Energies **2015**, *8*, 9029-9048; doi:10.3390/en8099029

OPEN ACCESS

energies

ISSN 1996-1073

www.mdpi.com/journal/energies

Article

Real-Time Recognition Non-Intrusive Electrical Appliance Monitoring Algorithm for a Residential Building Energy Management System

Kofi Afrifa Agyeman [1], Sekyung Han [1,*] and Soohee Han [2]

[1] Department of Electrical Engineering, Kyungpook National University, Daegu 41566, Korea; E-Mail: kofi.a.agyeman@ieee.org

[2] Department of Creative IT Engineering, Pohang University of Science and Technology, Pohang 37673, Korea; E-Mail: sooheehan@postech.ac.kr

* Author to whom correspondence should be addressed; E-Mail: skhan@knu.ac.kr; Tel.: +82-10-5689-2467; Fax: +82-53-950-6600.

Academic Editor: Neville Watson

Received: 10 June 2015 / Accepted: 19 August 2015 / Published: 26 August 2015

Abstract: The concern of energy price hikes and the impact of climate change because of energy generation and usage forms the basis for residential building energy conservation. Existing energy meters do not provide much information about the energy usage of the individual appliance apart from its power rating. The detection of the appliance energy usage will not only help in energy conservation, but also facilitate the demand response (DR) market participation as well as being one way of building energy conservation. However, energy usage by individual appliance is quite difficult to estimate. This paper proposes a novel approach: an unsupervised disaggregation method, which is a variant of the hidden Markov model (HMM), to detect an appliance and its operation state based on practicable measurable parameters from the household energy meter. Performing experiments in a practical environment validates our proposed method. Our results show that our model can provide appliance detection and power usage information in a non-intrusive manner, which is ideal for enabling power conservation efforts and participation in the demand response market.

Keywords: unsupervised disaggregation; demand response (DR); advanced metering infrastructure (AMI); current harmonics; hidden Markov model (HMM)

---

1. Introduction

Electricity is one of the most common and important commodities we use everyday. Electricity energy demands requested from the consumer sector in a smart grid (SG) are constantly increasing in recent times as a result of a proliferation of electric appliances in the market. However, the global reservation-to-production ratio of oil, natural gas and coal is 45.7, 62.8 and 119 years respectively [1]; and the annual growth rate of the main energy sources: oil and natural gas, stands at 2.3% and 2.2% respectively [2]. With the current rate of energy demand, dependency on these natural resources to meet the current energy demand growth rate cannot be sustained after a certain period in time. Moreover, with many countries aiming to considerably reduce their annual carbon dioxide (CO_2) emissions by 2050 due to the negative implications of CO_2 in the atmosphere, energy conservation has become an issue of global importance [3]. The critical nature of this situation has generated significant interest in greatly increasing energy efficiency measures or policies and, as a result, has opened discussions for a possible bailout.

One of such solutions is to build new power generation plants. However, this measure requires many years along with huge capital investment. In spite of this, the environmental concerns are also not favorable to building new power generation plants. This situation has led us to consider different ways of power production and utilization. As a result, much attention has been focused on microgrid and energy demand optimization. Energy demand optimization is a process of managing energy demand according to available generation resources so as to maintain a balance between demand and supply. One mechanism of energy optimization is to motivate consumers to adopt energy conservation behavior. For this purpose, demand-response (DR) participation is highly identified. DR participation can facilitate domestic energy management as well as mitigate frequency deviation by shedding some power generated (*i.e.*, microgrid) or loads on national grid in return for monetary incentives; while on the other hand, consumption feedback is provided as a self-learning tool.

To adapt such an energy demand optimization system, detailed end-user energy consumption information is an essential requirement. Residential buildings contribute significantly to national electricity consumption and energy. In the US, residential buildings account for 40% of primary energy and 73% of electricity consumption [4]. Prior studies indicate that, with an efficient energy management system, buildings' electricity consumption can be reduced by up to 10% to 15% [5]. There are basically three typical approaches to energy management in residential buildings: Efficiency of energy usage, energy curtailment [6], and DR participation of residential building. Efficiency of energy usage involves the use of more energy-efficient appliances. However, the purchase of such appliances may be comparatively costly compared with energy-inefficient appliances. Energy curtailment, on the other hand, requires the less usage of appliances, which may eventually reduce the cost of energy, but negatively affect the consumer energy usage behavior.

Information on building's electricity consumption can be acquired traditionally by the use of energy meter installed in consumer buildings. However, from the author's view point, knowledge about the energy consumption of individual appliances is key in demand energy optimization. This is because energy usage is an abstract concept to most energy consumers [7,8]. Moreover, energy consumers are unaware of the usage of energy by the individual appliances and thus which actions would be most beneficial for conserving energy [9,10].

To this end, building energy management that could detect and monitor individual appliance energy consumption is essential and paramount to any effective energy management system. There are basically two approaches of monitoring appliances energy consumption in buildings: Intrusive appliance load monitoring (IALM) and non-intrusive appliance load monitoring (NIALM). IALM is based on a set of measurement devices attached to each appliance. It is simple to measure the consumption of every appliance. However, this method is laborious, cost-intensive and time consuming. NIALM, on the other hand, assumes the installation of a single monitor normally in the main circuit panel which monitor, detect, and extract the essential information using only the measurements taken at the circuit panel level.

Numerous studies have identified diverse approaches to effective demand energy management. Notable among them is SG [11] with integrated home automation networks (HAN) [12]. With such a system, building an energy management system utilizes real time price information to schedule loads to minimize energy consumption bills and provide economic incentive by participating in the DR market. However, SG-HAN could face deployment impediments as a result of two potential barriers. Firstly, this scheme requires an intelligent power grid or SG system, which can provide bidirectional communication between consumers' electric appliances and utility companies at real-time. Whereas new electric appliances could be manufactured with the necessary communication components embedded, existing appliances would need to be modified. Secondly, traditional building energy management systems provide a centralized platform for managing building energy usage. With this system, an initial profile of all the electrical appliances and equipment in the building are recorded prior to the building energy management system installation. The system would require a reconfiguration when there are equipment changes or failure. An open issue is, how to provide an appliance-specific breakdown of energy use in a cost-effective manner without negatively impacting consumers' standard of living or their productivity. Without addressing this issue, residential energy management or conservation is unlikely to achieve widespread success.

In view of this problem, we propose a novel real-time recognition non-intrusive electrical appliance-monitoring algorithm for residential building energy management system (REMS). The novelty of our method is, that the electrical appliances or equipment are dynamically and seamlessly detected and identified from the energy meter (*i.e.*, REM) data. Appliance usage information is obtained at a single point (*i.e.*, circuit breaker). The underlying goal of our algorithm is to disaggregate appliance power load from the aggregated power load and to identify the appliance as well as its state of operation. Whilst NIALM topic has received attention since the early 1990s, there is still yet much to be done in the area of signal disaggregation in order to bring it to the commercial front. Our work has some distinguishing characteristics that contribute significantly to appliance detection and identification from an aggregated signal. In our work, we first consider the frequency measurement of currents (*i.e.*, current harmonics) since there are high increasing rate of non-linear appliances in most residential buildings [13]. Secondly, we propose a new power meter, which does not provide only power components of the appliances but rather the current harmonics, real power and reactive power. To this end, REM device is installed in the main panel board and monitors the current in real time. Based on the measured parameters, some transient state or "microscopic" features such as current harmonics (*i.e.*, 3rd, 5th and 7th) and steady state or "macroscopic" features such as real power, and reactive power are recorded. Thirdly, we adapt an unsupervised disaggregation; Markov model to create multiple conditional factorial hidden Markov model (MCFHMM) for signal disaggregation. Finally, the empirical data used in our model, consider a real world situation where we have over 100 different kinds of appliances and we want to estimate the actual appliances from this dataset.

