Unsupervised Disaggregation for NIALM

(DRAFT)

Sundeep Pattem

I. Introduction

Nonintrusive appliance load monitoring (NIALM) methods allow deduction of the kind of appliances used in homes and buildings, along with their individual energy consumption without the use of specialized hardware attached to each appliance. While the earliest developments in NIALM were made by George Hart and his associates in the early 1980s [1], [2], there has been a renewed surge of interest due to (a) growing ecological sustainability and efficiency concerns with use of energy, (b) impetus for the deployment of a 'smart grid' for energy distribution, and (c) advancements in hardware and wireless communications ('smart meters') for measuring and transmitting large volumes of energy consumption data. Research in academia and a growing number of companies is looking into (a) machine learning and pattern recognition techniques to extract and provide actionable insights for efficiency/savings to energy consumers and for targeted marketing to appliance manufacturers, (b) building 'big data' infrastructure to draw, store and process energy data from smart meters in millions of homes and buildings, and (c) create a new culture of energy awareness using social networking tools etc.

In this work, a method for unsupervised disaggregation of appliance signatures from smart meter data is presented. The primary features used for unsupervised learning relate to abrupt transitions in the power waveform. The method consists of a sequence of procedures for appliance signature identification, Hidden Markov Modeling (HMM) and the Viterbi algorithm, and residual analysis. The reasoning behind using these procedures and the particular sequence in which they are used, and how well known techniques are adapted for the disaggregation problem is discussed. Building on earlier work in the literature [3]–[5], we extend it with (a) careful handling of lower power signatures, (b) dimensionality reduction of input to the HMM (c) segmented Viterbi decoding exploiting intermittently occurring and known 'zero' state. and (d) detailed discussion of procedures, limitations, and hints from results for elements of a complete solution.

Results from an implementation using data from the REDD [6] show that the method is effective for unsupervised, magnitude-based disaggregation, and by illustrating limitations of the same, provide insights on what a more complete solution would take into account for fully capturing appliance signatures. Due to the limitations, even though ground truth is available from the REDD, the performance evaluation is currently more qualitative than quantitative.

II. Problem definition

The aggregate power consumption is assumed to be a discrete time, real valued waveform $\{x[i] : i \in [1, n]\}$. With m appliances contributing, the consumption of appliance j is denoted by $\{a_j[i] : i \in [1, n]\}$, and the aggregate,

$$x[i] = a_1[i] + a_2[i] + ... + a_m[i], \text{ for } i \in [1, n]. \tag{1}$$

The disaggregation problem involves obtaining the individual appliance contributions a_j given the aggregate x. The number of appliances (m) and their characteristics are, in general, unknown and need to be 'learned' from the aggregate.

The scope of this work is as follows:

1) The focus is on unsupervised disaggregation. The power waveform from a mains is taken to be the only input. No knowledge about the appliances (number, characteristics) is assumed, even though it is available in the REDD.
2) Exact solution to the disaggregation problem is known to be intractable. The approach taken is to identify and separate as many appliance signatures as possible.
3) The core disaggregation algorithm in the current method is designed for matching magnitudes of appliance ON/OFF power transitions, and does not attempt to model or disaggregate other kinds of signatures. It also does not currently exploit ON/OFF duration information even though it is available from the signature identification procedure described in the following section. This, and several other limitations and improvements possible are discussed in Section. VIII.

The following four stages in the method are now described: (i) Signature identification, (ii) Dimensionality reduction, (iii) Disaggregation using HMM and (iv) Residual analysis.

III. Signature identification

This stage involves pre-processing of the input by removing the always-on component and smoothing to remove transients while preserving lower power appliance signatures, followed by analysis of transition statistics for identifying appliance signatures.

A. Always-on component

It is observed that the aggregate power consumption almost always stays above a certain threshold, specific to the corresponding household/mains. In some cases, this threshold is close to the mode statistic for the aggregate time series waveform.

The input is taken to be a discrete time, real valued power consumption waveform $\{x[i] : i \in [1, n]\}$. The always-on threshold may be obtained as:

$$ao\_th(x) = 0.5 \text{ percentile value of input sequence } x \tag{2}$$

The always-on component (aoc_x), and variable component (vc_x) of the aggregate waveform x, are obtained as:

$$\begin{aligned} aoc\_x[i] &= ao\_th(x), \quad \text{if } x[i] > ao\_th(x) \\ &= 0, \qquad\qquad \text{else.} \end{aligned} \tag{3}$$

$$vc\_x[i] = x[i] - aoc\_x[i]$$

The total power consumption for the always on component is,

$$aoc\_power = \sum_n aoc\_x[i] \tag{4}$$

B. Smoothing

The primary feature we want to use for disaggregation is the magnitude of (a) abrupt, *almost instantaneous*, and (b) *significant* changes in the power consumption waveform. In reality, the change or ON/OFF transitions involve transients and can be spread out over a few seconds or even minutes. Smoothing is necessary to remove the transient spikes and obtain the closest approximation of the waveform with instantaneous ON/OFF transitions. What amounts to a significant change in power is relative to the magnitude of the operational power of the appliances. A transition for a lower power device might match, or even be dwarfed by, noise in a higher power appliance waveform.

Various standard procedures are available for smoothing, such as quantization, windowed central tendency statistics, and down-sampling. Each of these provides some benefit for accurate disaggregation, while introducing some undesirable artifacts. To maximize benefits and minimize loss of important information for the problem at hand, we employ a judicious sequence of operations for smoothing.

1) Quantization: We first perform quantization, since the quantization error waveform retains information that can be utilized to disaggregate low power signatures at a later stage much better than with the other smoothing techniques.

Uniform quantization is employed. A fixed quantization step (Δ)is used in the results presented here. It might be beneficial to choose the value of the step for each waveform (Δ_x), based on some initial analysis of magnitudes. The quantized version (q_vc_x) is obtained as:

$$q\_vc\_x[i] = \Delta_x \cdot floor\left[\frac{vc\_x[i]}{\Delta_x} + \frac{1}{2}\right] \tag{5}$$

2) Preserving lower power signatures: As discussed above, when there are devices running at a low power level that falls within the quantization step, their contribution can possibly be "zeroed out". If these devices stay on for long durations or have a high on/off duration ratio, the final results or insights obtained will be skewed and may not correspond to reality. To address this issue, the quantization error waveform is computed and retained for use in the residual analysis step (see Section VI):

$$res\_q\_vc\_x[i] = vc\_x[i] - q\_vc\_x[i] \tag{6}$$

3) Median smoothing: Transients pose a tricky problem. We would like to preserve the sharp edge/step when an appliance turns on/off, but smooth out just the transient portion of it. Any sort of averaging will introduce a slope and result in the sharp edge being lost. For this situation, median smoothing is most appropriate.

