Unsupervised Energy Disaggregation Using Conditional Random Fields

Panikos Heracleous, Pongtep Angkititrakul, Norihide Kitaoka, and Kazuya Takeda
Department of Media Science, Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8603, Japan
Tel: +81 52 789 3647
Fax: +81 52 789 3172
Email:panikos@g.sp.m.is.nagoya-u.ac.jp

***Abstract*—The task of energy disaggregation is to break up total energy usage into usage by its component electrical appliances. This can provide home owners with useful feedback about how they use electrical energy, and motivate them to save significant amounts of energy. Our ultimate goal is to provide ways to modify human energy consumption behavior in order to conserve and optimize the use of energy. This paper focuses on stochastic modeling and energy disaggregation based on conditional random fields (CRFs) using real-world energy consumption data. The proposed disaggregation method uses a clustering method and histogram analysis to detect the ON/OFF states of selected types of energy-using devices in the home. Data labeling is not required, because the label sequences are obtained by applying a clustering method, and decoding using all of the data. Long spans of data from 21 households were used in a binary classification experiment, in which an 86.1% average classification accuracy was achieved. The proposed method was also evaluated using hidden Markov models (HMMs), but significantly higher accuracy was obtained when CRFs were applied.**

***Index Terms*—Energy consumption, unsupervised energy disaggregation, CRFs, HMMs.**

I. Introduction

Increasing energy conservation is one of the main challenges facing modern society. Studies have shown that residential and commercial buildings consume 40% of our total energy resources [1], and that 20% of this consumption could be reduced [2] by changing human energy consumption behavior [3].

One effective method to reduce energy demand is to provide homeowners with information about the energy usage of their electrical appliances, as well as the times when their electrical appliances are in use. Using this information, people can change their energy usage behavior by using different electrical appliances at different times, or by selectively reducing energy consumption. Monitoring each electrical appliance would be an ideal way to provide occupants with detailed information about energy usage in their homes, however this would be costly and impractical. Disaggregating total energy consumption in each household would be easier and more practical. In this study we focus on energy disaggregation to particular energy consumption activities. Given the total energy consumption of a household, the proposed method can determine when a particular appliance is in use. In the case of overlapping use of multiple appliances, the occupants' behavior can be modified in order to optimize and reduce energy use.

The goal of energy disaggregation is to break up total energy usage into usage by its component appliances. This can provide home owners with useful feedback about how they use electrical energy, and can also motivate users to save significant amounts of energy. Several different energy disaggregation methods have been proposed, including sparse coding [4], detection and clustering based approaches [5], [6], pattern recognition [7], factorial HMMs, and difference HMMs and variants [8]–[10]. Our method differs from those of previously reported studies in that it uses CRF- [11] and HMM-based [12] energy-consumption modeling to classify energy usage, and to determine the times when each particular cabinet (i.e., appliances or groups of appliances) is in use, using only total energy consumption. In addition, ultra-low frequency (one sample per minute) energy consumption data are used.

Many researchers are also focusing recently on electricity price forecasting, which is an important issue for home owners, governments, and power system operators. Electricity price forecasting methods include time-series predictors [13], [14], neural networks [15]–[17], nearest-neighbor approaches [18], and QP with outage combinations [19].

II. Methods

A. HMMs and CRFs for modeling home energy consumption

CRFs and HMMs are natural ways to model the generation of the total energy consumption. The total energy consumption is assumed to be a hidden Markov chain with a finite number of states corresponding to the number of electrical appliances.

1) HMM modeling: HMMs are generative models widely used in speech recognition and computer vision tasks. HMMs maximize the joint probability of paired observation and label sequences. One disadvantage of HMMs is their assumption of frame independence. On the other hand, HMMs are simple, can be easily and quickly trained, and also globally maximize joint probability over the whole word sequence. Joint probability can be calculated as follows:

$$P(\mathbf{s}, \mathbf{x}) = \mathbf{P}(\mathbf{x}|\mathbf{s})\mathbf{P}(\mathbf{s}) \tag{1}$$

where **s** is the state sequence and **x** is the observation sequence.

