Annales UMCS Informatica AI 4 (2006) 161-168

Annales UMCS
Informatica
Lublin-Polonia
Sectio AI

http://www.annales.umcs.lublin.pl/

The use of self-organising maps to investigate heat demand profiles

Maciej Grzenda*

Faculty of Mathematics and Information Science, Warsaw University of Technology, Pl. Politechniki 1, 00-661 Warszawa, Poland

Abstract

District heating companies are responsible for delivering the heat produced in central heat plants to the consumers through a pipeline system. At the same time they are expected to keep the total heat production cost as low as possible. Therefore, there is a growing need to optimise heat production through better prediction of customers needs. The paper illustrates the way neural networks, namely self-organised maps can be used to investigate long-term demand profiles of consumers. Real-life historical sales data is used to establish a number of typical demand profiles.

1. Introduction

Modern utility companies transmitting heat to the consumers face new challenges. The development towards deregulated and more competitive energy markets means that many consumers may decide to resign from district heating services based on central heat sources and replace them with decentralised or individual boiler systems. In order to tackle this issue, utility companies have to improve service quality and minimise the price at the same time. For a broad overview of district heating see [1].

It is worth emphasising that the most significant part of district heating cost is usually the cost of producing the heat. Thus, by optimising heat supply significant cost reduction can be obtained. However, the latter objective can not be fulfilled without detailed analysis of consumers' demand profiles.

The work investigates yearly heat demand profiles and continues previous research on power demand prediction [2,3]. The objective is to devise a set of typical heat demand profiles that would correspond to typical groups of consumers. By obtaining such predicted yearly demand profiles, long-term optimisation of heat supply can be attempted. Still, the investigation of daily load profiles is also required to answer the needs of short-term optimisation.

E-mail address: grzendam@mini.pw.edu.pl

The paper presents the way self-organising maps [4,5] can provide valuable insight into typical heat demand profiles. The information available in the sales database system has been retrieved and transformed into historical data. The time series [6] of sales volume for each consumer has been obtained. As the heat sales strongly depends on space heating demands it varies in accordance with time of the year and average temperatures.

Finally, self-organised maps have been used to build a network of neurons representing typical sales profiles. The work concentrates on the application of neural networks to investigate these typical heat sale patterns.

The remainder of this paper is organised as follows:

– section 2 describes in details the data set of the problem,
– neural networks-based approach is outlined in section 3,
– the results of the computations are discussed in section 4,
– section 5 summarises the whole work.

2. Data set

The problem analysis is based on the data set from one of the Polish district heating utilities. The sales information from the billing system for the period of last four years has been used. For each consumer average monthly heat sale has been calculated. The heat sales information is the total heat consumption measured for each consumer separately. It is worth noting here that sales information is available for every single consumer. At the same time information regarding the type of a building, its technical attributes, number of rooms incl. rooms for rent and its thermal properties are available to different extent and for some consumers only. Therefore, to build up preliminary prediction models sales data should be used first and then for selected consumers with the same overall demand profile, detailed prediction models can be devised and tested.

It is worth emphasising that diverse thermal energy needs of consumers are answered by district heating system. Not only do they include space heating necessary to ensure human comfort, but also domestic hot water and industrial needs. Thus, while prevailing majority of demand is due to space heating requirements, other factors may strongly affect demand profiles and result in different demand profiles for different categories of consumers. Still, if the weather conditions were the only factors to influence heat demand, normalised profiles of different consumers would be the same.

Proper identification of consumers' needs in long term is necessary for overall heat demand prediction. Without such prediction, neither volume of sales nor volume of heat supply can be estimated. However, a district heating utility must assess both of them to evaluate the share of heat sources with different costs in the overall sales of heat. Moreover, in order to predict the impact of connecting new consumer to the system, its demand profile must be taken into account. The latter profile is supposed to answer the question whether the impact of connecting for instance a new hospital to the system is different in terms of normalised profile than the impact of connecting a new hotel.

In addition, the construction of prediction models can be treated as a task of creating a number of distinctive models for different consumers groups not to try to capture the behaviour of school, hotel, office, family house and restaurant in the same model. All the same, the latter objective can not be fulfilled without understanding of typical demand profiles and their relation to consumer categories.

After detailed analysis $N = 616$ sales profiles have been obtained. Every sales profile of consumer i, $i = 1,\ldots,616$ is represented by a vector $S_i = (s_{i,1},\ldots,s_{i,12})$, where $s_{i,m}$ denotes the average heat sale to consumer i during month m. Months are indexed in a traditional way i.e. 1 stands for January.

