficiency targeting[6], and non-technical loss (NTL) detection a scheme that leads to increased profits[7].

2. Related Work

Moreover, it can be used effectively in short and medium-term load and energy forecasting. The forecasts can be done for a single consumer, for a class or even for the total demand that is served by the retailer.

The core of load profiling is clustering which are of two types Direct clustering and Indirect clustering. The Direct clustering means clustering methods are applied directly on data and in Indirect clustering the dimension reduction take place in order to reduce the size of load data.Such Indirect clustering can be categorized into two sub-categories, feature extraction-based clustering and time series-based clustering. There are large number of clustering techniques that are widely used including k-means[9], fuzzy k-means [10],hierarchical clustering [10],self-organizing maps[11],support clustering [12],subspace clustering[13],ant colony clustering[14]. Feature extraction which transforms the data in the high dimensional space into a space of fewer dimensions is often used to reduce the scale of the input data.Moreover, as electricity consumption data are essentially a time series. A variety of mature analytical methods such as discrete Fourier transform(DFT),discrete wavelet transform (DWT)[16], symbolic aggregate approximation (SAX)[17] , and the hidden Markov model (HMM)[18]have been used which are capable to reduce the dimensionality of time series and maintain some originality about electricity consumption data.The existing studies on load profiling mainly focus on individual large industrial/commercial customer, a combination of small customers, load profiles of which shows much more regularity. It should be noted that although these dynamic characteristics are always “deluged” in a combination of customers, they could be described by several typical load patterns. However, with regard to residential customers, at least two new challenges will be faced. One challenge is the high variety and variability of the load patterns. The second problem is data related to high frequency of consumption data contained in load curves in order to be analyzed and to give a better results in a effective manner.

3. Proposed Methodology

To handle these two challenges, the paper implements a time-based Markov model to formulate the dynamics of customers electricity consumption behaviors, considering the state-dependent characteristics, which indicates that future consumption behaviors would be related to the current states. This assumption is reasonable as various electricity consumption behaviors would last for different periods of time before being capable of change, as could be abstracted from historical performances. The transitions and relations between consumption behaviors, or rather consumption levels, in adjacent periods are referred to as “dynamics” in this paper. These dynamics have been modeled by Markov model in several works. However, few papers consider the dynamics as a factor for clustering. Profiling of the dynamics could provide useful information for understanding the consumption patterns of customers, forecasting the consumption trends in short time periods, and identifying the potential demand response targets[15]. Moreover, this approach formulates the large data set of load curves as several state transition matrixes, greatly reducing the dimensionality and scale.

In addition to the Markov model, this paper tries to address the "data deluge" issue in three other ways. First, applying SAX to transform the load curves into a symbolic string to reduce the storage space and ease the communication traffic between smart meters and data centers. Second, a recently reported effective clustering technique by Fast Search and Find of Density Peaks (CFSFDP) is first utilized to profile the electricity consumption behaviors, which has the advantages of low time complexity and robustness to noise points [28]. The dynamics of electricity consumptions are described by the differences between every two consumption patterns, as measured by the Kullback–Liebler (K-L) [29] distance .Third, to tackle the challenges of big and dispersed data, the CFSFDP technique is integrated into a divide-and-conquer approach to further improve the efficiency of data processing, where adaptive k-means[9] is applied to obtain the representative customers at the local sites and a modified CFSFDP method is performed at the global sites. The approach could be further applied toward big data applications. Finally, the potential applications of the proposed method to demand response targeting, abnormal consumption behavior detecting and load forecasting are analyzed and discussed. Especially, entropy analysis is conducted based on the clustering results to evaluate the variability of consumption behavior for each cluster, which can be used to quantify the potential of price-based and incentive-based demand response. The contributions of this paper are as follows:

1) Time-based Markov model is applied to formulate the electricity consumption behavior dynamics instead of the shape of daily load profiles;
2) Customer segmentation is performed by a high efficient clustering algorithm named CFSFDP which is robust to noise and need no iterations;
3) A distributed clustering framework combining adaptive k-means and CFSFDP is proposed to tackle the large and distributed data set;
4) The application of the proposed modeling method and profiling algorithm are analyzed and discussed.

The proposed methodology for the dynamic discovery of the electricity consumption can be divided into six stages, as shown in Fig. The first stage conducts some load data preparations, including data cleaning and load curve normalization. The second stage reduces the dimensionality of the load profiles using SAX. The third stage formulates the electricity consumption dynamics of each individual

ISSN: 2454-4248
156 – 159

customer utilizing time-based Markov model.

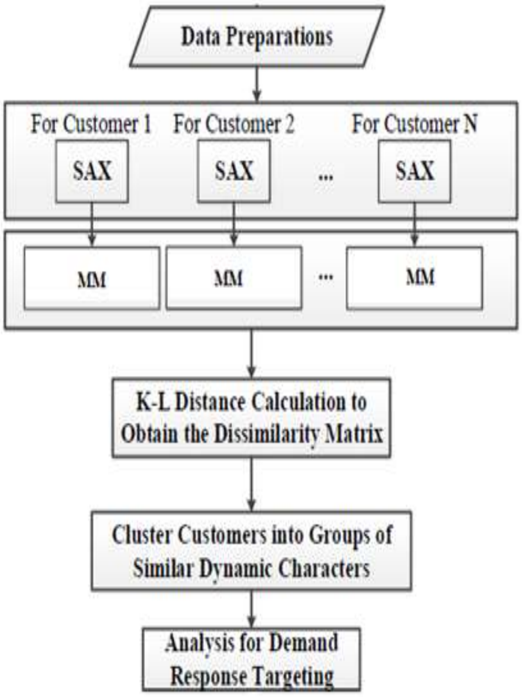

Fig.Featuring of electricity consumption behavior dynamics processes

The K-L distance is applied to measure the difference between any two Markov model to obtain the distance matrix in the fourth stage. The fifth stage performs a modified CFSFDP clustering algorithm to discover the typical dynamics of electricity consumption. Finally, the results of the analysis of the demand response targeting are obtained in the sixth stage.