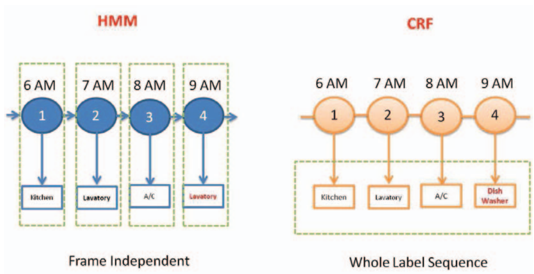

Fig. 1. Structure of HMM and CRF models.

Two assumption are made in HMM. First, it is assumed that each state s_t only depends on the a previous state s_{t-1}. The second assumption is that each observation x_i only depends on state s_i.

2) CRF modeling: CRFs is a modern approach similar to HMMs, but with a different nature. CRFs are discriminative models, which maximize the conditional probability of observation and state sequences $P(\mathbf{s}|\mathbf{x})$. CRFs assumes frame dependence, and as a result context is also considered. Figure 1 shows the structure of HMM and CRF models and the difference between them.

B. Database

In the current study, energy consumption data provided by the Toyota Motor Corporation was used. The data was collected in 42 households during three different collecting periods:

- 2011/10/01 00:00:59 2012/10/31 23:59:59, 397 days
- 2011/12/16 00:00:59 2012/10/31 23:59:59, 321 days
- 2012/05/31 00:00:59 2012/10/31 23:59:59, 154 days

For each household, energy data for 28 cabinets were available. Table I shows the 28 cabinets used during data collection.

Two variables for each cabinet are included in the data, namely, instantaneous power consumption, measured in Watts (w), and cumulative power consumption, measured in Watt hours (wh), with a sampling rate of once per minute (1 min). To calculate the energy consumption of each cabinet every minute, the following formula was used:

$$E[n] = P[n] - P[n-1] \ (Wh) \tag{2}$$

where E is the energy consumption, P is the cumulative power, and n is the time frame.

C. Proposed method for energy disaggregation

The proposed method assumes k hidden activities in the total energy consumption signal. Further, it assumes that only a single activity is defined within a time slot, and that each activity will generate feature vectors which form compact and disjointed clusters in feature space. Initially, a particular activity is hypothesized to be associated with each cabinet, however these activities are not yet defined. The proposed method then employs feature extraction, statistical modeling, clustering, and histogram analysis to associate activities with particular cabinets. A limitation of the proposed method is the requirement that all energy sub-signals be defined in the training phase and during histogram analysis. The training phase of the proposed method is as follows:

- Feature extraction and clustering
  In this step, energy consumption data is associated with labels, as this is a requirement for training the models. We refer to these labels as activity ID sequences. These are hypothesized, virtual activities corresponding to the classes used in clustering. The real cabinets will be obtained during the last step of the proposed method.
  From the total energy consumption signal and the energy consumption sub-signals of each cabinet, features are extracted for training the models and for clustering. Once a minute, Energy, ΔEnergy, and Time of day were extracted from the signals in a one day sequence. Therefore, each daily sequence consists of 1440 feature vectors. The length of the feature vectors differ, depending on whether they will be used for training the statistical models or for clustering. When used for training models, a feature vector length of 3 is used, which consists of total energy, the first derivative of total energy, and the time of day. For clustering, the features vectors are of length L:
  $$L = 2C + 1 \tag{3}$$
  where C is the number of cabinets considered in the experiment. In this case, the feature vectors consist of the sub-signals of each cabinet under consideration, the energy, and the time of day.
- Clustering
  Gaussian Mixture Model (GMM) clustering [20] is performed to classify the feature vectors into k classes. GMM clustering has the disadvantage that the number of classes should be defined a-priori. It is difficult, therefore, to select an optimal number of classes. Previous studies in the same framework, but using different data, found that 21 classes provided the highest performance [21]. In this study, different numbers of classes are used. For clustering, the features obtained from all sub-signals were used. This step results in a trained GMM with k Gaussian terms, where k is the hypothesized number of hidden activities.
- Decoding
  After clustering, scoring is performed, which results in activity ID sequences. For each time frame, the probability density functions are calculated using the feature vectors and the trained Gaussian terms of the GMM model. As a result, k scores are obtained, each one corresponding to a hidden virtual activity. The index of the maximum score is selected to be the ID associated with the feature vector. This step results in ID sequences for all of the feature vectors.
- Histogram analysis
  For each cabinet, an ID histogram is created using the corresponding labels of the ID sequence at the time when this cabinet was active. First, a small threshold is selected

TABLE I
CABINETS USED IN ENERGY DATA COLLECTION BY TOYOTA MOTOR CORPORATION.