Furthermore, each sales profile has been normalised so as to obtain average demand profile $D_i = (d_{i,1}, d_{i,2},\ldots,d_{i,12})$ out of original consumer profile S_i as follows:

$$d_{i,k} = \frac{s_{i,k}}{\sum_{m=1}^{m=12} s_{i,m}}, \quad k = 1,\ldots,12\,. \tag{1}$$

Thus, $N = 616$ normalised demand profiles have been obtained. These correspond to the average yearly consumption of heat at each consumer. Some of demand profiles are depicted in Fig. 1.

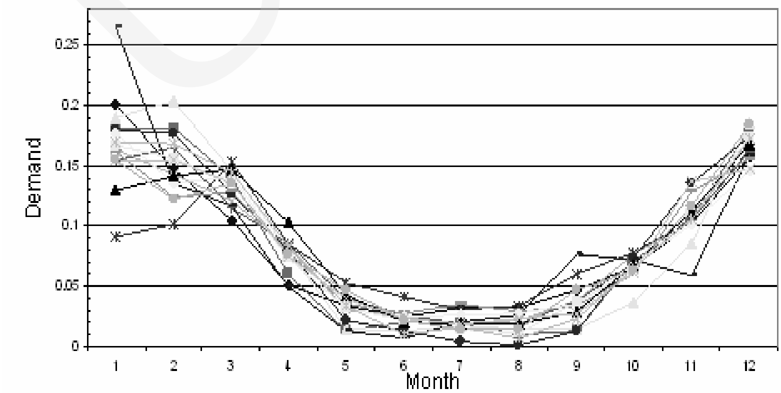

Fig. 1. Sample demand profiles Di of different consumers

While, in general, demand profiles reflect space heating needs resulting from average air temperatures, one can observe that minimal demand is reached by the consumers in the period between May and August. Similarly, peak heat consumption is attained between December and March. In other words, the exact time location of minimal and maximal heat consumption significantly varies among consumers. What is of outstanding importance even normalised monthly heat sales can differ as much as 300% among sample consumers randomly selected from the whole data set. These statements are further supported by statistic data analysis.

Fig. 2 depicts standard deviation of monthly demand $D_{i,m}$, $i = 1,\ldots,N$. It can be easily seen that the diversity among consumers in terms of heat demand reaches its peak in August, December, January and February and remains relatively low during remaining part of the year. This diversity may be due to the fact, there are many hotels and guest houses with rooms for rent among heat consumers. As a consequence, demand for heat is increased during high season i.e. in December, January, February, July and August. Therefore, whether one average demand profile could be used to approximate demands of all the consumers remains an open issue. However, if a number of different average load profiles could be devised, then they should differ most during the months listed above i.e. during the months with the largest standard deviation of $D_{i,m}$, $i = 1,\ldots,N$.

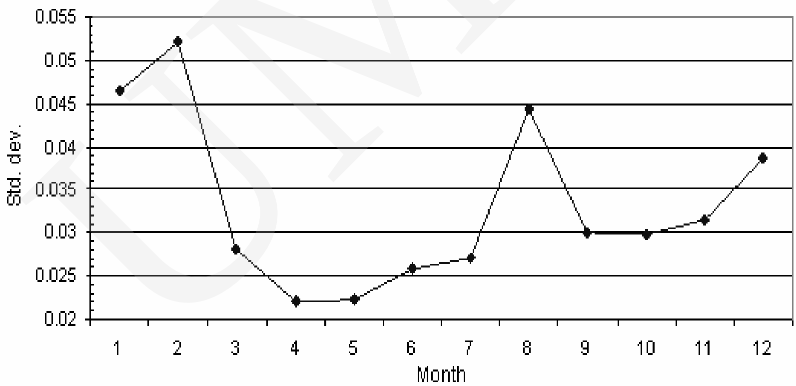

Fig. 2. Standard deviation of consumer's average demand during different months of the year

3. Neural networks and demand profile identification

While a separate demand profile D_i for each consumer could be devised, the problem remains to identify common profile groupings. The latter groupings are necessary for further market analysis, for instance to evaluate the impact of new consumers on the district heating system in view of prospective heat demand. The question is whether the observed variety of demand profiles can be explained by the existence of a number of diverse demand profiles among consumers.

