Hidden Markov Models for Nonintrusive Appliance Load Monitoring

Jacob A. Mueller[†], Anusha Sankara[‡], Jonathan W. Kimball[†], and Bruce McMillin[‡]
[†]Department of Electrical & Computer Engineering
[‡]Department of Computer Science
Missouri University of Science and Technology
Rolla, Missouri 65409
kimballjw@mst.edu

***Abstract*—A method of device modeling for nonintrusive appliance load monitoring (NIALM) is presented. The proposed method uses hidden Markov models to describe device behavior. Unlike previous methods of device modeling, observations are associated with instantaneous power measurements as opposed to step changes in power use or on-off transients. The training procedure for individual devices is discussed. Accuracies of seven different device models are assessed using k-fold cross validation. In this assessment, the correlations between sequences of known state transitions and calculated Viterbi sequences representing predicted transitions are determined. This process is repeated for power use profiles collected at different sampling rates. Individual devices' Viterbi sequences are shown to be able to accurately approximate the actual device power use.**

***Index Terms*—Nonintrusive load monitoring, appliance modeling, disaggregation**

I. Introduction

Nonintrusive appliance load monitoring (NIALM) methods are used to determine the individual contributions to energy use in a combined electric system. Accurate NIALM methods provide a detailed understanding of the devices active in a system from measurements of the composite energy use profile. In a residential setting, this profile can be a meter-level or main-breaker-level measurement. The breakdown of energy use provides a wealth of information that would otherwise be impossible or impractical to attain. It would, for example, be impossible for a utility company to measure the energy use of individual devices in the homes of their customers and would be impractical for the consumer to purchase the number of sensors this would require. Nonetheless, this information is of considerable value to both parties. For utility companies it provides insight into the nature of the loads they supply and how they may change over time. For consumers it provides a means to make informed decisions on appliance use or replacement. A consumer concerned with energy conservation would be able to determine which appliance upgrade would provide the most significant reduction of their carbon footprint.

The methods used for NIALM are diverse and numerous, but they share the same three basic principles of operation [1]. First, individual devices must be mathematically described, or modeled, according to some chosen load profile characteristics. Second, the power or energy use data must be collected at the time of operation using an appropriate sensor apparatus. Finally, the selected load profile characteristics must be recognized in this composite data collection and used to separate, or disaggregate, the appliances into their individual energy use profiles. Classification of NIALM variants may be succinctly accomplished according to differences at these three common stages. In this work, consideration is given to the existing methods of load modeling and an alternative method that uses hidden Markov models is proposed. This new method is intended to serve as the first stage in a full NIALM approach.

II. Background: Load Modeling in NIALM

The earliest implementations of NIALM were established in the 1980s. One early work in NIALM classified load types based on internal electric characteristics [2]. According to this method, devices could fall into one of six categories: purely resistive, one of two inductive-type classes, passive power-electronic, controlled power-electronic, or a special class for fluorescent lights. These distinctions were made based on the nature of the device's starting power transient and the harmonic content of its current—both functions of the electrical elements present in the device. For example, a motor-driven device would be characterized by an inrush current transient. This would classify the device into one of the inductive load categories. The rectifier on an electronic device's power supply would cause characteristic harmonics, which consequently place it in the passive power electronic category. This methodology has since been the subject of further research [3]. While harmonic content is a strong resource for aiding appliance description, there are drawbacks to this approach. The measured harmonic content of a residence is not independent of the harmonic quality of the distribution grid supplying its power. It is unclear whether this external factor would affect the predictive accuracy of harmonic-based load models. Additionally, measurement of

978-1-4799-5904-4/14/$31.00 ©2014 IEEE

harmonic content is only possible using relatively fast sampling rates, according to the Nyquist sampling theorem. Since harmonics up to the 11th are typically used, sampling rates on the order of 1 kHz or higher are required [1].