The remainder of the paper is organized as follows: Section 2 presents related works of the electric load recognition. Section 3 gives a description of the procedure of appliance disaggregation using MCFHMM. In Section 4, the experiment setup and evaluation is discussed. Finally the scope, limitation and further work of our proposed methodology are presented in Section 5.

2. Related Works

NIALM was developed as a low cost alternative to intrusive load monitoring. This method tries to analyze the usage of an appliance as well as its energy consumption. The load signature (*i.e.*, features) of a disaggregated power signal of an appliance, consist of power components that could be used to uniquely identify it. The NIALM methods, though based on different techniques, have several common underlining principles [14]: appliance features or signature selection; feature detection hardware device; and signal detection and disaggregation algorithm for aggregated signal.

Due to the importance of recognition accuracy of power signature, a lot of researches have gone on over the years to provide highly accurate and unique approach to load identification of power signature in NIALM. The initial approach to NIALM as proposed by Hart [15], identified appliances based on its steady state behavior. Hart conceptualized a finite-state machine to represent a single appliance where power consumption discretely varied with each step change [16]. Although the method performed well, the method had some drawbacks. For instance, Hart's method could not detect small energy consumption appliances, which were always on or had non-discrete change in power.

Apart from Hart, recent research efforts have attempted to improve the NIALM algorithm, often by proposing alternative techniques. The various proposed alternatives differ either in the signal sampling technique, classification or disaggregation algorithms and selection of features. Other recent steady state analysis could be seen in [17–20]. As an extension to Hart's work, Ducange *et al.* [17] proposed a twined load identification algorithm, where a fusion of finite-state machine (FSM) with fuzzy transitions algorithm was implored to identify appliance. A coarse description of this method involves an ad-hoc disaggregation algorithm that access a database for a change in the working state of the FSMs at any given time by analyzing a variation in a new coupled power parameters (*i.e.*, real and reactive power). This method could only produce two possible disaggregation solutions. Although this method gave a possible disaggregation solution, the result was always accompanied by a wrong solution as a pair. In [18], the use of particle swarm optimization (PSO) and backward propagation artificial neural network (BP-ANN) was adopted to improve the efficiency of load identification and computational time. The load signatures conceded under this work were only active and reactive power of appliances under steady state analysis. Even though research results showed significantly high recognition accuracy in less computational time, the authors failed to discuss or demonstrate analysis of appliances of similar or the same active and reactive power load signatures. In another steady state analysis as reported by Parson *et al.* [19], a non data training algorithm was proposed. Their work considers an approach in which prior models of general appliance types are turned into specific appliance type. They consider prior knowledge of appliance behavior and power demands as their main features. The researchers showed a live deployment of their work but failed to discuss the recognition accuracy and how the various appliances were identified. Again, little was known about appliances with similar features. Lin and Tsai [20] proposed a new technique with an automatic non-dominated sorting genetic algorithm (NSGA-II) in their recent work to estimate power consumed by each monitored major household appliance scheduled for DR participation. Their work showed very high significant recognition and identification accuracy. However, they had challenges with variable power demand appliances. For instance, an air conditioner, which is a typical example of variable power demand appliance, could not be monitored non-intrusively.

Contrary to the many steady-state approaches to NIALM, transient state analysis extract distinctive features such as shape, size, duration, and higher order harmonics by sampling current and voltage waveforms at a high frequencies to characterize appliance operations in its transient state. Recent papers [16,21–24] present new power signature analysis, load identification methods and feature selection approaches to recognize loads and to solve classification problems. For power signature analysis, most transient state analysis includes the use of turn-on transient energy algorithm. Since the envelopes of turn-on transient instantaneous power are closely linked to unique physical quantities, they could serve as reliable metrics for load identification [16]. For instance, in the case of [16,23], both turn-on transient energy algorithm and wavelet transform were adopted to analysis and capture the load signature or features. For load identification, both papers adopted the use of artificial neural network (ANN), adding to the many papers that have published to improve the performance of NIALM using ANN. The results showed by these papers were very much significant with a little drawback. In the case of [16], current and voltage waveform data should be sampled at a high frequency in order to capture the transient effect. However, modern energy meters are not equipped with such functionality, doing so will increase the cost of energy meters. Moreover, because of varying transient with these waveforms, it is essential that data set for load identification should have repeatable transient energy signature. Hence, much diligence is required to sampling of the instantaneous load profile of each turn-on transient load. In [23], though it is known that wavelet transform coefficient (WTC) works well than Fourier transform with respect to information acquisition for on/off transient signal identification of load events, WTC requires much longer computational time and larger machine resources such as memory usage. The paper approach to resolves these issues were not significantly evident.

Apart from pure steady state and transient state analysis, many other papers have sought to harness the benefits of both states. Wang *et al.* [25], in their paper, adopt mean-clustering algorithm together with multidimensional linear discriminates under both states to classify residential appliances. Their work showed a good result with the help of prior appliance information they acquired from a statistical agency. However, without any prior information, the identification accuracy of multi appliances was much lower. Moreover, their paper did not provide reliable and more accurate method for data acquisition. In [26], the paper considers ways in which smart meters could be equipped with NIALM functionality to predict the consumer energy usage behavior. However, their approach could predict appliances of higher power consumption only. Although the use of transient state analysis coupled with steady-state analysis provide an improved load disaggregation performance, nevertheless transient patterns are sensitive to wiring architecture, network topology and demand costly hardware for sampling the electrical signal at higher data rate.

One other load identification algorithm apart from ANN, which has gained most researchers attention, is the HMM. The HMM has been applied in many broad areas including NIALM. The most recent work in this area can be in seen in [24,27,28]. One such relevant study is from Kim *et al.* [27]. They used different variant of HMM to recognize the distinct electrical appliance of low power consumption. Their algorithms only consider one power feature of the appliance signature. The success of their approach motivated us to explore HMMs for developing appliance signature for residential power use.

The remainder of the paper is organized as follows: Section 2 provides a problem definition. Section 3 discusses appliance disaggregation using MCFHMM algorithm. Section 4 describes the model we use to identify the steady-state signatures of the household appliances. Section 5 presents our results, using both real power data and synthetic data.

