Non-Intrusive Load Monitoring Using Occupancy Information and Ambient Information*

Pengji Zhang
Carnegie Mellon University
Pittsburgh, Pennsylvania
pengjiz@andrew.cmu.edu

ABSTRACT

Non-intrusive load monitoring is of great importance in smart infrastructures. This project attempts to incorporate prior knowledge about the residents into traditional Hidden Markov Models for non-event based appliance load monitoring. We build a model by utilizing the information on temperature and the locations of the residents to estimate the initial probabilities and transition matrices of an extended Hidden Markov Model dynamically. Our experiments of the model on one single appliance TV show that incorporating the prior knowledge of the residents could greatly improve the performance of conventional models.

KEYWORDS

Non-intrusive Load Monitoring, Hidden Markov Model

ACM Reference Format:
Pengji Zhang. 2017. Non-Intrusive Load Monitoring Using Occupancy Information and Ambient Information. In *Proceedings of Data Driven Building Energy Management (12752).* ACM, New York, NY, USA, Article 4, 4 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

Non-intrusive load monitoring (NILM) is aimed to estimate the electricity consumption of each appliance from the aggregated consumption, without installing any additional sensors. [1] Because of the rapid growth in smart meter deployments all over the world—which provides an ideal platform for collecting electricity consumption data and communicating with consumers [3, 4]—we can expect NILM will be of great importance in the future for the management of the power system. [3]

Many approaches have been proposed to solve the NILM problem. Even though the methods of those solutions are very different from each other, most of them relies heavily on the power load data to train the model. For example, Parson et al. used the aggregated power load data to learn the transition matrix and the emission matrix for each appliance in the setting of a hidden markov model; [4] Hart et al. used a model to detect and attribute the significant changes in the power load. [1] However, given the rise of *smart cities* and *internet of things*, we may have a lot more information that can help us to improve the performance of existing models. For example, the traffic flow information of a certain day could tell us the times of people leaving or getting back home, thus could tell us much information about the usage of appliances. For another example, we may use the occupancy information from the smart thermostat systems to better predict the usage of certain appliances like TV, microwave, etc. Thus, for this project, we would like to explore the feasibility of building a NILM solution with *not only the power load and the data for the power system, but also the ambient information and occupancy information.*

2 DRED DATA SET

The data set we are going to use is the Dutch Residential Energy Dataset (DRED). [5] The data set contains the appliance level energy monitoring data, the ambient information (indoor and outdoor temperatures, wind speed, humidity and precipitation) and the room-level occupancy information of one household, over a period of 6 months from 5th July to 5th December 2015. [5]

Figure 1 shows the layout of the household as well as the locations of the appliances and the sensors.

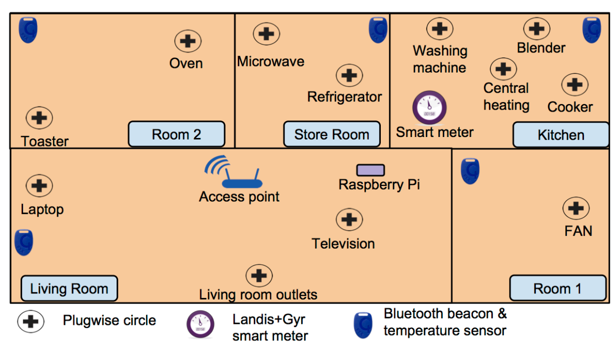

Figure 1: Household layout and the deployments of sensors in DRED. [5]

3 METHODS

There are two parts in our model: softmax regression for estimating the parameters of a hidden Markov model and an extended hidden Markov model for predicting the states of appliances according to the total power load of a household.

The whole process is shown in Figure 2.