Kitchen	Living-dining	Japanese room	Entrance	Lavatory	Main bedroom	Western room 1
Western room 2	Ecocute	Shelf 1	Shelf 2	Shelf 3	Socket (dining)	Cooking heater
A/C (living)	A/C (main bedroom)	A/C (Japanese room)	HEMS	Hall	Washing machine	Plug-in
A/C (Western room 1)	A/C (Western room 2)	Pulse counter	Bath dryer	Dish washer	Spare	Spare

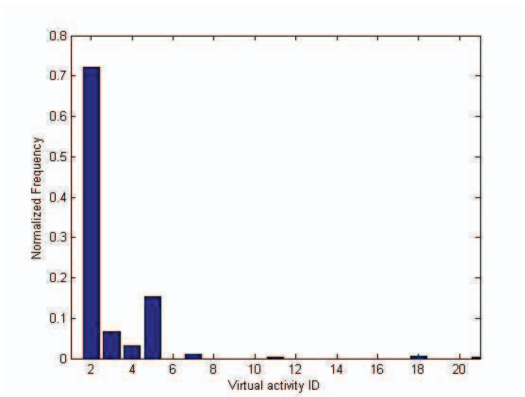

Fig. 2. ID histogram for the cabinet "kitchen".

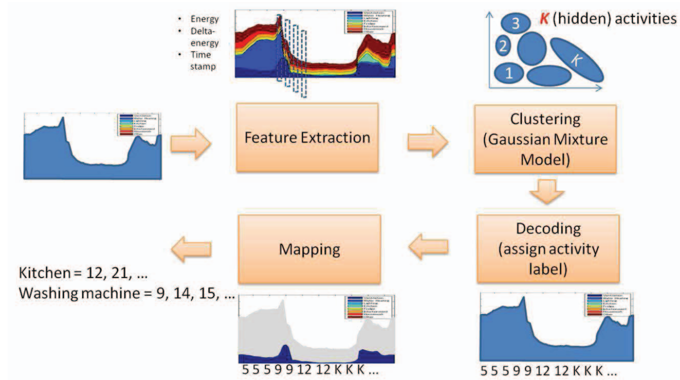

Fig. 3. Framework of the proposed energy disaggregation method.

empirically to remove some biases and exclude outlier data from the energy consumption signals of each cabinet. This threshold can also be adjusted in the classification phase to maximize accuracy. Figure 2 shows the ID histogram corresponding to the cabinet for the "kitchen" when using GMM clustering with 21 classes..

- Training CRFs and HMMs models
  Given the training data and the associated activity ID sequence, CRF models and HMM models with k states are trained. The models are trained to represent the activities sequence using the total energy. Training is performed in a maximum of 500 iterations. Depending on the training data and the computation environment, the training phase may last about 5-10 minutes.

In the evaluation phase, given an unknown total energy consumption sequence, an output ID label sequence is provided by the CRF or HMM models. The ON/OFF states of each cabinet are obtained by grouping the ID labels according to the created histograms. To group the histogram IDs with activities, we assign a threshold and track the frequency counts of activities of interest that are greater than the threshold. In the output sequence, all IDs belonging to a particular cabinet are considered to be active, and all other IDs are considered to be inactive. Specifically, after ID grouping, a sequence in the format of $1, 1, 0, 1, 0, 0, 1, 1, ..$ is obtained, which represents the ON (i.e., 1) and OFF (i.e., 0) states of a particular cabinet. Figure 3 shows the framework of the proposed method.

III. RESULTS

In this section, the results of a binary classification experiment are reported. For comparison purposes, results using a different database are also reported.

A. Experiments using data provided by Toyota Motor Corporation

In these experiments, energy consumption data from twenty-one households were used. The households were selected to include all three of the collection periods. For each household, twenty-eight cabinets were available, however, in the current study, only the following eight cabinets were used:

- Kitchen, Living-dining, Lavatory, Cooking heater, Dish washer, Air-condition (living room), Washing machine, Others.

