Unsupervised Disaggregation of Low Frequency Power Measurements

Hyungsul Kim* Manish Marwah† Martin Arlitt† Geoff Lyon† Jiawei Han*

Abstract

Fear of increasing prices and concern about climate change are motivating residential power conservation efforts. We investigate the effectiveness of several unsupervised disaggregation methods on low frequency power measurements collected in real homes. Specifically, we consider variants of the factorial hidden Markov model. Our results indicate that a conditional factorial hidden semi-Markov model, which integrates additional features related to when and how appliances are used in the home and more accurately represents the power use of individual appliances, outperforms the other unsupervised disaggregation methods. Our results show that unsupervised techniques can provide per-appliance power usage information in a non-invasive manner, which is ideal for enabling power conservation efforts.

1 Introduction

Concern over global climate change has motivated efforts to reduce the emissions of CO_2 and other GHGs (greenhouse gases). Energy use in the residential sector is a significant contributor of GHGs [49]. For example, the residential sector is responsible for over one third of all electricity use in the United States [2]. While information is available on the typical use of electricity in homes (e.g., space heating, space cooling, water heating and lighting account for about 50% of all residential electricity use [3]), it has not enabled most home owners to reduce their electricity consumption.

Two typical approaches to conserving energy are *efficiency* and *curtailment* [1]. The former involves one-time actions (e.g., upgrading to more energy-efficient appliances) that have a higher cost. The latter requires continuous participation (e.g., using less heating/cooling on a daily basis), with a smaller incremental cost. There are two general issues that inhibit consumers from applying these techniques. First, energy use is a very abstract concept to most consumers [24, 8]. Second, consumers are often mistaken about how energy is used in the home, and thus which actions would be most beneficial for conserving energy [15, 4, 38]. Numerous studies have identified the attributes of a solution to these issues: personalized, frequent, continuous, credible, clear and concise feedback that provides an appliance-specific breakdown of how energy is used in the home [5, 19, 7, 1, 11, 13, 15, 38]. Field studies showed that with proper feedback, residential electricity and/or gas use could be reduced by up to 50% [14], although typical savings were in the 9%-20% range [42, 20, 45, 1, 47]. Improved feedback can also help curtail peak use by up to 50% [27, 44].

Much of this research occurred decades ago, in response to the oil crisis in the 1970s [39]. At that time, computer hardware technology was not as advanced, so providing frequent feedback to home owners cost effectively seemed infeasible [19]. As the crisis subsided (and prices dropped), the financial incentive to conserve diminished [45]. The growing concern over climate change has revived the importance of conservation. Today, computer hardware technology is more advanced, so frequent feedback is now feasible. In particular, as old power meters are replaced with smart meters, more information will be available to consumers [38].

An open issue is how to provide an appliance-specific breakdown of energy use in a cost-effective manner. Without this, residential energy conservation efforts are unlikely to achieve widespread success. This paper investigates how to obtain this information via *power load disaggregation*. While this topic has received attention since the early 1990s [18], our work has three distinguishing characteristics. First, we assume only *low frequency measurements* are available. This makes our techniques more widely applicable since smart meters typically provide samples no more than once per second. Second, we use an *unsupervised disaggregation approach*, as this does not require the data to be labeled, which can be laborious and intrusive. Third, we use *empirical data* collected from seven homes over a six month period.

The specific problem we address is as follows. Given the aggregate power consumption for T time periods, $Y = \langle y_1, y_2, \ldots, y_T \rangle$, and the number of appliances, M, we want to infer the power load of each of the M appliances, that is,

$$\begin{aligned} Q^{(1)} &= \langle q_1^{(1)}, q_2^{(1)}, \ldots, q_T^{(1)} \rangle \\ Q^{(2)} &= \langle q_1^{(2)}, q_2^{(2)}, \ldots, q_T^{(2)} \rangle \\ &\vdots \\ Q^{(M)} &= \langle q_1^{(M)}, q_2^{(M)}, \ldots, q_T^{(M)} \rangle \end{aligned}$$

such that $y_t = \sum_{i=1}^{M} q_t^{(i)}$, where $q_t^{(i)}$ is the power load of appliance i at time t.

We achieve this using energy disaggregation meth-

*University of Illinois, Urbana-Champaign, IL
†HP Labs, Palo Alto, CA

ods based on extensions of a hidden Markov model (HMM). We use four HMM variants to model the data. Factorial HMM (FHMM) models the hidden states of all the appliances. Conditional FHMM (CFHMM) extends FHMM to incorporate additional features, such as time of day, other sensor measurements, and dependency between appliances. A third variant, factorial hidden semi-Markov model (FHSMM) extends FHMM to better fit the probability distributions of the state occupancy durations of the appliances. The fourth variant composes FHSMM and CFHMM, to consider the additional features together with the more accurate probability distributions of the state occupancy durations of the appliances. We refer to this variant as conditional factorial hidden semi-Markov model (CFHSMM).

Our paper makes two key contributions. First, we explore four unsupervised techniques for disaggregating low frequency power load data. Second, we provide a performance evaluation of the techniques using power load data from real homes. We find that CFHSMM outperforms the other variants, and demonstrate that unsupervised disaggregation techniques are feasible.

The remainder of the paper is organized as follows. Section 2 provides background information and related work. Section 3 discusses features that can be used for disaggregation of low frequency power measurements. Section 4 describes the four models we use to identify the stable-state signatures of household appliances. Section 5 presents our results, using power load data from actual homes. Section 6 summarizes our work.

2 Background and Related Work

2.1 Background Hidden Markov Models (HMM) are used for probabilistically modeling sequential data. HMMs are known to perform well at tasks such as speech recognition [37], problems in computational biology [28], etc.

A discrete-time hidden Markov model 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. More formally, an HMM is defined by the following:

- $S = \{S_1, S_2, \cdots, S_N\}$ the finite set of hidden states.
- the transition matrix $\boldsymbol{A} = \{a_{ij}, 1 \leq i, j \leq N\}$ representing the probability of moving from state S_i to state S_j,

$$a_{ij} = P(q_{t+1} = S_j | q_t = S_i), 1 \leq i, j \leq N,$$

  with $a_{ij} \geq 0, \sum_{j=1}^{N} a_{ij} = 1$, and where q_t denotes the state occupied by the system at time t.

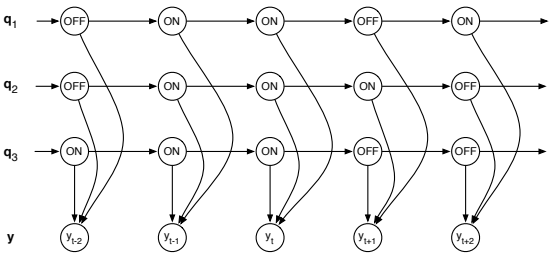

Figure 1: Graphical representation of factorial HMM.

- the emission matrix $\boldsymbol{B} = \{b(o|S_j)\}$, indicating the probability of emission of symbol $o \in V$ when system state is S_j; V can be a discrete or a continuous set, in which case $b(o|S_j)$ is a probability density function.
- $\boldsymbol{\pi} = \{\pi_i\}$, the initial state probability distribution,

$$\pi_i = P(q_1 = S_i), 1 \leq i \leq N$$

  with $\pi_i \geq 0$ and $\sum_{i=1}^{N} \pi_i = 1$.

