International Journal of Industrial Engineering and Management (IJIEM), Vol. 4 No 3, 2013, pp. 117-122
Available online at www.iim.ftn.uns.ac.rs/ijiem_journal.php
ISSN 2217-2661

UDK: 621.311:658

Pattern recognition of Load Profiles in Managing Electricity Distribution

Adonias Magdiel Silva Ferreira
Assistant Professor of Polytechnic School – Federal University of Bahia, Graduate Program in Industrial Engineering
Rua Aristides Novis St. Federação, Salvador, Bahia Brazil, adoniasmagdiel@ufba.br

Carlos Arthur Mattos Teixeira Cavalcante
Assistant Professor of Polytechnic School – Federal University of Bahia, Graduate Program in Industrial Engineering
Rua Aristides Novis St. Federação, Salvador, Bahia Brazil, arthurtc@ufba.br

Cristiano Hora de Oliveira Fontes
Assistant Professor of Polytechnic School – Federal University of Bahia, Graduate Program in Industrial Engineering
Rua Aristides Novis St. Federação, Salvador, Bahia Brazil, cfontes@ufba.br

Jorge Eduardo Soto Marambio
Executive manager of Norsul Engineering LTD, Tancredo Neves Avenue 1632 S 1802, Trade Center Building, South Tower, Caminho das Árvores Salvador, Bahia Brazil

Received (01 November 2012); Revised (03 January 2013); Accepted (17 March 2013)

Abstract

This works presents a method of selection, classification and clustering load curves (SCCL) which is able to identify a greater diversity of consumption patterns existing in the electricity distribution sector. The method was developed to estimate the features of a sample of load curves so as to identify the consumption behaviour of a population of consumers. The algorithm comprises four steps that extract essential features of a load curve of residential users, seasonal and temporal profiles in particular. The method was successfully implemented and tested in the context of an energy efficiency program developed by a company associated to the electricity distribution sector (Electric Company of Maranhão, Brazil). This program comprised the analysis of the impact of replacing refrigerators in a universe of low-income consumers in some towns in the state of Maranhão (Brazil). Patterns of load profiles using the typing method developed were applied and the results were compared with a well known method of time series clustering already established in the literature, the Fuzzy C-Means (FCM). Based on the main features of a load profile, the analysis confirmed that the SCCL method was capable for identifying a greater diversity of patterns, demonstrating the potential of this method for better characterization of types of demand.This is an important aspect for the process of making decision in the energy distribution sector. Furthermore, a well known index (Silhouette index) was also adopted to quantify the level of uniformity within and between clusters.

Key words: *Typing load profiles; clustering; electricity sector.*

1. INTRODUCTION

A making-decision process in an organization can be analyzed and improved using several methods or strategies. Methods of Data Mining (DM) that can extract useful information from data can be used to develop making-making-decision tools so as to improve production systems and management technology. In this context, the analysis of records of power measurements in substations and at customers' homes allows administrators in the power sector to identify opportunities for improvement in the load factor and energy efficiency of the distribution system, through a relationship with the client. This information also acts as support for making-decision [1, 2].

Some works describe pattern recognition of load curves based on clustering techniques. Gerbec et al.[3]used an hierarchical clustering method using the Ward technique [4], highlighting its convenience in the quantification of groups (and patterns). Gemignani et al. [5] combined the hierarchical and non-hierarchical methods to improve the efficiency of clustering in the recognition of different demand patterns from the same level of tension.

Zalewski [6] used fuzzy logic for clustering and typification of load curves. He performed the clustering of load profiles in order to classify substations in homogeneous groups according to consumption peak. Nizar [7] combined two methods, namely, Feature Selection and Knowledge Discover in Database (KDD) to get better patterns of load demand in a distribution system [7]. Nizaralso combined Feature Selection and KDD.

