Extracting Discriminative Features for Event-based Electricity Disaggregation

Om P. Patri
Department of Computer Science
University of Southern California
Los Angeles, CA 90089
Email: patri@usc.edu

Anand V. Panangadan, Charalampos Chelmis
and Viktor K. Prasanna
Ming-Hsieh Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089
Email: {anandvp,chelmis,prasanna}@usc.edu

Abstract—We describe a novel method for electricity load disaggregation based on the machine learning method of time series shapelets. We frame the electricity disaggregation problem as that of event detection and event classification from time series data. We use existing shapelet-based algorithms to separate appliance activity periods (caused by switching on/off of appliances and denoted as events) from time periods without any such activity. We then identify which type of appliances in a household correspond to the events detected within the power consumption data. Such appliance-level feedback is critical for end-users in managing their energy use efficiently. We use the BLUED dataset for experimental evaluation of the proposed method. This dataset is a fully labeled publicly available dataset of electricity consumption of a household in the United States for one week, the data being recorded at a very high frequency and externally labeled with the times when specific appliances were switched on or off. The proposed approach is able to achieve approximately 98% accuracy for event detection and between 77% to 84% accuracy for event classification. The data segments that were identified as being most discriminative for electricity disaggregation are visually interpretable, and the appliances identified to be responsible for heavy energy consumption can be reported to consumers to encourage reduction in energy consumption.

I. Introduction

Smart grids [1], [2] have enabled household electricity consumption monitoring at very fine time granularity. However, electricity consumption is typically recorded in the form of aggregate numbers. For encouraging energy conservation, it is more valuable to report energy usage at the level of individual appliances to consumers. Energy disaggregation refers to the technology that breaks down aggregate energy consumption into appliance level itemized measurements without any plug level sensors. Non-intrusive load monitoring (NILM) approaches attempt to perform such energy disaggregation, i.e., estimate the electricity consumption of individual appliances from aggregate power and/or voltage measurements [3]. Specifically, the NILM disaggregation problem is that of identifying events associated with individual appliance activity from a stream of measurements, typically active power and voltage, provided at a high temporal rate.

Several techniques for NILM disaggregation have been proposed [4], [5]. One approach to differentiate one load from another and estimate individual appliance usage is to examine the identifying signature that most consumer electronic appliances produce [6], [7]. Such signatures can be measured using special sensors [6] followed by mathematical analysis of current and voltage readings (e.g., spectral signature analysis or waveform-based analysis). The installation of specialized monitoring hardware however incurs hardware development and installation costs as well as associated software licensing fees. Furthermore, identifying electronic appliances signatures is not straightforward due to: (a) the large and diverse number of appliances in use, (b) the signature of an appliance can vary over time due to a variety of modes in which they are used (e.g., various programs of a washing machine), and (c) variations in human appliance usage patterns.

In this work, we propose the use of "shapelets" [8], a machine learning-based method for extracting discriminative subsequences within historic time series data, for electrical appliance use detection and classification. Our goal is not only to automatically detect and classify such events (i.e., identify the appliance that most likely caused the event), but also to extract interpretable features, i.e., signatures, from energy use data. The discriminative subsequences are indicative of the type of event – the specific appliance activity – and can therefore be used for real-time electricity disaggregation. The individual discriminative subsequences enable further analysis of the electrical characteristics of a specific appliance activity, and can be applied to disaggregation of datasets with a different mix of appliances.

United States utilities have recently being deploying millions of smart meters that collect energy consumption data from residential and commercial customers [9]. Despite the investment in infrastructure, energy savings and financial benefits have yet to reach their full potential. Access to appliance level data through disaggregation of smart meters' measurements has numerous benefits both from the consumer's and the utility's perspective. From the user's perspective, several studies suggest that great energy reductions can be achieved as a result of appliance-specific electricity consumption feedback (e.g., which specific appliance could most effectively reduce energy use for a given household) and automated personalized and cost-effective energy saving recommendations (e.g., what type of new appliance to purchase based on current use) [10]. From the utility's perspective, electricity disaggregation can lead to improved program design and diversification, as well as incentives recommendation. Perhaps the most important benefit for utilities is the development of enhanced prediction models for electrical energy consumption. Recent years have seen an increasing interest in developing load forecasting models;

however, to date, there has been little study on the differences in consumption characteristics of individual customers, and their impact on the accuracy of prediction models.

The rest of this paper is organized as follows. We first summarize the most relevant related work which has been undertaken in this area in Section V. We introduce shapelets and discuss the details of our approach for electricity disaggregation in Section II. We present our dataset in Section III. We describe evaluation of our approach and present our experimental results in Section IV. Finally, we discuss the findings of our work in Section VI and draw our conclusions in Section VII.