In order to address this issue, self-organising neural networks have been applied. The purpose of using self-organised neural networks, namely self-organised maps (SOMs) [5], is to provide measure of locating demand pattern groupings i.e. typical profiles $\hat{D}_k$, $k \ll N$ that would explain diversity in the population of D_i, $i = 1,\ldots,N$ demand profiles. The following solutions have been applied:

- two-dimensional square lattice of J neurons has been applied,
- demand profiles D_i, $i = 1,\ldots,N$ have been applied as input patterns.

Therefore, each neuron j located in the lattice is represented by a weight vector

$$w_j = \left[w_{j,1}, w_{j,2}, .., w_{j,12} \right]^T, \quad j = 1,2,\ldots,J \;. \tag{2}$$

During the learning process neurons are tuned to the input patterns. Moreover, the weight update algorithm tunes the weights of the winning neurons and some of its spatially close neighbour neurons. Thus, topographical map of input space is being created [4]. In other words, as a result of this cooperative process, both input pattern groupings can be identified and their relation. Therefore, SOM-based analysis is sometimes treated as a generalised form of standard Principal Component Analysis (PCA) [7].

4. Results

A number of computation series have been performed, using different weight update algorithms. Not only standard Winner-Takes-All (WTA) [4,5], but also Winner Takes All with Conscience (CWTA) [4] and Neural Gas (NGAS) [8] algorithms have been used to tune neuron weights. The latter algorithm aims to overcome deficiencies of standard WTA algorithm and has been used to construct self-organised maps described below. Among these deficiencies of standard WTA weight update algorithm, overrepresentation of regions with low input density is not of least importance [4].

Distinctive features of NGAS algorithm are:

- the neighbourhood function defined as follows:

$$G\left(j, D_i\right) = \exp\left(-\frac{m(j)}{\lambda}\right), \quad j = 1,\ldots,J \tag{3}$$

- where $m(j)$ defines the index of neuron j in the sequence of neurons ordered in accordance with growing distance of neuron j from pattern D_i i.e. $\left\| w_j - D_i \right\|$.
- a number of neurones are updated in each step, however due to the $G\left(j, x\right)$ definition it is enough to update only first K neurons in the sequence. The latter simplification results in substantial reduction of necessary computations comparing to standard algorithm.

In all cases, for each neuron j, $j = 1,\ldots,J$ in the lattice, at the end of tuning phase, the winning count Y_j has been calculated in the following manner:

$$Y_j = \sum_{i=1}^{N} \mathrm{eval}\left(i, j\right) \tag{4}$$

while

$$\mathrm{eval}(i,j)=\begin{cases}1 & \text{if } \left\|D_i-w_j\right\|=\min_{k=1,\dots,J}\left\|D_i-w_k\right\| \\ 0 & \text{otherwise}\end{cases} \tag{5}$$

Thus $Y_j : j=1,\dots,J$ denotes the number of input patterns that are the most similar to the weight vector w_j in terms of Euclidean distance.

For 49 out of $J = 100$ neurons, $Y_j = 0$. On the other hand, a number of neurons wins frequently. Fig. 3 depicts weight vectors $w_j : Y_j > 20$, normalised in the same manner as demand profiles. Each profile corresponds to a single neuron. Its (x,y) location on the neuron lattice is provided in the figure legend.

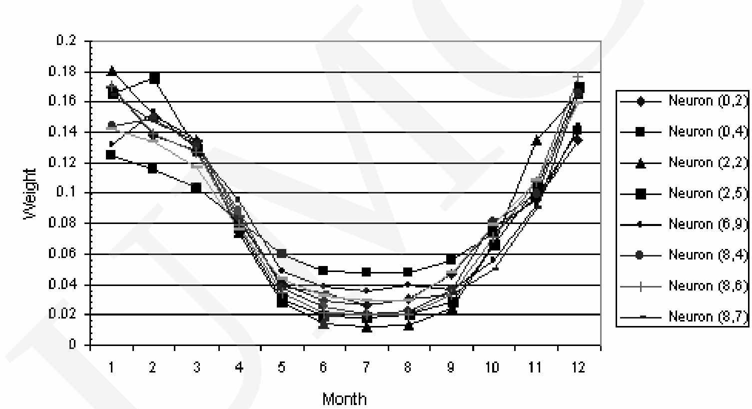

Fig. 3. Demand profiles represented by weight vectors of the winning neurons

The obtained results show that a number of distinctive demand profiles providing centroids of input consumer's profiles have been identified. This suggests that the observed diversity in the input patterns can be explained by underlying typical demand profiles. In addition, the fact that some 50% of neurones remain inactive, in spite of using NGAS algorithm that promotes diversity among neurones, points that significant simplifications can be made. In other words, when dealing with heat demand prediction, it is enough to concentrate on a limited number of consumers demand patterns.