Considering the methods of pattern recognition of load curves already consolidated in the literature, Pessanha [8] compared the development of some of them and highlighted that the Fuzzy C-Means (FCM) presents better quality of cohesion and distinction in problems of load curve clustering [8]. Gerbec et al. [9], Zakaria and Hadi [10], and Anuar and Zakaria [11] used the FCM method for the typification of load curves.

This work proposes a new method of selection, classification and clustering load curves (SCCL) based on a systematic extraction of features which is able to identify a greater diversity of demand patterns and also represents a potential tool for the improvement of making-decision through a better classification of heterogeneous consumption profiles in the electricity sector. The case study analyzed is an energy efficiency program carried out by the Electricity Company of Maranhão (CEMAR) (Brazil) whose aim was to analyse the impact of replacing the refrigerators of low-income consumers. Section 2 presents the SCCL method and Section 3 presents the case study and the results achieved through the application of Fuzzy C-Means and SCCL, demonstrating the potential of the latter to provide good quality information.

2. SCCL - A NEW METHOD FOR PATTERN RECOGNITION OF LOAD CURVES

The SCCL method (Figure 1) consists of two phases and uses consumption data sampled throughout the day. The first phase performs the classification of data (time series) and pattern recognition by successive iterations. The second phase performs the clustering of load curves according to the patterns recognized in the first phase.

The first phase is divided in four stages that each applies one concept (or feature) which is important for the pattern recognition in accordance with the requirements and indicators used by the electricity sector. The first three stages calculate the similarity between the load curves based on the curve with the highest peak demand (reference curve). The last stage calculates the similarity between the curves according to a criterion of seasonality.

The four stages in the first phase of the typification are described below:

1st stage: Typification by probability distribution of hourly demand.

Each curve is normalized in the interval [0; 1] dividing the hourly measurements by the peak demand curve. The consumption is quantified in the dimensionless value, referred to as power per unit (pu) [11]. The formation of groups is accomplished through the chi-square goodness of fit test which quantifies the degree of similarity between the distribution of the peaked curve and the distribution of other curves [12].

2nd stage: Typification by variation of hourly demand.

This step comprises the analysis of the correlation between the load curves belonging to the same group (groups obtained in the previous stage) and thereference curve of the group (peaked curve or curve with the highest peak demand). The formation of groups is accomplished through the t-test of correlation coefficient [13], obtained for each load curve in relation to the reference curve.

3rd stage: Typification by Load Factor (LF)

This stage comprises the analysis of similarity of LF between the curves of each group obtained at the end of the second stage and the respective reference curve. The formation of groups is accomplished through the t-test of differences in mean demands [13], obtained for each load curve in relation to the reference curve. The LF is a feature of the load curve which is calculated by dividing the mean demand by the maximum demand.

4th stage: Typification by seasonality at peak and off-peak times

The curves of each set generated in the third stage are undergo clustering by seasonal affinity which consists of calculating the median of the energy consumption before the edge time (9am to 6pm) (mA), at the edge time (6pm to 9pm) (mB) and at the empty time (without loading) (0am to 9am and 9pm to 12pm) (mC). Next, the curves are classified according to Table 1.

Table 1 Classification of patterns based on seasonal similarity.

Conditions	mB ≤ mC	mB>mC
mA ≤ mB	Type 1	Type 2
mA >mB	Type 3	Type 4

The first phase is composed of four successive stages involving different criteria. This phase is repeated several times (iterative process) to check if some prototypes (patterns) can be reassembled, and it finishes when there is convergence in the number of patterns obtained (prototypes). Thus, the number of prototypes is a result of the method itself and an initial estimate is not necessary.

The SCCL second phase (Figure 1) performs the clustering of the load curves of the initial sample using the smallest Euclidian distance (selection criteria) from the patterns obtained. The groups undergo a two step process. In the first, prototypes associated to few load curves are eliminated. The second step verifies if changes in the confidence level, used in the three first stages of the first phase, provides better quality of clustering. The metric used to qualify the clustering is the silhouette index which measures the level of cohesion and separation among the groups [14]. The median of the silhouette index of each load curve represents the General Silhouette Index (GSI).