Suppose we have sequential data $\boldsymbol{y} = \{y_1, y_2, \ldots, y_t, \ldots, y_T\}$. Every y_t is generated by a hidden state, q_t. The underlying states $\boldsymbol{q} = \{q_1, q_2, \cdots, q_t, \ldots, q_T\}$ form a Markov chain. Given the current state, the next state is independent of the past (Markov property).

$$P(q_{t+1}|q_t, q_{t-1}, \ldots, q_1) = P(q_{t+1}|q_t)$$

As an extension of HMMs, Ghahramani and Jordan [17] introduced *factorial HMMs* to model multiple independent hidden state sequences, as shown in Figure 1. In a FHMM, if we consider $\boldsymbol{Y} = \langle y_1, y_2, \ldots, y_T \rangle$ to be the observed sequence then $\boldsymbol{q} = \{\boldsymbol{q^{(1)}}, \boldsymbol{q^{(2)}}, \ldots, \boldsymbol{q^{(M)}}\}$ represents the set of underlying state sequences, where $\boldsymbol{q^{(i)}} = (q_1^{(i)}, q_2^{(i)}, \ldots, q_T^{(i)})$ is the hidden state sequence of the chain i. In general, factorial learning algorithms are used to discover multiple independent causes or factors underlying the data. FHMMs are preferred to HMMs for modeling time series generated by the interaction of several independent processes because using HMMs to model such processes requires exponentially many parameters to represent all the states.

2.2 Related Work The initial solution for disaggregating residential power load information was proposed by Hart [18]. Hart demonstrated how different electrical appliances generated distinct power consumption signatures, which often could be seen in the aggregated power load. He showed how on-off events were sufficient to

characterize the use of some appliances. For other appliances, Hart considered using Finite State Machines to develop signatures. Hart called this approach "Non-intrusive Appliance Load Monitoring"(NALM).

Other research efforts have attempted to improve NALM, often by proposing alternative signature identification techniques. Farinaccio and Zmeureanu [12] use a pattern recognition approach to disaggregate whole-house electricity consumption into its major end-uses. Prudenzi [36] proposes a neural net approach for identifying the electrical signatures of residential appliances. Laughman *et al.* suggest collecting data at higher frequencies (e.g., 8,000 Hz) to use higher harmonics in the aggregate current signal to generate appliance signatures [29]. Ito *et al.* [22] extract features from the current (e.g., amplitude, form, timing) to develop appliance signatures. Suzuki *et al.* [46] use an integer programming approach to disaggregate residential power use. Saitoh *et al.* [41] extract nine features from the measured current signal, and use them to classify the state of an appliance. Kato *et al.* [23] describe an "electric appliance recognition method". It uses Principal Component Analysis (PCA) to extract features from electric signals. These features are classified using a Support Vector Machine. For "unregistered" appliances, a one-class SVM is used. Lin *et al.* [31] use a dynamic Bayesian network to take user behavior into account, and a Bayes filter to disaggregate the data online. However, these methods have practical limitations which motivate the development of alternative techniques. Matthews *et al.* reflect on some of these works and describe the characteristics of a workable solution [32]. Our work focuses specifically on disaggregating low frequency power load data without the need for extra sensors, as these are important attributes of a cost-effective solution.

Several research efforts have prototyped tools for in-home use. Serra *et al.* built a prototype power meter, which included software to disaggregate the power consumption and automatically identify different appliances (as well as to detect malfunctioning appliances) [43]. However, they considered only a small number of appliances and used very simple signatures; thus the approach seems unsuitable for actual home environments. Kim *et al.* augment electricity usage data from a single power meter with ambient signals from inexpensive sensors placed near appliances [25]. They use three types of indirect sensors: magnetic, acoustic and light, to distinguish between multiple appliances that are simultaneously on and monitor variable power consumption. Unfortunately, the need for additional sensors is undesirable from a practical perspective.

An interesting variation on the NALM approach was proposed by Patel *et al.* [35]. They use a plug-in sensor to detect electrical events within a home. They leverage the fact that mechanical switches produce electrical noise [21], and that the noise characteristics can vary dramatically by appliance [48]. They apply machine learning techniques to recognize specific devices being turned on or off. More specifically, they perform a Fast Fourier Transform on the incoming signal to separate the component frequencies. They then use a Support Vector Machine to classify which appliance was turned on. In several trials, they found accuracies of 85–90% in classifying the events. However, they cannot determine the power consumed during each event from the noise. To address this, they developed a sensor that can be installed by the end user [34].

Disaggregating power data in commercial settings has additional challenges. For example, Norford and Leeb [33] present results for space-conditioning equipment in an commercial setting. Some of the challenges include more identical appliances, and more complex appliance signatures.

Lastly, hidden Markov models have been applied to a wide range of topics. One relevant study is from Yadwadkar *et al.* [50]. They use profile hidden Markov models to recognize distinct applications within a network file trace. The success of their approach motivates us to explore HMMs for developing appliance signatures for residential power use.

3 Disaggregation with Low Sampling Rates

There are two kinds of features for power disaggregation – transient signatures and stable-state signatures [18]. Transient signatures capture electrical events, such as high frequency noise in electrical current or voltage, generated as a result of an appliance turning on or off [35]. Although these features are good candidates for use in disaggregation, sampling data fast enough to capture them requires special instrumentation. For example, Patel *et al.* use a custom built device to measure at rates up to 100KHz [35]. However, most smart meters deployed in the U.S. have low sampling rates, typically 1Hz or less.

Stable-state signatures relate to more sustained changes in power characteristics when an appliance is turned on/off. These persist until the state of the appliance changes, which can be captured with low frequency sampling. But even for stable-state features, the frequency of sampling is important since at low sampling rates the probability of multiple on/off events occurring between two measurements increases, making the disaggregation task more difficult. In addition to the real power measurement, AC power meters typically provide several other metrics, such as, reactive power, frequency, power factor, etc., each of which could

Label	Location	Appliance	Power
fam_tv	Family Room	Television	73 W
fam_ps3	Family Room	Playstation 3	67 W
fam_stereo	Family Room	Home Theater	41 W
kit_ref	Kitchen	Refrigerator	82 W
liv_tv	Living Room	Television	177 W
liv_xbox	Living Room	Xbox 360	111 W
off_laptop	Office	Laptop	61 W
off_monitor	Office	Monitor	38 W

Table 1: Summary of the household appliances.

potentially be used as additional features depending on the set of appliances to be disaggregated.

In this work, we focus on stable-state features since these features can be more readily obtained, e.g., from smart meters, in which case *no* additional instrumentation is required in the homes. The most effective feature for disaggregation is the real power measurement. However, other power features may help distinguish appliances, so our approach is designed to allow multiple other features to be integrated into the model. Other useful features, unrelated to power metrics, are: duration on/off, date/time, dependency between appliances, daily schedule of the occupants, etc. Further, unlike past work, we develop unsupervised learning algorithms for disaggregating the appliances.

We collected detailed power measurements from 7 homes in California, for a period of six months. To enable us to know the ground truth, we installed extensive instrumentation in the home, collecting data at the individual appliance level. We then aggregate the data from multiple individual appliances to test the ability of the methods to disaggregate this data. We use the original traces of power use for each instrumented appliance to assess the performance of the disaggregation methods. It is important to clarify that if we can successfully disaggregate the aggregate power data, thorough (and expensive) instrumentation of homes will not be necessary to obtain per-appliance measurements. Further, laborious "labelling" of the collected data is not required. This is an important practical consideration, and the motivation for our focus on unsupervised techniques.

In the following subsections, we focus on one home, and investigate the possible stable-state features. Table 1 lists a subset of the monitored appliances in the home. Each "Label" is an abbreviation formed from the appliance type and its location. For example, "fam_tv" is the television located in the Family Room, while "liv_tv" is the television located in the Living Room.

