

Clients Segmentation According to their Domestic Energy Consumption by the Use of Self-Organizing Maps

Ignacio Benítez Sánchez, Ignacio Delgado Espinós, Laura Moreno Sarrión, Alfredo Quijano López
and Isabel Navalón Burgos

Abstract—The CENIT research project GAD (Demand Active Management), partly supported by the Spanish government, has as its objective the demand side management (DSM) of domestic end-users, smoothing the peaks of energy demand, therefore enhancing the conditions of transport and distribution networks and the quality of energy delivery and service. One of the objectives in the development of the project is the classification of users according to their patterns of daily energy load profiles. Based on daily hour measures from a sample of residential users, a self-organizing map (SOM) has been trained to classify users according to a specific number of resulting patterns of daily load profiles, attached to a number of indices that define the users' energy consumption. The so-trained SOM classifier allows the characterization of future users based on their load profiles, thus estimating their energy consumption habits and potentially manageable energy.

Index Terms—Active Demand Management, Classification Techniques, Self-Organizing Maps (SOM), Knowledge Discovery in Databases (KDD).

I. INTRODUCTION

THE new specifications that arise in the energy market make necessary an approach to an effective measurement and management of the end-user energy consumption and trends, not only concerning the traditionally supervised large consumption customer, but also the medium and high energy consumption residential user, whose consumption profile depicts unbalanced patterns of peaks of energy consumption, and valley or peak-off regions where the energy demand remains unsolicited. The Demand Side Management (DSM) tools allow a more effective interaction of energy production and consumption profiles, therefore providing the end-user a valuable interface to achieve different levels of energy management.

In this context, the GAD project has been created, with the objective to optimize energy consumption among medium and low voltage users by the research and development of new tools for the DSM. The GAD acronym stands for the term "Active Demand Management" in Spanish. It is sponsored by the CDTI (Technological Development Centre of the Ministry of Science and Innovation of Spain), and financed by the

INGENIO 2010 program. The promotion of the project comes from the National Strategical Consortium of the Electrical Active Demand Management; Iberdrola Distribución Eléctrica, S.A. is leading this group, and the rest of former companies are: Red Eléctrica de España, Unión Fenosa Distribución, Unión Fenosa Metra, Iberdrola, Orbis Tecnología Eléctrica, ZIV Media, DIMAT, Siemens, Fagor Electrodomésticos, BSH Electrodomésticos España, Ericsson España, GTD Sistemas de Información, Grupo Foresis - Acceda Mundo Digital and Airzone. More information is available at its web site, <http://www.gadproject.com>.

The GAD project works on the development of the necessary tools to achieve an exhaustive tool of communication with the user, to inform about the price of electricity and its origin on an hourly basis, and to facilitate the user the tools to manage and optimize the energy consumption at home, allowing the segmentation of peaks of energy consumption due to the aggregation of domestic appliances during a limited number of hours.

One of the tasks developed as an objective of the project is the segmentation of clients according to their consumption and load profiles, with the aim to determine different consumption patterns which may be of interest from the point of view of the GAD (and the DSM) objectives, i.e.: the average consumption per day (low, medium, high), the number of consumption peak periods and when these are produced.

Self-Organizing Maps (SOM) [1] have been chosen as the tool that allows classification of users according to their load profiles, by treating each daily load profile as a vector of 24 dimensions or hourly energy measures (expressed in Wh). The SOM algorithm allows an easy identification and group formation over a 2D grid displaying the complete scenario of load profiles based on the dataset used for training, comprised of daily energy consumption profiles from a sample of 625 monitored clients of the Spanish network operator Iberdrola, along the year 2008. Each measure includes the hourly load profile, and also a number of specific indices that define the consumption pattern and the characteristics of the client. The use of these indices and the corresponding values of the prototypes of each group, allow using the trained SOM to both classify new users according to their load profiles, and estimate their energy consumption characteristics.

