Disaggregating Appliance-Level Energy Consumption: A Probabilistic Framework

Sabina Tomkins
University of California Santa Cruz
satomkin@ucsc.edu

Lise Getoor
University of California Santa Cruz
getoor@soe.ucsc.edu

ABSTRACT

In this work we propose a probabilistic disaggregation framework which can determine the energy consumption of individual electrical appliances from aggregate power readings. Our proposed framework uses probabilistic soft logic (PSL), to define a *hinge-loss* Markov random field (HL-MRF). Our method is novel in that it can integrate a diverse range of features, is highly scalable to any number of appliances, and makes less assumptions than existing methods. As the residential sector is responsible for over a third of all electricity demand, and delivering appliance level energy consumption information to consumers has been demonstrated to reduce electricity consumption, our framework has the potential to make a significant impact on energy savings.

Keywords

Smart Meters, Probabilistic Graphical Models, Sustainability, Energy Signal Disaggregation, Non-intrusive load monitoring, Hinge-loss Markov Random Fields (HL- MRFs), Probabilistic Soft Logic (PSL)

1. INTRODUCTION

Reducing household energy consumption in the United States is critical not only for mitigating the negative health and environmental effects of air pollution, but for increasing energy security, resiliency, and stability. Households consume over one third of all electricity in the United States [1], and opportunities for decreasing this share abound [6]. However, these opportunities are impeded by the lack of information available to residential consumers.

Consumers have been found to be uninformed as to which appliances consume the most energy [7, 2], and which actions have the greatest savings potential. Furthermore, there is a growing body of evidence [4, 18] that detailed feedback about energy use can reduce consumption. Currently, this reduction is hampered by the fact that residents receive only aggregate energy information. Consider the case of receiving a shopping bill with a single figure, and being asked to spend less on the next shopping trip, from this it would be difficult to discern how to adjust purchasing habits. Itemized energy bills would afford more intimate knowledge of which appliances are responsible for how much energy, thus empowering consumers to take action towards reducing energy usage.

Frequent [19, 4] and specific information can aid consumers with making informed decisions about electricity usage. For example, greater savings are achieved when homeowners are told how much energy is used by each appliance [5]. Traditional electricity meters are limited by their inability to transmit power readings to remote data-storage systems, and therefore are inadequate for capturing data at the level of detail that consumers need.

Advanced Metering Infrastructure (AMI), such as smart meters, have the ability to transmit power readings wirelessly. Smart meters thus offer a unique opportunity to gather rich amounts of real-time data from which energy consumption patterns can be learned and ultimately utilized to offer actionable insight to consumers. While smart meters have been installed in homes across the United States, their potential in reducing consumption is far from realized. A critical challenge is with extracting appliance level information from aggregate power readings, a process referred to interchangeably as energy disaggregation, and non-intrusive load monitoring (NILM).

Energy disaggregation is the process of determining the energy consumption of individual appliances, given only an aggregated energy reading. With a successful disaggregation algorithm one would be able to give consumers an itemized energy bill, displaying how much energy is consumed by each appliance, rather than the aggregate monthly reading that they receive now. While there are existing approaches to this problem, no sound and complete solution is available.

Here we propose a probabilistic disaggregation framework. Our framework provides a flexible method for determining the proportion of total energy consumed by individual appliances. This framework can integrate data-driven learning with diverse forms of knowledge, from the expertise of domain specialists to information supplied by users. We analyze the benefit of two groups of features: temporal and appliance features. Temporal features can be mined directly from data, or given by a user, while appliance features require some knowledge about each of the appliances being disaggregated, such as their average power consumption. Preliminary results show that both types of features can be successfully applied to this task, however the characteristics of a particular home can influence their respective performances.

UrbComp'16, August 14, 2016, San Francisco, USA.

ACM ISBN 978-1-4503-2138-9.

DOI: 10.1145/1235

2. RELATED WORK

Hart's seminal nonintrusive load monitoring paper [8] introduced the problem of energy disaggregation. In NILM, the goal is to determine the consumption of individual appliances, *without* installing intrusive equipment, such as meters for each device. Broadly, there are two classes of solutions, those that require additional hardware and those that do not. As we are striving for a software solution, here we review only those approaches which do not require any additional equipment installation and are thus truly non-intrusive.