3.1 Power Consumption The real power consumption is the most significant feature. Table 1 shows the average values for each of the appliances. We assume that each appliance has two states (on and off) and

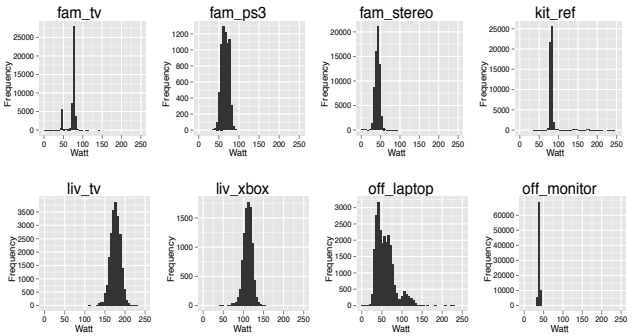

Figure 2: Histograms of appliance power consumption.

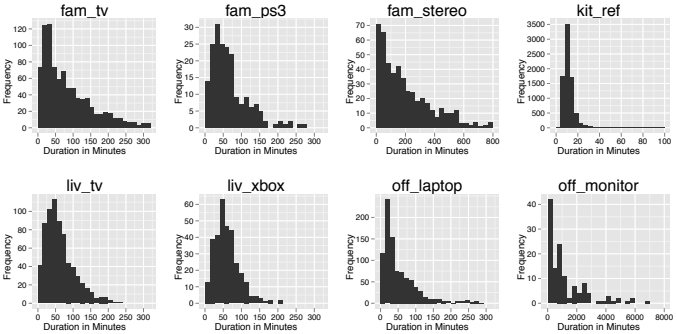

Figure 3: Histograms of appliance ON-durations.

its power consumption follows the Gaussian distribution when the appliance is on. As seen in Figure 2, this assumption is valid for most of the home appliances, except for the family room TV (fam_tv) and office laptop (off_laptop). fam_tv has a standby-mode in which it consumes less power. The power consumption of off_laptop varies depending on whether its battery is being charged, and its power state. Even though some appliances have multiple states, they can be considered to be composed of two or more two-state appliances.

3.1.1 ON-Duration Distribution Since we use HMMs to model the appliances, we want to determine what probability distribution function accurately captures the ON-durations. The geometric distribution is used for state occupancy in regular HMMs. However, the histograms of ON-durations shown in Figure 3 do not appear to be geometric. In geometric distributions, $Pr(d = x) \geq Pr(d = y) \iff x \leq y$. Thus, if we model the ON-state occupancy durations with a geometric distribution, it would mean that using an appliance for only one second occurs more frequently than using it for one minute. Obviously, this property does not hold for many household appliances. As Figure 3 shows, most of the peaks are not located in the first bin of the histograms. Thus, the ON-state occupancy dura-

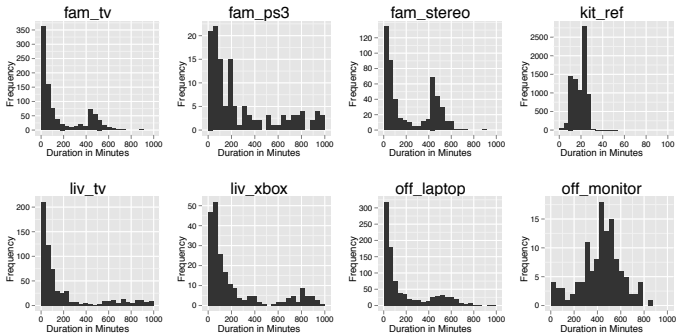

Figure 4: Histograms of appliance OFF-durations.

Label	λ	k	θ	LLR
fam_tv	0.00991	1.804	38.307	17.29
fam_ps3	0.01447	1.135	88.821	5.077
fam_stereo	0.00395	0.975	259.38	0.029
kit_ref	0.07783	5.895	2.1793	4151
liv_tv	0.01576	2.175	29.184	98.50
liv_xbox	0.01669	2.763	21.676	70.63
off_laptop	0.01840	1.371	39.633	26.73
off_monitor	0.00076	0.676	1945.2	7.143

Table 2: Estimated parameters for the exponential (λ) and gamma (k, θ) distributions, and LLR.

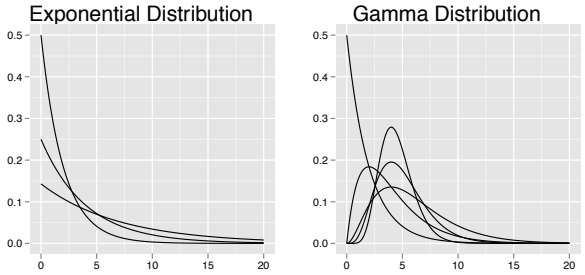

Figure 5: Exponential and Gamma distributions.

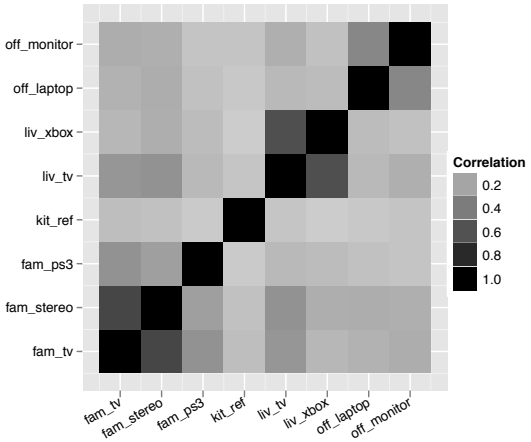

Figure 6: Correlations between the appliances.

tions need to be modeled with a different distribution.

We found that the gamma distribution is closer to most ON-duration distributions. Since the gamma distribution has two parameters, it has more freedom in terms of the distribution's shape. Figure 5 shows a set of exponential distributions, the equivalent of geometric distributions in the continuous domain, and a set of gamma distributions. We perform a quantitative comparison of the fitness of the gamma distribution with that of the exponential distribution.

For each appliance, we use maximum likelihood estimation (MLE) on the ON-durations to estimate the parameters for the exponential distribution and gamma distribution. The fitness of these distributions on the data is compared using log-likelihood ratio (LLR):

$$LLR = \log\left(\frac{\max_{k,\theta} P(durations|Gamma(k,\theta))}{\max_{\lambda} P(durations|Exp(\lambda))}\right)$$

Table 2 shows that all LLR values are positive, and most are large. This indicates that the gamma distribution is a better fit than the exponential distribution for all appliances.

3.1.2 OFF-Duration Shape As shown in Figure 4, there are generally two peaks in the OFF-duration distributions. The reason for the second peak is that most appliances are not used at night. This indicates the dependency between time of day and appliance use. If the second peaks are removed, the OFF-durations are approximated well by geometric distributions.

3.2 Dependency between appliances Usage patterns of some appliances show strong correlation with those of others. For example, an Xbox 360 cannot be used without a television, and a monitor cannot be used alone without a desktop or a laptop. We tested these dependencies in our dataset by measuring the correlations between every pair of appliances.

Figure 6 shows the Pearson's coefficients of all pairs of appliances as a heatmap. The figure shows four groups of strongly correlated appliances: {fam_tv, fam_stereo, fam_ps3}, {kit_ref}, {liv_tv, liv_xbox}, and {off_laptop, off_monitor}. Further, liv_tv and fam_tv are correlated, which implies that the family members in the house usually watch televisions at similar times. We also compute the conditional probabilities for every pair of appliances. The pairs with conditional probability greater than 0.9 are: P(fam_tv|fam_ps3) = 0.963, P(fam_stereo|fam_tv) = 0.944, P(fam_stereo|fam_ps3) = 0.998, P(liv_tv|liv_xbox) = 0.990, and P(off_monitor|off_laptop) = 1.000. Our results show that strong dependencies exist between appliances,

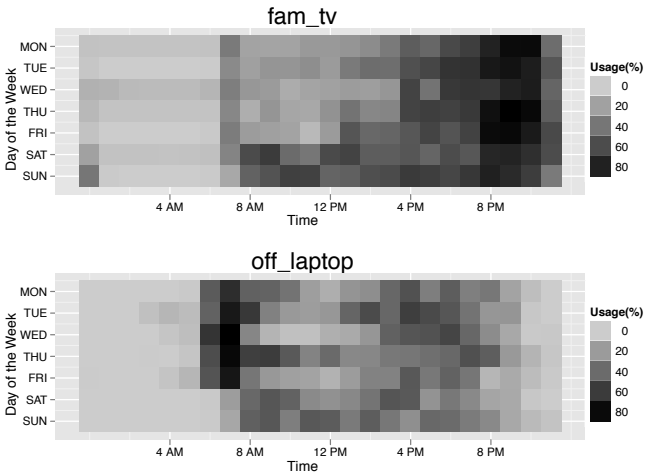

Figure 7: Daily and weekly usage patterns of appliances.