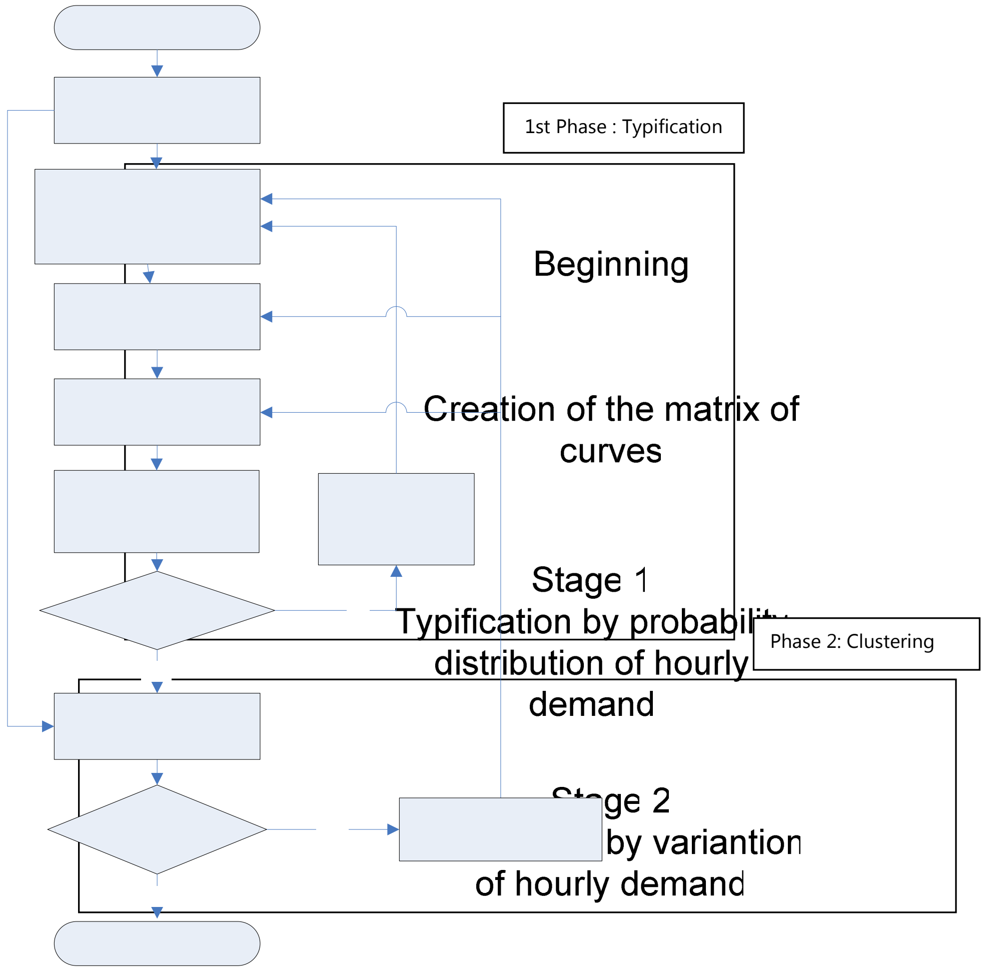

Figure 1. The SCCL method

3. CASE STUDY AND RESULTS

The SCCL method was applied to analyze possible changes in the consumption patterns in the context of an energy efficiency program implemented by the Electricity Company of Maranhão (CEMAR - Brazil) and developed during the period from November 2008 to July 2009. This program comprised the exchange of 5,250 old refrigerators for new ones in low-income communities. One sample of data with eighty load curves (old refrigerators) presented high electricity consumption while another sample of the same size with load curves after the exchange of refrigerators presented lower electricity consumption. This sample size represents an error level of 11% variation in sample means and a confidence level of 95% in the prediction of the population parameter.