3. Appliance Disaggregation Using Multiple Conditional Factorial Hidden Markov Model (MCFHMM)

3.1. Non-Intrusive Appliance Load Monitoring (NIALM) System Framework for Demand Response (DR)

Up until now, there has not been any particular standard established for DR design, monitoring, communication, control, *etc.* [21]. However, several conclusions on similar approaches can be summarized from [29]. Basically, different household appliances participate in DR in different ways. These appliances are categorized into two categories based on their characteristics [29]: High-power DR loads (*i.e.*, HVACs, Water pumps, PHEVs, *etc.*) and plug-in back-up DR loads (*i.e.*, TV, PC, printers, *etc.*). The connection of these appliances to the grid shows bi-directional framework of building DR control as shown in Figure 1. There are usually three main stakeholders in building energy control system: energy utility, grid operator and customer. Grid operator often acts as the medium to manage the bi-directional information flow between the customers and the utilities. The grid operator get the price and the system operation information form the utilities, and then passes it to the facility manager and the end-users to help them make the better decisions. All the actions taken by the facility manger or the end-user will also be delivered back to grid operation center at different time scales. In this way, buildings coordinate the bi-directional information flow and make an optimal DR control decision. In a residential building, the panel board or circuit breaker board usually acts as the converging point of the sub circuit branches at the main entrance. Loads are directly or indirectly connected to this outlet. Hence, non-intrusive measurement of the load signatures should be done at the circuit breaker level.

3.2. Load Signature

The specification of load signature for DR is another important aspect in NIALM. The load signature selection metrics is based on some steady-state features or “macro features”, transient features or “micro features” and comprehensive load survey features such as appliance penetration distribution (*APD*), appliances dependency distribution (*ADD*) and appliance cost rate. Our proposed state analysis of appliance specific load monitoring is based more on steady-state analysis. The research focuses on the use of power components such as active power (*P*), reactive power (*Q*), and current harmonics components (*h*). As reported in [30], as many as nine feature metrics such as supply voltage *V*, load current *I*, apparent power *VA*, power factor *pf*, energy kWh, harmonics, and phase could be extracted from appliance load signature. Our proposed approach is based on real power, reactive power and current harmonics. The performance of the power load disaggregation can be improved significantly if other additional inputs that indirectly relate to the state of an appliance are available. For the significant improvement of our model, we focus on other parametric properties such as APD, *ADD* and appliance cost rate.

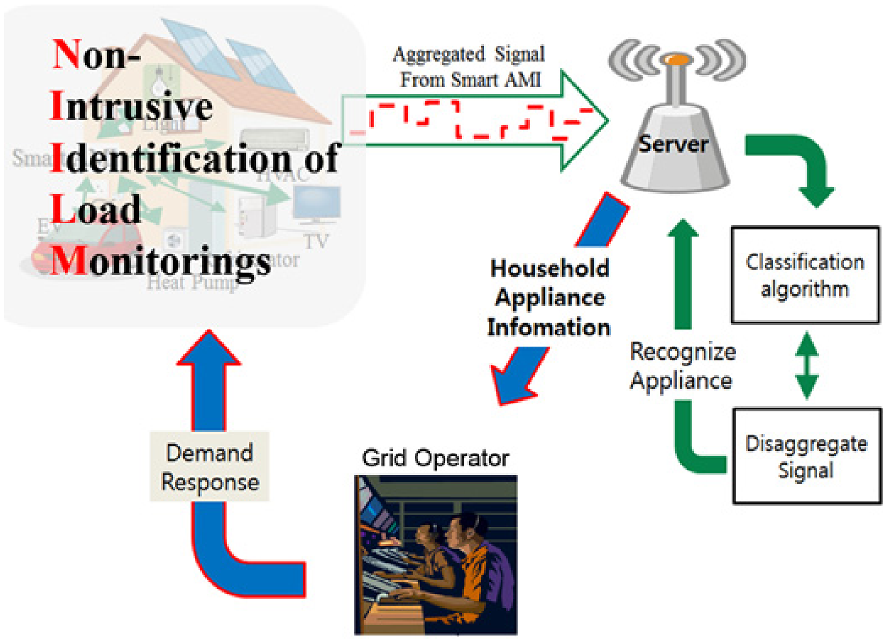

Figure 1. System architecture.

APD is the distribution of the usage of electronic appliances. The penetration rate is the average number of appliances per household in percentage. For example, a penetration rate of 90% for TV means that on average each household owns 0.9 TV (or 90% of households own one TV) and a penetration rate of 200% means that on average each household owns two TVs. Table 1 shows some product cases with their penetration rate. The penetration rate of the appliance is used as feature to disaggregate aggregated load power.

Table 1. Appliance penetration rate.

Product case	2005	2010	2020
	Rate (%)	Rate (%)	Rate (%)
Mobile phone	406	446	492
Lighting	93	107	143
Radio	60	60	60
Electric toothbrush	22.4	22.5	25.9
Electric oven	38	38	38
TV+	144	202	210
Washing machine	96.6	97.9	100.0
DVD	75	90	130
Audio mini system	60	60	60

The usage pattern over a period of time can provide the frequency distribution of an appliance's usage penetration. For instance, if at time t, the living room light is turn on over a period of time, T, the probability of the living room light being on at any time will be higher than the other appliances, which are rarely used.

ADD describes the strong correlation in the usage pattern of some appliances with others. For instance, Playstation 4 cannot be used without a television and a stabilizer cannot be used without other appliances such as fridge, TV and audio system. We tested these dependencies with our dataset by measuring the correlation between every pair of appliances using Pearson's coefficients. We subsequently computed the conditional probability of the appliance pairs that shows strong correlation. The conditional probabilities of the correlated pairs are used as a feature in disaggregating power.

3.3. Problem Definition

The specific problem we seek to address could be defined mathematically as follows: given the aggregated power consumption, Y with measured features T, $Y = \{y_1, y_2, \cdots, y_T\}$ and the number of appliance, M we want to infer from the load power signature Q, of each of the M appliances, that is:

$$\begin{aligned} Q^1 &= \{q_1^1, q_2^1, \cdots, q_M^1\} \\ Q^2 &= \{q_1^2, q_2^2, \cdots, q_M^2\} \\ &\vdots \\ Q^T &= \{q_1^T, q_2^T, \cdots, q_M^T\} \end{aligned} \tag{1}$$

Such that:

$$y_t = \sum_{i=1}^{M} q_i^t; 1 \leq i \leq M \tag{2}$$

We achieved this using energy disaggregation method based on extension of HMM. Our variant of HMM considers the addition of other features together with more accurate probability distributions of the state occupancy durations of the appliances. We refer to this variant as MCFHMM.

3.4. Data Acquisition and Pre-Processing

Figure 2 represents the power distribution layout of the household. In this simulated circuit, the voltage source supplying the household along with its internal impedance is designated as V_s and Z_s respectively. The impedance reflects the voltage drop due to the loading of the electrical circuit.

The measurement acquisition system includes an instantaneous current and voltage recorders together with the feature parameters. They are connected next to the voltage source in order to measure the total current of the household. The measured current and voltage instantaneous values are recorded and fed into the recognition algorithm. During this simulation, one appliance was connected at a time to measure its current and voltage within a specified time period. The measured parameters were stored in a database for processing. In a similar way, multiple appliances were connected at the same time and their parameters measured and stored.