Algorithm 1 The Generative Approach with Hidden Variables.

1: $\lambda \leftarrow$ Initial parameters
2: **repeat**
3: $\quad \lambda' \leftarrow \lambda$
4: $\quad \lambda \leftarrow \arg\max_\lambda E\left[\log P(\boldsymbol{Y}, \boldsymbol{q}|\lambda)|\boldsymbol{Y}, \lambda'\right]$
5: **until** λ converges
6: $\boldsymbol{q}^* \leftarrow \arg\max_{\boldsymbol{q}} P(\boldsymbol{q}|\lambda, \boldsymbol{Y})$

which can be used as features for disaggregation.

3.3 Additional Features The performance of power load disaggregation can be improved if additional inputs that indirectly relate to the state of an appliance are available. We focus on inputs that do not require additional instrumentation. For example, people tend to have daily and weekly patterns in their activities. Thus, we expect usage of appliances to also have temporal patterns. Figure 7 shows the usage of fam_tv and off_laptop for each day of a week, aggregated over 6 months. The figure shows that the TV is watched more at night and on weekends; the laptop is used every weekday morning. Other appliances also exhibit temporal usage patterns (not shown). Thus, time of day and day of the week are useful features. In this work, we consider only time of day and day of week as additional features, as this information does not require additional instrumentation to be used. However, the models developed in the next section could integrate other features, if the information were available. For example, the outside temperature would strongly correlate with the use of heating or air conditioning. Similarly, sound, light or vibration sensors can help identify a variety of appliances [25].

4 Appliance Models

In this section, we develop probabilistic models of appliance behavior. These models integrate the stable-state features described earlier. Further, learning the parameters of these models is unsupervised. This is highly desirable for residential power disaggregation, as labeled data is not required, simplifying deployment.

Being variants of HMM, our models are generative, that is, we define a probabilistic model that explains the generating process of the observed data. These models can contain hidden variables that are not observed. In our case, the states of appliances are the hidden variables, and the aggregate power load is the observation.

The models have several parameters that can be learned from data. The learning process consists of estimating the parameters from the observations such that the model can best describe the observations. Then, using the model with these parameters, we estimate the hidden variables, which are the states of the appliances. Specifically, this algorithm is described in Algorithm 1. We first initialize the parameters. For a given observation $\boldsymbol{Y}$, we estimate the parameters in a model by an Expectation-Maximization algorithm (EM: Line 2-5). Then, we estimate the hidden states by using Maximum Likelihood Estimation (MLE: Line 6).

As our base model we chose a factorial hidden Markov model (FHMM), which is described in Section 2. Based on the observations from Section 3, we create three variants, which we describe next.

4.1 FHSMM An inherent problem in FHMMs is that a state occupancy duration is constrained to be geometrically distributed. However, as shown in Section 3.1.1, the ON-durations are modeled better with a gamma distribution. Modeling state occupancy durations in HMMs has been studied in [40, 30]. The models are called Hidden Semi-Markov Model (HSMM) or Non-Stationary Hidden Markov Model (NSHMM). We define a Factorial Hidden Semi-Markov Model (FHSMM) as the model obtained by combining the method of modeling state occupancy durations in HSMM with FHMM.

4.2 CFHMM FHMMs do not consider additional features such as time of day, day of week, or input from other sensors. To use these, we propose a Conditional Factorial Hidden Markov Model (CFHMM), where the transition probabilities are not constant but are conditioned on the extra features. This model is similar to a coupled hidden Markov model (CHMM) [6]. However, CFHMMs have a more general form, as they consider the dependencies between hidden state sequences and the additional input sequences.

Figure 8 shows the relationship of these two models with FHMM. Next, we combine FHSMM and CFHMM

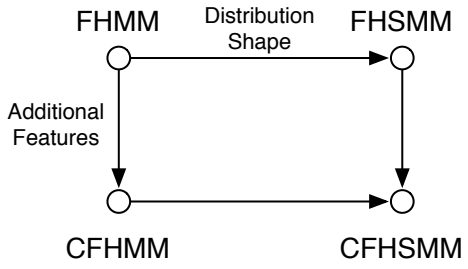

Figure 8: Relationships between the various models.

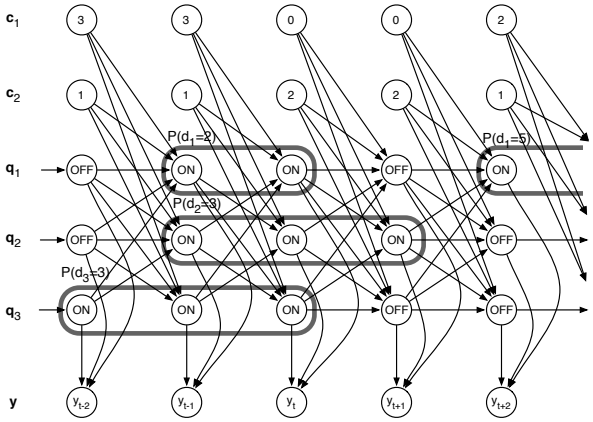

Figure 9: The graphical representation of CFHSMM.

to create the Conditional Factorial Hidden Semi-Markov Model (CFHSMM).

4.3 CFHSMM We extend the FHMM model to create the Conditional Factorial Hidden Semi-Markov Model (CFHSMM). This new model has the advantages of both FHSMM and CFHMM. Figure 9 shows the graphical representation of CFHSMM. $c_1, c_2, \ldots, c_K$ represent the additional features. Further, the model uses a gamma distribution for ON-durations. Lastly, the state of an appliance at time t also depends on the states of other appliances, and the additional features at time $(t-1)$. This extension allows the model to consider the dependencies between appliances and the dependencies on additional features.

4.3.1 Parameter Estimations There are several parameters in the model.

- $\pi_j^{(i)}$, the initial probabilities, $P(q_1^{(i)} = j)$
- $f_{jkl}^{(i)}$, the conditional probability for feature k of value l, $P(c_{t-1}^{(k)} = l | q_t^{(i)} = j)$
- $m_{jkl}^{(i)}$, the conditional probability for appliance k of state l, $P(q_{t-1}^{(k)} = l | q_t^{(i)} = j)$
- $\mu^{(i)}$, the mean of the power consumption for the appliance i
- $\kappa^{(i)}$ and $\theta^{(i)}$, the parameters for the gamma distribution of ON-state duration

For a given set of parameter λ, the joint probability of the observation sequence $\boldsymbol{Y}$ and the set of the state sequences $\boldsymbol{q}$ is the product of the initial probability, the emission probability, and the transition probability.