The versatility of SOM or Kohonen networks rely mainly on an easier interpretation of the classification, based on a 2D grid projection of multidimensional objects. The SOM training

Ignacio Benítez Sánchez, Ignacio Delgado Espinós, Laura Moreno Sarrión and Alfredo Quijano López are with the Institute of Electric Technology (ITE), Parque Tecnológico, Paterna, Valencia, Spain (e-mail: {ignacio.benitez}{ignacio.delgado}{laura.moreno}{alfredo.quijano}@ite.es). Isabel Navalón Burgos is with Iberdrola Distribución S.A., Spain (e-mail: isabel.navalon@iberdrola.es)

Manuscript received February 25, 2009.

algorithm makes use of unsupervised competitive learning [2], a type of learning procedure where the input object is compared with the weights or codebook of all the neurons, choosing the one with the highest similarity to the object. This “winning” neuron is called the Best Matching Unit or BMU. The similarity is usually computed as the minimum Euclidean distance (1), although other configurations and similarity measures might be used [1].

$$d_{\vec{x}\vec{y}} = \sqrt{\sum_{i=1}^m |x_i - y_i|^2} = \sqrt{(\vec{x} - \vec{y})^T (\vec{x} - \vec{y})} \quad (1)$$

A learning rule is then applied, which updates the weights of the BMU, and also the weights of its closest neighbors, following a formula like the one described in (2), where D is a distribution function (usually Gaussian) which decreases as the distance between the neuron and the BMU gets higher.

$$w_{i(k+1)} = w_{i(k)} + D[x_{(k)} - w_{i(k)}] \quad (2)$$

In the latest years, the SOM techniques have been applied to databases of consumption or load profiles, to serve as classification [3], [4] or prediction tools [5]. The present document describes how the SOM and other data mining techniques [6] have been applied to the dataset gathered by the GAD project, in order to accomplish the objectives of end-users classification. The methodology designed is detailed in Section II, whereas an analysis and the results obtained are explained in Section III. The Conclusion Section outlines the main observations that can be extracted from the results obtained. A References Section is also included.

II. SOM CLASSIFICATION APPLIED TO THE GAD PROJECT FOR CLIENTS' SEGMENTATION

In this paper, the SOM technique is applied on data from daily load profiles of residential users, with the objective to extract useful information on the consumption habits and potentially manageable consumption of either previously monitored or new clients. The SOM accepts for classification a vector of 24 characteristics or dimensions, being these dimensions the 24 measures of energy consumption per hour, expressed in Wh. The analysis has been performed on a sample of 625 clients that have been monitored throughout a year. These data have been divided in subsets according to seasonality (four seasons) and the type of day (working / non-working day), obtaining eight subsets of data. Each of the samples is structured as an object with 24 measures of energy consumption, with a set of qualitative and quantitative indices added. These indices describe the characteristics of the user and the load profile, and can be divided in three categories:

- Information on the client's contract and geographical situation. Provided by the Spanish operator Iberdrola, these variables comprise: a unique code to identify the user, the climate area (from a number of eight different climate areas in Spain), and five qualitative variables, computed from the consumption values of the electricity bill. Three of them determine the level of energy consumption throughout the year (High consumption,

Medium consumption and Low consumption), and two qualitative variables define the increment of energy consumption in Winter (Winter peak) or in Summer (Summer peak). The five variables take only two possible values: zero (not applicable) or one (applicable).

- Quantitative information on the load profile. A number of indices is computed and added to each sample. The information describes the load profile in a set of variables: daily average consumption per hour (in Wh), maximum and minimum values of energy consumption per hour, the time of the maximum consumption, the peak/valley mathematical relation, the number of consumption peaks, and an index of peaks pattern, obtained as a six-bits digit where each bit stands for a different section of the day (early morning – breakfast time – morning – lunchtime – afternoon – dinner / evening). If a bit is set to one, it means that at least one consumption peak period is given at that section of the day.
- Quantitative and qualitative information obtained from questionnaires. The sample of clients were requested to answer a questionnaire regarding the type of home, number of residents, appliances owned, and consumption habits. The results have been computed to assign values to certain indices. These indices comprise the active management possibilities (percentage of high, medium and low penetration appliances at home, estimation of potentially manageable power, expressed in kWh per week); the familiar and residential characteristics (number of children, characteristics of the building); and the concern with environmental issues, the renewable energies and the knowledge about the current fee of electrical consumption.