The profile groupings identified in the form of weight vectors reflect the diversity of sample demand profiles illustrated in Fig. 1. In particular, one may notice that peak consumption is reached in January, February or December, depending on the resulting $\hat{D}_k$ profile considered. Thus, typical demand profiles obtained from a self-organising map reflect diversity of original demand profiles.

In order to investigate the relation between consumer category and its typical demand profile, additional computation series have been performed. For each neuron j the total number of closest consumers' profiles $Y_j = \sum_{c=1}^{C} Y_j(c)$ has been calculated. $c = 1,\dots,C$ denotes the index of a consumer category and C denotes the total number of categories used. Sample categories are family houses (FH), commercial buildings (C), hotels (H), hospitals (HS). The following conclusions have been made when analysing the number $Y_j(c)$ for the neurons of the lattice:

- hospital demand patterns are closest to the two neighbouring neurons on the lattice,
- more than 80% of hotel consumers can be related to just 4 neurons grouped in a distinctive spatial cluster on the lattice,
- also commercial buildings form a spatial cluster of 6 neurons on the lattice, however some 35% of these consumers can not be related to this cluster,
- demands of family houses are very different, the neuron with the highest win rate for FH consumers represents only 8.9% of all FH consumers.

As a consequence, the SOM-based demand analysis suggests that some groups of consumers, like hospitals, hotels and to some extent commercial buildings share similar demand profiles. On the other hand, some new subcategories for FH consumers would need to be established. Still, one may investigate different subgroups of consumers associated with the neurons of the lattice and take into account their spatial relation to other consumers. For instance, the fact some FH resemble hotels may be due to the fact that some of these consumers offer many rooms for rent, thus increase their demand for heat significantly during high season. Consequently, the context map created by assigning $Y_i(c)$ values to the neurons helps understand the relation between different consumer classes. In addition, it can be used by the marketing department to analyse and select the best matching demand pattern for a new consumer basing on its category and other attributes. As a consequence, the impact of a new consumer on the district heating system in terms of averaged demand profile can be evaluated and used for overall demand and supply analysis.

5. Conclusions

The paper shows the way the SOM neural networks can provide for time-series data analysis. Results of the computation suggest that diversity of heat demand profiles can be explained by a number of different base demand profiles. A number of such base profiles has been obtained in the form of weight vectors of the neurons with the highest winning rate. The results strongly suggest that a limited number of typical demand profiles can be used to provide valid representation of all consumers. Moreover, the method helps validate existing consumer classes in view of consumer demand. It has been shown that for some consumer categories typical demand profile can be associated, while remaining categories require more detailed classification to ensure appropriate demand assignment.

Future works will include the identification process of consumer's profile through supervised learning based on descriptive attributes, like ordered heat volume, consumer's category (domestic, public use, company office, school etc.). In addition, similar profiles will be created for short time consumption to answer the requirements of short-time prediction. Thus, hourly demand will be analysed.

References

[1] Pierce M.A., *Environmental Benefits from District Heating in the Nordic Capitals*. 16th Congress of the World Energy Council, Tokyo, (1995).
[2] Grzenda M., Macukow B. *Evolutionary Model for Short Term Load Forecasting*. Proc. of MENDEL 2001 conference. Brno, (2001) 119.
[3] Grzenda M., Macukow B., *Evolutionary Neural Network-Based Optimisation for Short-Term Load Forecasting*. Control and Cybernetics, 31 (2002) 371.
[4] Haykin S., *Neural Networks: a Comprehensive Foundation*. Prentice-Hall Inc., (1999).
[5] Jang J.-S.R., Sun C.-T., Mizutani E., *Neuro-Fuzzy and Soft Computing. A Computational Approach to Learning and Machine Intelligence*. Prentice Hall, (1997).
[6] Brandt S., *Data Analysis. Statistical and Computational Methods for Scientists and Engineers*. Springer Verlag, (1999).
[7] Lattin J., Carroll J., Green P., *Analyzing Multivariate Data*. Thomson Learning, (2003).
[8] Martinetz M., Berkovich S., Schulten K., *"Neural-gas" Network for Vector Quantization and Its Application to Time Series Prediction*. Trans. Neural Networks, 4 (1993) 558.