Load Recognition for Automated Demand Response in Microgrids

Adeel Abbas Zaidi, Friederich Kupzog (Member IEEE), Tehseen Zia
Institute of Computer Technology
Vienna University of Technology
Gusshausstrasse 27-29, 1040 Vienna, Austria
{zaidia, kupzog, zia}@ict.tuwien.ac.at

Peter Palensky (SM IEEE)
Austrian Institute of Technology
palensky@ieee.org

***Abstract*—Microgrids are well-suited for electrification of remote off-grid areas. This paper sketches the concept of a plug-and-play microgrid with a minimum of configuration effort needed for setup. When the load of such an off-grid microgrid grows over the generation capacity and energy storage is not sufficient, demand has to be reduced to prevent a blackout. In order to decide which loads are inessential and can be shedded, automated load recognition on the basis of measured power consumption profiles is needed. Two promising approaches from the area of speech recognition, Dynamic Time Warping and Hidden Markov Models, are compared for this application. It is found that a key feature to achieve good recognition efficiency is a careful selection of the features extracted from the measured power data.**

I. INTRODUCTION

Rapidly increasing energy demand all over the world has forced the utilities and their consumers to look at power production and utilization in a different way. While building up of new power plants is impeded by huge capital investment, microgrids are penetrating very promptly into the world's energy business [1]. Moreover, microgrids are best suited for the electrification of the rural areas, where connection to public grid is too expensive and impractical [2].

Microgrid is a paradigm of defining the operation of distributed generation, in which different microsources operate as a single controllable system that serves a cluster of loads in the local area [3]. It can be connected to the utility grid through a tie line or can work in isolated mode depending on the requirement of the customers as shown in Fig. 1. However, in case of remote areas, microgrids are always operated in an island mode. Similar to large power systems, the main challenge for secure operation of the microgrid is to maintain the demand-supply balance. This balance is traditionally maintained by taking measures on the generation side either by exploiting generation reserves or by exchanging power with the utility grid. However, approaches of on-line energy management including peak-load reduction can support demand reduction during situations of energy imbalance [4]. This article proposes a new concept for microgrid stability bearing on demand side management.

Due to the polluting nature and increasing cost of fossil fuel, microgrids are compelled to utilize renewable energy resources (RES) for generation purpose [5], [6]. Whereas, the potential of RES contribution depends on their generation physics. Photovoltaic panels can produce energy only during the day, whereas wind turbines can provide a significant share but wind power is highly volatile. Due to the deviations from the predicted power output of fluctuating sources, it becomes more challenging to maintain the stability of microgrids. Although controllable power generation (such as diesel generator and/or gas turbines) can be utilized to compensate these variations [7], it is very difficult to handle the situation with limited resources, especially in an island operation of microgrid. In this situation, considering the potential of demand side to uphold the stability of microgrids is expedient.

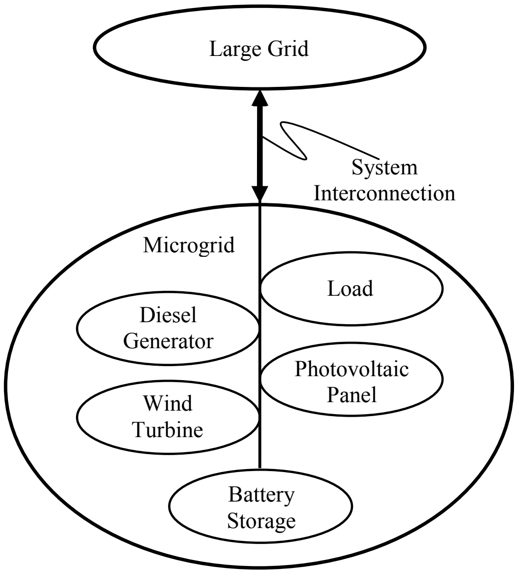

Fig. 1. Basic concept and typical elements of a microgrid [4].

Demand side management (DSM) can simply be described as *any actions taken on consumers' side to optimize energy consumption*. It is the process of maximizing the end-use efficiency of electricity to optimize the available generation resources. DSM always emphasizes on reducing power consumption instead of supplying additional energy. A smart DSM strategy can automatically switch off the selected loads to reduce the overall demand on the system during imbalance situations. Levels of automation in demand side management can be manual, semi-automated, or fully-automated [4].