Many of the proposed approaches make use of probabilistic graphical models. Factorial hidden Markov models (FHMM)s and variants thereof, have been a popular choice for disaggregation algorithms. In this setting each appliance is represented with a single HMM, where the discrete hidden state variables correspond to the state of the appliance, and the observed continuous random variables correspond to the power readings. FHMMs allow multiple HMMs to be joined through a single observed variable in such a way that approximate inference is tractable. We highlight several variants of this approach below.

Kolter and Jaakola, employed an additive FHMM for disaggregation, where each observed output is an additive function of the states of the hidden variables [11]. Their method asserts the strong assumption that only one appliance changes state at a time, which allows for an efficient convex formulation. Parson [15] et al., employ this additive FHMM, and take positive steps away from requiring detailed information about the parameters for the on-duration distribution. They start with general prior models of appliances, and tune these models to fit each home.

Perhaps the most exhaustive exploration and adaption of FHMMs was completed by Kim et al., who used a conditional factorial hidden semi-Markov model (CFHSMM) in disaggregation. Their CFHSMM is able to remove several assumptions required by other models, such as: the assumption that on duration distributions are geometric. Additionally, Kim et al. [10] incorporate non-traditional features such as time of day. Johnson and Willsky[9], introduce a hierarchical Dirichlet process hidden semi-Markov model, which also relaxes the restriction that the state durations be encoded with geometric distributions. Their model differs from others in that it can have multi-state hidden variables while other models assume that these variables are binary. Additionally, Johnson et al. [9] found general parameter values for the appliance on-state distributions which applied to a set of homes, rather than learning these values for each home.

Finally, Makonin et al. [14], uses a super-state HMM, where the hidden variable can take on any of 2^{C} states, where any state is a combination of the on/off states of each appliance. A unique advantage of their approach is that it can perform exact inference, unlike other approaches for which inference is approximate. Unlike some of these approaches our model does not assume that only one appliance changes state at a time. Additionally the model is agnostic to the distributions of the on-durations. We can incorporate multiple states, and additional features easily.

The largest shortcoming with previous approaches, and which we address with our probabilistic framework, is the ability to be flexible to real-world demands and constraints. For example, it is unclear the extent to which available training data can be applied to new homes and more broadly, what kind of data guarantees are necessary to ensure quality disaggregation results. Our framework is flexible in that it can assimilate and learn from both appliance-level power readings and data attributes, and human supplied knowledge.

3. DISAGGREGATING APPLIANCES

A disaggregation algorithm takes total power consumption over a given time period as input and returns the consumption of individual appliances over that time period. Suppose we have a set of appliances, $A_1 \dots A_M$. Let $A_{i,t}$ be the consumption of the i-th appliance at time t. We are given a sequence of aggregate power readings

$$S = \{S_1, ..., S_N\},$$

where each reading is the sum of the power consumed by all appliances at that time, $S_t = \sum_{i=1}^{M} A_{i,t}$. For any appliance, we would like to determine the proportion of total power consumed by the i-th appliance,

$$P_i = \frac{\sum_{t=1}^{N} A_{i,t}}{\sum_{t=1}^{N} S_t}.$$

Furthermore, certain constraints must be met for a disaggregation algorithm to be useful to consumers [20]. These constraints are designed to evaluate an algorithm not just on its performance, but on its applicability to home energy systems. A disaggregation algorithm with perfect accuracy is not going to be useful for consumers if it requires that submeters be installed in homes, or if it can only return results once a month, or if it is only able to disaggregate refrigerators. Therefore, deployability is a key metric in evaluating disaggregation algorithms.

Additionally, all current disaggregation algorithms allow for the assumption of prior knowledge of the power draw of the appliances. For each appliance which is being disaggregated, one must know how many possible states the it has, and approximately how much power the appliance consumes in each of its states. For example, a fan may have three states: low, medium, and high, while a television can either be on or off. However, the general statistical model which describes an appliance, such as the distribution of its power draw in an on state, is not necessarily given as input to a disaggregation algorithm.