The SCCL method was also used to define the number of groups to be considered in the FCM method. The confidence level adopted at the 3 first stages of SCCL first phase were 97% in both cases 1 (before the replacement of refrigerators) and 2 (after the replacement of refrigerators).

The application of SCCL method in case 1 was capable of recognizing the existence of two groups of demand profiles. The FCM method recognized different patterns of demand, with an inferior quality of clustering in relation to those identified by the SCCL method. The GSI obtained by FCM was 0.31 and by the SCCL method was 0.52 (Figure 2).

For case 2, the SCCL method also identified two groups with different consumption patterns. The FCM method recognized two patterns with similar profiles of demand which implies that FCM was capable of recognizing only one pattern of consumption in the homes after refrigerator replacement. The GSI obtained by FCM was slightly lower (0.39) compared to SCCL method which was 0.40 (Figure 3).

The value of GSI associated to the SCCL method was lower in case 2, demonstrating that after the replacement of old refrigerators with new ones the electricity consumption profiles became more similar.

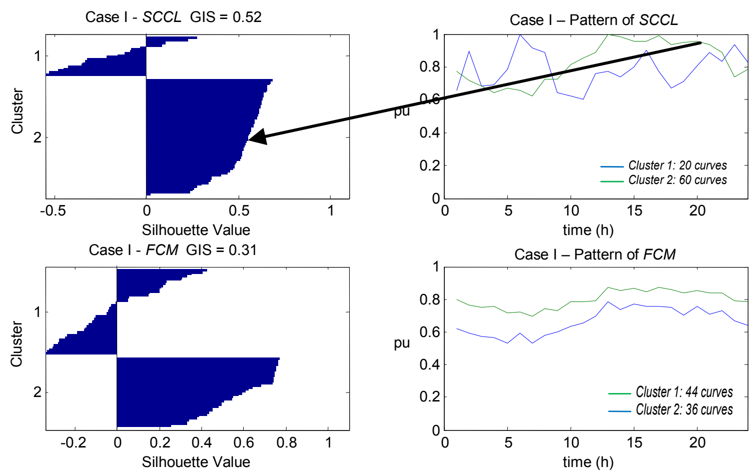

Figure 2. Indices silhouette and patterns recognized by SCCL and FCM methods (case I)

Despite the replacement of refrigerators making the load curves more homogeneous, the SCCL was still able to recognize different demand patterns, revealing the existence of two distinct types of consumers present in case 2 (after the replacement of refrigerators).

According to SCCL, the consumption averages associated to the generator group (group with the largest number of load curves) was 80.62 kWh and 57.79 kWh in cases 1 and 2, respectively, showing a 35.76% reduction in consumption and attesting the success of energy efficiency program.

The patterns recognized by SCCL associated to the minority of the refrigerators presented a less uniform profile in both cases 1 and 2. The improper use is an external factor that contributes to increase the maximum demand reducing the load factor and also increasing the heterogeneity of the profile. The patterns recognized by SCCL associated to the minority of consumers can reveal that some households open their refrigerator more often throughout the day (improper use). Even considering a reduction in the energy consumption, this information provides managers of the efficiency program with more specific information so as to focus their action to improve consumption habits.

Generally, the replacement of refrigerators makes the load curves more homogeneous. The sample consumption data measured from old refrigerators is related to equipments that exhibit inefficiencies to maintaining the internal temperature over a period. Thus, the load curves in this case present a performance more sensitive to consumption habits. On the other hand, in the case of new refrigerators, the consumption data are measured from the same type of equipment which due to its technological conditions maintains the internal temperature by a longer time and hence consumes less energy, generating load curves less sensitive to consumption habits. In the new refrigerators, the best thermal conditioning contributes to a less frequent engine work reducing the average power consumption in load curve and also the load factor.