In [4], another early implementation of NIALM, the author takes an entirely different approach to appliance modeling and proposes four basic types of residential loads: on-off devices, finite state machines (FSM), constantly active devices, and continuously variable loads. On-off devices, such as light bulbs or microwaves, transition only between on and off states. FSM loads may have multiple states and multiple transition possibilities. Constantly active devices are simply on and have only one state, using a nominal amount of power. Continuously variable loads do not transition between states of operation discretely, and may be active anywhere in a continuous range of power use up to some device-specific maximum power. This approach to modeling devices uses the changes in steady-state power use characteristics that correspond to transitions between states. Using the changes in steady-state values, as opposed to the actual transient responses themselves, allows for a lower sampling rate to be used and decouples the measurements from the external factors that influence transient or harmonic-based approaches. However, using only transition characteristics means that only the on-off and FSM type appliances can be accurately modeled, and the latter only with added complexity [5]. Consequently, more recent attempts to apply this method have considered only on-off type appliances [6]-[7].

The transition characteristics observed in [4] to model appliances include both active (P) and reactive (Q) power changes. These observations are then grouped according to their position on the complex plane. By observing the local clusters of these step changes and matching similar magnitude positive and negative power change events, the appliances present in the measurements can then be identified. The device identification process matches the measured differences with known devices' power changes using a detection algorithm. The collected durations between state changes are fed back into the models to refine probability distributions describing device activity durations.

A similar modeling process is used in [8] for step changes in P by characterizing devices using transition probabilities. A single probability distribution is used to represent both the probability that a certain device will be active at a given time and the probability that this activity will correspond to a certain change in P. A set of probability distributions for each of the on and off states of different devices in the system is then used to find the most likely states for a measured power use waveform using the Viterbi algorithm [9]. To do this, though, two assumptions must be made. First, it is assumed that no two devices' probability distribution functions overlap in a given system. This places a limit on the similarity of devices for which this method will be accurate. The second assumption is that no two devices start or stop simultaneously. In practice, these assumptions are obviously not reasonable. In a later work, one of the authors of [8] identifies the overlap between these device activity probability distributions as the primary reason behind low accuracies in NIALM methods, and proposes a new disaggregation method attempting to remedy this issue [6]. However, this revised approach is only able to model on-off type devices. In [10], similar transition probability models are used to represent continuously variable and FSM appliances, in addition to simple on-off devices, using probability mass functions generated from discretely sampled current change measurements.

Hidden Markov models have been used to represent device behavior as well [7],[11]-[13]. In Markov models, the probability of a transition between states and the probability of a certain change in power being observed are treated as separate quantities, and both probabilities are assumed to be dependent only on the current state of the device. Since the actual states of the devices cannot be determined directly from the combined power use measurement, these are said to be hidden Markov models. The methods that use this structure consist of a transition probability matrix (T) and an observation probability matrix (O). The process of creating these models by generating appropriate T and O matrices is referred to as training, and can be either supervised or unsupervised. An unsupervised training procedure for a simple on-off representation of multiple devices in a combined system is given in [7]. While unsupervised, this method requires preprocessing to remove transient spikes that otherwise decrease the accuracy of the models. In this case, the T matrix represents all possible combinations of transitions between the on and off states of the devices in the system, so the number of devices must be known a priori. The O matrix contains the probabilities of a change in power being observed. For a system of simple on-off states, this is approach is questionable. A dryer, which consists of a heating element and tumble motor, will be observed to have a change in power equal to the consumption of both internal components when starting. While the tumble motor remains on, the heating element may turn on and off to regulate the internal temperature of the basket. In order to account for all of these transitions in an on-off model, the internal components of the dryer would need to be modeled independently. This is addressed briefly in [12], in reference to a similar on-off type representation. In that system, however, only low power (less than 200W) electronic devices are used, including TVs, computer monitors, and gaming consoles. In terms of load modeling, this only represents a fraction of the loads commonly found in a residential context. When tested on an accurate representation of residential loads, this on-off approach to device modeling has been shown to have fairly poor accuracy at the disaggregation stage, around 50% [13]. A more practical approach, in which devices are able to be modeled with an arbitrary number of states, was used in [11] on the same test data set. In this case, total accuracy was not given, but individual device accuracy was as high as 77%.