However, only automated approaches are viable to ensure the stability of microgrids in real time. The proposed

978-1-4244-5226-2/10/$26.00 ©2010 IEEE

microgrid set-up utilizes the potential of communication technology for direct load control. The proposed microgrid system is self-configurable in which the microgrid central controller (MGCC) can communicate with the loads and directs them to isolate from the grid and vice versa. As the individual loads are not directly accessible by the MGCC, it is proposed to add a simple element (control node) to the microgrid loads: a combined power meter and switch, which can be read and operated remotely. The communication between the central controller and control nodes (CN) can be established using telecommunication lines, power line communication, or wireless sensor networks. The architecture of the proposed microgrid system is illustrated in Fig. 2.

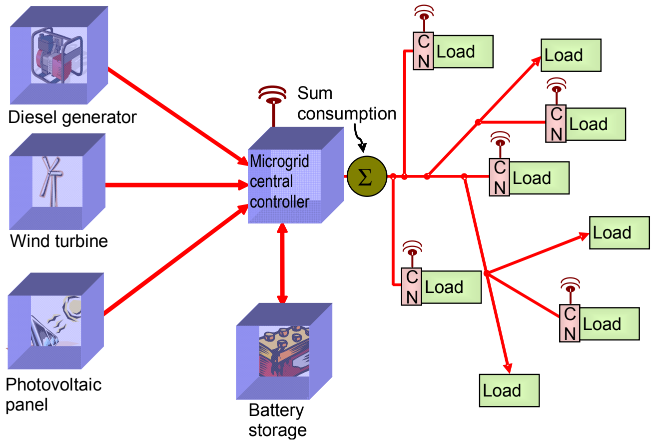

Fig. 2. Architecture of the proposed microgrid system [8].

The control node is ideally installed between every load and the central controller. MGCC receives information from the control node, on the basis of which it identifies the type of connected load and decides whether to switch off that load or not in overload situation. The process of load detection and isolation should be done automatically in order to have a simple microgrid set-up.

Assuming that the microgrid operation is managed by a central controller (MGCC), one of the main tasks of this controller is to find out the types of connected loads. This article specifically deals with the task to identify the load type from the measured power consumption data, so that the loads are assigned some priorities and shedding is performed in a precedence scheme using non-essential loads first.

Investigation of distinct attributes in power consumption profiles of different loads such as peak power consumption, daily operation timings, and the manner of operation (like duty cycle is case of a thermostatically control devices) is viable for load recognition. Based on these features, this paper presents a detailed analysis of two techniques applied for load identification; *Dynamic Time Warping* (DTW) [8] and *Hidden Markov Model* (HMM). DTW is used to match the templates of power consumption profiles of different appliances and decision is made on the basis of shortest path, whereas HMM – being a probabilistic model – is used to estimate the state-change probabilities of different appliances and recognition is made on this estimation.

The purpose of this paper is to compare the performance of these two approaches for the application of automated load recognition. The remaining paper is organized as follows: Section II gives an overview of related work in the area of load recognition and motivates further research in this area, Section III introduces the concepts that build the basis for the described work, and Section IV discusses the implementation of the proposed recognition algorithms as well as achieved recognition results. In Section V, the two approaches are compared.

II. Related Work

Various techniques are available in scientific literature for load recognition but most of them are focused on load classification to identify residential, commercial or industrial loads for load forecasting and tariff calculation [9], [10].

Some research is being carried out in the field of non-intrusive load monitoring (NILM) where the total consumption profile is monitored at the service entry point. The acquired consumption profile is then disaggregated to estimate the individual appliances. George Hart introduced one of the earliest approaches for NILM known as steady state approach [11], in which individual loads and their schedule are monitored by identifying changes in the aggregated power. These changes are reflected by switching of load from one steady state to another. Although the method provides basis for the commercial utilization of NILM technique, there are some limitations such as the appliances operating with infrequent change of state are hard to detect.

Another approach is proposed by Powers to disaggregate the consumption profile [12]. It is a rule based technique, which utilizes pre-collected information such as customer behavior of operating loads, customer load ownership and energy consumption levels of load, to develop recognition rules. The approach is mainly used by the utilities to disaggregate household load profiles into individual appliances. However, the data collection of customer behavior is often time-consuming and expensive process.