(4.1)
$$P(\boldsymbol{Y}, \boldsymbol{q}|\lambda) = \psi_{in}(\boldsymbol{Y}, \boldsymbol{q}|\lambda) \cdot \psi_e(\boldsymbol{Y}, \boldsymbol{q}|\lambda) \cdot \psi_t(\boldsymbol{Y}, \boldsymbol{q}|\lambda)$$

The initial probability is

$$\psi_{in}(\boldsymbol{Y}, \boldsymbol{q}|\lambda) = \prod_{i=1}^{M} \pi_{q_1^{(i)}}^{(i)}$$

The emission probability is

$$\psi_e(\boldsymbol{Y}, \boldsymbol{q}|\lambda) = \prod_{t=1}^{T} b_{\boldsymbol{q_t}}(y_t)$$

The transition probability is

$$\begin{aligned}
&\psi_t(\boldsymbol{Y}, \boldsymbol{q}|\lambda) \\
&= \prod_{i=1}^{M} \prod_{t:q_t^{(i)}=0} \left(\left(\prod_{j=1}^{M} m_{q_{t+1}^{(i)} j q_t^{(j)}}^{(i)} \right) \left(\prod_{j=1}^{K} f_{q_{t+1}^{(i)} j c_t^{(j)}}^{(i)} \right) \right) \\
&\prod_{t:q_t^{(i)}=1} \left(\left(\prod_{j=1:i\neq j}^{M} m_{q_{t+1}^{(i)} j q_t^{(j)}}^{(i)} \right) \left(\prod_{j=1}^{K} f_{q_{t+1}^{(i)} j c_t^{(j)}}^{(i)} \right) \right) \\
&\prod_{t:q_t^{(i)}=1, q_{t-1}^{(i)}=0} P(d = \ell_t^{(i)} | \kappa^{(i)}, \theta^{(i)})
\end{aligned}$$

where $\ell_t^{(i)}$ is the length of the ON-state subsequence of the appliance i starting at time t. All these parameters can be estimated using the Expectation Maximization (EM) algorithm. EM iteratively re-estimates the parameter values using an "auxiliary function" until convergence to a local maximum occurs.

The auxiliary function to be maximized is

$$\phi(\lambda, \lambda') = \sum_{\boldsymbol{q}} P(\boldsymbol{Y}, \boldsymbol{q}|\lambda') \log P(\boldsymbol{Y}, \boldsymbol{q}|\lambda)$$

where λ' is the set of the parameters in the previous iteration.

In each iteration, the EM algorithm performs the E-step and M-step. In the E-step, the conditional distribution $P(\boldsymbol{Y}, \boldsymbol{q}|\lambda')$ is determined. Then, in the M-step, the parameters are updated to be $\arg\max_\lambda \phi(\lambda, \lambda')$.

We first look at the M-step, and then explain the E-step.

By Equation 4.1, the auxiliary function becomes:

$$\begin{aligned}\phi(\lambda,\lambda') = \quad & \textstyle\sum_{\boldsymbol{q}} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')\log\psi_{in}(\boldsymbol{Y},\boldsymbol{q}|\lambda) \\ & + \textstyle\sum_{\boldsymbol{q}} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')\log\psi_{e}(\boldsymbol{Y},\boldsymbol{q}|\lambda) \\ & + \textstyle\sum_{\boldsymbol{q}} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')\log\psi_{t}(\boldsymbol{Y},\boldsymbol{q}|\lambda)\end{aligned}$$

Since all the three terms do not have parameters in common, they can be maximized separately. For the first term,

$$\begin{aligned}&\sum_{\boldsymbol{q}} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')\log\psi_{in}(\boldsymbol{Y},\boldsymbol{q}|\lambda) \\ &= \sum_{\boldsymbol{q}} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')\sum_{i=1}^{M}\log\pi^{(i)}_{q_1^{(i)}} \\ &= \sum_{i=1}^{M}\sum_{\boldsymbol{q}}\log\pi^{(i)}_{q_1^{(i)}} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')\end{aligned}$$

Now, we can maximize the term of each appliance separately. For $i\in\{1,2,\ldots,M\}$,

$$\sum_{\boldsymbol{q}}\log\pi^{(i)}_{q_1^{(i)}} P(\boldsymbol{Y},\boldsymbol{q}|\lambda') = \sum_{j\in\{0,1\}}\log\pi^{(i)}_{j} P(\boldsymbol{Y},q_1^{(i)}=j|\lambda')$$

by using marginal expression for time $t=1$ in the right hand side. Adding the Lagrange multiplier, using the constraint that $\pi_0^{(i)}+\pi_1^{(i)}=1$, and setting the derivative equal to zero, we get:

$$\pi^{(i)}_j = \frac{P(\boldsymbol{Y},q_1^{(i)}=j|\lambda')}{P(\boldsymbol{Y}|\lambda')},\ \forall j$$

Similarly, for $i\in\{1,2,\ldots,M\}$, we get:

$$m^{(i)}_{jkl} = \frac{\sum_{t=1}^{T-1} P(\boldsymbol{Y},q_t^{(k)}=l,q_{t+1}^{(i)}=j|\lambda')}{\sum_{t=1}^{T-1} P(\boldsymbol{Y},q_{t+1}^{(i)}=j|\lambda')},\ \forall j,k,l$$

$$f^{(i)}_{jkl} = \frac{\sum_{t=1}^{T-1} P(\boldsymbol{Y},c_t^{(k)}=l,q_{t+1}^{(i)}=j|\lambda')}{\sum_{t=1}^{T-1} P(\boldsymbol{Y},q_{t+1}^{(i)}=j|\lambda')},\ \forall j,k,l$$

For the emission probability, as mentioned earlier, we use the gaussian distribution. However, we assume that the variance of the power consumption for appliances are the same. When we left the variances as free variables, we found overfitting problems. One possible explanation is that most of the errors or noise are caused by a sensor, not by appliances. This assumption also make it much simpler to estimate the emission parameters. We use σ to denote the fixed variation. The updating equation for $\boldsymbol{\mu}$ shown here is equivalent to the one found in [17].

$$\begin{aligned}&\phi_e(\lambda,\lambda') \equiv \sum_{\boldsymbol{q}} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')\log\psi_e(\boldsymbol{Y},\boldsymbol{q}|\lambda) \\ &= \sum_{\boldsymbol{q}} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')\sum_{t=1}^{T}\log b_{\boldsymbol{q_t}}(y_t) \\ &= \sum_{\boldsymbol{q}} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')\sum_{t=1}^{T}\left(\frac{\log 2\pi\sigma^2}{2} - \frac{(y_t-\sum_{i=1}^{M} q_t^{(i)}\mu^{(i)})^2}{2\sigma^2}\right)\end{aligned}$$

Then,

$$\begin{aligned}\frac{\partial\phi_e(\lambda,\lambda')}{\partial\mu^{(i)}} = \quad & \sum_{t=1}^{T} y_t q_t^{(i)} P(\boldsymbol{Y},\boldsymbol{q}|\lambda') \\ & - \sum_{t=1}^{T}\sum_{j=1}^{M}\mu^{(j)} q_t^{(i)} q_t^{(j)} P(\boldsymbol{Y},\boldsymbol{q}|\lambda') = 0\end{aligned} \tag{4.2}$$

Let $\langle q_t^{(i)}\rangle = \sum_q q_t^{(i)} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')$, and $\langle q_t^{(i)} q_t^{(j)}\rangle = \sum_q q_t^{(i)} q_t^{(j)} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')$. Then, Equation 4.2 becomes:

$$\frac{\partial\phi_e(\lambda,\lambda')}{\partial\mu^{(i)}} = \sum_{t=1}^{T} y_t\langle q_t^{(i)}\rangle - \sum_{t=1}^{T}\sum_{j=1}^{M}\mu^{(j)}\langle q_t^{(i)} q_t^{(j)}\rangle = 0$$

These can be solved by the normal equations

$$\boldsymbol{\mu} = \left[\sum_{t=1}^{T}\langle \boldsymbol{q_t}\boldsymbol{q_t}^T\rangle\langle \boldsymbol{q_t}\boldsymbol{q_t}^T\rangle\right]^{-1}\left[\sum_{t=1}^{T}\langle \boldsymbol{q_t}\boldsymbol{q_t}^T\rangle\langle \boldsymbol{q_t}\rangle y_t\right]$$

where $\boldsymbol{q_t} = [q_t^{(1)} q_t^{(2)}\ldots q_t^{(M)}]$, $\langle \boldsymbol{q_t}\boldsymbol{q_t}^T\rangle = \sum_{\boldsymbol{q}} \boldsymbol{q_t}\boldsymbol{q_t}^T P(\boldsymbol{Y},\boldsymbol{q}|\lambda')$ and $\langle \boldsymbol{q_t}\rangle = \sum_{\boldsymbol{q}} \boldsymbol{q_t} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')$.