II. SHAPELETS FOR ELECTRICITY DISAGGREGATION

A. Background on Shapelets

Time series shapelets [8] are discriminative subsequences within a labeled time series collection, which differentiate instances of the positive class from those of the negative class. They are identified based on an information gain criteria — the subsequence with the highest information gain is identified to be the most discriminative between the classes and thus is the shapelet. A decision tree classifier is built based on Euclidean distance from the shapelet to the given sequence. Classification can then be performed using the decision tree. When a new time series instance is provided to the classifier, the Euclidean distance to the previously extracted shapelets is calculated, and depending on which branch of the tree it falls in (distance lesser or greater than distance threshold of the shapelet), it is assigned a class label. For the case of multiple shapelets from one time series, the same decision tree can be extended to have multiple nodes. Various extensions to time series shapelets have been proposed in the literature, the state-of-the-art being the Fast Shapelets [11] algorithm which we use in this work. Other extensions include a unsupervised method for finding shapelets [12] and a method for finding 'local' shapelets [13], which might not be optimal but are more useful for early classification while streaming through the time series data.

The shapelet method has specific advantages relevant to the problem of electricity disaggregation. Shapelet-based event classification is computationally fast since it does not require the full training data for classification, only the extracted shapelet subsequence(s) are needed. The method does not makes assumptions on the nature of the data (which suits real world data such as energy use) unlike conventional time series analysis approaches such as autoregressive methods or Autoregressive integrated moving average (ARIMA) models. They are also visually interpretable which is key to understanding the patterns of energy use.

B. Proposed Approach

In this work, we present a two-stage approach to electricity disaggregation. In the first stage, the stream of electricity load measurements is analyzed to detect the occurrence of an *event*, which corresponds to a change in some load-drawing appliance's state. In the second stage, any detected events are further analyzed to identify the specific class of appliances that most likely created the event. These stages correspond to the *Event detection* and *classification* steps in a typical event-based

NILM stage, as described by Anderson et al. [4]. Both stages are implemented using the machine learning-based shapelets approach, i.e., they require a training phase using labeled data.

For the first stage, fixed length subsequences of load measurements that are known to be free of any events along with subsequences that include a load change event are used as training data to create a Fast Shapelets classifier. After training is complete, a new subsequence can be presented to the classifier and it's output indicates whether this new subsequence contains an event or not.

For the second stage, training data contains fixed length subsequences of load measurements that are known to contain a load-change event of a specific class of appliances. Since there can be many such classes, we use the multiclass classification capability of the Fast Shapelets algorithm. Given a segment of load measurements, the resulting shapelets classifier's output indicates which of the appliance classes (seen during the training phase) is most likely to contribute to an event in the sequence.

Together, the two shapelets classifiers determine if a segment of load measurements contains an appliance-level event, and if so, which appliance class caused that event. The training of the two shapelet classifiers is performed independently for simplicity and speed of training. We also evaluate the performance of the classifiers separately. Note that the shapelets classifiers expect the input to be fixed length sequences. Thus, the continuous load measurements have to be segmented before they can be input to these classifiers (the specific preprocessing steps are described Section IV-A). These steps for electricity load disaggregation are illustrated in Figure 1.

III. DATASET

We use the Building-Level fUlly-labeled dataset for Electricity Disaggregation (BLUED) dataset [14] which contains voltage, current and power measurements for a single family in the United States for one week. Every state transition for each appliance in the house is labeled and time-stamped, thus providing the ground truth for the energy disaggregation task. Every single switching on/off of any appliance is recorded and called as an 'event'. The data available is downsampled to 60 Hz (from the collection sampling rate of 12 kHz). Real and reactive power, for both phase A and phase B are available in the dataset. We use the real power values from both phase A and B in our experiments. This dataset has been previously used for event detection using an approach based on Hidden Markov Models [4].