For every time instant n and window $\left[n - \frac{w_{ms}}{2}, n - \frac{w_{ms}}{2}\right)$,

$$ms\_q\_vc\_x[i] = median\Big\{q\_vc\_x[n - \frac{w_{ms}}{2}], q\_vc\_x[n - \frac{w_{ms}}{2} + 1], ..., q\_vc\_x[n + \frac{w_{ms}}{2} - 1]\Big\} \tag{7}$$

Expected issues and available measures:

- Introduces artifacts (as do other filters)
  - the edges get shifted, by a maximum of $\frac{w_{ms}}{4}$
  - if the size of the window is too wide many edges might be completely lost
- choose window size w_{ms} carefully for individual waveforms/households, based on analysis of transition times in original waveform
- if important, after performing disaggregation on smoothed waveform, look into ways of correcting for artifacts introduced by this procedure.

4) Down-sample/window-wise constant smoothing: While median smoothing removes most of the transients, the ON/OFF transitions might still not be instantaneous. To obtain the sharp edges, down-sample smoothing (piecewise constant over a small, fixed window w_{ds}) is employed.

$$ds\_ms\_q\_vc\_x[l \cdot w_{ds} + q] = ms\_q\_vc\_x[l \cdot w_{ds}] \ \ \forall l \in [0, floor(\frac{n}{w_{ds}})], j \in [1, w_{ds}) \tag{8}$$

Expected issues and available measures:

- can result in "combining" events that happen within the window, and hence introduce errors in disaggregation. Choose window w_{ds} carefully, based on duration analysis..

C. Statistics for signature identification

Using the method described here, only appliances with 'perfectly rectangular' signatures or portions of their signatures exhibhiting this property in a clearly discernible manner after the pre-processing operations can be properly identified and disaggregated.

For convenience, we use the notation $pp\_x \equiv ds\_sm\_q\_vc\_x$ for the result of the pre-processing operations. A differentials waveform is obtained as follows:

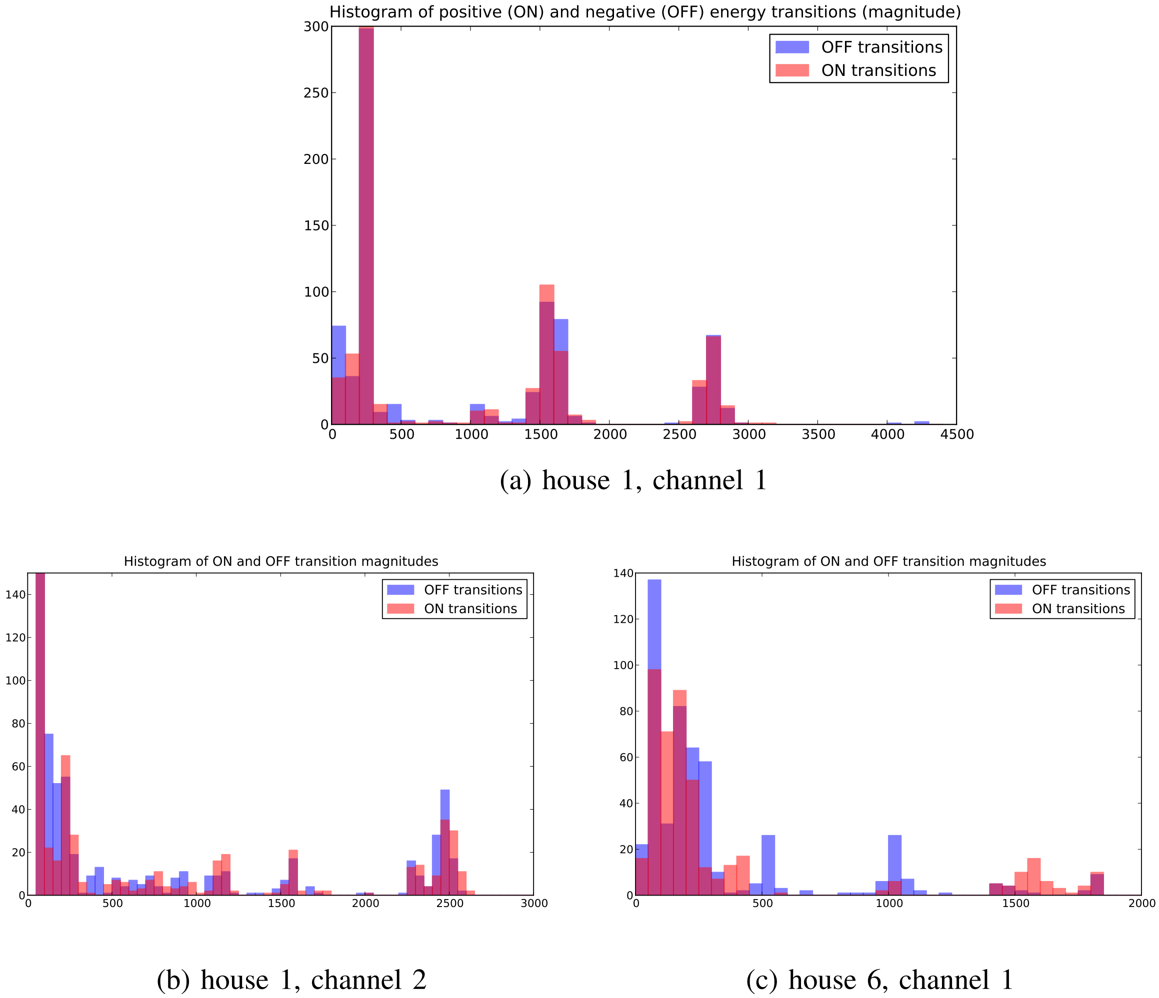
$$diff\_pp\_x[i] = pp\_x[i] - pp\_x[n-1] \tag{9}$$

Figure. 1 illustrates the identification of appliance signatures via the histogram (the CDF can be used as well for this purpose) for the non-zero differentials for different house/mains combinations. It shows examples of cases in which the current method can recover appliance signatures (a) to a large extent, (b) only partially, and (c) very little or not at all.

'Valid' signatures for the current method are obtained as follows:

1) Matching peaks in distributions of 'ON' and 'OFF' transition magnitudes. The peaks can be identified by checking for persistence in k-means clustering with several values of k.
2) For each signature, the peak is an indicator of the central tendency. The variance around it is also computed, which reflects natural variations, smoothing/quantization errors etc. At this point, it is possible to fit a model such as a Gaussian distribution for each magnitude signature with parameters as computed.
3) Using the variance or model parameters for every pair of signatures that are adjacent in the histogram/CDF, compute the best possible cut-off thresholds to separate them:

$$[threshold\_lo(a), threshold\_hi(a)] \ \ \forall a \in [1, K].$$



(a) house 1, channel 1

(b) house 1, channel 2

(c) house 6, channel 1

Fig. 1. Good match in positive (ON) and negative (OFF) transition magnitudes identifies appliance signatures that can be disaggregated by current method. (a) Well matched peaks identifiable. (b) Combination of matched peaks and spread (between 500 and 1000). Current method cannot incorporate signatures with such spread. The matched peak signatures can be disaggregated. (c) No match in peaks, current method not applicable.