4. DISAGGREGATION FRAMEWORK

We propose a flexible framework which can disaggregate individual appliances from aggregate power readings. The framework is designed to adapt to multiple categories of information. Here, we show how to incorporate knowledge on appliance behavior, and temporal patterns extracted from data.

Our framework integrates diverse sources of information into a joint probability distribution over the energy usage for all M appliances. This distribution is a graphical model, whose potential functions are defined using weighted logical rules. The weights associated with these rules can be learned from data. To predict the probability that the i-th appliance is on at time t, given the evidence, we use *hinge-loss* Markov random fields (HL-MRFs)[3].

4.1 Hinge-loss Markov Random Fields

Hinge-loss Markov random fields are a general class of conditional, continuous probabilistic models, parametrized with a set of weighted *hinge-loss* functions. Hinge-loss functions can model a rich diversity of relationships, and critically, admit highly scalable inference. For example logical rules can be expressed with *hinge-loss* functions through Łukasiewicz relaxations, a property which we will exploit in the energy setting. Formally, a HL-MRF describes the following conditional probability density function over continuous random variables, $\mathbf{X}, \mathbf{Y} \in [0, 1]$:

$$P(\mathbf{Y}|\mathbf{X}) \propto exp\left(-\sum_{j=1}^{m} w_j \phi_j(\mathbf{Y}, \mathbf{X})\right)$$

Where ϕ_j is a *hinge-loss* potential,

$$\phi_j = max\{l_j(\mathbf{Y}, \mathbf{X}), 0\}^p$$

$p \in \{1, 2\}$, l_j is a linear function of $\mathbf{X}$ and $\mathbf{Y}$ and w_j is the positive weight associated with ϕ_j.

4.1.1 Probabilistic Soft Logic

To generate a HL-MRF for our probabilistic disaggregation model we use probabilistic soft logic (PSL). PSL is a templating language for HL-MRFs which has been successfully deployed in a diverse range of settings, from recommender systems [13] to stance prediction in online forums [17]. PSL allows us to write weighted rules which can express the various types of information one might have in the energy analytics setting. When a set of weighted rules and observed data is input to a PSL program, a specific HL-MRF is defined.

A rule in PSL consists of terms, predicates, and weights. A term is either a variable or a constant, and a predicate is a relation between terms. To define a predicate one must specify its name, and the number of arguments it takes. Weights are positive values associated with each rule. An example rule is,

$$\lambda : \text{Friends}(A, B) \wedge \text{Likes}(A) \rightarrow \text{Likes}(B)$$

where λ is a weight, Friends, and Likes are predicates, and A, B are variables. By substituting constants, a and b, for the variables A and B respectively, one obtains three ground atoms: Friends(a, b), Likes(a), and Likes(b), such that each ground atom takes a value in $[0, 1]$. We would like to infer the value of our target unobserved variable, the probability that appliance A is on at time T.

With Łukasiewicz logic, the truth values of logical statements can be relaxed from Boolean values, to the interval $[0, 1]$. This function can then be input to the *hinge-loss* potential functions of a HL-MRF. Suppose that we let the atoms Friends(a, b), Likes(a), and Likes(b) correspond to three random variables: $f, l_1, \text{and } l_2$ respectively. We will use the Łukasiewicz rule, which given two continuous truth values $q, r \in [0, 1]$ defines a conjunction of q and r as $q \wedge r = max\{q + r - 1, 0\}$. Finally, using the formula $q \rightarrow r = q \wedge \neg r$, we arrive at the weighted *hinge-loss* potential corresponding to the above rule,

$$\lambda \cdot max\{f + l_1 - l_2 - 1, 0\}.$$

The HL-MRF returned by PSL can be used to infer the maximum a posteriori (MAP) assignments the unobserved variables, such as the joint probability that a collection of appliances is on. Finally, PSL enables weights to be learned from data. This is an especially desirable feature, when prior knowledge can be used to set only approximate relative weights, but further information is required to discern the true importance of proposed rules.