Farinaccio focuses the rule based approach in a different way [13]. In this method, rules are developed to detect the variations that occur during on and off events. As compared to steady state approach, this method includes more sophisticated rules. The technique not only measures power consumption between on and off events but also the variations during this interval. Moreover, it requires no information about the life style of the user. However, the approach is based on assumption that on and off events of different appliances will not occur simultaneously, a fact that can result in mixing-up of the variations.

Leeb proposed another NILM method to recognize loads using their start-up transients [14]. The start-up transients of different appliances are diverse, depending upon the nature of appliance. In this approach, the spectral contents of loads are recorded as templates, which are then compared with test loads. This method can provide useful information to identify the individual loads. But the detection of start-up transients is a difficult task because it requires high sampling rate for data recording [15].

The load recognition techniques described above are based on identification of an individual load contained in the total (combined) load profile of all the appliances. Whereas, the requirement of the proposed system is to identify a load from a set of individual load profiles of different loads. Therefore, two approaches from speech recognition domain are utilized which are described in the following section.

III. LOAD RECOGNITION

In order to perform load recognition for automated demand response, load profiles of different appliances are considered. The process is performed in two phases; pre-processing or feature extraction, followed by identification.

In the pre-processing phase, data is manipulated to extract different features such as *average energy consumption* (AEC), *edge counts* (EC), *percentage energy consumption* (PEC) and *discrete Fourier transformation* (DFT). However, DFT is found to be inefficient due to the mostly permanent character of loads with no or very less transitions. Hence, DFT is eliminated for next phase.

The processed data is then exploited for the identification. In the identification phase, two techniques are utilized; *dynamic time warping* (DTW) and *hidden Markov model* (HMM). Both the techniques are commonly used in speech recognition applications [16], [17].

A. Dynamic Time Warping

The DTW is a dynamic programming algorithm used for template matching or pattern matching of two time series (the sample, which is to be identified, and the template, which serves as reference, one for each load type to be detected). For instance, two vectors of pre-processed measured load profiles ($a_1, a_2, \ldots, a_n$) and ($b_1, b_2, \ldots, b_m$), which may be of different lengths, are considered. The algorithm calculates the local distance between both the vectors using Euclidean distance metric. The resultant (d_{ij}) will be a local distance matrix of size n x m.

$$d_{ij} = |a_i - b_i|, \quad i = 1, 2, \ldots, n \quad , \quad j = 1, 2, \ldots, m$$

The local distance matrix is then optimized for the computation of minimal distance matrix using following criterion.

$$a_{ij} = d_{ij} + min(a_{i-1,j-1}, a_{i-1,j}, a_{i,j-1})$$

where, a_{ij} is the minimum distance between the two vectors (sample and template). Finally, the algorithm decides which sample matches best to the known template. The decision is made on the basis of the nearest neighbor rule.

B. Hidden Markov Model

The hidden Markov model is a probabilistic model, which is based on a first-order Markov chain. The model construction is usually performed in two steps; structure modeling and parameter estimation. The first step deals with determining the number of states and interconnection between them. As no optimal method is available for the structure modeling, the step is performed manually. On the other hand, parameter estimation step estimates transition and emission probabilities. The process is carried out by using maximum likelihood estimation on the basis of training data.

The model construction requires two types of vectors; emission vector and state vector. The emission vector reflects the observable output symbols, which, for instance, are the pre-processed measured load profiles whereas the states vector represents the number of states in the model, defined on the basis of emission vector.

Once the model is constructed, the recognition process is carried out either by estimating the probability of emission vector or by finding the most likely state sequence. The probability estimation is carried out by using forward algorithm; whereas, the state sequence is obtained by utilizing well-known Viterbi algorithm.

IV. IMPLEMENTATION AND RESULTS

A. Data Acquisition

Daily power consumption data of different household appliances have been collected at Institute of Computer Technology for several days with a sampling rate of 10 sec. (8640 samples/day). The appliances include:

- Fridge
- Microwave
- Dishwasher
- Coffee machine
- Computer
- Printer

The appliances are selected by characterizing them on the basis of different parameters such as *thermostatically controlled appliances* (e.g. fridge), *fixed operation appliances* (e.g. dishwasher and coffee machine) and *usage dependant appliances* (e.g. microwave, computer and printer).