IV. Dimensionality reduction

It is observed that appliance ON/OFF transitions are relatively less frequent events compared to the time period for which they stay ON or OFF. As a result, the differentials have a distribution heavily concentrated close to zero. This is more so for differentials computed following quantization and smoothing, i.e. for the $diff\_pp\_x$ waveform. We see that the number of instances for which the differentials are non-zero is orders of magnitude smaller than the length of the input. Since the probabilistic inference that follows requires knowledge solely of the non-zero differentials, the dimension of the input to it can be significantly reduced. The sequence of time instants for these non-zero differentails is also computed for the final reconstruction/disaggregation. The sequence of aggregate power value $pp\_x[j]$ at these time instants is also computed for use in the 'segmented' Veterbi decoding procedure.

We now describe a Hidden Markov Model for maximum likelihood separation of the signatures identified by processing the lower dimension input.

V. Hidden Markov Modeling

A Hidden Markov Model (HMM) is appealing for the disaggregation problem, with the state evolving due to appliances switching between ON and OFF states "behind the curtain". For the Markov property, the assumption is that the overall state at discrete time instant T is dependent only on the state at time $T-1$ and the occurrence of an appliance ON/OFF transition at time T, independent of the states at times before $T-1$. While independence between appliances might hold in many cases, appliances usually have typical ON/OFF duration times and the Markovian property does not hold. The HMM and its variants have been employed with success for the disaggregation problem, and this is a subject of a number of recent works [3]–[5]. In this work, a standard HMM is currently employed, with a view to using it as a benchmark on expanding the work to variants that incorporate more information and are scalable (see Section. VIII-B).

The standard HMM formulation in brief [7], [8]: Given a sequence of n states $S = \{s_1, s_2, ..., s_n\}$ and n observations $O = \{o_1, o_2, ..., o_n\}$, the joint probability,

$$\begin{aligned} P(S,O) &= P(s_1, s_2, ..., s_n, o_1, o_2, ..., o_n) \\ &= P(s_1) \cdot \prod_{j=2}^{n} P(s_j/s_{j-1}) \cdot P(o_j/s_j), \end{aligned} \tag{10}$$

where $P(s_j/s_{j-1})$ is the probability of transitioning from state s_{j-1} to state s_j and $P(o_j/s_j)$ is the probability of making observation o_j in state s_j. $P(S,O)$ is maximized to obtain the maximum-likelihood solution for the set of states.

We now establish the states, observations, and associated probabilities for our setting.

A. An HMM for the disaggregation problem

Assume that K distinct appliance signatures were identified as previously described, and given labels $\{1, 2, 3, ..., K-1, K\}$. For these individual appliance signatures i.e. the transition magnitude peaks and range around them, we define a set of K "pure" states, identifying appliance with label l as state 2^{l-1}:

$$\Omega_p = \{2^i : i \in [0, 1, ...K-1]\} \tag{11}$$

A complete set of pure and composite states is obtained as all possible combinations of the K distinct appliances. With the identifier for each state taken as the sum of identifiers of all appliances in the particular combination, we get a set of 2^K states:

$$\Omega = \{0, 1, 2, ..., 2^K - 2, 2^K - 1\} \tag{12}$$

1) State transition probabilities: A procedure for obtaining approximate values for the state transition probabilities is as follows:

1) Each differential in Eqn. 9 is mapped to an appliance a and "pure" state 2^{a-1} based on magnitude,

$$ps(differential) = sgn(differential) \cdot 2^{a-1}, \tag{13}$$

   for appliance a s.t. $abs(differential) \in [threshold\_lo(a), threshold\_hi(a)]$
2) the length of the $diff\_pp\_x$ waveform is traversed in sequence and consecutive positive and negative transitions corresponding to each appliance are matched.
3) 'level' waveforms for each appliance are obtained. If a positive transition for appliance a at time instant i is followed by a negative transition for the same appliance at $i+d$, $level_(a, k) = 2^{a-1} \forall k \in [i, i+d)$
4) the sum of all level waveforms is computed, $sum\_level(k) = \sum_{a \in [1,K]} level(a, k)$
5) the transition counts and probabilities can be computed from the level sum waveform.

Note that this simple procedure is erroneous - otherwise there would be no need for using a HMM. It is assumed that the sequence is long enough that, for the purpose of computing transition probabilities, the error is within an acceptable limit. The estimates of the transition probabilities and threshold parameters for the appliances/states the can be improved using techniques such as the forward-backward and Expectation Maximization (EM) algorithms (currently not employed in this work).

2) Observation probabilities: In our setup, the observation at each time instant is also the differential in our pre-processed waveform,

$$o_j = pp\_x[j] - pp\_x[j-1] \tag{14}$$

With the pure state mapping $ps(o_j)$ obtained in the same manner as described above, the state-conditional observation probabilty,

$$P(o_j/s_j) = P([ps(o_j) + s_j]/s_j) \iff [ps(o_j) + s_j] \in \Omega \tag{15}$$

B. Viterbi algorithm for most likely state sequence

Equation 10 is equivalently expressed as:

$$logP(S,O) = logP(s_1) + \sum_{j=2}^{n} \Big[logP(s_j/s_{j-1}) + logP(o_j/s_j) \Big] \tag{16}$$

For a dynamic programming solution, edge weights are defined for every pair of states $s, s' \in S$ s.t $P(s/s') > 0$ as,

$$d(s,s',j) = logP(s_j = s/s_{j-1} = s') + logP(o_j/s_j = s)$$

the best (or maximum weight) path to state s at time j,

$$D(s,j) = \max_{s' \in S, P(s/s')>0} \big[D(s', j-1) + d(s,s',j) \big] \tag{17}$$

and the best path weight for the full sequence of length n,

$$D^* = \max_{s \in S} D(s,n) \tag{18}$$

and the optimal sequence of states $\{s_j^* : j \in [1,n]\}$ is obtained by starting with the state s_n^* in Eqn. 18 and backtracking on solutions to Eqn.17.

C. Segmented Viterbi sequence decoding

A special feature of the sequence decoding for the current setting is that the 'zero' state i.e the state in which none of the appliances identified are ON, is easily identifiable and with a high degree of certainty. We exploit this for improved accuracy in disaggregation.

When $pp\_x[i] < \min_{a \in [1,K]} threshold\_lo(a)$, we set $s_i^* = 0$. This means that the Viterbi algorithm is effectively applied separately to segments of the original waveform bookended by consecutive visits to the 'zero' state. This prevents the errors from one part of the original sequence from carrying on into others and causing further errors. It needs to be further investigated whether not using this 'zero correcting' at every occurrence and allowing longer sequences is beneficial.

The output of the above procedures is a set of K disaggregated waveforms:

$$disagg\_x[a,i] \forall a \in [1,K], i \in [1,n]$$

A sample of the output and comparisons with the original waveform are illustrated in Figure 2.