Lastly, we have $\kappa^{(i)}$ and $\theta^{(i)}$ parameters to be optimized. Since there are no closed-form equations for estimating $\kappa^{(i)}$ and $\theta^{(i)}$, we need to estimate them numerically by the Newton-Raphson method [9].

Let

$$\begin{aligned}s^{(i)} = \quad & \log E[d^{(i)}|\boldsymbol{Y},\lambda'] - E[\log d^{(i)}|\boldsymbol{Y},\lambda'] \\ = \quad & \log\sum_{\boldsymbol{q}}\sum_{t:q_{t-1}^{(i)}=0,q_t^{(i)}=1}\ell_t^{(i)} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')/P(\boldsymbol{Y}|\lambda') \\ & - \sum_{\boldsymbol{q}}\sum_{t:q_{t-1}^{(i)}=0,q_t^{(i)}=1}\log\ell_t^{(i)} P(\boldsymbol{Y},\boldsymbol{q}|\lambda')/P(\boldsymbol{Y}|\lambda')\end{aligned}$$

where $d^{(i)}$ is the random variable for the ON-state occupancy duration and $\ell_t^{(i)}$ is the length of the ON-state subsequence of the appliance i starting at time t.

Then, we initialize $\kappa^{(i)} = s^{(i)}$, and iteratively update $\kappa^{(i)}$ by the following equation:

$$\kappa^{(i)} = \kappa^{(i)} - \frac{\log\kappa^{(i)} - \psi(\kappa^{(i)}) - s^{(i)}}{1/\log\kappa^{(i)} - \psi'(\kappa^{(i)})}$$

where ψ is the digamma function and ψ' is the trigamma function.

After iteratively estimating $\kappa^{(i)}$, we set

$$\theta^{(i)} = E[d^{(i)}|\boldsymbol{Y}, \lambda'](\kappa^{(i)})^{-1} = \left(\sum_{\boldsymbol{q}} \sum_{t:q_{t-1}^{(i)}=0, q_t^{(i)}=1} \ell_t^{(i)} \frac{P(\boldsymbol{Y}, \boldsymbol{q}|\lambda')}{P(\boldsymbol{Y}|\lambda')} \right) (\kappa^{(i)})^{-1}$$

These updating equations complete the M-step in our EM algorithm. In contrast to the M-step, the exact inference of the conditional distribution $P(\boldsymbol{Y}, \boldsymbol{q}|\lambda')$ in the E-step is computationally intractable as mentioned in [17]. There are alternative ways to approximate the inference, including Gibbs sampling and the mean field approximation [17]. Here, we use Gibbs sampling [16], one of the Monte Carlo methods, because of its simplicity. Since Gibbs sampling is a well-known tool and easy to adapt to any model, we omit its details.

4.3.2 Hidden State Estimation The goal of the energy load disaggregation is to discover the states of appliances. We are more interested in the sequences of the hidden variables in the CFHSMM than the parameters in the model. After learning the parameters, we need to use Maximum Likelihood Estimation (MLE) to estimate the sequences of the hidden variables.

In other words, we want to find $\boldsymbol{q}^*$ such that

$$\boldsymbol{q}^* = \arg\max_{\boldsymbol{q}} P(\boldsymbol{Y}, \boldsymbol{q}|\lambda)$$

The Viterbi algorithm can efficiently estimate the hidden states for HMMs. It uses dynamic programming to solve the optimization problem. However, dynamic programming for CFHSMMs is computationally intractable [17]. Thus, we use simulated annealing (SA) [26] to find $\boldsymbol{q}^*$. For the same reason as with Gibbs sampling, we omit the explanation of SA.

5 Experimental Results

5.1 Experiment Setup Our experimental setup monitors power consumption from seven residential homes. At each residence we have installed a mix of sensing nodes, each containing a Zigbee (`www.zigbee.org`) radio transceiver, collectively forming an in-home wireless sensor network using Digi (`www.digi.com`) components. Figure 10 shows our residential deployment topology. It includes a whole-home meter to determine overall electrical energy use (a smart meter proxy), several individual energy monitoring nodes (typically attached to larger appliances), and several clustered energy monitoring nodes to capture the aggregate consumption from grouped devices, such as an entertainment center. Power data is collected every 3 seconds.

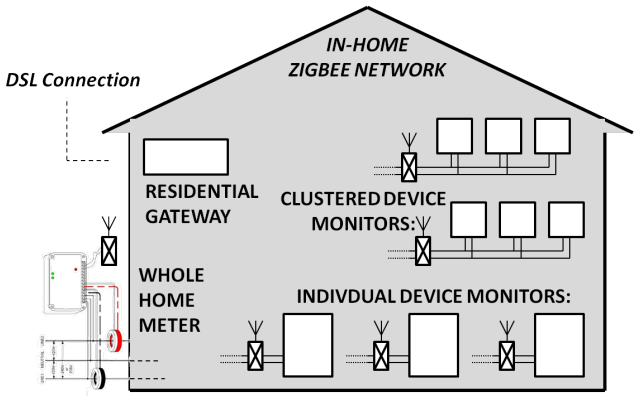

Figure 10: The in-home sensing topology.

A residential gateway connected to a DSL line enables remote management of the devices and collection of the power measurements. We combine data from individual device monitors to create our datasets. This approach provides us with the ground truth to evaluate the performance of our models.

5.2 Evaluation Metrics Accuracy is a commonly used evaluation metric. However, with power disaggregation the state distribution is very skewed because using an appliance is a relatively rare event. Therefore, accuracy is not an appropriate metric for evaluating power load disaggregation because a model that always says all the appliances are off will achieve high accuracy.

Instead, we adapt a metric from the information retrieval domain, *F-measure*. In the information retrieval domain, the common task is to classify relevance of documents for a given query. Because relevant documents are relatively rare, evaluation metrics in the information retrieval consider skewed classes.

F-measure is widely used in this type of evaluation. In binary classification tasks, there are four possible outcomes from a binary classifier: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). *F-measure* is the harmonic mean of *Precision* and *Recall*. *Precision* is defined as $\frac{TP}{TP+FP}$ and *Recall* is defined as $\frac{TP}{TP+FN}$. Thus,

$$F\text{-}measure = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

We use the following process to apply *F-measure* to our work. We convert our method to a binary classifier such that if the power consumption of an appliance is greater than 0, the output label is positive, and otherwise it is negative. However, our task is not only classifying the states of an appliance, but predicting how much power it consumes. Therefore, among true pos-

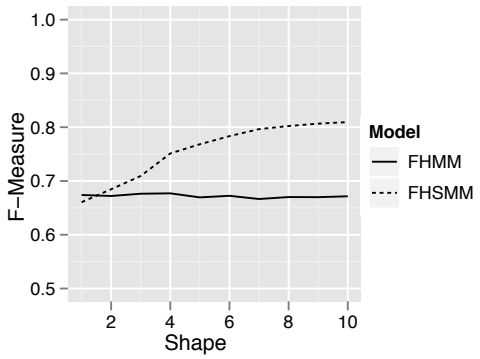

Figure 11: The effect of ON-duration shape.