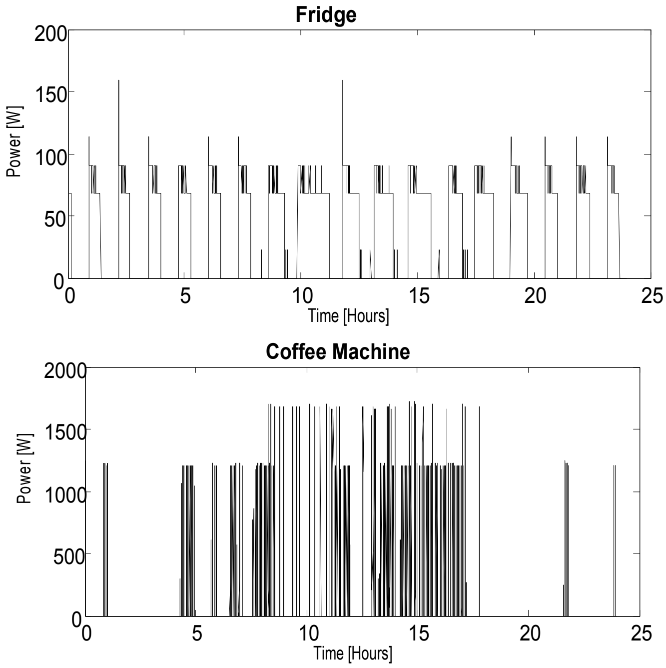

Fig. 3. Examples for measured load profiles.

A couple of examples of the collected power profiles are shown in Fig. 3. On these load profiles, pre-processing is performed to extract the useful features. The pre-processed vectors are tested for identification by using both the approaches (dynamic time warping and hidden Markov model) with the outcome of satisfactory results.

B. Implementation of Dynamic Time Warping

Dynamic time warping was tested for the pre-processing features; energy consumption, rising edge counts and percentage energy consumption. The detailed implementation of dynamic time warping is described in [8]. The results of the DTW algorithm, applied to the pre-processed data are presented in Table I. The Table shows that when the pre-processed load profiles of different appliances were compared to the template of fridge, the recognition was quite well on the basis of shortest distance. Moreover, both the features have proven to be the worthy of their selection.

TABLE I
SHORTEST DISTANCES BASED ON "ENERGY CONSUMPTION", "RISING EDGE COUNT" AND "PERCENTAGE ENERGY CONSUMPTION" FEATURES

TEMPLATE: FRIDGE			
PRE-PROCESSING / **SAMPLE**	ENERGY CONSUMPTION	RISING EDGE COUNT	PERCENTAGE ENERGY CONSUMPTION
FRIDGE 1	3.26E+07	300	29571
FRIDGE 2	7.83 E+07	444	217501
FRIDGE 3	5.08 E+07	426	244719
COFFEE MACHINE	3.71E+09	1200	798288
MICROWAVE	3.34E+09	2551	1202427
DISHWASHER	1.01E+11	2117	1375281
COMPUTER	1.19E+09	1633	3472036
PRINTER	1.32E+09	2593	532978

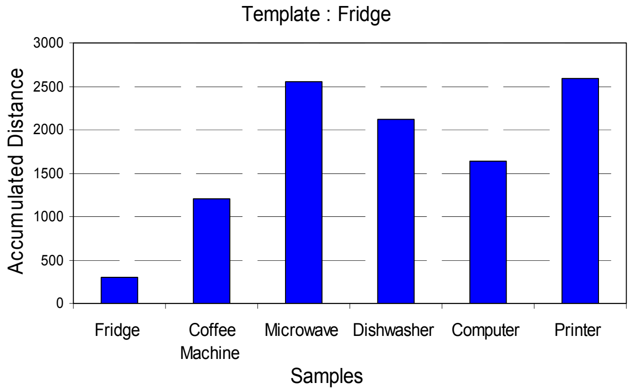

Fig. 4. Example for DTW (edge count feature).

Results of DTW are more clearly observable in Fig. 4, in which comparison of different samples (fridge, coffee machine, dishwasher, and printer) to a single template (fridge) is graphically shown. The bars represent the accumulated distance of edge count feature.

C. Implementation of Hidden Markov Model

The pre-processing features utilized for the implementation of hidden Markov model are energy consumption, edge count and percentage energy consumption. In order to build the HMM, the emission vector is taken as pre-processed data emitted by the load, and the state vector is defined separately for each load. The number of states for each model depends upon the nature of load and pre-processing as shown is Table II.