4.2 Probabilistic Disaggregation Model

Our goal is to the find the probability, that the ith appliance, A_i, is on at time t. We introduce the predicate IsOn(T, A_i), whose continuous truth values we infer for each appliance i from $[1, M]$, and each time t from $[1, N]$. We utilize the information that is both available in the power readings, and which users may be conveniently able to supply. To do so we create temporal rules, appliance rules, and inter-appliance rules. Temporal rules describe the probability that an appliance is on given the time of day and day of the week. Appliance rules utilize the information which is available purely from the supplied appliance signatures. Inter-appliance rules describe the relationships between appliances. An exciting feature of our approach is that additional rules can be added as they become available, for example temperature can be used to predict the use of heating and cooling appliances, and rules to describe temperature can easily be integrated into the model.

4.2.1 Temporal Rules

Whether or not some appliances are on depends on the time of day, and day of the week. For example, it is more likely that a cooking appliance, such as a microwave will be used in the evening, than in the middle of the night. Thus we introduce two predicates: Hour(T, H) and DayOfWeek(T, D), where $H \in \{0, 23\}$ and $D \in \{Sunday...Monday\}$. We then use these predicates in the following rule,

$$w_t : \text{DayOfWeek}(T, D) \wedge \text{Hour}(T, H) \rightarrow \text{IsOn}(T, A_i)$$

This template is used to generate a set of rules for each appliance, one for each day of the week and hour pair. Thus we may arrive at different rules as:

$$10 : \text{DayOfWeek}(T, Sunday) \wedge \text{Hour}(T, 09:00) \rightarrow \text{IsOn}(T, Microwave)$$

$$.01 : \text{DayOfWeek}(T, Sunday) \wedge \text{Hour}(T, 04:00) \rightarrow \text{IsOn}(T, Microwave)$$

This full set of rules allows us to model a rich variety of temporal dependencies. Additionally we create a set of rules to capture the persistence of an appliance being on. To do so we use a predicate Precedes(T_i, T_j), which is true if index value i, directly precedes index value j, or $i = j - 1$. The following rules allow us to express the probability that an appliance will be on at time t, if it is on at time $t - 1$.

$$w_t : \text{IsOn}(S, X) \wedge \text{Precedes}(S, T) \rightarrow \text{IsOn}(T, A_i)$$

$$w_t : \text{IsOn}(S, X) \wedge \text{Precedes}(S, T) \rightarrow \neg\text{IsOn}(T, A_i)$$

4.2.2 Appliance Rules

These rules utilize the information which can be extracted from prior knowledge of the appliance signatures. Here we introduce the additional predicates: TotalEnergy(T, B), Duration(T, D), and PotentiallyOn(T, X), and SwitchedOn(T, X).we have These predicates are explained in detail below, beginning with TotalEnergy(T, B). These rules are designed to predict if an appliance is contributing to an aggregate reading over a time period, given the persistence and magnitude of the total energy over that same period.

We partition the energy readings into buckets, such that these buckets cover the entire range of possible energy reading values. TotalEnergy(T, B) takes a time, and a bucket id b, where b corresponds to an interval of possible energy values. We can now view the input data as a sequence of bucket ids. To find the a suitable number of intervals we clustered each home into k clusters from 1 to 16(the total number of possible combinations of appliances), using k-means clustering implemented with the python package scikit learn[16]. We chose the best k by first assigning each point to a cluster, and then assigning that point a power value by drawing a value from a normal distribution fit to that cluster. The k which generates the least mean squared error is then selected. The best k fluctuated between 6 and 8 for each home, so we chose 7 as the optimal partition size. To demarcate the intervals we used the appliance signatures. The lowest partition corresponds to no target appliances being on. The highest partition corresponds to a situation where all appliances could be on.