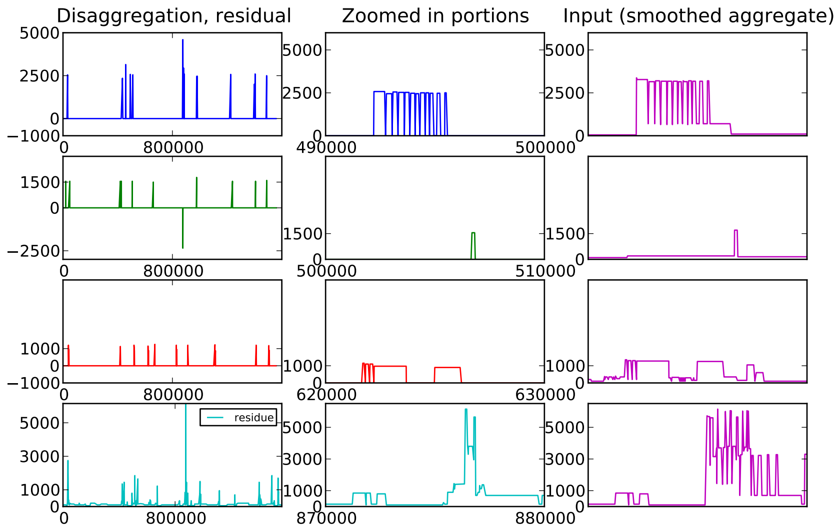


Fig. 2. Disaggregation results for house 1, mains 2. Note that the disaggregated signature with magnitude 2500W corresponds to only a portion of a real appliance signature. The other portion remains in the HMM stage residual.

VI. Residual analysis

The residual after disaggregation is obtained as:

$$disagg\_res\_x[i] = pp\_x[i] - \sum_{a \in [1,K]} disagg\_x[a,i] \tag{19}$$

Using the quantization error computed in Eqn. 6, the unquantized residual is:

$$uq\_disagg\_res\_x[i] = disagg\_res\_x[i] + res\_q\_vc\_x[i] \tag{20}$$

The histogram and CDF of the residuals give clues to:

- whether disaggregation in the HMM step was successful to a satisfactory degree. If significant peaks for the same magnitudes remain, it indicates that the states were not identified correctly or the transition probabilities may be off etc. and the earlier steps may need to be revisited.
- signatures of low power devices that either stay on for long periods of time, or have a magnitude/periodicity pattern that can be disaggregated following the same procedure as for the high power devices, with an appropriate set of parameters for quantization and smoothing.

In studying the residual waveforms ($uq\_disagg\_res\_x$) on the REDD for individual homes/mains, it is observed that they could be very different and the appropriate set of operations for disaggregating residuals needs to be chosen on a case by case basis. In fact, from the overall results, it appears that this might hold for the general approach to the disaggregation problem.

VII. Results

We present results of the method described for a set of house/mains combinations from the REDD (Figures 4- 9, appear after references). The original consumption waveforms exhibit a range of appliance signatures, and the current method is seen to be successful in extracting magnitude signatures. The HMM stage is able to provide disaggregation of between 56% and 67% of the total energy. As seen for house 1, mains 1 (89%), this can be improved further with residual analysis (results not shown for all house/mains combinations, work in progress). Due the limitations in identifying full appliance signatures (as shown by the results themselves), even though ground truth is available from the REDD, the performance evaluation is currently more qualitative than quantitative.

Figure 4 shows the complete results for the disaggregation and residual analysis steps for house 1, mains 1. With the magnitude signatures identified from the histogram shown in Figure 1 (a), the disaggregated waveforms in Figure 4 (a) are seen to comply well, with some errors in the sequence detection also showing up. Some of the obvious errors (by magnitude) can be easily corrected for as well. The residual shows that relatively few portions of the waveform corresponding to the identified signatures are left behind. The zoomed in portion of the residual highlights a situation where the appliance has multiple stages/modes of operation and only a part of its signature could be disaggregated. A more complete solution will have to incorporate such appliance models and perhaps correlate portions of the same signature extracted separately, as discussed in Section. VIII.

Figure 4 (b) shows an ”inverted CDF” like graph for the residual from the HMM procedure, with and without the quantization error being added. It can be seen that (a) more than 95 percentile of the residual has magnitude less than 200W, and a substantial portion of it around 75-80W, and (b) the quantized version would lead to substantial overestimation of the energy consumed by the underlying signature.

Figure. 4 (c) illustrates the pattern hinted at in the flat portions of the CDF for the residual. Using smoothing techniques described earlier with a long window size, we extract the signature for an appliance operating around 75-80W.

Figures. 4 (d) and (e) show the fractions of energy consumption for each of the identified signatures following the HMM and residual analysis steps respectively. While close to 90% of the total consumption could be assigned to the different magnitude signatures, the problem of mapping them to real appliances remains. More than one of these magnitude signatures could potentially belong to the same appliance.

Also, as observed in the HMM residue here, and earlier in Figure. 2, significant portions of appliance signatures that do not fit the 'rectangular' rise and fall signatures either end up in the residue or contribute to errors and artifacts in the sequence detection. They could have leaked into the signature obtained from residual analysis, which needs more attention.

For house 1, mains 2 (Figure. 5), the residual energy is relatively large at 44%. This indicates that the appliance signature seen in the histogram (Figure. 5 (b)) between 500W and 1000W makes a substantial contribution to the energy consumption. Also, as observed in Figure. 2, the disaggregated signature with magnitude 2500W corresponds to only a portion of a real appliance signature. The other portion remains in the HMM stage residual.

For house 3, mains 1 (Figure. 7), it is observed that a signature identified with magnitude 900W contibutes only .4% of the energy. This could be a rarely used appliance, but more likely is a small part of the signature of another appliance. In such cases, it would make sense to go back to the signature identification step and come up with a more plausible set of signatures to run the HMM and sequence detection over. The residual shows portions of appliance signatures that could not be disaggregated.

For house 4, mains 2 (Figure. 8), a larger incidence of errors for signatures of magnitude 1100 and 500 were observed. This probably indicates errors in the range thresholds established in the signature identification step, that might need to be repeated for more accurate results.

VIII. Current limitations and further work

We discuss the limitations of the current method and implemetation, and improvements possible. Some of these were highlighted by the results as described in the previous section, and are put together here to present a clear picture for further progress.

A. *Limitations of appliance signature modeling*

The current method works for single state/mode appliances with short-lived on/off transients. For appliances with multiple modes or long-lived transients or slow ramp-up and ramp-down in power, the following behaviors were observed:

1) If there is a mode with a good match in positive and negative transition magnitudes, this mode is taken to be an appliance signature and disaggregated. The other mode either remains in the residue, or depending on the power level, some portions erroneously "leak" into other appliance waveforms.
2) Different modes of the same appliance having different transition magnitudes, but good match in on/off transitions for each one are disaggregated as separate appliances.

3) If there is no match in sequence of on/off transitions or in case of slow ramp-up and ramp-down of power, the appliance waveform is ignored in the signature identification stage and mostly remains in the residue after the HMM stage, apart from occasional leaks into other waveforms.