*Fall 2017, 12752 Data Driven Building Energy Management Project

ACM ISBN 123-4567-24-567/08/06…$15.00
https://doi.org/10.475/123_4

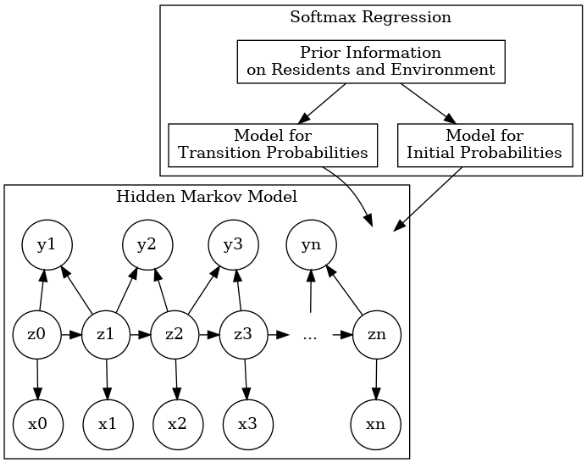

Figure 2: Methods for this project. Two softmax regression models are fitted to predict the parameters for an extended HMM model proposed by Parson et al. [4], which will be used for predicting the states of appliances from the total power demand of the household.

3.1 Softmax Regression for Estimating Transition/Initial Probabilities

Softmax regression is one of the multi-class version for conventional logistic regression. It is capable to output the probabilities of each class, which makes it a good choice for estimating the probabilities in a hidden Markov model. The form of softmax regression model is shown in Equation 1. [2]

$$P(y = k|\mathbf{x}; \theta) = \frac{\exp(\theta^{(k)T}\mathbf{x})}{\sum_{i=1}^{K}\exp(\theta^{(i)T}\mathbf{x})} \tag{1}$$

The features we will use in the model for initial probabilities include

(1) Temperatures at different locations as shown in Figure 1.
(2) Occupancy information of each room. For this feature, we use the percentage of the residents being detected in one room out of total number of records in one minute. For example,

$$O^k_{\text{kitchen}} = \frac{\text{Number of records in kitchen in the } k\text{th minute}}{\text{Total number of records in the } k\text{th minute}} \tag{2}$$

(3) Hour of the day. This feature is to capture the daily patterns of appliance usage.
(4) Day of the week. This feature is to capture the weekly patterns of appliance usage.

Note that in this project the last two features are encoded with one-hot-encoder.

3.2 Extended Viterbi Algorithm for Decoding the States of Appliances

The algorithm we choose to predict the states of one appliance from the total energy consumption is an extended Viterbi algorithm based on the work of Parson et al. [4] Parson et al. proposed an extended Viterbi algorithm for decoding the hidden states in the hidden Markov model shown in Figure 3.

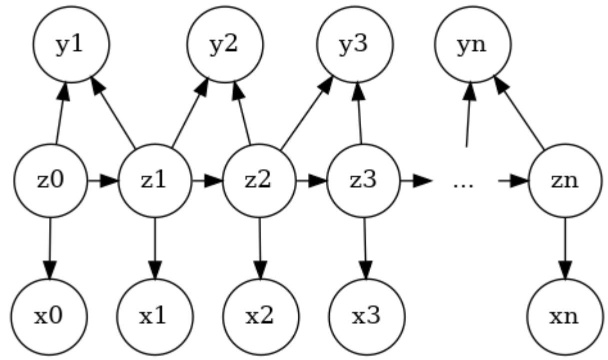

Figure 3: The hidden Markov model with two observed variables x and y.

In their model,

$$\begin{aligned} p(\mathbf{x}, \mathbf{y}, \mathbf{z}|\theta) = p(z_1|\pi)\prod_{t=2}^{T} p(z_t|z_{t-1}, \mathbf{A}) \\ \prod_{t=1}^{T} p(w_{z_t} \le x_t|z_t, \phi)\prod_{t\in S} f(y_t|z_t, z_{t-1}, \phi), \end{aligned} \tag{3}$$

where

$$p(w_{z_t} \le x_t|z_t, \phi) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x_t - \mu_{z_t}}{\sigma_{z_t}\sqrt{2}}\right)\right] \tag{4}$$

$$f(y_t|z_t, z_{t-1}, \phi) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{y_t - (\mu_{z_t} - \mu_{z_{t-1}})}{2\sigma^2}\right). \tag{5}$$

We further extend this model by injecting information of occupancy and environment in the transition probabilities and initial probabilities,

$$\pi = \mathrm{g}(I_1) \tag{6}$$

$$\mathbf{A_t} = \mathrm{g}'(I_t), \tag{7}$$

where g and g′ are the two models from the softmax regression on the training samples. Our algorithm does not assume one single parameter for the whole hidden Markov model, instead the parameters change with time according to the external information. We suppose this would improve the performance because it captures more variation in the process.

Besides, we will also use moving window smoothing to smooth the predicted state sequences.