4. CONCLUSION

This work introduces a new method for pattern recognition of load curves that uses criteria and characteristics inherent to the electricity sector and represents a potential tool for the recognition of patterns of consumption in a given population sample. Among other applications, the method of selection, classification and clustering load curves (SCCL) is capable of evaluating the impact of energy efficiency programs, promoted by the sector. This kind of program is encouraged annually by dealerships of Brazilian electric sector that imposes specific requirements to the distribution sector such as the application of at least 60% (sixty percent) of 0.50% of the net operating revenue in energy efficiency programs intended for low-income consumers. Therefore, the method proposed represents an important support tool for making-decision at management level.

The real case analysed in this work comprised an energy efficiency program carried out by the Electricity Company of Maranhão (CEMAR) (Brazil) which analyzed the impact of replacing 5,250 refrigerators in low-income consumers. The results obtained through SCCL, compared to another well-known method of clustering (Fuzzy C-Means, FCM), reveal the viability and potential of the former to recognize patterns and generate accurate information to support the implementation of efficient management actions, based on the real features of the consumer market.

The algorithm developed can be used to recognize patterns of loads curves of any equipment or consumer unit. Typical applications comprise the adjustment of the fare, capacity planning, planning the management of the electrical sector based on the demand side or increase the knowledge about the short-term demand to assess the impact of the implementation of specific projects.

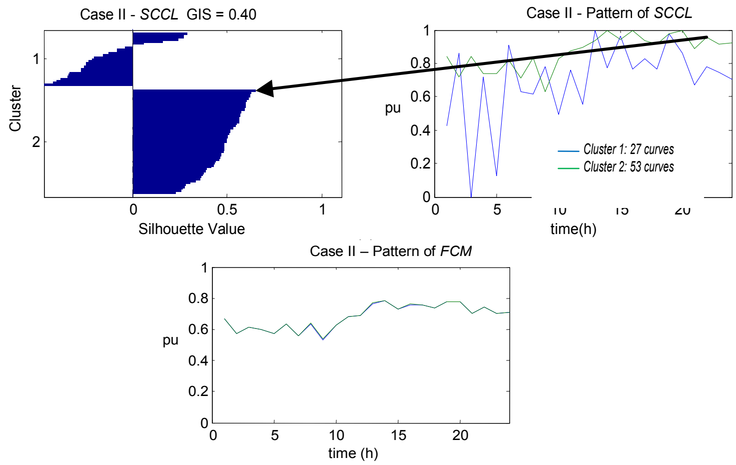

Figure 3: Silhouette and patterns using SCCL and FCM methods (case II).

5. REFERENCES

[1] J. K. Lin, S. K. Tso, H. K. Ho, C. M. Mak, K. M. Yung, Y. K. Ho. Study of climatic effects on peak load and regional similarity of load profiles following disturbances based on data mining. Electrical Power and Energy Systems 28 2006, pp. 177–185.

[2] G. Piatetsky. Data mining and knowledge discovery 1996 to 2005: overcoming the hype and moving from "universty" to "business" and "analytics", Data Mining Knowledge Discovery 15 2007, p. 99-105.

[3] D. Gerbec, S. Gasperic, I. Smon, F. Gubina. A methodology to classify distribution load profiles. Presentedat the IEEE, 2002, pp. 848-851.

[4] A. K. Jain, M. N. Murty, P. J. Flynn. Data Clustering: A Review. ACM Computing Surveys, Vol. 31 No. 3 1999, pp. 264-323.

[5] M.M.F. Geminagnani, C. C. B Oliveira, C. M. V. Tahan. Proposition and Comparative Analysis of Alternative Selection and Classification of Load Curve for Defining Types for tariff studies. Décimo Tercer Encuentro Regional Iberoamericano de Cigré – XIII ERIAC 2009, pp. 1-6.

[6] W. Zalewski. Aplication of Fuzzy Inference to Electric Load Clustering. IEEE International Conference on Power Systems, New Delhi 2006, pp. 1-5.