With the objective of having the same number of samples per client, and in order to avoid an overtraining of the SOM, a preprocessing of the data is carried in form of a clustering algorithm, the *K-means* [7]. Clustering algorithms allow the division of objects of a dataset into a number of groups, according to similarity among them in a mathematical sense, usually computed as the minimum Euclidean distance, although other quantitative and qualitative measures have been described and could be applied [8]–[12]. From the subset of data segregated by the season and the type of day, a clustering algorithm is applied to obtain only two prototypes of each client, therefore the training dataset is reduced to a maximum number of objects of twice the number of clients, achieving an equal representation of all the clients from the sample, and preventing an overtraining of the SOM.

Since the level of energy consumption is one of the main objectives of the classification task, the data have not been normalized, as considered to be objects of 24 measures of the same variable (energy consumption, in Wh) [1]. However, analysis with normalized data have been performed, in two ways: data normalized between zero and one (referenced to the highest value of consumption at each case), and data normalized to null average and unit variance. Both results improve the classification concerning the shape of the profile, but fail to properly classify clients according to their level of energy

consumption (in low, medium or high energy consumption users per day); therefore non-normalized data have been used. The results obtained allow a classification based on the two perspectives: shape of the profile (number of peaks and when these are formed) and average consumption per day (low–medium–high consumption profile users).

Once the SOM has been trained, a *K-means* clustering algorithm is applied on the map, in order to group the different regions in a specific number of clusters or prototypes. The number of clusters has been chosen based on initial tests and values of the Davis-Boulding (DBI) index [3], and an index which computes the number of users whose load profiles are assigned to a unique cluster at least in more than 50% of its available samples, giving thus a measure of the accuracy in classification. An optimal number of clusters was obtained to be between 8 and 11 clusters. A number of ten clusters has been chosen by default.

The so trained SOM classifier allows a correspondence between new users' load profiles, and one of the obtained patterns, thus defining the users' energy consumption based on a single measure. The information of a classified client is related to the indices attached to the pattern that his consumption best fits to. The values of the indices for each obtained pattern or prototype are computed as the average or modal values of the indices' values from all the users assigned to that cluster.

The final step of the analysis is the classification of the total subset of data with the trained SOM, yielding a membership to one of the obtained clusters for each sample or load profile. The assignment of a client to a pattern or prototype is made choosing the cluster with the highest number of the user's load profiles assigned. The ten different patterns of consumers obtained, allow identifying segments of clients who meet the target of the GAD project: high average daily consumption, high number of peaks per day, a specific peaks pattern, and a high value of potentially manageable power of the electrical appliances at home, among other features. An example of this methodology is described in the following Section.

III. ANALYSIS AND RESULTS

The following analysis describes the results of the described SOM methodology applied to the data of 625 clients that are spread out all over Iberdrola electricity network in Spain, monitored daily through the season of Summer, 2008, on working days. A total of 37756 load profiles are available for training the SOM.

The SOM training algorithm implements a 4-fold cross-validation [13], where the training set has been divided in four mutually exclusive subsets, and the SOM has been trained four consecutive times. Each time, three subsets have been used for training and the remaining one has been used to validate the training. At the end of the training procedure, the SOM has been trained and validated with four different subsets of data.