Ideally, each bucket would be perfectly mapped to some collection of appliances being on, however that may not be the case, as appliances have similar distributions and no perfect partition exists. For example, we may have a bucket b which corresponds to the interval $1300 \leq energy \leq 2000$, which could be mapped to either, a microwave and a dishwasher, or a microwave and a lighting appliance. To further differentiate appliances we would like to use information about the amount of time an appliance is used for. If we know that the length of the sequence of consecutive b's spans an interval of 3 hours, this may lead us to believe that the combination of appliances is the microwave and the lights, rather than the microwave and the dish washer, as dishwashers normally run for one hour or less. To capture this duration information we introduce the predicate, Duration(T, D), that takes a time t, and a duration id d, which corresponds to the duration of a given total energy bucket, b. We choose four types of intervals, in order to separate different kinds of appliances. The shortest interval has length less than 10 minutes, for appliances such as the microwave, and the longest interval has a length of several days. PotentiallyOn(T, X), predicts an appliance A_i being on at time t given the current total energy, S_t using the following rule:

$$\text{PotentiallyOn}(t, A_i) = \begin{cases} 0, & \text{if } \mu_{A_i} + \sigma_{A_i} > S_t \\ 1, & \text{otherwise} \end{cases}$$

where μ_{A_i} is the mean of the distribution of the on-state power draw of the i-th appliance, and σ_{A_i}, is the standard deviation of that distribution. This allows us to disqualify appliances if their consumption would exceed the total consumption at time t.

SwitchedOn(T, X) is designed to catch the event of a certain appliance turning on. Whenever the difference in two consecutive total energy readings is within one standard deviation of the average power of an appliance, SwitchedOn(T, X) of that appliance is set to true. Denoting $\Delta S_t = S_t - S_{t-1}$,

$$\text{SwitchedOn}(t, A_i) = \begin{cases} 1, & \text{if } |\Delta S_t - \mu_{A_i}| \leq \sigma_{A_i} \\ 0, & \text{otherwise.} \end{cases}$$

5. EXPERIMENTS

We evaluate our framework on a publicly available Reference Energy Disaggregation Dataset (REDD) [12]. The goal is to determine P_i, the proportion of energy consumed by the i-th appliance, for each i. Additionally we evaluate the ability of the model to correctly classify the on and off states of each appliance.

5.1 Data

To preprocess the dataset we used an open source script written by Makonin et al. for this dataset [14]. We select four homes, homes 1, 2, 3, and 6. Each home has roughly two weeks of data. For each home we split the data into three parts: training, validation and test. Parameter tuning is done using only the validation set, withholding the test set until the final evaluation.

For disaggregation we choose four of the five appliances used by both Makonin [14], and Johnson [9]: microwave, dish washer, fridge, and lights. We refrain from using the furnace as it is only included in one home, and for a short length of time. The power consumption of these appliances is shown in figure 1.

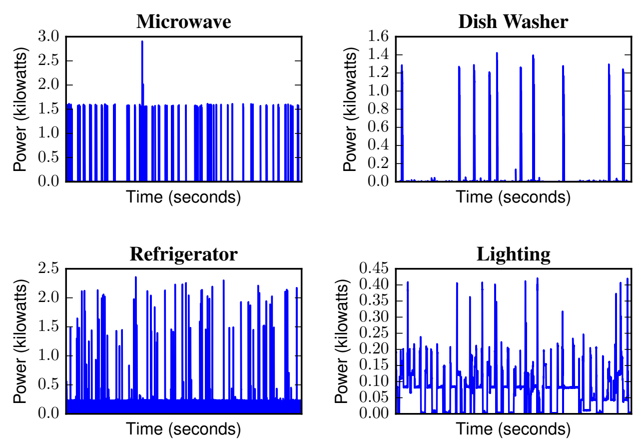

Figure 1: Each appliance has a unique signature.

We modeled each appliance with only two states: on or off, though this choice is not imposed by the framework, which can accept appliances with any number of states. For the off state, we assumed the appliance consumed no power, which is not necessarily true, and could easily be changed in future versions.

The on state of the appliance is assumed to have a Gaussian distribution, with parameters μ and σ, representing the mean and standard deviation. While all appliances of a given type, such as a refrigerator, use similar amounts of energy, the values of μ and σ can vary considerably across homes, as can be seen in figure 2. We used two approaches to find values for μ and σ for each home.