In these approaches to appliance modeling, devices are almost always characterized by changes in P and Q, requiring some level of edge detection. Using these changes as device signatures, whether as changes in steady-state power use or the actual transient responses themselves, is accompanied by a set of challenges that has crippled the accuracy of these models in the disaggregation stage. These issues include problems handling multiple simultaneous device transitions, added complexity or preprocessing required to handle FSM devices, and an inability to distinguish between devices with similar power change characteristics.

III. Proposed Method of Device Modeling

The proposed approach to device modeling is similar in structure to previous methods using hidden Markov models, but differs in its description of observations and in the procedure used to generate and train the transition and observation matrices. All devices are treated as FSM appliances. The simplest devices, on-off type appliances, are simply FSM devices with only two operational modes. Devices with more internal processes have more states. Using the dryer example, the two internal process—heating element and tumble motor—can be either off or on, for a total of four states. However, since dryer operation never involves an active heating element without active tumble motor, the device can be said to have three states. It is mentioned in [12] that each internal binary process in a device may be modeled as a single on-off device. While this is true, matching the two processes to a single device provides more information to distinguish between similar devices. While the dishwasher and dryer may both have heating elements and motors, the sequence of transitions through their operational modes is very different.

The main difference between the proposed method and previous methods is the treatment of observations. In the proposed method, observations are associated with instantaneous power use characteristic of a device state. In previous methods, observations were associated with step changes in power characteristic of transitions between states. In the proposed method an observation is defined as a power measurement in time. A sequence of observations is the sequence of active power measurements grouped into discrete, fixed-size buckets. Separating the measurements into buckets prevents the O matrix from being infinite in size. There are several advantages to this approach. A lower sampling rate may be used without concern for missing device transitions. No edge detection or preprocessing is required. Devices may be identified not only by their power use, but by their characteristic patterns of transition between operational modes as well. Using commercially available home power monitoring devices, active power is the simplest and most readily available electrical characteristic appliance characteristic. This method of modeling, however, is not bound to active power and can use any electrical characteristic to describe device behavior.

A supervised approach is used to train device models. Historical data, or recordings of individual power use for a specific device, is used to shape the T and O matrices of the device models. Once this is done for a variant of a device, though, it does not need to be repeated for that variant, meaning that no semi-invasive means of model training are required. The training process itself involves using maximum likelihood estimation (MLE) to determine the most appropriate T and O matrices for a given sequence of observations and a corresponding sequence of known state transitions. The supervised aspect of this process is the determination of state transitions for the sequence of observations given as training data. For two state (on-off type) devices, this is as simple as labeling observations above a noise floor as state two (on-state) and those below it as state one (off-state). For devices with more states, a histogram of the observations is used to find some separation between clusters of observations. The clusters are then labeled incrementally. The sequence of states determined according to this method corresponds to an ideal Viterbi sequence calculated at 100% accuracy, and is used later as a benchmark for comparison to calculated Viterbi sequences.

Using the dryer as an example, a visual representation of the modeling and training process is shown in Figure 1. Figure 1a shows the sequence of observations for a dryer. This waveform is the instantaneously sampled active power of the dryer after being separated into discrete buckets of 10 W. The right y-axis scaling shows the measurements in kW, and the left in their corresponding buckets. From the sequence of observations, it is clear that two processes are occurring in the device. The periodic pulses correspond to a 2.5 kW heating element and the lower plateau corresponds to the constant power used by the tumble motor. Figure 1b shows a histogram of the observations. The observations around bucket 250 represent the 2.5 kW heating element pulses. Superimposed on these two graphs are dotted lines denoting the boundaries