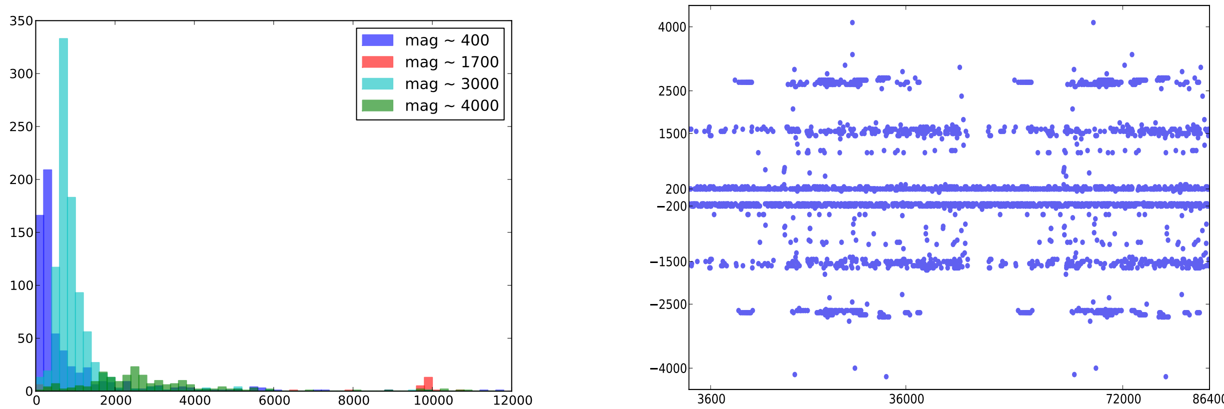
These issues maybe addressed as follows:

1) Since signatures are of different types, a complete solution might be one in which the first step is to characterize various types of signatures and run different disaggregation preocedures on them. Existing/new methods [5], [9] for characterizing appliances that do not fit the rectangular pulse signature description are needed.
2) For different magnitude signatures/modes belonging to same appliance, a consistent proximity in time or an ”always occur together” pattern might need to be checked for.

B. Performance improvements within current model

The performance obtained using the current model can be improved in several ways:

1) Use HMM with state duration modeling [10] with ON/OFF time duration statistics: The level waveforms described in the signature identification step also provide information on ON/OFF time duration statistics for each appliance. A sample of histograms is illustrated in Figure. 3 (a). Along with the magnitude, this is a distinctive feature that can be employed to improve accuracy of the unsupervised learning procedure.
2) Use time of day modeling: Figure. 3 (b) shows that there are distinctive patterns for the appliances based on time of day. The waveform can be segmented into parts of day and the state transition probabilities can be obtained separately for each segment for improved sequence decoding.
3) Use the forward-backward and Expectation Maximization techniques: to improve the state transition probabilities and parameters of the appliance signatures.
4) Use full power of HMM: employ the any path method with multiple reference HMM models [8] to verify and accurately choose between possible sets of signatures/states.
5) Use variants of the HMM for scalability: The standard HMM formulation used has a scalability limitation, since the number of states grows exponentially with the number of appliance signatures. The factorial HMM is a well known alternative that addresses this issue and it has been explored recently for the disaggregation problem as well [4], along with semi-Markov models that include timing and knowledge of appliances. The additive factorial HMM [5] has a formulation to closely match the nature of the disaggregation problem, and explicitly models for multiple modes/stages that each appliance operates in.



(a) (b)

Fig. 3. (a) Histograms of ON duration time for appliances (b) Scatter plot of power transitions by time of day for house 1, mains 1. The time series of differentials (for 18 days) was plotted with a wrap around of 86400 seconds (= 1 day). Clear trends seen for combination of appliances by time of day. The state space and transition probabilities can be adapted accordingly for greater accuracy.

6) Establish a sequence/palette of procedures for residual analysis that can be automated to deal with most cases.
7) Fully automate signature identification and computation of parameter settings.

* *Results appear after references*

REFERENCES

[1] G. W. Hart, "Residential energy monitoring and computerized surveillance via utility power flows," in *IEEE Technology and Society Magazine, Volume 80, Issue 12, pp. 1870-1891*, 1989.

[2] G. W. Hart, "Nonintrusive appliance load monitoring," in *Proceedings of the IEEE, Volume 8, Issue 2, pp. 12-16*, 1992.

[3] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "Using hidden markov models for iterative non-intrusive appliance monitoring," in *Neural Information Processing Systems, Workshop on Machine Learning for Sustainability*, 2011.

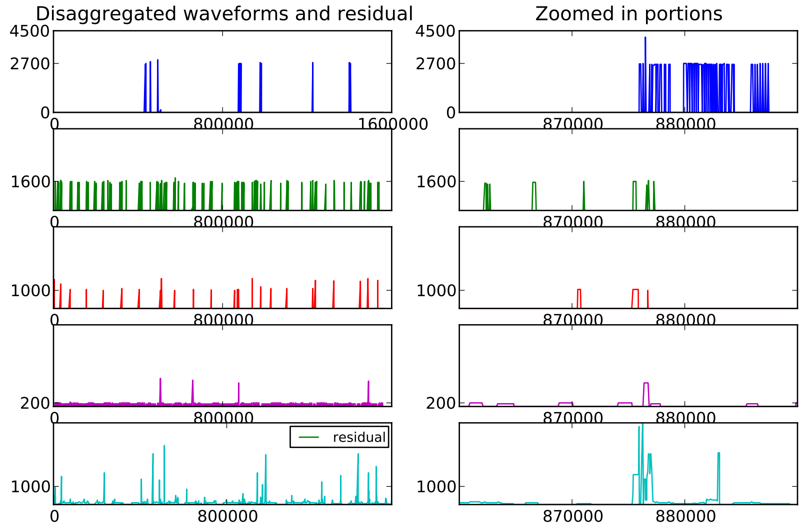
[4] H. Kim, M. Marwah, M. Arlitt, G. Lyon, and J. Han, "Unsupervised disaggregation of low frequency power measurements," in *11th International Conference on Data Mining, pp. 747-758*, 2011.

[5] Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial hmms with application to energy disaggregation.," in *15th International Conference on Artificial Intelligence and Statistics, pp. 1472-1482*, 2012.

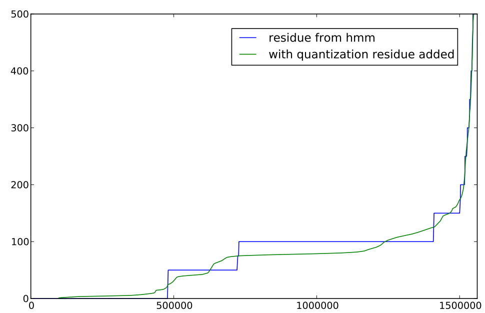
[6] Z. Kolter and M. Johnson, "Redd: A public data set for energy disaggregation research," in *http://redd.csail.mit.edu/ and SustKDD workshop on Data Mining Applications in Sustainability*, 2011.