4 RESULTS

4.1 Data Preprocessing

In the DRED data set, the minimum sampling frequency in the data set is 1 Hz for the temperature data, so in the experiment, we choose to take the mean values within each minute of all the variables to align the sampling frequencies.

The records of the appliance 'TV' are used to evaluate our model. This data set contains the power demand of each appliance, but does not contain the states. Therefore, we set the states as 'off' when the power demand is near zero, and 'on' otherwise (See Figure 4). This is not a very precise way for this. We may use some automated process with the current and voltage data to better classify the states. However, this manual way should be enough for testing our model in this project.

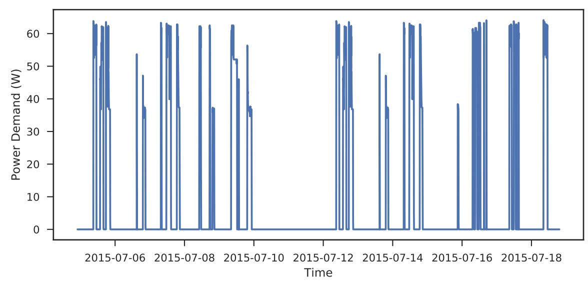

Figure 4: Mean power demand of TV in one minute.

Then we need to split the data set into two parts: one for the softmax models and the other one for the HMM. Because there are many gaps in the data set (Figure 5) and large gaps will greatly influence the performance of hidden Markov models but not the performance of softmax regression models, which do not rely on time, so we choose a range of samples that has small gaps for the HMM model and other samples for the softmax models.

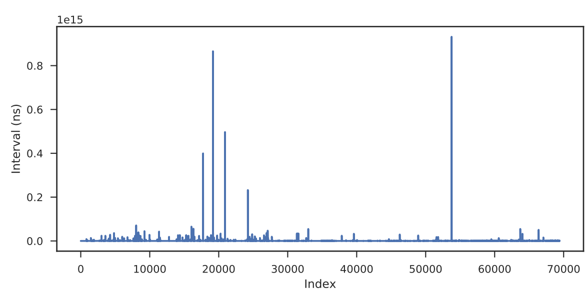

Figure 5: Intervals between consecutive samples (in nano seconds). The 42,000th to 47,000th samples are used in the HMM, other samples are used to fit the softmax models.

4.2 Softmax Regression

In our experiment, we pick a few models with different regularizers and strengths of regularization. Figure 6 shows the results of 5-fold crossing validation of those models and the performance of the baseline model (one single set of probabilities obtained by counting).

We can see that the models do improve the accuracy of estimating the probabilities, even though for predicting initial probabilities those models are not very ideal—the AUC of the PR curve is about 0.2 only.

We know that by nature the usage data of many appliances are imbalanced—many appliances are at the 'off' state most of the time,

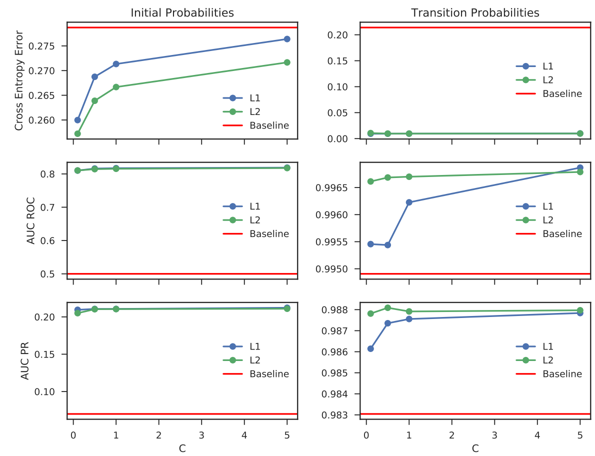

Figure 6: Cross entropy losses, AUC ROC and AUC PR of different models.

like the toaster, while some other appliances are at the 'on' state most of the time, like the fridge. So we choose the model with good AUC ROC as well as AUC PR scores as our final model, i.e. the model with L2 regularizer and $\lambda = \frac{1}{C} = 1$.

4.3 Decoding the States

With the fitted model, we will use the extended Viterbi algorithm to predict the states of the appliance of interest—TV. The window size of the smoother is 3, and the threshold for accepting $f(y|z_t, z_{t-1})$ as defined by Parson et al. [4] is 0.033. Figure 7 shows the results.