[7] A. H Nizar, Z. Y. Dong, J. H. Zhao. "Load profiling and data mining techniques in electricity deregulated market". Presented at the IEEE Power Engineering Society (PES) General Meeting 2006, Montreal, Quebec, Canada, June 2006, pp. 1-7

[8] J. F. M. Pessanha,R. M. C.Velásquez, A. C. G.Melo. Cluster Analysis Techniques in the Building of Load Curves Classification. XV NATIONAL SEMINARY OF ELECTRIC DISTRIBUTION (November, 2002, Salvador, Bahia). Anais. Bahia, 2002.

[9] D. Gerbec, S. Gasperic, I. Smon, F. Gubina. Determining the load profiles of consumers based on fuzzy logic and probability neural networks. IEE Proc.-Gener. Transm. Distrib., Vol. 151, No. 3, May 2004, pp. 395-400.

[10] Zuhaina Zakaria, K L Lo, Hadi Mohamad Sohod. Application of Fuzzy Clustering to Determine Electricity Consumers' Load Profiles First International Power and Energy Conference, Putrajaya, Malaysia, 2006, pp. 99-103.

[11] N. Anuar, Z. Zakaria. Cluster Validity Analysis for Electricity Load Profiling. IEEE International Conference on Power and Energy, Kuala Lumpur Malaysia, 2010,pp. 35-38.

[12] Joseph. Janes. Categorical relationships: chi-square. Library Hi Tech, Vol. 19 Issues: 3 2001, pp. 296-298.

[13] T. W. O'Gorman. A comparison of an adaptive two-sample test to the t-test, rank-sum, and log-rank tests. Communications in Statistics - Simulation and Computation, Vol. 26 1997, pp. 1393-1411.

[14] P. J. Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics. Vol. 20 1987, pp. 53-65.

Prepoznavanje shema profila opterećenja u distribuciji električnom energijom

AdoniasMagdiel Silva Ferreira, Carlos Arthur Mattos Teixeira Cavalcante, Cristiano Hora de Oliveira Fontes, Jorge Eduardo Soto Marambio

Primljen (01. Novembar 2012.); Recenziran (03. Januar 2013.); Prihvaćen (17. Mart 2013.)

Rezime

Ovaj rad predstavlja metod selekcije, klasifikacije i klasteringa krivi opterećenja (SCCL) koji može da identifikuje veću raznolikost shema potrošnje koje postoje u sektoru distribucije električne energije. Metod je razvijen kako bi se procenile osobine uzorka krivi opterećenja da bi se identifikovalo *ponašanje potrošnje potrošačke populacije. Algoritam sadrži četiri koraka koja izdvajaju najvažnije osobine krive opterećenja rezidencijalnih korisnika, sezonskih i povremenih profila posebno. Metod je uspešno primenjen i testiran u kontekstu programa energetske efikasnosti koji je razvila kompanija povezana sa sektorom distribucije električne energije (Elektronska kompanija u Maranhaou, u Brazilu). Ovaj program sadrži analizu uticaja promene frižidera u univerzumu potrošača malih primanja u nekim gradovima u državi Maranhao (Brazil). Sheme profila opterećenja koje koriste razvijeni metod uzorka primenjene su i rezultati su upoređeni sa dobro poznatim metodom klasteringa vremenskih nizova koji je već utvrđen u literaturi, poznat kao FCM (Fuzzy C-Means). Na osnovu glavnih osobina profila opterećenja, analiza je potvrdila da je SCCL metod sposoban da identifikuje veći broj različitih shema, demonstrirajući potencijal ovog metoda za bolju karakterizaciju tipova potražnje. Ovo je značajan aspekat za proces donošenja odluka u sektoru distribucije energije. Nadalje, dobro poznat indeks (silueta indeks) je takođe prilagođen kako bi kvantifikovao nivo uniformnosti unutar i između klastera.*

Ključne reči: *Uzorci profila opterećenja; klastering; sektor električne energije.*