In the first approach, we found the mean and standard deviation of the power readings for each appliance, with the

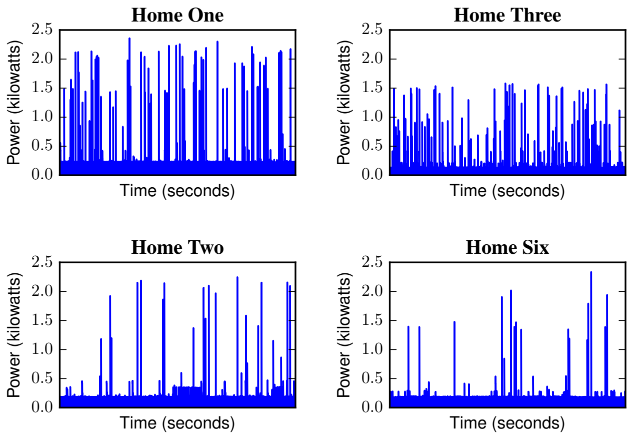

Figure 2: The duration and power used by the refrigerator varies by home.

single modification that we only considered readings greater than a small value of 50 watts, which is approximately the average minimum aggregate value across homes. In the second approach, the power readings of each appliance were grouped into two clusters (one for each state), using the python package SciKit Learn's [16] implementation of k-means clustering. Then, we found the mean and standard deviation of each cluster, and set the mean and standard deviation of the on-state distribution to equal those of the cluster with the higher mean. This latter approach was more successful when tested on the validation data, and was thus chosen for the remaining experiments. We include these values in the table 1.

Home	Microwave	Dish Washer	Lighting	Refrigerator
House 1	$\mathcal{N}(1520, 61^2)$	$\mathcal{N}(1085, 94^2)$	$\mathcal{N}(160, 79^2)$	$\mathcal{N}(201, 57^2)$
House 2	$\mathcal{N}(1839, 78^2)$	$\mathcal{N}(1199, 21^2)$	$\mathcal{N}(149, 38^2)$	$\mathcal{N}(171, 48^2)$
House 3	$\mathcal{N}(1716, 53^2)$	$\mathcal{N}(734, 10^2)$	$\mathcal{N}(317, 147^2)$	$\mathcal{N}(126, 50^2)$
House 6	NA	NA	$\mathcal{N}(125, 44^2)$	$\mathcal{N}(149, 31^2)$
Average	$\mathcal{N}(1692, 64^2)$	$\mathcal{N}(1006, 62^2)$	$\mathcal{N}(188, 77^2)$	$\mathcal{N}(162, 46^2)$

Table 1: Though similar, individual household parameter values can differ.

5.2 Evaluation Metrics

Evaluating the effectiveness of the disaggregation algorithm involves inspecting both the quality of the prediction of the on/off states, and of the estimation of the total energy consumed. To evaluate the quality of the prediction task we look at the F-measure for each appliance and each home, finding surprising results. Although the estimation accuracy is a popular metric we decided not to use it here, as it requires the assumption that the total predicted energy always equal the total true energy, which we do not enforce.

There are two testing protocols for disaggregation, the total energy being disaggregated can be noisy or de-noised. In the de-noised setting, the total energy is set to be the sum of the appliances being disaggregated, in the noisy setting the true total energy is unadjusted. We use the noisy setting, as it is the more realistic of the two.

5.3 Results

We evaluate our framework in two ways. First, in figure 3 we show the estimated relative consumption of the four appliances being disaggregated, and next we evaluate the ability of the model to predict the on/off states of the appliances. Additionally, we look at the effect of integrating temporal features into the appliance feature model.

Here we examine the percentage of total energy allocated to each appliance by the combined, appliance and temporal feature model. These percentages are over the test set, which is 25% of all data for each home, ranging from one and a half to three and a half days. We see that the lighting is error is greatest for the lighting appliance, where the proportion of true energy is 44.4%, but only 41.4% is assigned.