The "Ecocute", "Plug-in", and "HEMS" cabinets were excluded from the experiments because of the difficulty in associating them with a particular activity. The energy consumption of the remaining 18 cabinets were summed and included in the cabinet Others. As previously described, GMM-based clustering was performed using 12, 16, 21, and 30 Gaussian terms. Note that these numbers are not optimal and that different numbers of classes may result in higher accuracy. Energy, ΔEnergy, and Time of day values for each minute were used as features. For CRF and HMM training and testing, total energy, Δtotal energy, and time of day were used (i.e., three features). For clustering and histogram analysis, features were extracted from each energy sub-signal. Half of the data available were used for training, and the other half for testing. Depending on the household, 73,440 to 233,280 feature vectors were used for training and testing.

Figure 4 shows the average classification accuracy for the 8 selected cabinets in all of the households, in relation to different numbers of GMM clustering mixtures. As is

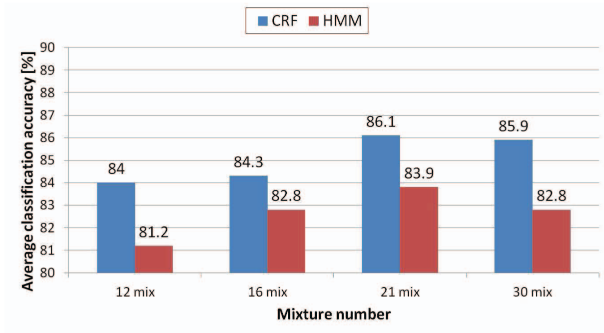

Fig. 4. Average classification accuracy using different numbers of mixtures.

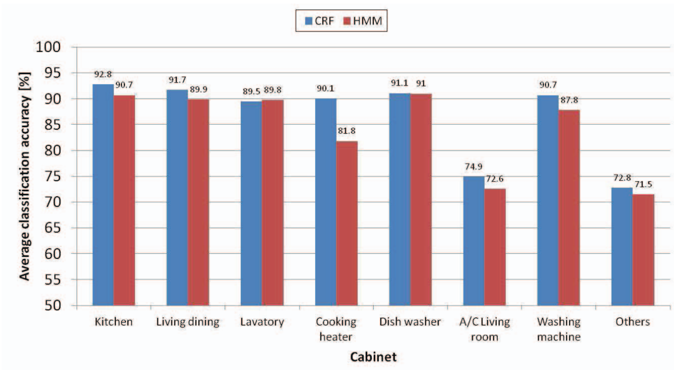

Fig. 5. Average classification accuracy for eight cabinets (using 21 classes).

shown, average accuracy increases up to 21 mixtures and then decreases for both HMMs and CRFs. These results suggest that an optimal number of mixtures may exist. We can also see that CRFs outperform HMMs in all cases. Specifically, when using 21 mixtures and CRF modeling, an 86.1% average classification accuracy was achieved. This is a very promising result, which illustrates the effectiveness of the proposed method. The number of mixtures is a critical factor and has an effect on classification accuracy. GMM-based clustering does not guarantee the optimal number of mixtures. Dirichlet

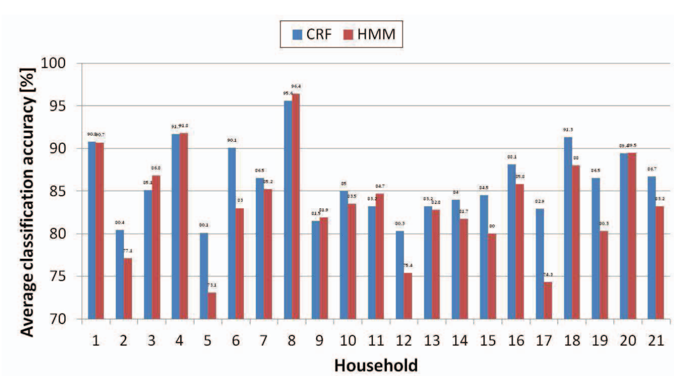

Fig. 6. Average classification accuracy for twenty-one households.