The results can be seen in Table I. The set of variables listed are obtained from the information available of both the load profile and the consumer, as was detailed in Section II. The values are provided as quantitative, qualitative or percentage

values, according to the nature of the information gathered. The variables are the following:

- Number of peaks: average value of the number of consumption peaks detected at each load profile assigned to the prototype.
- Hourly average consumption per day: average value of all the average values of the load profiles assigned to the prototype.
- Peaks pattern: qualitative index that defines the shape of the load profile, or when consumption peaks are produced, dividing the duration of the day in six sections, as previously described in Section II. The values are obtained as the modal values per day section of all the load profiles assigned to the prototype.
- Maximum consumption per day: the measure of maximum energy consumption per hour, in Wh, obtained from the prototype.
- Hour of maximum consumption: obtained from the prototype.
- Minimum consumption per day: the measure of minimum energy consumption per hour, in Wh, obtained from the prototype.
- Peak/valley relation: obtained from the prototype, as the division of maximum and minimum energy consumptions (adimensional).
- Cost knowledge: a qualitative index that refers to the knowledge of the user concerning the electricity bill. The value is obtained as the average of all the customers at each prototype.
- Environment: a qualitative index that refers to the knowledge of the user concerning the environmental issues related to the energy resources and electricity production. The value is obtained as the average of all the customers at each prototype.
- Family: a qualitative index that refers to the characteristics of the home, and the number of residents. The value is obtained as the average of all the customers at each prototype.
- Low, Medium and High penetration percentages: these three indices are a measure, between 0 and 1, of the percentage of domestic appliances owned which are considered of a high penetration or presence (TV set, washing machine, microwave oven, fridge and oven), medium (DVD reader, computer, dishwasher, music player, vitroc ceramic and printer), and low (air conditioning, water heater, dryer and electric heating). Typically, the percentage of appliances of low penetration determine to a great extent the potential of manageable energy consumption.
- Manageable consumption: this a quantitative index obtained from the estimated energy consumption of the amount of appliances considered manageable, obtained from the answers of each of the customers assigned to the prototype. The value is given in kWh/week.
- High, Medium and Low consumption variables: a value is given in percentage of customers of High, Medium and Low consumption per year, as obtained from the information of the electricity bill (see Section II).

TABLE I: Indices obtained for each prototype, numbered from 1 to 10, along with the percentage of clients assigned.