[7] C. M. Bishop, *Chapter 13, Pattern Recognition and Machine Learning*, Springer, 2006.

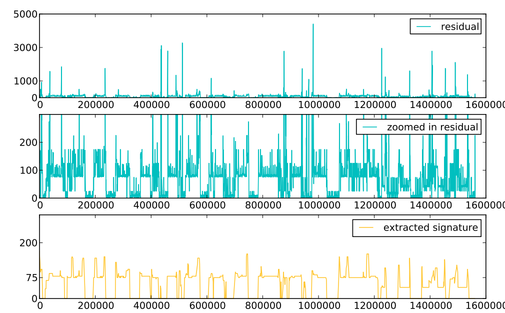
[8] S. Theodiridis and K. Koutroumbas, *Chapter 9, Pattern Recognition*, Academic Press, 2009.



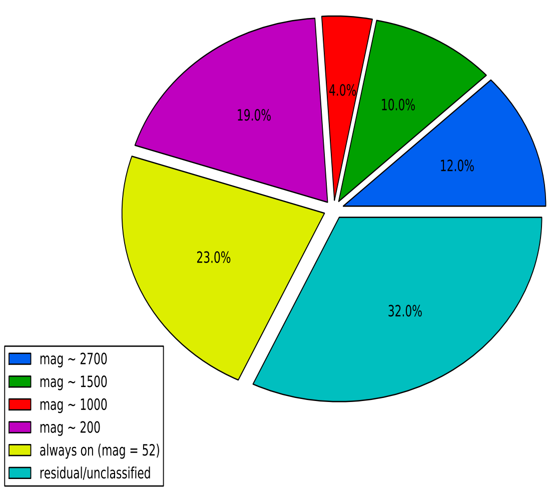
(a)



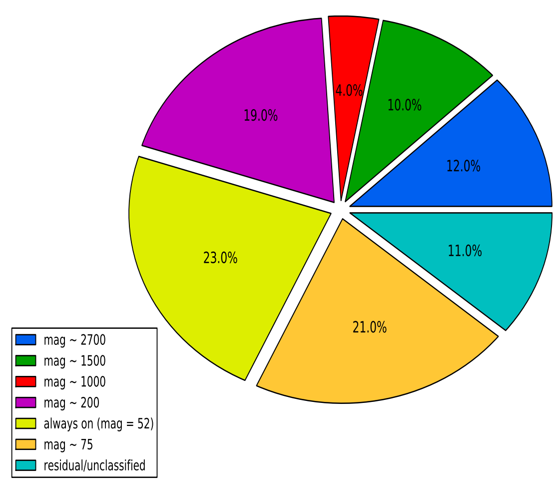
(b)



(c)

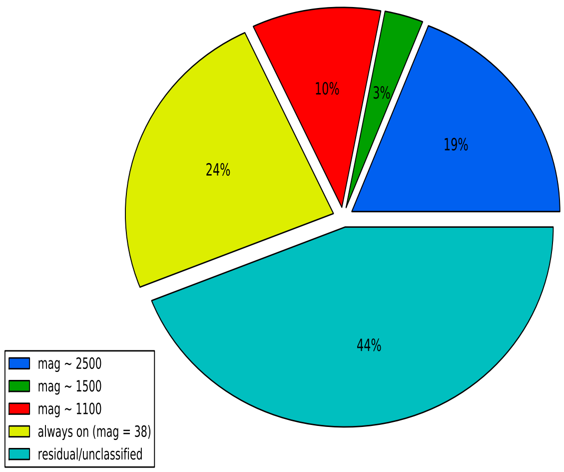


(d)

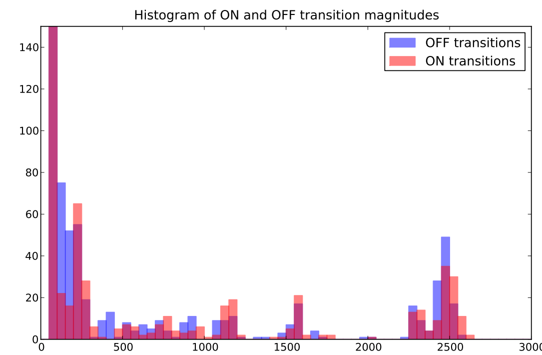


(e)

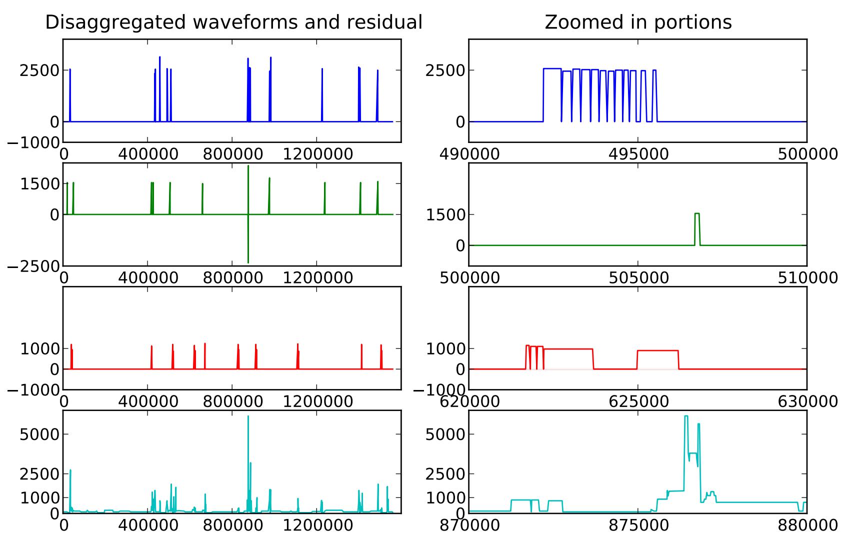
Fig. 4. House 1, mains 1. (a) Disaggregated signatures. (b) “inverted CDF” of residue from HMM step. (c) Residual analysis and signature extraction. Fraction of energy consumption for appliances (d) prior to residual analysis, and (e) post residual analysis



(a)