TABLE II
NUMBER OF STATES FOR EACH LOAD CORREPONDING TO PRE-PROCESSING FEATURE

PRE-PROCESSING / **LOAD TYPE**	ENERGY CONSUMPTION	EDGE COUNT	PERCENTAGE ENERGY CONSUMPTION
FRIDGE	4	3	3
MICROWAVE	3	4	3
COFFEE MACHINE	3	4	5
DISH WASHER	3	4	3
COMPUTER	6	4	3
PRINTER	2	3	3

Due to the variety of pre-processing features, defining the state vectors is relatively complex. It is a manual process, in which pre-processed data is thoroughly observed to decide the number of states for each load. After that, the specific state of the model is determined for each value of the emission vector. For example Fig. 5 illustrates the pre-processed data of a coffee machine over a week. The chosen pre-processing type is energy consumption with a window size of 2 minutes. Although the detection process works with the daily load profiles, the model is trained over the data of one week in order to have better recognition.

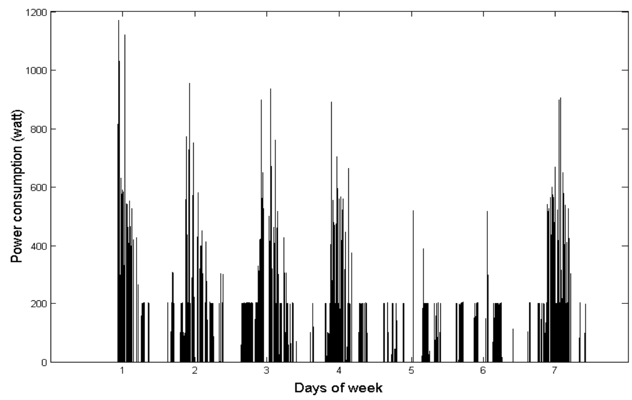

Fig. 5. Pre-processed (energy consumption feature) data of coffee machine for a week.

On the basis of pre-processing shown in Fig. 5, the coffee machine is assigned three states. However, the number of states may be increased. For instance, the first state represents the off status where no or very little power is consumed, the second state is at about 200 watts and the third state is defined for the power consumption more than 250 watts.

Once the emission vector and state vector are obtained, the transition and emission probabilities are estimated using maximum likelihood algorithm. The whole procedure is carried for each load to construct their HMMs. The initial state probability is not calculated during this procedure because the off state is considered as the initial state. As the off state is considered as a starting point for all the loads, initial state has the highest probability.

After constructing the model for each load, an observation sequence (i.e. pre-processed load profile of one day which is to be recognized) is presented to the model. The forward probability for every state is estimated at discrete point in time corresponding to each value in sequence vector. The observation sequence is presented to each model and the same process is performed. The model with an outcome of highest state probability is considered to be the best match.

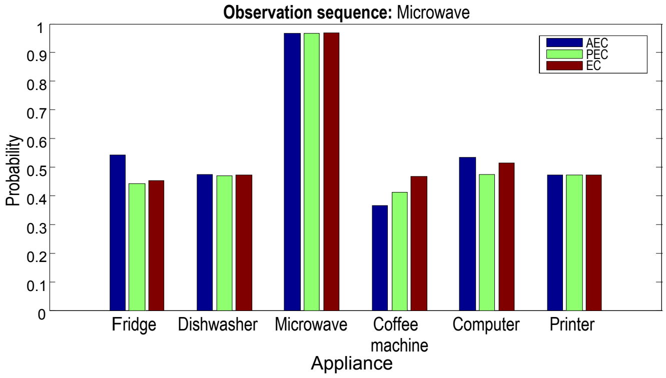

Fig. 6. Estimated probabilities for microwave for different pre-processing types.

Fig. 6 represents the estimated probabilities, when the observation sequence of microwave with different pre-processing types is presented to the models of all the appliances.

The results of recognition accuracy for different appliances based on applied pre-processing features are shown in Fig. 7. While performing these experiments, several load profiles of each load are presented to the models for recognition, and performance in terms of recognition efficiency, of each feature is measured. Recognition efficiency is the ratio of the number of samples (observation sequences), recognized by the models to the total number of samples (observation sequences), presented to the models.