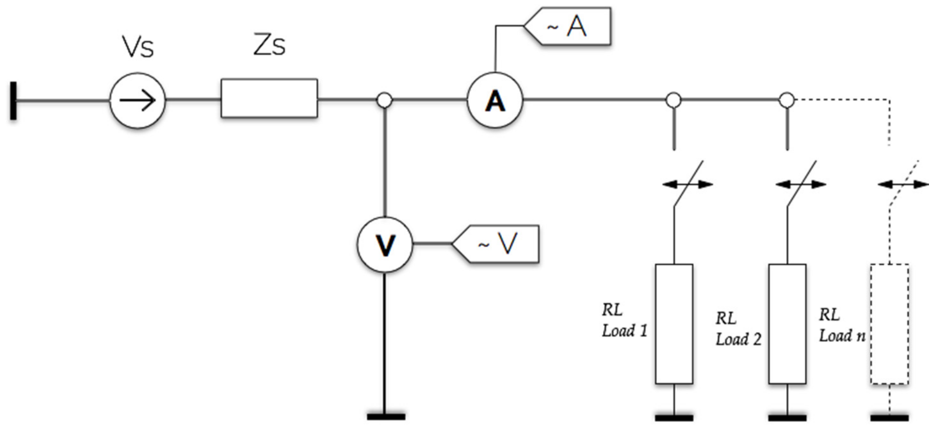

Figure 2. Single-phase power network.

Six household electrical appliances were used as loads in this study: Fridge_new (Load 1), Fridge_old (Load 2), LCD_TV (Load 3), Desktop lamp (Load 4), Standing heater (Load 5) and wall fan (Load 6). These electrical appliances have the following operational states: two-state, multi-states and continuously variable state. Two measurement procedures were performed to collect data for each load. Table 2 provides a detailed power rating of the listed appliances.

Table 2. Appliance power ratings.

No.	Appliance	*I* (A)	*V* (v)	*P* (W)	Frequency (Hz)
1	Fridge_new	2.72	220	500	60
2	Fridge_old	2.72	220	500	60
3	LCD TV	0.27	220	60	60
4	Desk lamp	0.1	220	20	60
5	Standing heater	3.63	220	800	60
6	Iron	4.55	220	1000	60

The measured data were stored in a database for processing. The structure of the database was normalized to ensure data integrity and faster queries.

The measured data forms a pool of appliance signatures that could be represented by the matrix below:

$$\begin{bmatrix} \text{Index} \\ 1 \\ 2 \\ \vdots \\ \vdots \\ \vdots \\ m \end{bmatrix} \begin{bmatrix} P & Q & H^3 & H^5 & H^7 \\ p_1 & q_1 & h_1^3 & h_1^5 & h_1^7 \\ p_2 & q_2 & h_2^3 & h_2^5 & h_2^7 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_m & q_m & h_m^3 & h_m^5 & h_m^7 \end{bmatrix} \begin{bmatrix} \text{Appl.} \\ \text{App}_1 \\ \text{App}_2 \\ \vdots \\ \vdots \\ \text{App}_m \end{bmatrix} \quad (3)$$

For a given measured aggregated power *mv*, the measured signal can be modeled a row vector:

$$[mv] = [\hat{P} \quad \hat{Q} \quad \widehat{H^3} \quad \widehat{H^5} \quad \widehat{H^7}] \quad (4)$$

where:

$$\hat{P} \quad \coloneqq \quad \text{measured real power} \quad (5)$$

$$\hat{Q} \quad \coloneqq \quad \text{measured reactive power} \quad (6)$$

$$\widehat{H^3} \coloneqq \text{3rd harmonic component} \quad (7)$$

$$\widehat{H^5} \coloneqq \text{5th harmonic component} \quad (8)$$

$$\widehat{H^7} \coloneqq \text{7th harmonic component} \quad (9)$$

The initial step that was considered before finding the best combination that represents the data was to sort the appliance signatures. The sorting order of the signatures was based on the least value of the measured signal. After the data set has been sorted based on the sorting order, the data set was trimmed to eliminate records whose feature values are greater than the measured values. The new matrix has a record less than the original dataset.

3.5. Model Description

HMM is used for probabilistically modeling sequential data. HMMs are known to perform well at tasks such as speech recognition [31]. A discrete-time HMM can be viewed as a Markov model whose states are not directly observed: instead, each state is characterized by a probability distribution function, modeling the observations corresponding to that state. Our model is based on HMM. We defined a probabilistic model that explains the generating process of the observed data. This model contains hidden variable that are not observed. With regards to our work, the states of the appliances are the hidden variables, and the aggregated features (*i.e.*, P, Q, H_3, H_5, H_7) are the observations.

The model has several parameters that can be learned from the captured data. The learning process consists of estimating the factorial observations of the appliances. The parameters from the observation, such as the initial probability of selecting a given state and the transition probability and the observation probability estimate the model that best describe the observation. In order to achieve this, the parameters of the model are adjusted so as to maximize the efficiency of describing the model that best describe the observation. Subsequently, using this model with the estimated parameters, we can estimate the hidden variables, which are the states of the appliances. The pseudo-code for our algorithm is described below:

(1) Generate factorial states
(2) $\lambda \leftarrow [A, B, \pi]$; Initial parameters
(3) Repeat
(4) $Q \leftarrow [q \mid \lambda]$; Generate sequence
(5) $\lambda' \leftarrow \lambda$; Generate new parameters
(6) Until λ converges
(7) $q^* \leftarrow argmax_q P[Q, Y]$; The mostly like sequence

In Figure 3, the network topology of MCFHMM where each of state consists of a list of appliances based on the factorial of the number of total appliances is shown. Each appliance has observation symbols, which contributes to the total measured observation symbols.

As stated in the earlier section, given the parameter estimator λ, where λ represent A, B, π and the Observation symbol, $Y = \{y_1, y_2, y_3, \ldots, y_T\}$, How do we choose a corresponding state sequence $Q = \{q_1, q_2, q_3, \ldots, q_M\}$ which is optimal to the appliance state of operation and identification.

Mathematically, this quest can be express in the Equation (10):

$$\text{argmax}_q P(Q|Y,\lambda) = \frac{\text{argmax}_q \; P(Y|Q,\lambda)P(Q)}{P(Y)} \tag{10}$$

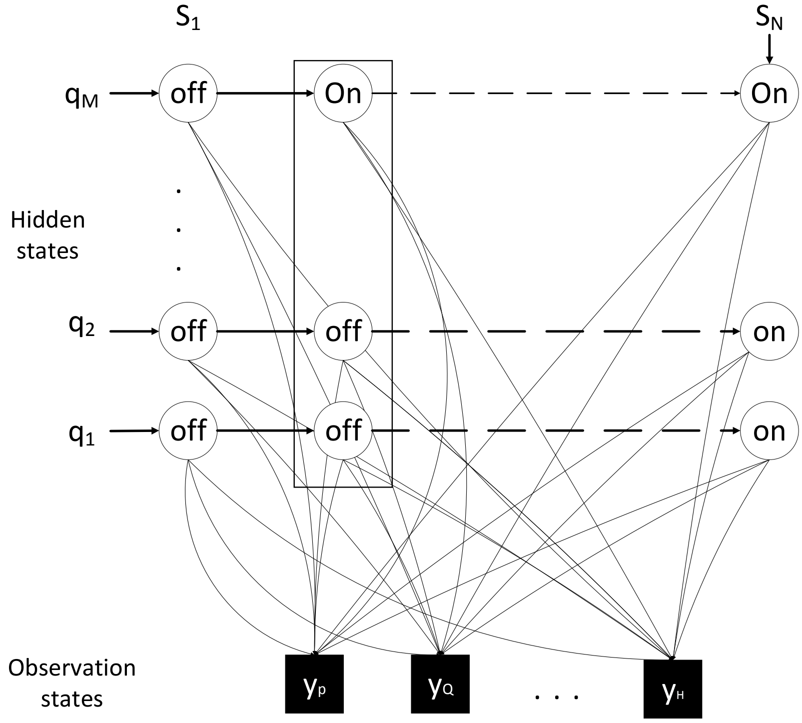

Figure 3. Network topology of multiple conditional factorial hidden Markov model (MCFHMM).