Variable	P. 1	P. 2	P. 3	P. 4	P. 5	P. 6	P. 7	P. 8	P. 9	P. 10
Number of peaks	6,93	4,45	7,25	5,49	2,99	5,66	6,37	4,56	4,30	5,87
Hourly average consumption per day (Wh)	734,07	706,65	1754,82	399,48	104,42	515,95	933,37	280,88	189,91	363,17
Peaks pattern	000111	100001	000111	000111	000001	000111	000111	000101	000101	011101
Maximum consumption per day (Wh)	1486,94	1029,67	2246,22	621,61	152,60	904,39	1551,01	425,22	312,39	667,63
Hour of maximum consumption	16	23	23	23	23	22	22	23	23	12
Minimum consumption per day (Wh)	254,76	486,63	590,51	191,01	86,52	250,19	464,59	151,18	115,51	173,15
Peak/valley relation	5,83	2,11	3,80	3,25	1,76	3,61	3,33	2,81	2,70	3,85
Cost knowledge	0,98	0,93	0,75	0,93	0,92	0,90	0,93	0,90	0,84	0,85
Environment	2,29	2,21	1,88	1,99	1,95	1,97	2,07	2,06	1,98	1,90
Family	1,88	1,62	1,88	1,83	1,48	1,84	2,07	1,72	1,43	1,69
Low penetration appliances (%)	0,26	0,28	0,41	0,19	0,17	0,26	0,35	0,16	0,15	0,15
Medium penetration appliances (%)	0,66	0,48	0,44	0,58	0,40	0,60	0,60	0,54	0,42	0,46
High penetration appliances (%)	0,94	0,90	0,88	0,90	0,84	0,93	0,90	0,89	0,88	0,87
Manageable consumption (kWh/week)	33,41	31,17	51,48	25,58	24,08	31,49	33,28	25,68	22,02	29,03
High consumption (%)	57,14	89,66	87,50	34,71	9,04	49,32	77,78	11,63	6,35	17,31
Medium consumption (%)	38,10	6,90	12,50	55,37	37,29	36,99	22,22	68,37	41,27	69,23
Low consumption (%)	4,76	3,45	0,00	9,92	53,67	13,70	0,00	20,00	52,38	13,46
High Consumption Off-Peak(%)	2,38	3,45	0,00	0,83	0,00	1,37	0,00	0,93	0,00	0,00
High C. Summer peak (%)	11,90	17,24	25,00	3,31	1,13	13,70	18,52	3,26	1,59	1,92
High C. Winter peak (%)	16,67	37,93	37,50	15,70	6,78	26,03	25,93	6,51	4,23	9,62
High C. Winter and Summer peak (%)	26,19	31,03	25,00	14,88	1,13	8,22	33,33	0,93	0,53	5,77
Medium Consumption Off-Peak(%)	0,00	0,00	0,00	3,31	2,26	0,00	3,70	0,93	1,06	1,92
Medium C. Summer peak (%)	7,14	0,00	0,00	11,57	3,39	4,11	7,41	8,84	6,88	9,62
Medium C. Winter peak (%)	14,29	6,90	12,50	19,83	20,90	17,81	7,41	31,16	20,63	36,54
Medium C. Winter and Summer peak (%)	16,67	0,00	0,00	20,66	10,73	15,07	3,70	27,44	12,70	21,15
Low Consumption Off-Peak (%)	0,00	0,00	0,00	0,83	3,39	0,00	0,00	0,47	3,17	3,85
Low C. Summer peak(%)	2,38	0,00	0,00	2,48	7,34	2,74	0,00	2,33	7,94	1,92
Low C. Winter peak (%)	2,38	3,45	0,00	3,31	25,99	6,85	0,00	8,84	16,40	1,92
Low C. Summer and winter peak (%)	0,00	0,00	0,00	3,31	16,95	4,11	0,00	8,37	24,87	5,77
B3 (%)	23,81	34,48	50,00	26,45	22,60	26,03	37,04	16,74	25,40	17,31
B4 (%)	14,29	10,34	12,50	12,40	14,12	9,59	11,11	11,63	10,58	7,69
C1 (%)	0,00	13,79	0,00	6,61	18,64	10,96	0,00	14,88	16,40	26,92
C4 (%)	14,29	6,90	25,00	4,13	2,82	9,59	18,52	3,26	2,12	1,92
D1 (%)	2,38	0,00	0,00	5,79	4,52	5,48	0,00	12,56	7,41	21,15
D2 (%)	2,38	6,90	0,00	7,44	10,17	2,74	3,70	8,84	10,58	3,85
D3 (%)	38,10	24,14	12,50	28,10	18,08	31,51	29,63	22,33	15,34	11,54
E1 (%)	4,76	3,45	0,00	9,09	9,04	4,11	0,00	9,77	12,17	9,62
Customers assigned (%)	3,52	2,72	1,44	7,52	25,28	9,12	0,64	25,28	21,76	2,72

K-means, non-normalized data, 10 clusters

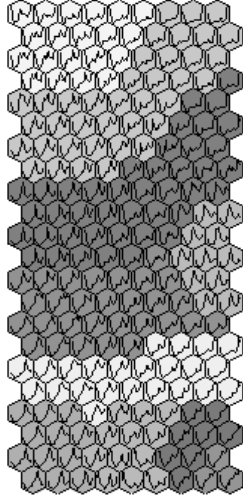


Fig. 1: Clustering algorithm applied on the trained SOM.