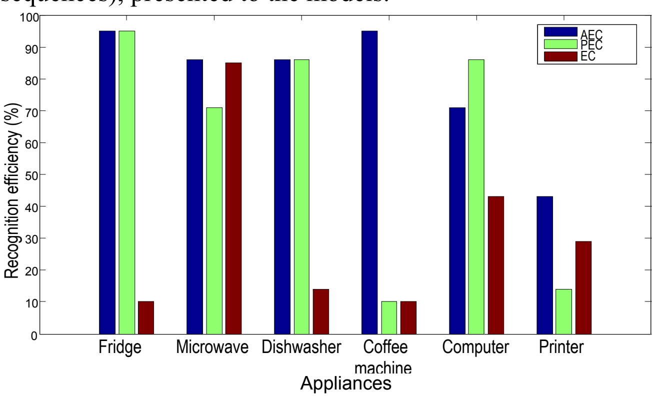

Fig. 7. Recognition accuracy of different pre-processing features.

It can be noticed from the bars of different pre-processing features, shown in Fig. 7 that the energy consumption is a reliable feature for recognition of almost all the loads, whereas edge count feature performs inefficiently in most of the cases.

V. Discussion

Since dynamic time warping is a template matching technique, the load recognition strictly depends on template database, captured during training phase. To accommodate all the changes in load profiles, a large template database is required. In speech recognition, the number of templates can be reduced by taking the average of all templates corresponding to a word to form a single template [18]. As the nature of changes in load profiles is totally different than that of words, the technique of reducing the number of template is not applicable in load recognition. In case of speech recognition, identical words can differ in their length but the pattern remains almost the same. Whereas, in the case of load recognition, an appliance can have entirely different power consumption profiles depending on their usage. For instance, during the same operational timings (e.g. 24 hours), the power profile of a coffee machine for making 15 cups will be different from that for making 30 cups. As the features (e.g. number of rising edge or energy consumption per window) are totally different in both cases of coffee machine usage, two different templates are required to recognize them. Since the usage behavior of an appliance is irregular that yield a variety of power consumption profiles, it is very difficult to manage the templates in such a huge number.

In the case of HMM, recognition process does not require any template database. Once the model of a load is built, the usage behavior of that load does not affect the recognition process. It is due to the fact that in HMM, the internal dynamics of the load are captured in terms of the states at the time of model construction. The observation sequence (i.e. pre-processed data) is used to estimate the states in contrast to the DTW, where the observation sequence is matched directly. Table III shows the effect of usage behavior of an appliance on the recognition process, performed through DTW.

TABLE III

Recognition of three different usage profiles of a microwave through DTW

Sample: Microwave			
Template	Shortest Distance		
Microwave	377	542	552
Dishwasher	292	474	1171
Printer	713	379	1063
Coffee machine	734	1519	3694
Fridge	1987	2122	3343

It can be noticed in Table III that when three samples (showing different usage behavior during a day) of the same microwave are matched with the templates of different appliances (such as dishwashers, printer, coffee machine, fridge and microwaves), the results based on shortest path distance depict that some times the samples of a microwave are better recognized as a dishwasher or a printer (see highlighted entries in Table III) than that of the microwave.

Similar sort of recognition experiment is performed using HMM; three observation sequences, showing different usage behavior, of a microwave are presented to the hidden Markov models of different loads and the estimated probabilities are tabulated in Table IV.

TABLE IV
RECOGNITION OF THREE DIFFERENT USAGE PROFILES OF A MICROWAVE THROUGH HMM

OBSERVATION SEQUENCE : MICROWAVE			
MODEL	PROBABILITY		
MICROWAVE	0.97	0.99	0.97
DISHWASHER	0.47	0.65	0.40
PRINTER	0.47	0.65	0.40
COFFEE MACHINE	0.41	0.57	0.35
FRIDGE	0.44	0.61	0.37

It is evident from the probability outcome listed in Table IV that every time an observation sequence of a microwave is best recognized by the model of the same. It is due to the fact that the probability of occurrence of an appliance is not influenced by its usage behavior.

VI. CONCLUSION

Although considerable related work is available in the scientific literature in the field of automatic load detection, the concept of automated demand response in microgrid demands different requirements. Here, every interruptible load has its own power measurement device and therefore can individually be analyzed. Since the system is self-organizing, load identification is the pre-requisite for automated demand side management. This work has utilized two approaches, DTW and HMM, to determine the type of load. Out of these two approaches, HMM is found to be more promising because load recognition through HMM is not influenced by the usage behavior. Whereas, in the case of DTW, the recognition process requires a large template data base because the recognition is affected by the usage behavior of the appliance. Since this work is conducted in the context of household loads, only a restricted number of possible loads is considered. As mentioned in Table III, if recognition results of DTW are dubious with such a small number of loads, its uncertainty will be increased with an increase in the number of loads.