4. Distributed Algorithm For Large Data Sets

The electricity consumption data is become huge for the collection of high density populated area although the time -based Markov model and SAX applied on both hands which are used to reduce for dimensionality reduction of load profile is not been effective for the large data. As, it means the electricity consumption data for the consumers will have tocome from different sites to the central site in which we apply the techniques in order to give the results. As it become more problem,time consuming and also require high cost inorder to gather the information from all sites to central state. On the other hand, the analysis and clustering of large data sets gathered from each distributed site need a very large time and memory overhead. When applying the CFSFDP, the dissimilarity matrix of all the customers should first be obtained, which accounts for most of the computation time.

5. Conclusion

The electricity consumption towards large data sets have been proposed. SAX and time-based Markov model are utilized to model the electricity consumption dynamic characteristics of each customer. A density-based clustering technique, CFSFDP, is performed to discover the typical dynamics of electricity consumption and segment customers into different groups. Finally, a time domain analysis and entropy evaluation are conducted on the result of the dynamic clustering to identify the demand response potential of each group's customers.

References:

[1] USA department of Energy, Smart Grid/Department of Energy, http://energy.gov/oe/technology-development/smart-grid,2014.

[2] I. P. Panapakidis, M. C. Alexiadis, G. K. Papagiannis, "Load profiling in the deregulated electricity markets: A review of the applications", *Proc. 9th Int. Conf. Eur. Energy Market (EEM)*, pp. 1-8, 2012.

[3] R. Granell, C. J. Axon, D. C. H. Wallom, "Impacts of raw data temporal resolution using selected clustering methods on residential electricity load profiles", *IEEE Trans. Power Syst.*, vol. 30, pp. 3217-3224, Nov. 2015.

[4] N. Mahmoudi-Kohan, M. P. Moghaddam, M. K. Sheikh-El-Eslami, E. Shayesteh, "A three-stage strategy for optimal price offering by a retailer based on clustering techniques", *Int. J. Elect. Power Energy Syst.*, vol. 32, no. 10, pp. 1135-1142, 2010.

[5] P. Zhang, X. Wu, X. Wang, S. Bi, "Short-term load forecasting based on big data technologies", *CSEE J. Power Energy Syst.*, vol. 1, no. 3, pp. 59-67, Sep. 2015.

[6] N. Mahmoudi-Kohan, M. P. Moghaddam, M. K. Sheikh-El-Eslami, S. M. Bidaki, "Improving WFA k-means technique for demand response programs applications", *Proc. IEEE Power Energy Soc. Gen. Meeting (PES)*, pp. 1-5, 2009.

[7] C. Leon et al., "Variability and trend-based generalized rule induction model to NTL detection in power companies", *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 1798-1807, Nov. 2011.

[8] R. Li, C. Gu, F. Li, G. Shaddick, M. Dale, "Development of low voltage network templates—Part I: Substation clustering and classification", *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3036-3044, Nov. 2015.

[9] K.-L. Zhou, S.-L. Yang, C. Shen, "A review of electric load classification in smart grid environment", *Renew. Sustain. Energy Rev.*, vol. 24, pp. 103-110, Aug. 2013.

[10] G. J. Tsekouras, P. B. Kotoulas, C. D. Tsirekis, E. N. Dialynas, N. D. Hatziargyriou, "A pattern recognition methodology for evaluation of load profiles and typical days of large electricity customers", *Elect. Power Syst. Res.*, vol. 78, no. 9, pp. 1494-1510, 2008.

[11] S. V. Verdu, M. O. Garcia, C. Senabre, A. G. Marin, F. J. G. Franco, "Classification filtering and identification of electrical customer load patterns through the use of self-

organizing maps", *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1672-1682, Nov. 2006.

[12] G. Chicco, I. S. Ilie, "Support vector clustering of electrical load pattern data", *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1619-1628, Aug. 2009.

[13] M. Piao, H. S. Shon, J. Y. Lee, K. H. Ryu, "Subspace projection method based clustering analysis in load profiling", *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2628-2635, Nov. 2014.

[14] G. Chicco, O. M. Ionel, R. Porumb, "Electrical load pattern grouping based on centroid model with ant colony clustering", *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1706-1715, May 2013.

[15] J. Torriti, "A review of time use models of residential electricity demand", *Renew. Sustain. Energy Rev.*, vol. 37, pp. 265-272, Sep. 2014.

[16] Y. Xiao, J. Yang, H. Que, M. J. Li, Q. Gao, "Application of wavelet-based clustering approach to load profiling on AMI measurements", *Proc. IEEE China Int. Conf. Elect. Distrib.(CICED)*, pp. 1537-1540, 2014.

[17] A. Notaristefano, G. Chicco, F. Piglione, "Data size reduction with symbolic aggregate approximation for electrical load pattern grouping", *IET Gener.Transm.Distrib.*, vol. 7, no. 2, pp. 108-117, Feb. 2013.

[18] A. Albert, R. Rajagopal, "Smart meter driven segmentation: What your consumption says about you", *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4019-4030, Nov. 2013.