For a single input, *Y* will be constant and therefore so will *P(Y)*, so we only need to find:

$$\text{argmax}_q P(Y|Q,\lambda)P(Q) \tag{11}$$

$P(Y|Q)$: the probability of the observation sequence given the state sequence is the model likelihood, whereas *P(Q)* the probability of the states is the prior probability of the sequence. It can be seen from the above algorithm that a complete specification of HMM requires specification of observation symbol, the specification of the state and the three probability measures A, B and π. For convenience, we use the compact notation:

$$\lambda = (A, B, \pi) \tag{12}$$

Given that:

$$Y = (P_m;\ Q_m;\ 3H_m;\ 5H_m;\ 7H_m) \tag{13}$$

$$q_i = (P_i;\ Q_i;\ 3H_i;\ 5H_i;\ 7H_i) \tag{14}$$

Then:

$$\begin{aligned} Err_i &= |Y - Q_i| \\ &= \sqrt{(P_m - P_i)^2 + (Q_m - Q_i)^2 + (3H_m - 3H_i)^2 + (5H_m - 5H_i)^2 + (7H_m - 7H_i)^2} \end{aligned} \tag{15}$$

where:

$$Err_{\text{Total}} = \sum_{i=1}^{N} Err_i \tag{16}$$

The probability distribution for an appliance, q_i being ON can be defined as one minus the ratio of the error of the state to the total error of the model.

$$P_{\text{on}}(q_i) = 1 - \frac{Err_i}{Err_{\text{Total}}} \geq \varepsilon; 1 \leq i \leq N, \varepsilon \approx 0.9 \tag{17}$$

On the other hand, the probability distribution for an appliance, q_i being OFF can be defined as the ratio of the error of the state to the total error of the model:

$$P_{\text{off}}(q_i) = \frac{Err_i}{Err_{\text{Total}}};\ 1 \leq i \leq N \tag{18}$$

The probability distribution for the observation, y_i can be defined as the observation over the total observations:

$$P(y_i) = \frac{y_i}{\sum_{j=1}^{M} y_j};\ 1 \leq i \leq M \tag{19}$$

The initial state distribution, $\pi = \{\pi_i\}$ where:

$$\pi_j = \prod_{i=1}^{M} p(q_i|S_j^i);\ \ 1 \leq j \leq N \tag{20}$$

The *APD* = $\{P_{\text{rs}}\}$ where:

$$P_{\text{rs}} = P(q_t = r|q_t = S_i, \text{stock}]; 1 \leq i \leq N \tag{21}$$

The *ADD* = $\{d_{ij}\}$ where:

$$d_{ij} = P\big(d = \ S_i\big|q_t = S_j\big); 1 \leq j \ \leq N, 1 \leq k \leq M \tag{22}$$

The Transition probability distribution, $A = \{a_{ij}\}$ is the movement from a state S_j to another state, S_i:

$$a_{ij} = p[S_i|S_j;\ [i,j|1 \leq i,j \leq N]] \tag{23}$$

The observation symbol probability distribution in state j, $B = \{b_j(k)\}$ where:

$$b_j(t) = p(y_t\ |q_i^t = S_j]; 1 \leq j \ \leq N, 1 \leq t \leq T \tag{24}$$

i.e., the probability distribution is based on the dependency distribution between the appliances, the *APD* and the observation distribution. Accordingly:

$$B = \prod_{t=1}^{T} \frac{\sum_{i=1}^{N} q_i^t}{y_t}; 1 \leq i \leq N \tag{25}$$

$$a_{ij_{\text{minor}}} = ADD \cdot APD \cdot AUP \tag{26}$$

The *APD* for a given state *j*, *APD* = $\{P_j\}$ where:

$$P_j = \prod_{j=1}^{V} q_j = r\,; 1 \leq j \leq V, V \in M \tag{27}$$

The *ADD* = $\{d_{ij}\}$ for all the appliances in a given state j can be estimated by finding the Pearson correlation coefficient between each pair of appliances. The conditional probability of all the pair greater 0.9 is used as an additional feature, *i.e.*,

$$\text{Pearson correlation, } r_{ij} = \frac{\sum_{t=1}^{T}[q_i^t - \overline{q_i}][q_j^t - \overline{q_j}]}{\sqrt{\sum_{t=1}^{T}[q_i^t - \overline{q_i}]^2}\sqrt{\sum_{t=1}^{T}[q_j^t - \overline{q_j}]^2}}, 1 \le i, j \le \mathrm{N} \tag{28}$$

$$d_{ij} = P\,(q_i \mid q_j, r_{ij} > 0.9\,, S_j); 1 \le i, j \le \mathrm{N} \tag{29}$$

The appliance usage penetration (AUP) = U_{iT} for all the appliances in a given state j, can be estimated as the sum of the probabilities of the occurrence of the appliances:

$$a_{ij} = \left(\prod_{t=1}^{T} \frac{\sum_{i=1}^{N} q_i^t}{y_t}\right) + \left(\sum_{i=1}^{N} \frac{\sum_{t=1}^{T} q_i^T}{\sum_{i=1}^{V}\sum_{t=1}^{T} q_i^T}\right)\left(\sum_{i=1}^{N} \frac{\sum_{t=1}^{T} q_i^T}{\sum_{i=1}^{V}\sum_{t=1}^{T} q_i^T}\right) P\left(q_i \middle| q_j, r_{ij} > 0.9\,, S_j\right) \tag{30}$$

For a given set of parameter λ initially estimated, there is now the need to find how the choice of the sequence of the state has on the observation sequence, whether it could represent the given model. The joint probability of the observation sequence, Y and the set of the state sequence Q could be estimated by using the Forward-Backward procedure. Considering the forward variable, $\alpha_t(i)$. We defined $\alpha_t(i)$ as:

$$\alpha_t(i) = P(y_1 y_2 \cdots y_t, Q_t = S_i \mid \lambda) \tag{31}$$

where $Q_t = \{q_1, q_2, \cdots, q_N\}$.