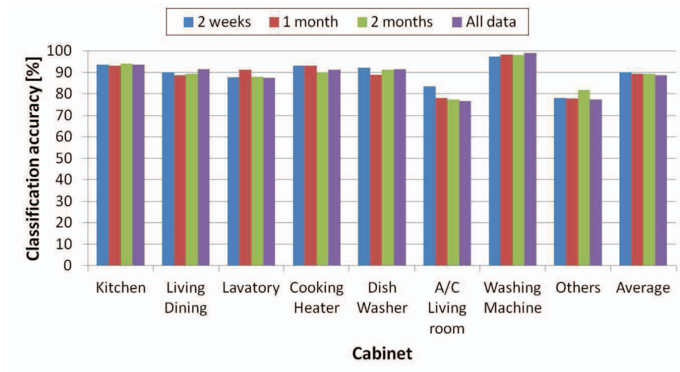

Fig. 7. Average classification accuracy for ten households using training data of different lengths.

Process Mixture Models (DPMM)-based clustering may offer a more efficient clustering approach.

Figure 5 shows the average classification accuracy for the eight cabinets using clustering with 21 classes. As can be seen, both CRFs and HMMs are highly accurate, with CRFs achieving higher performance. The ”air-conditioner” cabinet was classified with less accuracy, however. This may be because power usage for this cabinet varies widely depending on the season. In the current study, seasonal effects were not considered, so a mismatch between training and test conditions is possible. The ”Others” cabinet was also classified with less accuracy. This cabinet represents the summation of several cabinets with different characteristics, therefore statistical models may fail to correctly capture their characteristics. Finally, a two-tailed, paired t-test between CRFS and HMMs was performed which resulted in a P value of 0.0431, which, by conventional criterion, is considered to be statistically significant.

Figure 6 shows the average classification accuracy for each of the 21 households. As is shown, the classification accuracy rate varies among the households. Using CRFs, the highest accuracy rate was 95.6% and the lowest was 80.1%. Using HMMs, the highest accuracy rate was 96.4% and the lowest was 73.1%. In most cases (i.e., 17 out of 21), CRFs achieved better or very similar accuracy to HMMs. This result was achieved by using 21 mixtures for all households. The differences among the households may be due to the number of mixtures used during clustering. The authors believe that the best number of mixtures may vary, depending on the household. It is also possible that some households have more regular energy usage patterns or less simultaneous cabinet usage than others, making energy usage easier to classify.

An experiment was conducted to investigate the effect of the size of the training data on classification accuracy. Training data of 2 weeks, 1 month, 2 months, and all training data were used. The test size was 2 weeks of data in all cases. Figure 7 shows the results obtained using CRFs and GMM-clustering with 21 classes. The results are very interesting, showing no significant differences in accuracy. These results indicate that

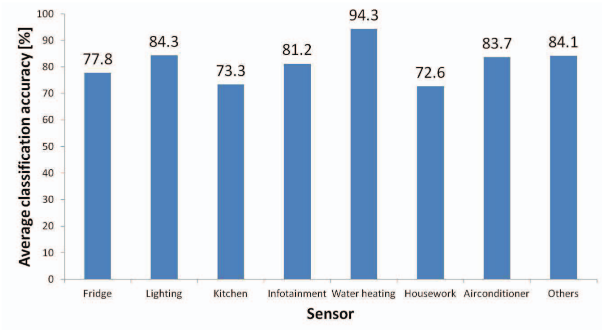

Fig. 8. Average classification accuracy for fifty-four households using publically available data.

training data of one or two months is sufficient to achieve high performance.

In the current study, a model for each household was trained. For training, big data were used and 28 cabinets were considered. In the case of energy human behavior changes in a particular house, new model or adapted model may be necessary.

B. Experiments using publically available data

The publically available database used in this study was collected by researchers at Niigata University, Japan [22]. Energy usage data was collected in six different areas of Japan from November, 2002 to March, 2005. The six areas selected for data collection were Hokkaido, Tohoku, Kanto, Hokuriku, Kansai, and Kyushu. In each of the six areas, eight sensors were used in nine households. The following sensors were used: ventilation, water heating, lighting, kitchen, refrigerator, infotainment, housework, and others. A sampling rate of 15 minutes was used for data collection. Figure 8 shows classification accuracy for all energy-consumption activities using the proposed CRF modeling method. As shown, the proposed method also achieved a high rate of classification accuracy when using this data. For instance, in the case of water heating, the disaggregation accuracy rate was 94.3%. The average classification accuracy rate for all activities was 81.4%, which is a very promising result.