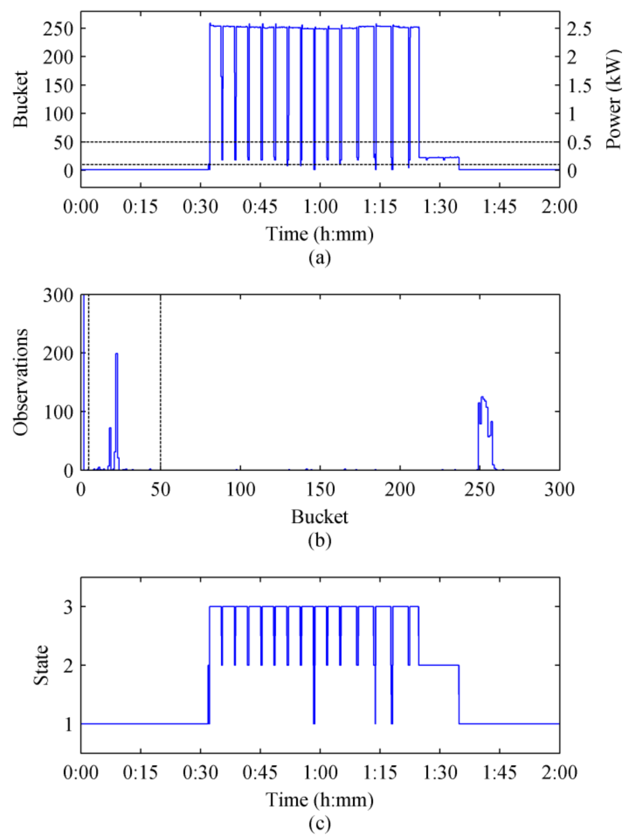

Fig. 1. Dryer active power measurement after bucketing (a), histogram of discrete observations (b), and artificial sequence of states (c).

between state definitions. Observations below bucket 5 (or power draw less than 50 W) are defined to be state 1 operation (off), observations between buckets 5 and 50 (power draw between 50 W and 500 W) are defined to be state 2, and anything above bucket 50 is defined as state 3. Figure 1c shows the sequence of states determined by breaking down the observations according to the state definitions.

The maximum likelihood estimate of the parameter T is determined from this sequence of states according to

$$T_{i,j} = \left(\frac{nT_{i,j}}{nT_i} \right), \tag{1}$$

where $T_{i,j}$ is the probability of transitioning from state i to state j, $nT_{i,j}$ is number of transitions from state i to state j, nT_i is the number of transitions from state i to any state. The O matrix estimate is determined as

$$O_{i,k} = \left(\frac{nO_{i,k}}{nO_k} \right), \tag{2}$$

where $O_{i,k}$ is the probability of an observation in bucket k at state i, $nO_{i,k}$ is the number of observations in bucket k at state i, and nO_k is the total number of observations in bucket k.

Since appliances are trained as generic models, each new variant of the device trained into the model changes the T and O parameters. Each different variant influences the model parameters equally. For a new device variant with parameters T_{new} and O_{new} added to an existing model with parameters T_{old} and O_{old}, the new parameters are found according to

$$T_{model} = \left(\frac{m \times T_{old} + T_{new}}{m+1} \right) \tag{3}$$

$$O_{model} = \left(\frac{m \times O_{old} + O_{new}}{m+1} \right) \tag{4}$$

where T_{model} and O_{model} are the updated model parameters and m is the number of device variants already used to train the model.

IV. Device Data Acquisition

The appliance data used was a combination of locally collected power use profiles and devices featured in the Tracebase repository, which is documented in [14]. Locally collected profiles were measured using a Power Standards Lab PQube measurement device, which is shown as Figure 2. The PQube is a flexible power and energy monitoring device that measures line voltages and currents and internally calculates relevant power characteristics. It is able to automatically adjust for different measurement configurations. For a residential setting, this included only single and split phase configurations, depending on the type of device profiled. Measurements from the PQube were logged at 1 second intervals using Modbus over TCP/IP. Multiple variants of appliances were collected to ensure accurate device representation. Device variants were defined by differences in manufacturer or operational cycle. For example, two microwaves made by different manufacturers, both capable of operating in full-power or defrost modes, would provide a total four different device variants for the microwave appliance type.