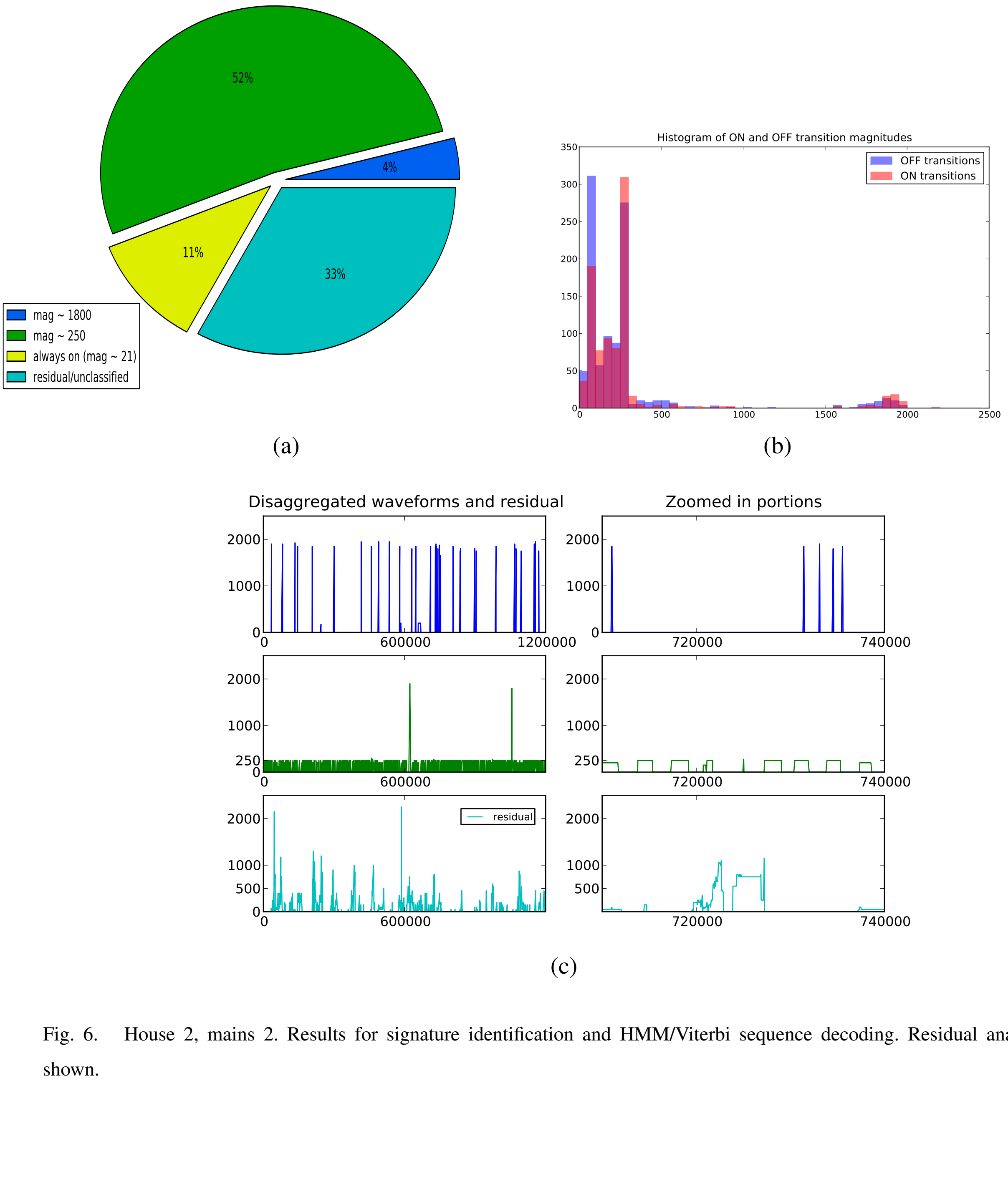


(b)



(c)

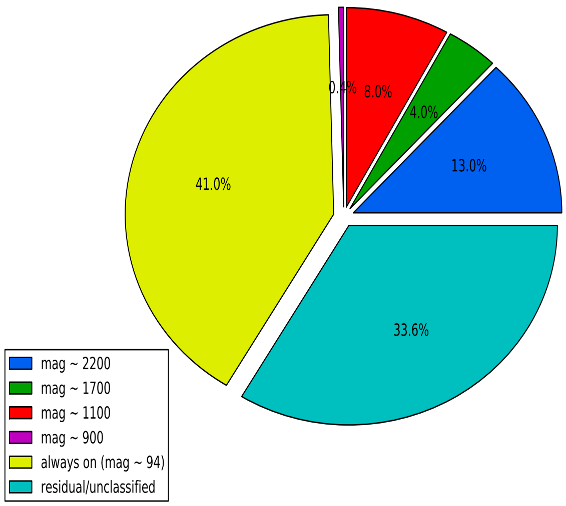
Fig. 5. House 1, mains 2. Residual energy is relatively large at 44%. Appliance signature seen in the histogram (b) between 500W and 1000W could be making a substantial contribution to energy consumption. Also, as observed in Figure 2, the disaggregated signature with magnitude 2500W corresponds to only a portion of a real appliance signature. The other portion contributes to residual. Residual analysis not shown.



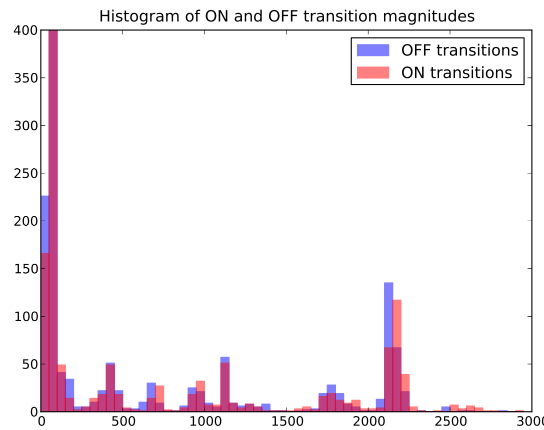
(a) (b)

(c)

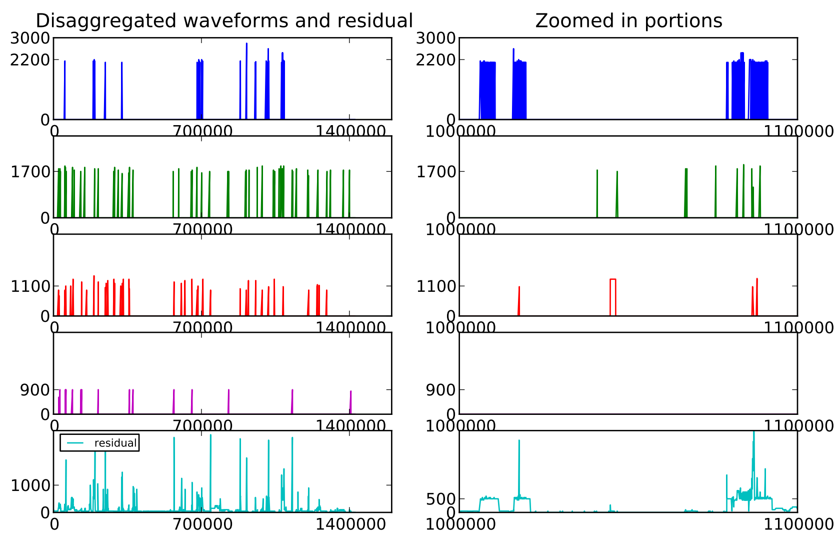
Fig. 6. House 2, mains 2. Results for signature identification and HMM/Viterbi sequence decoding. Residual analysis not shown.