We estimated the probability of the partial observation sequence y_1, y_2, ⋯, y_t until time t and the state S_i at time t, given the model λ. What this means is that we solve $\alpha_t(i)$ inductively. We initially estimated the forward variable by multiplying the initial observation symbol probability with its initial state distribution.

$$\alpha_1(j) = \prod_{i=1}^{M} \begin{matrix} p\left(q_i \middle| S_j^i\right) P\left(y_1 \middle| q_i^1 = S_j\right); \\ 1 \le j \le N; 1 \le t \le T \end{matrix} \tag{32}$$

For N states, the initial forward variable generates a vector matrix, which consists of the probability of an appliance set, Q_1 occupying states S_i and the probability of its observation distribution. Subsequently, at time $t > 1$, the forward variable could be estimated by induction as:

$$\alpha_{t+1}(i) = \left(\sum_{j=1}^{N} \alpha_t(i) a_{ij}\right) b_j(y_{t+1}), 1 \le t \le T - 1.1 \le j \le N \tag{33}$$

Similarly, we defined $\beta_t(i)$ as:

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij}\, b_j(y_{t+1}) \beta_{t+1}(i); 1 \le t \le T - 1.1 \le j \le N \tag{34}$$

Finally, the joint probability of the observation sequence and the state sequence given a parameter could be defined as:

$$P(Q, O \mid \lambda) = \sum_{i=1}^{N} \alpha_T(i) \tag{35}$$

The main objective of the energy disaggregation algorithm is to discover the states of the appliances, which contribute to the observation sequence. Our focus is on the hidden states of the MCFHMM model that matches the observation sequence. After learning the parameter, λ, we estimated the most likely state sequence, q^* that maximize the model:

$$q^* = \mathrm{argmax}_{\mathrm{q}}\, P(\mathrm{Y}, \mathrm{Q} \mid \lambda) \tag{36}$$

4. Experimental Setup and Evaluation

In case I, the appliances were connected individually whilst in case II; the appliances were connected in groups in different combinations. Figures 4 and 5 represent the configuration of single appliance and multiple appliances configuration respectively.

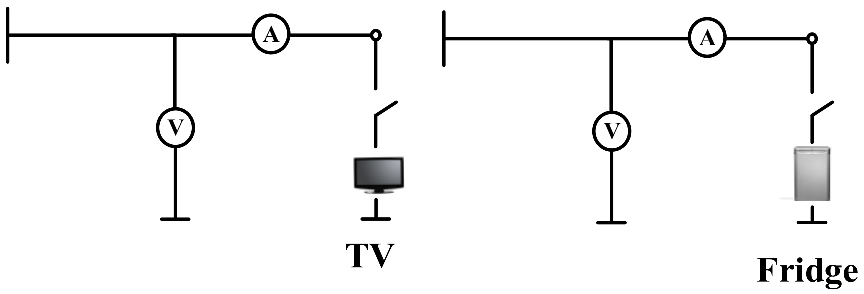

Figure 4. Single appliance configuration.

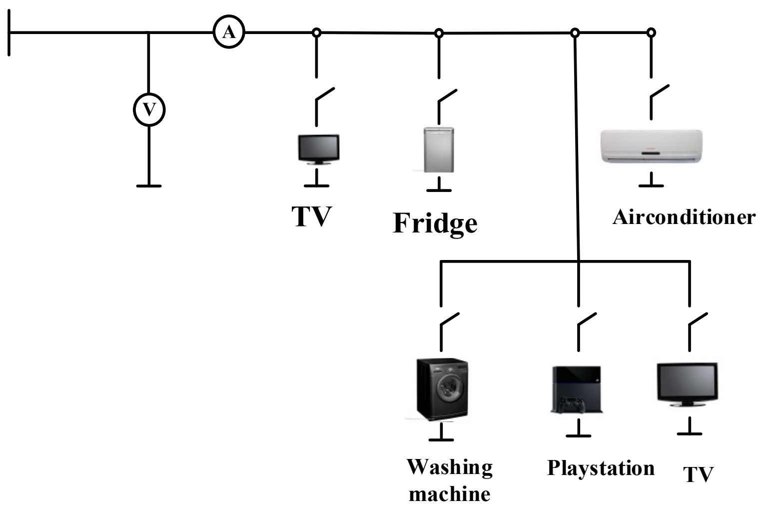

Figure 5. Multiple appliance configurations.

In order to validate the proposed model, we performed an experiment in the practical environment. Six sampled electrical appliances were used as loads in this case study: Printer, Vacuum cleaner, LCD TV, Desktop computer, Standing heater, and Electric Iron. These electrical appliances are assumed to be operated as a two-state (A), multi-state (B) and continuously(C) variable loads. Two measurement procedures were performed to collect data for each load. Table 3 below provides detailed power rating of the under listed devices.

In each case, the instantaneous current and voltage were measured and the power features such as real power, reactive power and the current harmonics were computed. Two different devices did the computations of these features: an Oscilloscope and a device we built by ourselves called REM (*i.e.*, recognition energy meter). Figures 6 and 7 show REM and the measurements of the individual appliances, respectively.

Table 3. Appliance power ratings.

Load	Appliance name	Appliance type	State	Power ratings
Load 1	LCD TV	32" SEETIV	A	60 W, 0.272 A
Load 2	Vacuum cleaner	Daewoo electronics RC-715	B	1100 W, 5 A
Load 3	Desktop computer	Samsung DM-V100-AA230G	A	150 W, 1.5 A
Load 4	Printer	Samsung ML-6080	A	250 W, 2.1 A
Load 5	Standing heater	SUO 7514-9006	B	3.63 A, 800 W
Load 6	Iron	Tefal pressing iron	C	4.55 A, 1000 W

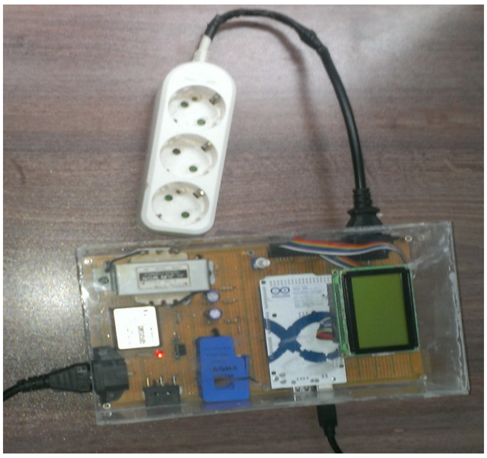

Figure 6. Recognition energy meter.

Figure 7. Appliance measurement.

Tables 4 and 5 show the estimated measurements and the actual measurements when one load was connected at a time.

Form the tables, it can be seen that pure resistive appliance show no harmonic components while non-linear loads with high power rating shows the highest harmonic components. The error associated with the various measurements is shown in Table 6.

Table 4. Estimated feature parameters.

Load	Power		Current harmonics		
	P	***Q***	**3rd**	**5th**	**7th**
Load 1	59.37	79.8	0.37	0.13	0.04
Load 2	1100.19	707.59	0.86	0.33	0.30
Load 3	69.79	82.01	0.47	0.13	0.063
Load 4	249.59	279.83	0.30	0.16	0.061
Load 5	800.5	2.1	0.00	0.00	0.00
Load 6	1001	1.9	0.00	0.00	0.00

Table 5. Actual feature parameters.

Load	Power		Current harmonics		
	P	***Q***	**3rd**	**5th**	**7th**
Load 1	49.37	76.8	0.32	0.15	0.06
Load 2	1100.19	707.53	0.85	0.35	0.30
Load 3	68.76	82.31	0.46	0.12	0.062
Load 4	237.29	279.85	0.33	0.15	0.06
Load 5	812.62	2.3	0.00	0.00	0.00
Load 6	1090	1.7	0.00	0.00	0.00

Table 6. Error estimation.