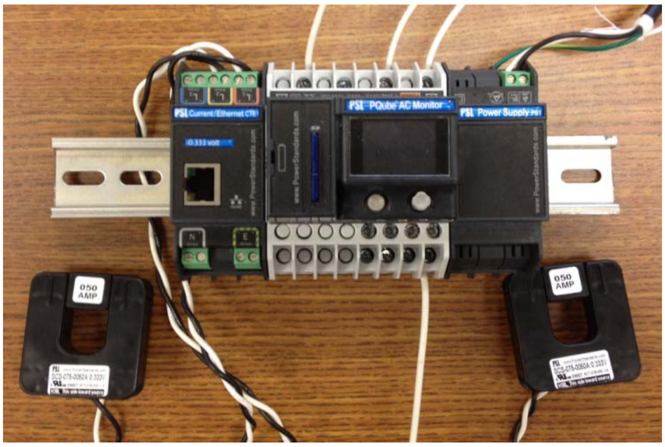

Fig. 2. PQube measurement device

The Tracebase data set contains similar profiles of individual devices. The profiles' sampling periods are variable, between 1 second and 10 seconds. Since the database is fairly extensive, containing over 1,000 device profiles, it was possible to select only those with the shortest sampling periods. This minimized the amount of reconstruction necessary when using generating fixed 1 second sampled observation sequences. When appropriate, missed samples were reconstructed using zero-order holds. In addition to the Tracebase data set, other public data sets are available for NIALM testing as well [10],[13]-[15]. For this work, however, the information in the Tracebase repository was sufficient to fill the gaps in the locally collected data.

V. Verification of Individual Models

In order to verify the accuracy of the modeling process, k-fold cross validation was used for eight different appliance models. For each appliance, five variants were used, four as training data and one as testing data. A Viterbi sequence was calculated for each test sequence using the models trained on the four training sequences. This calculated Viterbi sequence was then compared to the artificially generated sequence of known state transitions. This was done five times, with each device variant acting as test set once. The entire process was repeated for sampling periods of 5 seconds and 30 seconds. A comparison of the known state transitions and calculated Viterbi sequences for a dishwasher is shown as Figure 3. The known state transitions are shown in Figure 3a, the 5 second sampled Viterbi sequence in Figure 3b, and the 30 second sampled Viterbi sequence as Figure 3c.

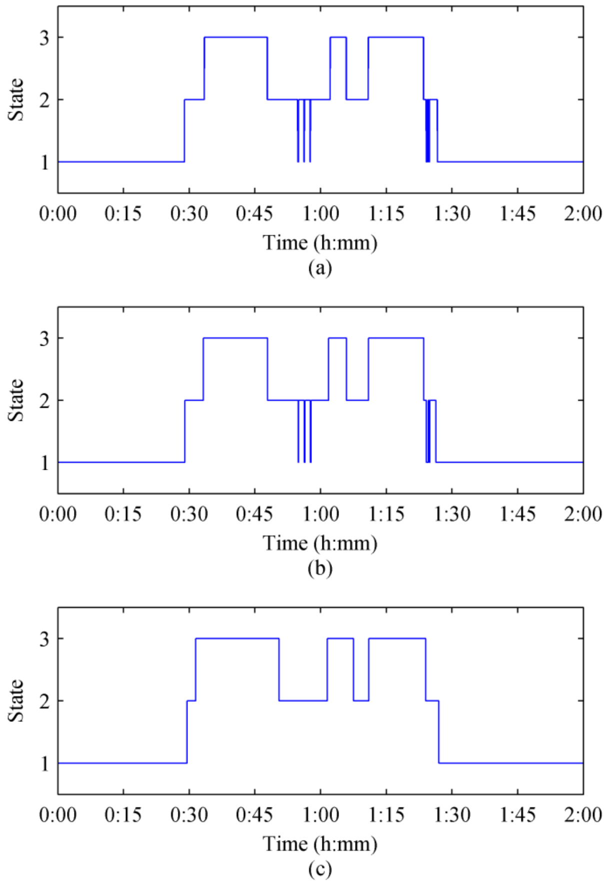

Fig. 3. Comparison of dishwasher known state transitions (a), 5s sampled Viterbi sequence (b), and 30s sampled Viterbi sequence (c).