IV. Conclusion

In this paper we proposed a novel, unsupervised method for the disaggregation of energy usage activities based on CRFs and HMMs. The proposed method was evaluated using two different data sets. Energy usage activities were disaggregated and classified with average classification accuracy rates of 86.1% and 81.4% for each data set, respectively, when disaggregating eight different energy usage activities, demonstrating its generality. As future work, the authors plan to integrate additional information in the feature extraction phase.

Acknowledgments

This work was partially supported by the Core Research for Evolutional Science and Technology (CREST) Program of the Japan Science and Technology Agency (JST). The authors would also like to thank the Toyota Motor Corporation for providing us with their energy usage data.

References

[1] Perez-Lombard, J. Ortiz, and C. Out, "A review on buildngs energy consumption information," *Energy and buildings*, vol. 40, pp. 394–398, 2008.

[2] J. Creyts, A. Derkach, S. Nyguist, K. Ostrowksi, and J. Stephenson, "Reducing u.s. greenhouse gas emissions: How much at what cost?," *U.S. Greenhouse Gas Abatement Mapping Initiative, Technical Report*, 2007.

[3] G. Crabtree, "Energy future report: 'energy future: think efficiency'," *American Physical Society, Technical Report*, 2008.

[4] J. Z. Kolter and A. Y. Nq, "Energy disaggregation via discriminative sparse coding," *Neural Information Processing Systems*, 2010.

[5] S. Drenker and A. Kader, "Nonintrusive monitoring of electric loads," *IEEE Computer Applications in Power*, vol. 12, no. 4, pp. 47–51, 1999.

[6] D. Rahayu, B. Narayanaswamy, S. Krishnaswamy, C. Labbe, and D.P. Seetharam, "Learning to be energy-wise: discriminative methods for load disaggregation," *in Third International Conference on Future Energy Systems: Where Energy, Computing and Communication Meet (e-Energy)*, pp. 1–4, 2012.

[7] L. Farinaccio and R. Zmeure, "Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses," *Energy and Buildings*, vol. 30, no. 3, pp. 245–259, 1999.

[8] H. Kim, M. Marwah, M. F. Arlitt, G. Lyon, and J. Han, "Unsupervised disaggregation of low frequency power measurements," *SDM'11*, pp. 747–758, 2011.

[9] J. Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial hmms with application to energy disaggregation," *Speech Communication, Special Issue on Speech Under Stress*, vol. Proc. of the International Conference on Artificial Intelligence and Statistics, 2012.

[10] S. Pattem, "Unsupervised disaggregation for non-intrusive load monitoring," *11th International Conference on Machine Learning and Applications (ICMLA)*, vol. 2, pp. 515–520, 2012.

[11] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," *18th International Conference on Machive Learning*, pp. 282–289, 2011.

[12] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *in Proc. of the IEEE*, vol. 77 (2), pp. 257–286, 1989.

[13] J. Contreras, R. Espinola, F.J. Nogales, and A.J. Conejo, "Arima models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1014–1020, 2003.

[14] A.J. Conejo, M.A. Plazas, R. Espinola, and A.B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and arima models," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1035–1042, 2005.

[15] A.M. Gonzalez, A.M.S. Rogue, and J.G. Gonzalez, "Modeling and forecasting electricity prices with input/output hidden markov models," *IEEE Trans. Power Sys.*, vol. 20, no. 1, pp. 13–24, 2005.

[16] G. Li, C.C. Liu, C. Mattson, and J. Lawarree, "Day-ahead electricity price forecasting in a grid environment," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 266–274, 2007.

[17] L. Wu and M. Shahidenpour, "A hybrid model for day-ahead price forecasting," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1519–1530, 2010.

[18] A. T. Lora, J.M.R. Santos, A.G. Exposito, J.L.M. Ramos, and J.C.R. Santos, "Electricity market forecasting based on weighted nearest neighbors techniques," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 1294–1301, 2007.

[19] Q. Zhou, L. Tesfatsion, and C.-C. Liu, "Short-term congestion forecasting in wholesale power markets," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2185–2196, 2011.

[20] G.M. McLaughlan, "Mixture models: inference and applications to clustering," *Dekker*, 1988.

[21] P. Heracleous, P. Angkititrakul, and K. Takeda, "Stochastic modeling and disaggregation of energy-consumption behavior," *2014 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, p. (to appear), 2014.

[22] Niigata University, ," http://tkkankyo.eng.niigata-u.ac.jp/HP/HP/database/japan2/index.htm/, 2003.