We can see that the predicted sequence is more unstable compared to the true values.

Figure 8 shows the confusion matrix. There are a lot of false 'off' states. The F1 score of the model is only 0.14, which is not very satisfying.

5 DISCUSSION

With a hidden Markov model, injecting information of the residents and the environment will help improve the accuracy of the parameters, but the performance of the whole model on non-intrusive appliance load monitoring tasks is still not very impressive. We suppose this is because the noise from other appliances. In the hidden Markov model with two observed variable structure, much information on the state change of the appliance comes from the variable y, which is the differences between two consecutive power demands of the whole household. However, some appliances like fridges may interfere the prediction process of the model. From Figure 7 we can see that when the television is at the 'On' or 'Off' state, the power demand of the whole house is still changing, which may make the model to produce false state changes.

For the future work, we will bundle certain appliances as a group, and predict their states with the factorial hidden Markov model. We think this will help avoid the 'dancing' of predictions but will not require as much computation as the full factorial hidden Markov model that includes all the appliances.

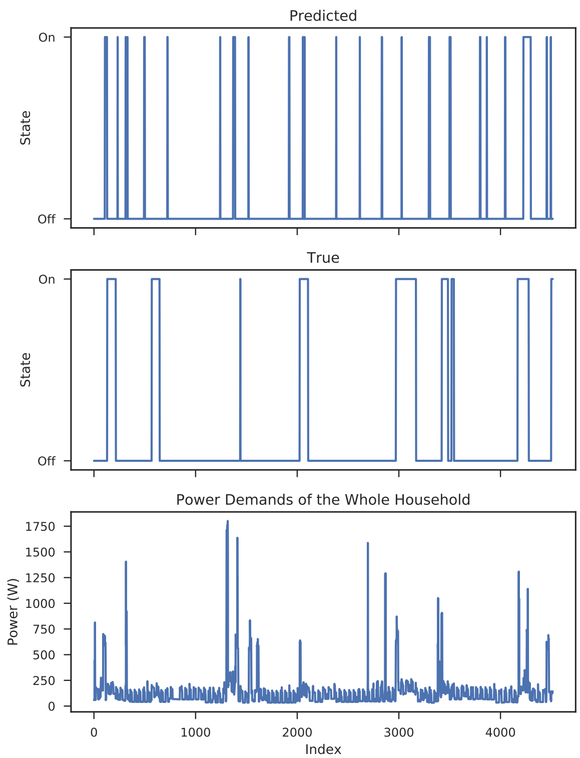

Figure 7: Predicted states, true states and power demands of the household.

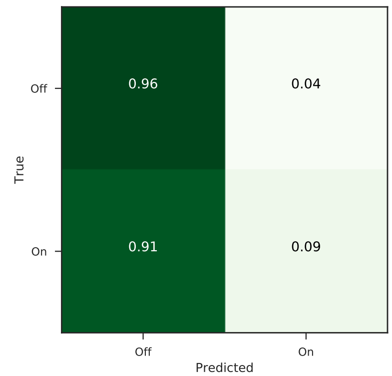

Figure 8: Normalized confusion matrix of the model on the test set.

We may also try to collect more data on the residents and environment, such as the age, number of residents in the house, etc. We think such information will further improve the accuracy of our model.

REFERENCES

[1] George William Hart. 1992. Nonintrusive appliance load monitoring. *Proc. IEEE* 80, 12 (1992), 1870–1891.

[2] David Heckerman and Christopher Meek. 1997. Models and selection criteria for regression and classification. In *Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 223–228.

[3] Hyungsul Kim, Manish Marwah, Martin Arlitt, Geoff Lyon, and Jiawei Han. 2011. Unsupervised disaggregation of low frequency power measurements. In *Proceedings of the 2011 SIAM International Conference on Data Mining*. SIAM, 747–758.

[4] Oliver Parson, Siddhartha Ghosh, Mark J Weal, and Alex Rogers. 2012. Non-Intrusive Load Monitoring Using Prior Models of General Appliance Types.. In *AAAi*.

[5] Akshay SN Uttama Nambi, Antonio Reyes Lua, and Venkatesha R Prasad. 2015. Loced: Location-aware energy disaggregation framework. In *Proceedings of the 2Nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*. ACM, 45–54.