The metric of accuracy used in the k-fold cross validation was the correlation between the sequences of known state transitions and the calculated Viterbi sequences. To ensure fair comparison of accuracies, the known Viterbi sequence was kept at a sampling interval of 1 second. If a device completed a full cycle of operation between samples, as was often the case for the microwave at 30s sampling, this was treated as 0% accuracy. These resulting correlations were then averaged together for each device, and are shown in Table 1.

The high correlations indicate that the device models are effective at determining the sequences of states as defined for measurements on individual devices even without being trained on observations from that device. The 5s sampled sequences were found to have higher correlation to the known state sequences. Figure 3, though, shows that the general load profile shapes of the state sequences are still intact when the Viterbi sequences are calculated from 30s sampled data, which is significantly slower than most commercially available home monitoring devices are able to offer.

From these Viterbi sequences and the devices' O matrices, approximations of the original test sequences may be obtained. Figure 4 shows a comparison of the raw active power measurements for a dishwasher and approximations based on the 5s and 30s Viterbi sequences. While this process of transforming a power consumption profile to and from

TABLE I. CORRELATIONS OF VITERBI SEQUENCES IN K-FOLD CROSS VALIDATION

Device	**5s Samples**	**30s Samples**
Dryer	0.9894	0.9488
Washer	0.9622	0.7981
Dishwasher	0.9963	0.958
Stove	0.9781	0.9279
Microwave	0.7075	0.3235
Toaster	0.973	0.8470
Desktop PC	0.8892	0.8268
Refrigerator	0.9961	0.9714

representation as a sequence of states has little practical use on its own, it shows the ability of the models to retain information. This is important because it demonstrates the suitability of this approach to a larger process of load monitoring. A system of devices can be modeled as a combination of their independent states in the same way that a device is modeled as a combination of independent processes. A single Viterbi path can be found for a given measurement of the devices' combined power consumption, and can then be disaggregated into the individual device components. The similarity of the original and approximated load profiles in Figure 4 indicates

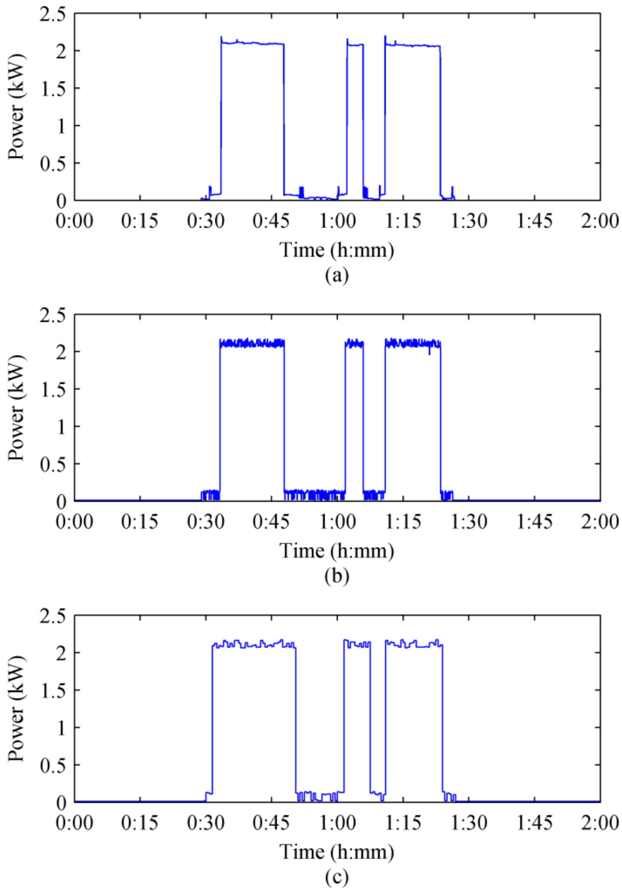

Fig. 4. Comparison of original active power measurement (a) and power use approximations generated using Viterbi sequences calculated from 5s (b) and 30s (c) sampled observations.

that the device models contain sufficient information to determine power and energy use of individual devices if their individual Viterbi sequences can be found. This forms the basis of a hidden Markov model based NIALM approach.