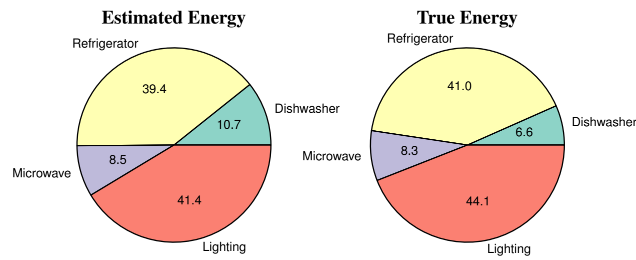

Figure 3: Our framework can determine the relative contributions of each appliance.

It is also useful to see the results which might be shared with a user. Instead of a plot of total power over time, users would receive a similar plot but for each appliance. For example, here are the plots which might be received for the dish washer and microwave.

5.3.1 *Disaggregating with Appliance Features*

We use prior information about the distribution of the power consumed by each appliance to predict if an appliance is on given total energy. This prior information is essential in disaggregating heavily consumptive appliances, which can be identified by the current amount of energy being consumed. However disaggregating smaller loads can be more difficult as these appliances can be confused for each other, and a combination of these lightly consumptive appliances can have the same consumption as a single larger load. We see in table 2 that some homes are easier to disaggregate than others.

	Home One	Home Two	Home Three	Home Six	Average
Accuracy	81.3%	85.6%	78.6%	72.1%	79.3%
F-measure	50.3%	57.0%	45.8%	80.7%	58.4%
Precision	36.1%	50.2%	32.6%	72.1%	47.7%
Recall	87.8%	84.9%	100.0%	100.0%	93.1%

Table 2: The appliance feature model does well on recall.

5.3.2 *Effect of Temporal Features*

Here we add temporal features to the appliance model, and see a slight boost in F-measure. The extent to which residents follow regular temporal patterns in their every day lives is varied. Exploring the effect of using temporal features in disaggregation allows us to learn both how reliable these features can be, and to evaluate homes based on their temporal regularity. The extent to which temporal patterns can be found in energy usage can be useful for utility companies. Homes which do not display temporal patterns may be good candidates for real-time pricing programs; as residents' schedules are less structured they may be better able to respond to flexible pricing structures.

	Home One	Home Two	Home Three	Home Six	Average
Accuracy	80.9%	85.6%	75.4%	72.1%	78.5%
F-measure	53.4%	59.9%	43.5%	80.7%	59.3%
Precision	37.3%	49.2%	29.6%	72.1%	47.0%
Recall	100.0%	100.0%	100.0%	100.0%	100.0%

Table 3: Adding temporal features improves recall, and the F-measure scores.

Though both models earn similar F-measure scores, they are not actually predicting the same values as being on or off. This can clearly be seen in figure 4 below, which shows the true power used by the microwave in home one, as well as the power estimated by the model trained only with temporal features, and the power estimated by the model trained only with appliance features.

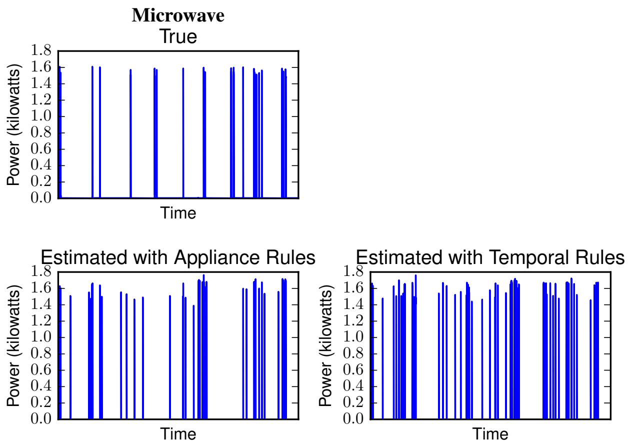

Figure 4: The power consumed by the microwave is over-estimated by both methods, but slightly more by the model which incorporates temporal patterns.

6. FUTURE WORK

A considerable weakness for disaggregation algorithms is the extent to which their quality depends on good prior knowledge of appliance behavior. In future work we would like to address this shortcoming. Additionally, current algorithms need access to sub-metered data. Discovering the full extent to which such data can be replaced with strong prior knowledge, or user supplied information, is imperative for determining the feasibility of disaggregation algorithms when limited disaggregated training data is available. To the best of our knowledge no such feasibility study has been done.