Testdata	FHMM	CFHMM
fam_tv, fam_ps3, fam_stereo	0.717	0.985
fam_tv, liv_tv, liv_xbox	0.621	0.862
fam_tv, fam_ps3, liv_tv, liv_xbox	0.524	0.718
fam_tv, fam_stereo, liv_tv, liv_xbox	0.680	0.867
fam_tv, fam_ps3, liv_tv	0.562	0.744
fam_tv, fam_ps3, fam_stereo, liv_tv	0.621	0.803
fam_tv, fam_stereo, liv_xbox	0.724	0.881
All 5 appliances	0.594	0.751
liv_tv, liv_xbox	0.854	0.999
fam_tv, fam_ps3, fam_stereo, liv_xbox	0.590	0.731

Table 3: The top 10 most improved testdata.

itives, we consider predictions that differ significantly from ground truth as incorrect. More specifically, we split the true positives into two categories, accurate true positive (ATP), and inaccurate true positive (ITP). We distinguish the predictions as follows. Let x be the predicted value, and x_0 be the ground truth value.

- When $x = 0$ and $x_0 = 0$, the prediction is true negative (TN).
- When $x = 0$ and $x_0 > 0$, the prediction is false negative (FN).
- When $x > 0$ and $x_0 = 0$, the prediction is false positive (FP).
- When $x > 0$, $x_0 > 0$, and $\frac{|x-x_0|}{x_0} \leq \rho$, the prediction is an accurate true positive (ATP).
- When $x > 0$, $x_0 > 0$, and $\frac{|x-x_0|}{x_0} > \rho$, the prediction is an inaccurate true positive (ITP).

where ρ is a threshold.

We redefine *Precision* and *Recall* such that $Precision = \frac{ATP}{ATP+ITP+FP}$ and $Recall = \frac{ATP}{ATP+ITP+FN}$. *F-measure* remains the harmonic mean of the new *Precision* and *Recall*. We use the new *F-measure* as our metric with $\rho = 0.2$ in the evaluation. Most appliances in our evaluation have standard variations of around 20% of their means. For example, the power consumption of kit_ref has standard deviation of 15W, where its mean is 82W.

Since the output of the unsupervised models do not have labels on each appliance, we compute *F-measure* for all possible mappings, and take the maximum values as their performance.

5.3 ON-Duration Distribution In this section, we test the effectiveness of ON-duration shape as a feature. For this test only, we create two synthetic datasets. We generate two independent time-series data with the same power consumption, ON-duration mean, OFF-duration mean, OFF-duration shape, but different ON-duration shape.

Each synthetic data set has a power consumption of 100 W, mean ON-duration of 30 time units, mean OFF-duration of 60 time units, and OFF-duration shape parameter of 1. The first data set has ON-duration shape parameter of 1, while the second has various ON-duration shape parameters from 1 to 10. The shape of a gamma distribution changes from that of an exponential distribution to that of a Gaussian distribution as its shape parameter increases. Thus, as the value of the shape parameter gets larger, the difference between the two shapes of ON-durations increases. Figure 11 shows that FHSMM performs better as the shape parameter increases, but FHMM shows no change.

5.4 Dependencies Next, we evaluate the gains resulting from modeling the appliance dependencies and additional features. We chose two groups of appliances that have strong correlations to other appliances – {fam_tv, fam_ps3, fam_stereo}, and {liv_tv, liv_xbox}. We scaled the appliances to have the same power consumption, and generated all the possible combinations of these five appliances for the testdata. We scaled the power so that power level becomes ineffective as a feature for disaggregation. There are 26 testdata with at least two appliances. For each testdata, we evaluate the *F-measure* of FHMM and CFHMM. The averages are 0.734 for FHMM and 0.838 for CFHMM. Table 3 lists the top 10 test cases where maximum improvement was seen through use of CFHMM.

These evaluations show the effectiveness of modeling the dependencies between appliances and the additional features. For {liv_tv, liv_xbox} testdata, CFHMM disaggregated the load perfectly because the model inferred the appliance dependency of liv_xbox to liv_tv (i.e., an Xbox needs to be used with a TV).

Home ID	Num. of Appliances	FHMM	CFHSMM
Home 1	4	0.983	0.998
Home 2	6	0.899	0.930
Home 3	6	0.859	0.881
Home 4	7	0.625	0.693
Home 5	8	0.713	0.781
Home 6	8	0.641	0.722
Home 7	10	0.796	0.874

Table 4: The evaluations on several homes.

5.5 Overall Performance We tested the performance of our models on all the seven homes from where we collected data. Table 4 shows the results. The results in Sections 3 and 5.4 use Home 6's data.

Even though we are monitoring more than 20 appliances in each house, we have much fewer appliances in the data sets because the other appliances were not *active*, that is, either they were never turned on, or were always on. The always-on loads form part of the base load (also called vampire load). Most of the power load disaggregation algorithms (including ours) cannot disaggregate base load since disaggregation is based on the characteristics of the appliance power state changes.

Figure 12 shows the *F-measure* of the four models versus the number of appliances. There are several important observations. First, disaggregation using low frequency data becomes more challenging as the number of appliances increase. Further, the plot shows the effectiveness of additional features. CFHSMM performs better in all cases although the difference is more pronounced for larger number of appliances (7 and 8). The difference between the performance of CFHSMM and CFHMM is minimal indicating that for this data set most of the gain in performance of CFHSMM comes from additional features considered rather than use of the gamma distribution for ON-durations. Thus, for dealing with more appliances, it is desireable to integrate other additional features into our models.

6 Conclusions

In this paper, we investigated how effective unsupervised disaggregation of low frequency power measurements is. This is an important topic, as an effective method of this type could facilitate residential electricity conservation efforts. We considered a existing model FHMM and three new models (FHSMM, CFHMM and CFHSMM). Using low frequency measurements from real homes, we showed that CFHSMM outperformed the other unsupervised methods, and was capable of accurately disaggregating power data into per-appliance usage information.

We plan to extend this work in multiple ways. First, our results revealed that the tested methods work well

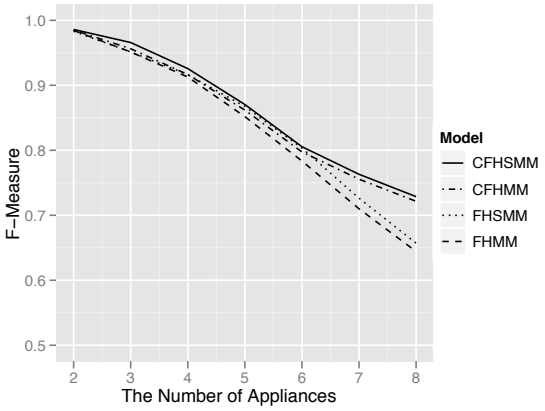

Figure 12: Comparison of model performance.

for appliances with simple or modestly complex power signatures, but less well for more complex signatures. Handling this subset of signatures is an important topic. Second, we need to develop more extra features like vibrations from sensors to enhance our method to deal with more number of appliances. Third, we need a method to estimate the number of appliances in the whole-home power measurements. Fourth, we intend to monitor residential gas and water usage, to facilitate conservation of those resources too.

References

[1] A. Abrahamse, L. Steg, C. Vlek, and T. Rothengatter. A review of intervention studies aimed at household energy conservation. *Environmental Psychology*, 25:273–291, 2005.