Load	Error %	Load	Error %
Load 1	0.104	Load 4	0.123
Load 2	0.001	Load 5	0.121
Load 3	0.011	Load 6	0.040

However, in case II, the appliances were connected in twos and threes in different combinations and the results measured and recorded. The labels of the loads were taken into consideration when storing the data. Table 7 shows the aggregated features of the appliances.

Table 7. Appliances aggregated features.

Load	Power		Current harmonics		
	P	***Q***	**3rd**	**5th**	**7th**
$Load_{12}$	1149.59	784.34	1.175	0.505	0.365
$Load_{123}$	1218.33	866.65	1.635	0.625	0.427
$Load_{43}$	306.07	362.17	0.795	0.275	0.127
$Load_{135}$	930.76	161.42	0.785	0.275	0.127
$Load_{56}$	1817.65	4.01	0.001	0.002	0.001
$Load_{246}$	2342.49	989.09	1.185	0.505	0.36

In a practical environment, there could be over a million different kinds of appliances. In order to test the robustness of the algorithm, we generated a million simulated data and stored them the same way as the actual data. We decided to test our classification algorithm on both the synthetic data and the actual data. For evaluation metric accuracy, we adapted a metric from the information retrieval domain, F-measure. Using this metric, we converted our method to a binary classifier such that if an appliance is turned on, it is considered as 1, otherwise 0. Since most appliances in our evaluation have a standard deviation 10%, we computed *F-measure* with a ρ = 0.1 as:

$$F\text{-}measure = \frac{2 \cdot precision \cdot Recall}{Precision + Recall} \quad (37)$$

where:

$$Recall = \frac{True\ positive\ (TP)}{True\ positive\ (TP) + False\ negative\ (FN)} \quad (38)$$

$$Precision = \frac{True\ positive\ (TP)}{True\ positive\ (TP) + Fasle\ positive\ (FP)} \quad (39)$$

We subsequently computed F-measure for all possible combinations and their maximum values used as their performance. Table 8 shows the performance of the entire model against the aggregated loads.

The performance of MCFHMM is compare to one of the know optimization algorithm called particle optimization swarm (PSO). PSO optimizes a problem by iteratively trying to improve a candidate solution with regards to a given measure of quality. PSO optimizes a problem by having a population of candidate solutions, here dubbed particles, and moving these particles around in search-space according to simple mathematical formulae over the particle's position and velocity. Each particle's movement is influenced by its local best-known position but is also guided toward the best-known positions in the search-space, which are updated as better positions are found by other particles. This is expected to move the swarm toward the best solutions. For a limited number of appliances remaining, our objective with PSO was to formulate a mathematical model that will minimize the cost associated the measured values and the expected whiles satisfying the constraint. For each experiment, we chose a particle size of 50 and the lowest error gradient of 1e-99. We set the maximum number of iteration to 1000 and maximum iteration without change to 100. Our objective here is to build a multiple of solutions with the same minimum cost and estimate the best solution from the candidate solutions.

Table 8. Model performance.

Load	Appliance	MCFHMM
Load$_{12}$	LCD TV, Vacuum cleaner	0.899
Load$_{123}$	LCD TV, Vacuum cleaner, desktop computer	0.975
Load$_{43}$	Printer, desktop computer	0.869
Load$_{135}$	LCD TV, desktop computer, standing heater	0.980
Load$_{56}$	Standing heater, Iron	0.635
Load$_{246}$	Vacuum cleaner, printer, Iron	0.723

It was observed from experiment, that for a small dataset, PSO does very well as compared to MCFHMM as seen in Figure 8. The candidate solutions generated by both methods were constant and most cases, the most optimal solution was obtained. However, as the number appliance increase, the performance of PSO falls. The candidate's solution generated by PSO is no more constant and in most times, the optimal solution could not be obtained. For a number of appliances, more than 50 PSO could not get the optimal solution because of the wide search space and the combinatorial solution required. However, in all these cases, MCFHMM did fairly well irrespective of the number of appliances to search from. MCFHMM performs better when it has history of appliance to consider and the appliance usage pattern.

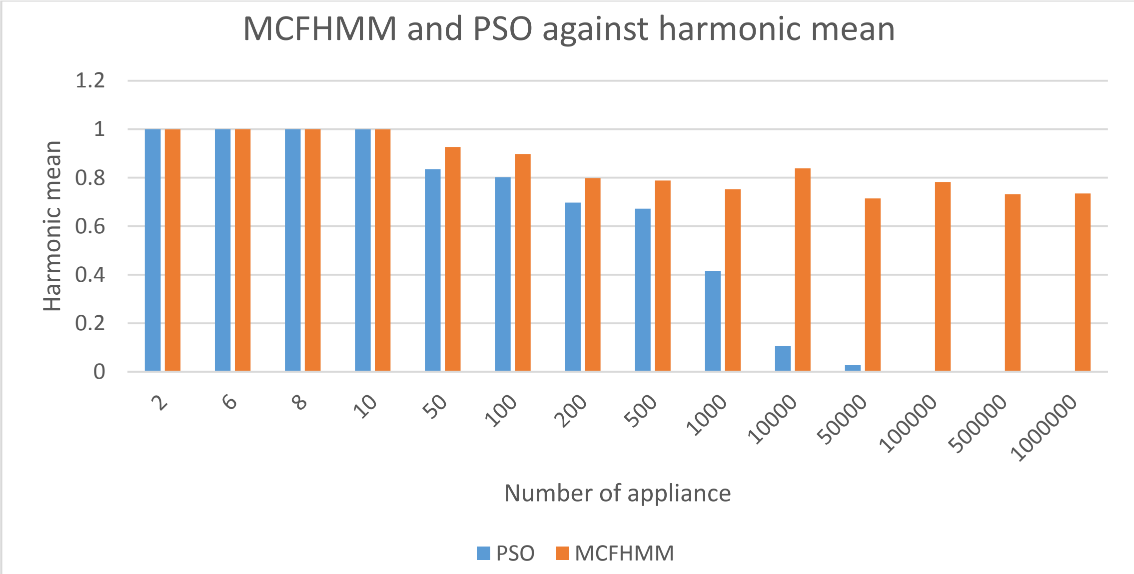

Figure 8. Harmonic mean of MCFHMM and particle optimization swarm (PSO).

5. Conclusions

In this paper, we investigated how electrical appliances could be detected and identified from an aggregated power based on residential load using MCFHMM algorithm. Our algorithm is capable of accurately disaggregating power data into individual appliance profile. This new classification of residential appliances uses the main power consumption (*i.e.*, real and reactive power), current harmonics (*i.e.*, 3rd, 5th, 7th) and working behavior as the features of each appliance. With this, new appliance identification is designed and implemented. The first step of appliance recognition is event detection, which combines the advantages of steady-state features (*i.e.*, real and reactive features) and transient features (*i.e.*, current harmonics). The sampling of these signals was 1.2 KHz so as to capture the signal harmonics. Based on these measured features, a new classification and identification algorithm, a variant of HMM was adopted. The algorithm identifies an appliance based on appliance identification information already stored in the database. Some of the identification information, such as appliance penetration rate, *ADD* and usage penetration is collected by an information agency whilst the remaining information is at the circuit breaker level of the residential building.