7. CONCLUSION

We have introduced a probabilistic disaggregation framework for non-intrusive load monitoring. Our framework can determine the relative consumption of individual appliances to a total energy bill. We have introduced two categories of features and shown that together they have the ability to predict if an appliance is on.

Energy disaggregation is still an open problem. Results are sensitive to good prior information about the amount of power an appliance may use, which may be difficult to obtain in the real world. We have proposed a framework which is flexible to these constraints. We can easily incorporate additional features into our framework, and require less assumptions to be met than existing solutions. Additionally, we have the ability to disaggregate across homes, simultaneously learning inter-home relationships, and disaggregating appliances. An adaptive framework for identifying appliance energy consumption has great potential to decrease energy demand, and we propose the first steps towards such a solution.

References

[1] U. E. I. Administration. Electric power annual. *http://www.eia.gov/electricity/annual/html/epa_01_01.html*, Accessed: 2016.

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

[3] S. H. Bach, M. Broecheler, B. Huang, and L. Getoor. Hinge-loss markov random fields and probabilistic soft logic. 2015.

[4] G. Brandon and A. Lewis. Reducing household energy consumption: A qualitative and quantitative field study. *Journal of Environmental Psychology*, 1999.

[5] S. Darby. The effectiveness of feedback on energy consumption. A Review for DERFA of the Literature on Metering, Billing, and direct Displays. 2006.

[6] T. Dietz, G. T. Gardner, J. Gilligan, P. C. Stern, and M. P. Vandenbergh. Household actions can provide a behavioral wedge to rapidly reduce us carbon emissions. *Proceedings of the National Academy of Sciences*, 2009.

[7] G. T. Gardner and P. C. Stern. The short list: The most effective actions u.s. households can take to curb climate change. *Environment: Science and Policy for Sustainable Development*, 2008.

[8] G. W. Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, Dec 1992.

[9] M. J. Johnson and A. S. Willsky. Bayesian nonparametric hidden semi-markov models. *Journal of Machine Learning Research*, 14(1), Feb. 2013.

[10] H. Kim, M. Marwah, M. Arlitt, G. Lyon, and J. Han. *Unsupervised Disaggregation of Low Frequency Power Measurements.*

[11] J. Z. Kolter and T. Jaakkola. Approximate inference in additive factorial hmms with application to energy disaggregation. In *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics (AISTATS-12)*, 2012.

[12] J. Z. Kolter and M. J. Johnson. Redd: A public data set for energy disaggregation research. *Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA*, 2011.

[13] P. Kouki, S. Fakhraei, J. Foulds, M. Eirinaki, and L. Getoor. Hyper: A flexible and extensible probabilistic framework for hybrid recommender systems. In *ACM Conference on Recommender Systems (RecSys)*, 2015.

[14] S. Makonin, F. Popowich, I. V. Bajic, B. Gill, and L. Bartram. Exploiting hmm sparsity to perform online real-time nonintrusive load monitoring. *IEEE Transactions on Smart Grid*, 2015.

[15] O. Parson, S. Ghosh, M. Weal, and A. Rogers. Non-intrusive load monitoring using prior models of general appliance types. In *Twenty-Sixth Conference on Artificial Intelligence (AAAI-12)*, 2012.

[16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 2011.

[17] D. Sridhar, J. Foulds, M. Walker, B. Huang, and L. Getoor. Joint models of disagreement and stance in online debate. In *Annual Meeting of the Association for Computational Linguistics (ACL)*, 2015.

[18] P. C. Stern. Contributions of psychology to limiting climate change. *American Psychologist*, 1989.

[19] J. H. Van Houwelingen and W. F. Van Raaij. The effect of goal-setting and daily electronic feedback on in-home energy use. *Journal of Consumer Research*, 1989.

[20] M. Zeifman. Disaggregation of home energy display data using probabilistic approach. *Consumer Electronics, IEEE Transactions on*, 2012.