[2] U. E. I. Administration. Electric power annual 2008. `http://www.eia.doe.gov/cneaf/electricity/epa/epaxlfilees1.pdf`, 2010.

[3] U. E. I. Administration. Electricity faq. `http://www.eia.doe.gov/ask/electricity_faqs.asp`, 2010.

[4] S. Attari, M. DeKay, C. Davidson, and W. de Bruin. Public perceptions of energy consumption and savings. *Proceedings of the National Academy of Sciences*, 2010.

[5] L. Becker. Joint effect of feedback and goal setting on performance: a field study of residential energy conservation. *J. of Applied Psychology*, 63(4):428–433, 1977.

[6] M. Brand, N. Oliver, and A. Pentland. Coupled hidden markov models for complex action recognition. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, pages 994–999, 1997.

[7] G. Brandon and A. Lewis. Reducing household energy consumption: a qualitative and quantitative field study. *Environmental Psychology*, 19:75–85, 1999.

[8] M. Chetty, D. Tran, and R. Grinter. Getting to green: understanding resource consumption in the home. In *UbiComp*, Seoul, Korea, 2008.

[9] S. C. Choi and R. Wette. Maximum likelihood estimation of the parameters of the gamma distribution and their bias. *Technometrics*, 11(4):683–690, 1969.

[10] S. Darby. The effectiveness of feedback on energy consumption. Technical report, 2006.

[11] L. Farinaccio and R. Zmeureanu. Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end uses. *Energy and Buildings*, 30:245–259, 1999.
[12] C. Fischer. Feedback on household electricity consumption: a tool for saving energy? *Energy Efficiency*, 1:79–104, 2008.
[13] R. Fitch. New meter helps buyer save energy. *House and Home*, 51(5):58, 1977.
[14] G. Gardner and P. Stern. The short list: the most effective actions U.S. households can take to curb climate change. *Environment Magazine*, 2008.
[15] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.
[16] Z. Ghahramani and M. I. Jordan. Factorial hidden markov models. *Machine Learning*, 29:245–273, 1997.
[17] G. Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(2):1870–1891, 1992.
[18] S. Hayes and J. Cone. Reduction of residential consumption of electricity through simple monthly feedback. *Applied Behavior Analysis*, 14(1):81–88, 1981.
[19] J. V. Houwelingen and F. V. Raaij. The effect of goal-setting and daily electronic feedback on in-home energy use. *Consumer Research*, 16:98–105, 1989.
[20] E. Howell. How switches produce electrical noise. *IEEE transactions on electromagnetic compatibility*, EMC-21(3):162–170, 1979.
[21] M. Ito, R. Uda, S. Ichimura, K. Tago, T. Hoshi, and Y. Matsushita. A method of appliance detection based on features of power waveform. In *IEEE Symposium on Applications and the Internet*, Tokyo, Japan, 2004.
[22] T. Kao, H. Cho, D. Lee, T. Toyomura, and T. Yamazaki. Appliance recognition from electric current signals for information-energy integrated network in home environments. In *International Conference on Smart Homes and Health Telematics*, Tours, France, 2009.
[23] W. Kempton and L. Layne. The consumer's energy analysis environment. *Energy Policy*, 22(10):857–866, 1994.
[24] Y. Kim, T. Schmid, Z. Charbiwala, and M. Srivastava. Viridiscope: design and implementation of a fine grained power monitoring system for homes. In *UbiComp*, Orlando, FL, 2009.
[25] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):pp. 671–680, 1983.
[26] R. Kohlenberg, T. Phillips, and W. Proctor. A behavioral analysis of peaking in residential electrical-energy consumers. *Applied Behavior Analysis*, 9(1):13–18, 1976.
[27] A. Krogh, M. Brown, S. Mian, K. Sjölander, and D. Haussler. Hidden Markov models in computational biology. *Molecular Biology*, 235:1501–1531, 1994.
[28] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong. Power signature analysis. *IEEE Power and Energy*, 1(2):56–63, 2003.
[29] S. Levinson. Continuously variable duration hidden markov models for automatic speech recognition. *Computer Speech and Language*, 1(1):29 – 45, 1986.
[30] G. Lin, S. Lee, J. Hsu, and W. Jih. Applying power meters for appliance recognition on the electric panel. In *IEEE Industrial Electronics and Applications*, Taichung, Taiwan, 2010.
[31] H. Matthews, L. Soibelman, M. Berges, and E. Goldman. Automatically disaggregating the total electrical load in residential buildings: a profile of the required solution. In *Intelligent Computing in Engineering*, Plymouth, UK, 2008.
[32] L. Norford and S. Leeb. Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algoritms. *Energy and Buildings*, 24:51–64, 1996.
[33] S. Patel, S. Gupta, and M. Reynolds. The design and evaluation of an end-user-deployable, whole house, contactless power consumption sensor. In *Human factors in computing systems*, Atlanta, GA, 2010.
[34] S. Patel, T. Robertson, J. Kientz, M. Reynolds, and G. Abowd. At the flick of a switch: detecting and classifying unique electrical events on the residential power line. In *UbiComp*, Innsbruck, Austria, 2007.
[35] A. Prudenzi. A neuron nets based procedure for identifying domestic appliances pattern-of-use from energy recordings at meter panel. *IEEE Power Engineering Society Winter Meeting*, 2:491–496, 2002.
[36] L. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.
[37] Y. Riche, J. Dodge, and R. Metoyer. Studying always-on electricity feedback in the home. In *Human factors in computing systems*, Atlanta, GA, 2010.
[38] B. Ritchie, G. McDougall, and J. Claxton. Complexities of household energy consumption and conservation. *Consumer Research*, 8(3):233–242, 1981.
[39] M. Russell and R. Moore. Explicit modelling of state occupancy in hidden markov models for automatic speech recognition. volume 10, pages 5 – 8, 1985.
[40] T. Saitoh, Y. Aota, T. Osaki, R. Konishi, and K. Sugahara. Current sensor based non-intrusive appliance recognition for intelligent outlet. In *Int. Technical Conference on Circuits/Systems, Computers and Communications*, Shimonoseki City, Japan, 2008.
[41] C. Seligman and J. Darley. Feedback as a means of decreasing residential energy consumption. *Applied Psychology*, 62(4):363–368, 1977.
[42] H. Serra, J. Correia, A. Gano, A. de Campos, and I. Teixeira. Domestic power consumption measurement and automatic home appliance detection. In *IEEE Intelligent Signal Processing workshop*, Algarve, Portugal, 2005.
[43] R. Sexton, N.Johnson, and A. Konakayama. Consumer response to continuous-display electricity-use monitors in a time-of-use pricing experiment. *Consumer Research*, 14(1):55–62, 1987.
[44] P. Stern. What psychology knows about energy conservation. *American Physchologist*, 47(10):1224–1232, 1992.
[45] K. Suzuki, S. Inagaki, T. Suzuki, H. Nakamura, and K. Ito. Nonintrusive appliance load monitoring based on integer programming. In *Int. Conference on Instrumentation, Control and Information Technology*, Tokyo, Japan, 2008.
[46] T. Ueno, F. Sano, O. Saeki, and K. Tsuji. Effectiveness of an energy-consumption information system on energy savings in residential houses based on monitored data. *Applied Energy*, 83:166–183, 2006.
[47] R. Vines. Noise on residential power distribution circuits. *IEEE transactions on electromagnetic compatibility*, EMC-26(4):161–168, 1984.
[48] R. Watson, M. Zinyousera, and R. M. (editors). Technologies, policies and measures for mitigating climate change. Technical report, 1996.
[49] N. Yadwadkar, C. Bhattacharyya, and K. Gopinath. Discovery of application workloads from network file traces. In *USENIX FAST*, San Jose, CA, 2010.