- High, Medium and Low Consumption profiles with Off-Peak, Summer peak, Winter peak and both Summer and Winter peaks energy consumption: these qualitative variables, as described in Section II, define the consumer's yearly energy consumption, and the presence or absence of increments due to seasonality. The values are given as percentages of the total amount of clients assigned to the prototype.
- B3, B4, C1, C4, D1, D2, D3 and E1: these are the codes for the eight different climate areas where users belong to. The code indicates the characteristics of Winter, with a letter, and in Summer, with a number, e.g., D3 defines an area with harsh Winter and quite hot Summer. For each variable, the amount of users at each prototype belonging to each area is indicated as a percentage.
- Customers assigned: the amount of customers assigned to each prototype is given, in percentage (%).

Ten clusters, each with a prototype and a set of indices attached, have been identified. The ten clusters can be located on the SOM, as can be seen in Fig. 1, whereas the prototypes that define the different energy consumption patterns can be seen in Fig. 2. As can be seen, the classification made by the SOM is influenced by the shape of the load profile and the level of energy consumption. This classification, therefore, allows to identify groups of consumers which are of great interest from the point of view of the GAD project objectives, i.e., groups of consumers with a high level of energy consumption, a high number of peaks of energy consumption per day, a high value of the peak/valley relationship, and a high value of the estimated index of potentially manageable consumption. In this analysis of the Summer season, working days, prototypes 1, 3 and 7 and the clients assigned to each one are the groups which best fit with these objectives (see Table I).

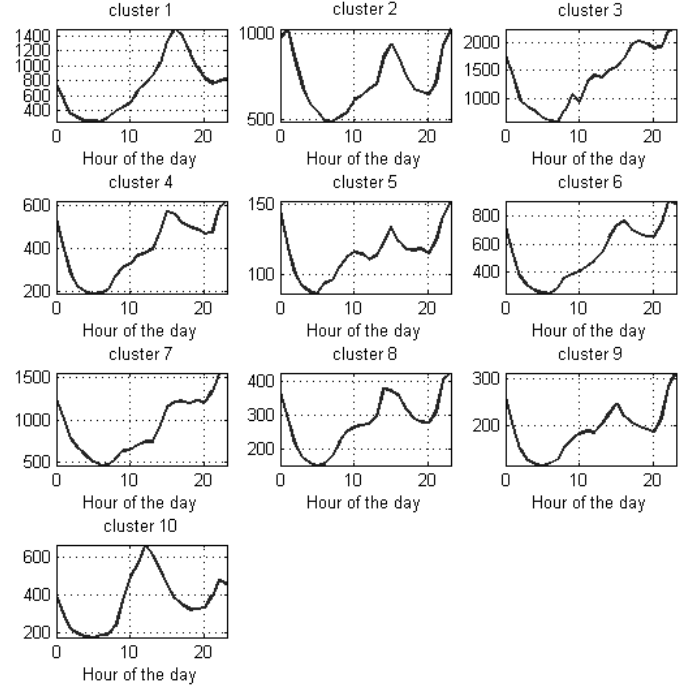


Fig. 2: Prototypes obtained from the SOM classification.

IV. CONCLUSION

The SOM has been trained with a sample of clients from the different regions of Spain. The indicators attached to the prototype's load profile serve also as an estimator of the characteristics of new users, provided their load profiles are available. New users can thus be classified and an estimation of its characteristics according to energy consumption can be obtained.

A correlation can be observed from the results, among the variables that define the consumers' energy consumption (amount of energy consumed per day, energy consumption profile for the year, number of peaks) and the variables of low, medium, and high penetration appliances and the estimation of manageable energy consumption. The three identified prototypes of interest are integrated by users of mainly Medium or High energy consumption, with a high number of peaks and daily average energy usage, and high values of low penetration appliances and manageable consumption.

The influence of climate regions can also be observed, mainly in prototype three, formed in a 50% of customers from the climate area B3, a code that indicates mild Winter and quite hot Summer. Moreover, all the users assigned to prototype 3 belong to climate areas that end in "3" or "4", numeric codes which designate very hot Summer seasons.