(a)

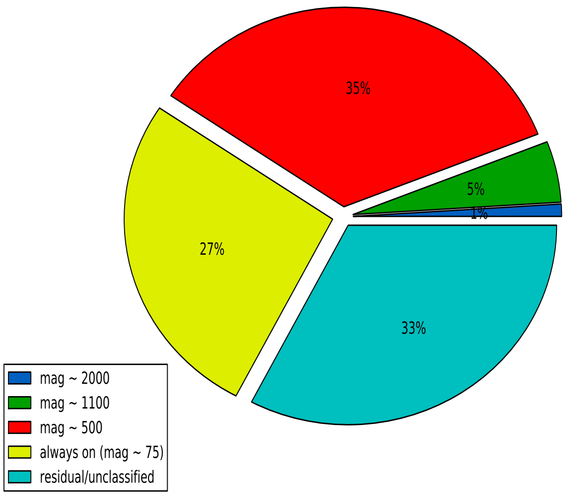


(b)

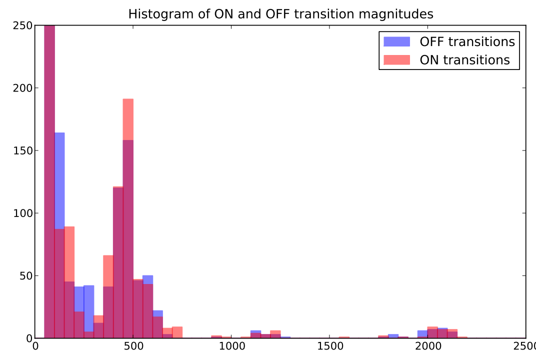


(c)

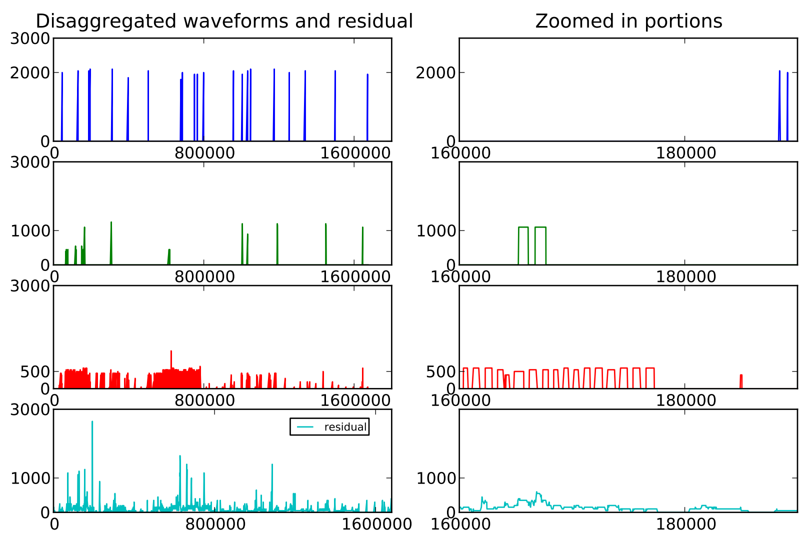
Fig. 7. House 3, mains 1. Signature identified with magnitude 900W contibutes only .4% of the energy. Likely a small part of the signature of another appliance. Return to the signature identification step for a more plausible set of signatures to run the HMM and sequence detection over. Residual shows portions of appliance signatures that could not be disaggregated. Residual analysis not shown.



(a)

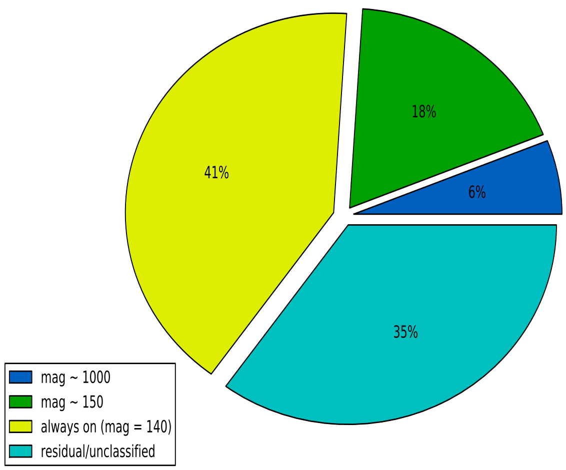


(b)

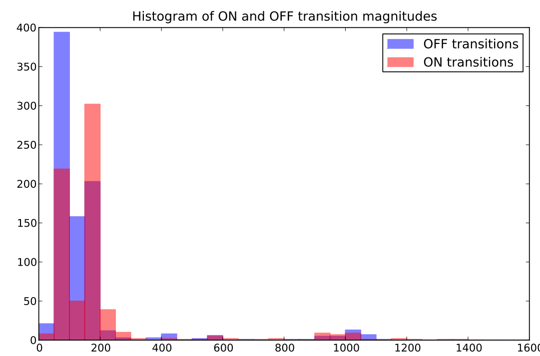


(c)

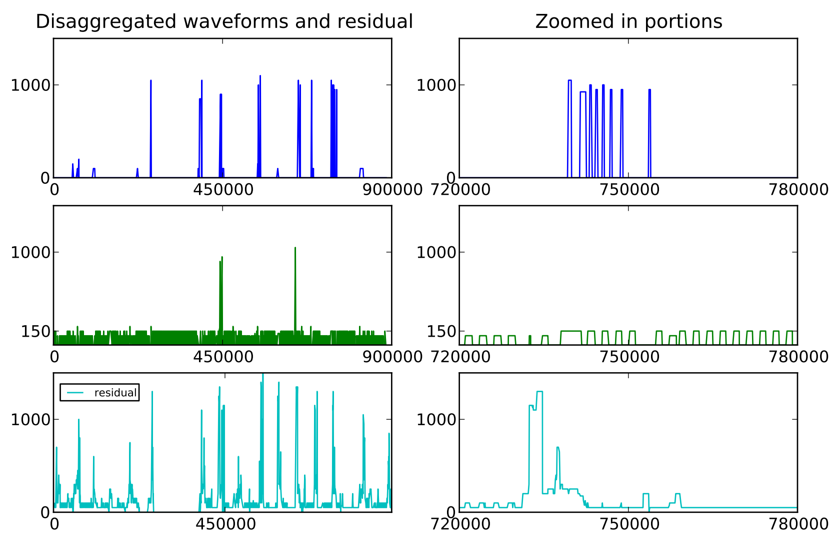
Fig. 8. House 4, mains 2. A larger incidence of errors for signatures of magnitude 1100 and 500. This probably indicates errors in the range thresholds established in the signature identification step, which might need to be repeated for more accurate results. Residual analysis not shown.



(a)



(b)



(c)

Fig. 9. House 6, mains 2. Results for signature identification and HMM/Viterbi sequence decoding. Residual analysis not shown.

[9] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "Non-intrusive load monitoring using prior models of general appliance types," in *26th AAAI Conference on Artificial Intelligence (to appear)*, 2012.

[10] M. Russell and R. Moore, "Explicit modelling of state occupancy in hidden markov models for automatic speech recognition," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp 5-8*, 1985.