VI. Conclusion and Future Work

This work presents a method of modeling appliances using hidden Markov models. The method is shown to be able to determine the most likely sequence of operational states for a FSM type device from a collection of measurements of active power consumption. The proposed method of modeling appliances requires no preprocessing or edge detection and is able to maintain accuracy at very low sampling rates. This approach to device modeling forms the foundation for a method of nonintrusive load monitoring using hidden Markov models. Future work in this project will focus on expanding this method of modeling appliances into a full NIALM procedure, including representing multiple devices as a single combined hidden Markov model and disaggregating combined Viterbi sequences into individual device components.

References

[1] M. Zeifman and K. Roth, “Nonintrusive appliance load monitoring: Review and outlook,” Consumer Electronics, IEEE Transactions on, vol. 57, no. 1, pp. 76–84, 2011.

[2] F. Sultanem, “Using appliance signatures for monitoring residential loads at meter panel level,” Power Delivery, IEEE Transactions on, vol. 6, no. 4, pp. 1380–1385, 1991.

[3] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong, “Power signature analysis,” Power and Energy Magazine, IEEE, vol. 1, no. 2, pp. 56–63, 2003.

[4] G. W. Hart, “Nonintrusive appliance load monitoring,” Proceedings of the IEEE, vol. 80, no. 12, pp. 1870–1891, 1992.

[5] G. W. Hart and A. T. Bouloutas, “Correcting dependent errors in sequences generated by finite-state processes,” Information Theory, IEEE Transactions on, vol. 39, no. 4, pp. 1249–1260, 1993.

[6] M. Zeifman, “Disaggregation of home energy display data using probabilistic approach,” Consumer Electronics, IEEE Transactions on, vol. 58, no. 1, pp. 23–31, 2012.

[7] S. Pattem, “Unsupervised Disaggregation for Non-intrusive Load Monitoring,” in Machine Learning and Applications (ICMLA), 2012 11th International Conference on, 2012, pp. 515–520.

[8] M. Zeifman and K. Roth, “Viterbi algorithm with sparse transitions (VAST) for nonintrusive load monitoring,” in Computational Intelligence Applications In Smart Grid (CIASG), 2011 IEEE Symposium on, 2011, pp. 1–8.

[9] G. D. Forney Jr, “The viterbi algorithm,” Proceedings of the IEEE, vol. 61, no. 3, pp. 268–278, 1973.

[10] S. Makonin, F. Popowich, L. Bartram, B. Gill, and I. V. Bajic, “AMPds: A Public Dataset for Load Disaggregation and Eco-Feedback Research,” in Electrical Power and Energy Conference (EPEC), 2013, pp. 1–6.

[11] O. Parson, S. Ghosh, M. Weal, and A. Rogers, “Non-Intrusive Load Monitoring Using Prior Models of General Appliance Types.,” in Proceedings of the 26th Conference on Artificial Intelligence (AAAI-12), Toronto, CA, 2012, pp. 356-362.

[12] H. Kim, M. Marwah, M. F. Arlitt, G. Lyon, and J. Han, “Unsupervised Disaggregation of Low Frequency Power Measurements.,” in SDM, 2011, vol. 11, pp. 747–758.

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

[14] A. Reinhardt, P. Baumann, D. Burgstahler, M. Hollick, H. Chonov, M. Werner, and R. Steinmetz, “On the accuracy of appliance identification based on distributed load metering data,” in Sustainable Internet and ICT for Sustainability (SustainIT), 2012, pp. 1–9.

[15] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe, and M. Berges, "BLUED: A Fully Labeled Public Dataset for Event-Based Non-Intrusive Load Monitoring Research," in Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD), Beijing, China, 2012.