Moreover, due to the satisfactory performance, the recognition process through HMM may be extended for non-intrusive load detection. It has been found that the energy consumption feature is more trustworthy pre-processing feature for the recognition, whereas edge count feature performs inefficiently in most of the cases. However, the robustness of the HMM approach proposed here has to be verified in future work by applying it to a wider range of appliances.

REFERENCES

[1] Hernandez-Aramburo, C.A., Green, T.C., Mugniot, N., “Fuel consumption minimization of a microgrid”, IEEE Transactions on Industry Applications, vol. 41 (3), pp. 673-681, 2005

[2] Venkataramanan, G., Marnay, C., “A larger role for microgrids”, IEEE Power and Energy Magazine, vol. 6 (3), pp. 78-82, 2008

[3] Hatziargyriou, N., Asano, H., Iravani, R., Marnay, C., “Microgrids: An overview of ongoing research, development, and demonstration projects”, IEEE Power and Energy Magazine, vol. 5 (4), pp. 78-94, 2007

[4] Zaidi, A. A., Kupzog, F., “Microgrid automation – A self-configuring approach”, 12th IEEE – International Multitopic Conference, pp. 565-570, 2008

[5] Arai, J., Iba, K., Funabashi, T., Nakanishi, Y., Koyanagi, K., Yokoyama, R., “Power electronics and its applications to renewable energy in Japan”, IEEE Circuits and Systems Magazine, vol. 8 (3), pp. 52-66, 2008

[6] Kapetanovic, T., Buchholz B. M., Buchholz. B., Buehner, V., “Provision of ancillary services by dispersed generation and demand side response needs, barriers and solutions”, elektrotechnik und informationstechnik, Spriinger-Verlab GmbH, Wien, Austria, pp. 452-459, 2008

[7] Rikos, E., Tselepis, S., Hoyer-Klick, C., Schroedter-Homscheidt, M., “Stability and power quality issues in microgrids under weather disturbances” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 1 (3), pp. 170-179, 2008

[8] Kupzog, F., Zia, T., Zaidi, A. A., “Automatic electric load identification in self-configuring microgrids”, IEEE Africon, September 2009

[9] Tsekouras, G. J., Hatziargyriou, N. D., Dialynas, E. N., “Two-stage pattern recognition of load curves for classification of electricity customers”, IEEE Transactions on Power Systems, 2007

[10] Gerbec, D., Gasperic, S., Smon, I., Gubina, F., “Consumers’ load profile determination based on different classification methods”, IEEE Power Engineering Society General Meeting, 2003

[11] Hart, G., W., “Nonintrusive appliance load monitoring”, Proceedings of IEEE journal, vol. 80, pp. 1870-1891, 1992

[12] Power, J., Margossian, B., Smith, B. “Using a rule-based algorithm to disaggregate end-use load profiles from premise-level data”, IEEE Computer Application in Power, pp. 42-47, 1991

[13] Farinaccio, L., Zmeureanu, R. “Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses”, Energy and Building Journal, vol.30 (3), pp. 245-259, 1999

[14] Leeb, S., B., Kirtley, J., L., “A multiscale transient event detector for nonintrusive load monitoring”, Proceedings of the IECON, vol. 1, pp. 354-359, 1993

[15] Pihala, H., “Non-intrusive appliance load monitoring system based on a Modern kWh-meter”, Thesis for Official Degree of Licentiate in Technology, Technical Research Center of Finland, ESPOO, 1998

[16] Rabiner, R., “A tutorial on hidden Markov models and selected applications”, Proceedings of the IEEE journal, vol. 77, no. 2, pp. 257-286, 1989

[17] Rabiner, R., Rosenberg, A., Levinson, S., “Considerations in dynamic time warping algorithms for discrete word recognition”, IEEE Transitions on Acoustics, Speech and Signal Processing, vol. 26 (6), pp. 575-582, 1978

[18] Yu, K., Mason, J., Oglesby, J., “Speaker Recognition using Hidden Markov Models, Dynamic Time Warping and Vector Quantization”, IEE Proceedings Vision, Image and Signal Processing, vol. 142(5), pp. 313-318, 1995