The identification accuracy of multi appliances without any prior information is above 80%. From the performance measure, it could be seen that non-linear appliances have higher predictability as compared with the linear load. This suggests that inclusion of more features could lead to more accuracy in the detection and identification of appliances. Our future goal is to find a more accurate method of load identification based on a dynamic appliance feature list and time. We believe this will improve the application identification accuracy and computational time.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. *Review of World Energy*; British Petroleum: Houston, TX, USA, 2010.
2. Bureau of Energy. *Long-Term Demand Forcast and Energy Report*; Technical Report for Ministry of Economic Affairs: Taipei, Taiwan, 2011.
3. Armel, K.C.; Gupta, A.; Shrimali, G.; Albert, A. Is disaggregation the holy grail of energy effiiciency? The case of electricity. *Energy Policy* **2013**, *52*, 213–234.
4. *Building Energy Data Book*; United States Department of Energy: Washington, DC, USA, 2009.
5. Ehrhardt-Martinez, K.; Donnelly, K.; Laitner, J. *Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities*; American Council for an Energy-Efficient Economy: Washington, DC, USA, 2010.
6. Farinaccio, L.; Zmeureanu, R. Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses. *Energy Build.* **1999**, *30*, 245–259.
7. Kempton, W.; Layne, L. The consumer's energy analysis environment. *Energy Policy* **1994**, *22*, 857–866.
8. Brandon, G.; Lewis, A. Reducing household energy consumption: A qualitative and quantitative field study. *Environ. Psychol.* **1997**, *19*, 994–999.
9. Attari, S.; Dekay, M.; Davidson, C.; de Bruin, W. Public perceptions of energy consumption and savings. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 16054–16059.
10. Gardner, G.T.; Stern, P.C. The short list: The most effective actions U.S. households can take to curb climate change. *Environ. Mag.* **2008**, *50*, 12–25.
11. Park, S.; Kim, H.; Moon, H.; Heo, J.; Yoo, S. Concurrent simulation platform for energy-aware smart metering systems. *IEEE Trans. Consum. Electron.* **2010**, *56*, 1918–1926.
12. Balasubramanian, K.; Cellatoglu, A. Improvements in home automation stategies for designing apparatus for efficient smart home. *IEEE Trans. Consum. Electron.* **2008**, *54*, 1681–1687.
13. Priyadharshini, A.; Devarajan, N.; Uma saranya, A.; Anitt, R. Survey of Harmonics in Non Linear Loads. *Int. J. Recent Technol. Eng.* **2012**, *1*, 92–97.
14. Zeifman, M.; Roth, K. Non-intrusive appliance load monitoring: Review and outlook. *IEEE Trans. Consum. Electron.* **2011**, *57*, 76–84.
15. Hart, G.W. Non-Intrusive Appliance Load Monitoring. *IEEE Proc.* **1992**, *80*, 1870–1891.
16. Chang, H.H.; Chen, K.L.; Tsai, Y.P.; Lee, W.J. A New Measurement Method for Power Signature of Non-intrusive Demoand Monitoring and Load Identification. In Proceedings of the Industry Applications Society Annual Meeting (IAS), Orlando, FL, USA, 9–13 October 2011.
17. Ducange, P.; Marcelloni, F.; Michela, A. A Novel Approach Based on Finite-State Machines with Fuzzy Transitons for Noninstrusive Home Appliance Monitoring. *IEEE Trans. Ind. Appl.* **2014**, *10*, 1185–1197.
18. Chang, H.H.; Lin, L.S.; Chen, N.; Lee, W.J. Particle-Swarm-Optimization-Based Nonintrusive Demand Monitoring and Load Identification in Smart Meters. *IEEE Trans. Ind. Appl.* **2013**, *49*, 2229–2236.
19. Parson, O.; Siddhartha, G.; Weal, M.; Rogers, A. Non-Intrusive Load Monitoring Using Prior Models of General Appliance Types. In Proceedings of the Twenty-Sixth Conference on Artificial Intelligence (AAAI-12), Toronto, ON, Canada, 22–26 July 2012.

20. Lin, Y.H.; Tsai, M.S. An Advanced Home Energy Management System Facilitated by Nonintrusive Load Monitoring with Automated Multiobjective Power Scheduling. *IEEE Trans. Smart Grid* **2015**, *6*, 1839–1851.
21. Dawei, H.; Weixuan, L.; Nan, L.; Ronald, G.H.; Thomas, G.H. Incorporating Non-Intrusive Load Monitoring into Building Level Demand Response. *IEEE Trans. Smart Grid* **2013**, *4*, 1870–1877.
22. Lin, Y.H.; Tsai, M.S. Non-Intrusive Load Monitoring by Novel Neuro-Fuzzy Classification Considering Uncertainties. *IEEE Trans. Smart Grid* **2014**, *5*, 2376–2384.
23. Chang, H.H.; Lian, K.L.; Su, Y.C.; Lee, W.J. Power-Spectrum-Based Wavelet Transform for Noninstrusive Demand Monitoring and Load Identification. *IEEE Trans. Ind. Appl.* **2014**, *50*, 2081–2089.
24. Ahmed, Z.; Alexender, G.; Michele, N.; Muhammad, A.I. Low-Power Appliance Monitoring Using Factorial Hidden Markov Model. In Proceedings of the 2013 IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Melbourne, Australia, 2–5 April 2013; pp. 527–532.
25. Wang, Z.; Zheng, G. Residential appliances indentificaiton and monitoring by a non-intrusive method. *IEEE Trans. Smart Grid* **2012**, *3*, 80–92.
26. Benyoucef, D.; Klein, P.; Bier, T. Smart Meter with Non-Intrusive Load Monitoring for Use in Smart Homes. In Proceedings of the 2010 IEEE International Energy Conference and Exhibition (EnergyCon), Manama, Bahrain, 18–22 December 2010; pp. 96–101.
27. Kim, H.; Marwah, M.; Arlitt, M.; Lyon, G.; Han, J. Unsuperivised Disaggregation of Low Frequency Power Measurement. In Proceedings of the Eleventh SIAM International Conference on Data Mining (SDM), Mesa, AZ, USA, 28–30 April 2011.
28. Tehseen, Z.; Dietmar, B.; Adeel, Z. A Hidden Markov Model Based Procedure for Identifying Household Electric Loads. In Proceedings of the IECON 2011—37th Annual Conference on IEEE Industrial Electronics Society, Melbourne, Australia, 7–10 November 2011; pp. 3101–3106.
29. Ning, Z.; Ochoa, L.F.; Kirschen, D.S. Investigating the impact of demand side management on residential customers. In Proceedings of the 2011 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies (ISGT Europe), Manchester, UK, 5–7 December 2011.
30. Prasad, R.S.; Semwal, S. A Simplified New Procedure for Identification of Appliances in Smart Meter Applications. In Proceedings of the 2013 IEEE International Systems Conference (SysCon), Orlando, FL, USA, 15–18 April 2013; pp. 339–344.
31. Rabiner, L. A tutorial on hidden Marko v models and selected applications in speech recognition. *IEEE Proc.* **1989**, *77*, 257–286.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).