More conclusions can be obtained from these results and the analyses performed on the remaining subsets of data (Winter, Autumn and Spring, working and non-working days). As can be extracted from the prototypes' load profiles, most of the prototypes for working days have the highest consumption peaks at night, between 21 and 24 hours (see Fig. 2). The Analysis performed on the data from Summer season, non-working days, yields a higher number of consumption pro-

totypes where the peaks of maximum energy consumption are mostly produced at a time frame between morning and afternoon (from 10 to 16 hours). Similar results have been obtained in this sense when analyzing the other seasons.

Another interesting conclusion is the fact that there is a relation between the hourly average consumption per day and the level of annual energy consumption. Most of the consumers with a low annual consumption belong to clusters with low hourly average consumption and vice versa.

This paper details the results of the analysis on the subset of data from Summer season, working days. The remaining seven subsets have also been analyzed, and a comparison study among seasons and type of days has been made. One of the conclusions obtained indicate that up to the 80% of customers' energy consumption habits can be described with 4 or 5 prototypes per season and type of day. Concerning this analysis for the Summer season, for instance, the prototypes defining the majority of consumers would be prototypes 5, 8 and 9, which cover the 72,32% of the sample of monitored customers, who are mostly of Medium and Low yearly consumption profile (see Table I). Despite this fact, the search for small groups of high consumption and potentially manageable energy usage consumers has proven to be appropriate to improve in energy savings and efficiency, since the highest percentage of collective response is obtained from these small groups of customers [14].

REFERENCES

- [1] T. Kohonen, *Self-Organizing Maps*. Springer, 2001.
- [2] S. Haykin, *Neural Networks, a Comprehensive Foundation*. MacMillan, 1994.
- [3] G. Chicco, R. Napoli, and F. Piglion, "Comparisons Among Clustering Techniques for Electricity Customer Classification," *IEEE TRANSACTIONS ON POWER SYSTEMS PWRS*, vol. 21, no. 2, p. 933, 2006.
- [4] S. Verdu, M. Garcia, C. Senabre, A. Marin, and F. Franco, "Classification, Filtering, and Identification of Electrical Customer Load Patterns Through the Use of Self-Organizing Maps," *Power Systems, IEEE Transactions on*, vol. 21, no. 4, pp. 1672–1682, 2006.
- [5] A. Lendasse, M. Cottrell, V. Wertz, and M. Verleysen, "Prediction of electric load using Kohonen maps-Application to the Polish electricity consumption," in *American Control Conference, 2002. Proceedings of the 2002*, vol. 5, 2002.
- [6] J. Jackson, "Data mining: A conceptual overview," *Communications of the Association for Information Systems*, vol. 8, pp. 267–296, 2002.
- [7] A. K. Jain and R. C. Dubes, *Algorithms for clustering data*. New Jersey: Prentice Hall, 1988.
- [8] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, 1981.
- [9] D. Barabará, Y. Li, and J. Couto, "Coolcat: an entropy-based algorithm for categorical clustering," in *CIKM '02: Proceedings of the eleventh international conference on Information and knowledge management*, (New York, NY, USA), pp. 582–589, ACM Press, 2002.
- [10] S. Guha, R. Rastogi, and K. Shim, "Rock: A robust clustering algorithm for categorical attributes," *icde*, vol. 00, p. 512, 1999.
- [11] V. Ganti, J. Gehrke, and R. Ramakrishnan, "Cactus - clustering categorical data using summaries," in *KDD '99: Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, (New York, NY, USA), pp. 73–83, ACM Press, 1999.
- [12] J. V. de Oliveira and W. Pedrycz, eds., *Advances in Fuzzy Clustering and its Applications*. John Wiley & Sons, Ltd., 2007.
- [13] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in *INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE*, vol. 14, pp. 1137–1145, LAWRENCE ERLBAUM ASSOCIATES LTD, 1995.
- [14] A. Faruqui, R. Hledik, S. Newell, and J. Pfeifenberger, "The Power of Five Percent," *The Electricity Journal*, vol. 20, no. 8, pp. 68–77, 2007.