A snippet of the data containing the first quarter of the real power data (44 hours from the week-long total) for phase A is shown in Figure 2. This portion contains more than 9 million measurements and over 200 events representing multiple appliance activity. Some extreme values have been eliminated to maintain scale of the figure. The events occurring during this period of time are marked by red asterisks. Each event denotes appliance activity at the specified timestamp (switching on/off). From the figure, we can observe spikes in power use, which represent the event that an appliance is switched on/off. Our task is twofold: (i) to identify the event time locations from the non-event ones, and (ii) to trace back which appliance is responsible for a newly observed event.
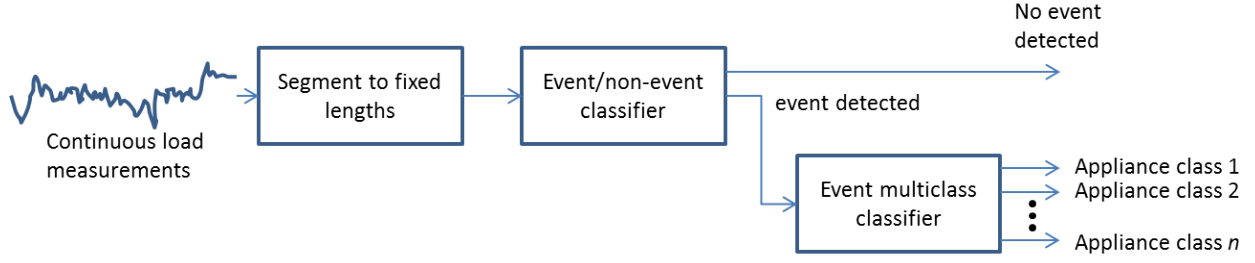
Fig. 1.   Block diagram of the proposed two-stage classifier-based electricity load disaggregation approach.
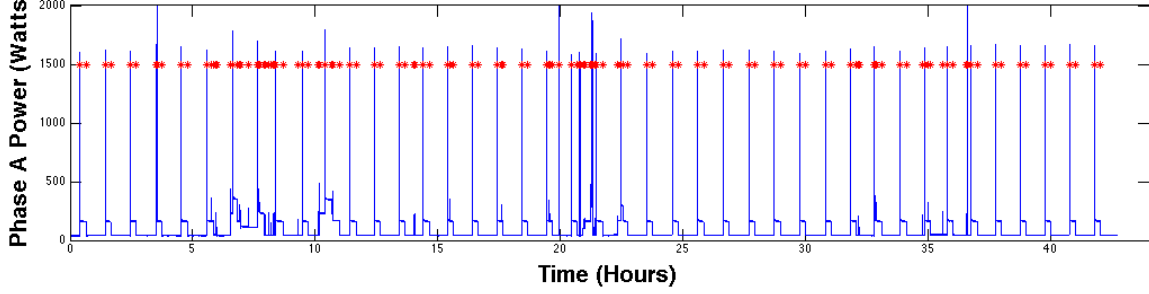

Fig. 2.   Power consumption in phase A over time (first quarter of the dataset). Red asterisks indicate location of appliance activity as reported in the BLUED labeled dataset [14].

IV.   EVALUATION

In this section, we describe our complete evaluation process, including pre-processing, parameter selection, experimental tasks performed and also provide results of evaluation. The full BLUED dataset is about one week long, consisting of nearly 37 million data points. In all our experiments, we use the first half of this period for training our classifier, and report results on the second half, thus ensuring a 50%-50% train-test split. The testing instances were independent from the training instances as ensured by this split. Also, as the splits are in consecutive time intervals, this provides us with intuitive results about disaggregation of observed historical data to predict future power use behavior.

A.   Pre-processing

The machine learning-based approach that we take for electricity disaggregation requires training data segments, each labeled with event information, for extracting the shapelets. Since the dataset is large and not available in the form of labeled time series instances, we need to process it to into a labeled format suitable for machine learning. We consider a window around each event, such that multiple time series instances can be extracted from the dataset for training. The parameters for this window can be varied. As the BLUED data suggests that each event lasts for at least 5 seconds, we define the window around event to include all data points within 1 second before the event and upto 5 seconds after the event. Since each second of data contains 60 values, this fixes the length of all time series segments in our training data at 360. This parameter can be varied to get a different set of events based on the requirements.

B.   Detection and Classification Tasks

We evaluate the performance of the Fast Shapelets [11] algorithm on detecting events and classifying appliance activity from aggregate electrical power data. The first task is to *detect* events, i.e. differentiate event data segments from the non-event segments. To achieve this, we extract all event instances, according to the pre-processing steps described earlier. Then, we extract several non-event instances of the same length from the power consumption data (ensuring that no events overlap at all with any of the non-event segments). We perform this sampling of non-event instances in a balanced manner (1:1 ratio of events:non-events) as well as an imbalanced manner (1:4 ration of events:non-events). This experiment is performed for the power data in both phase A and phase B.

The second task is to perform actual disaggregation, i.e. to *classify* which appliance is actually responsible for which event. This is achieved by a multiclass shapelet-based classifier. We divide the appliances into groups, based on the nature of the appliances and their average power consumption. This gives us multiple classes of appliances for training and classification. The classes we use, along with instances of appliances of those classes are mentioned in Table I.

The refrigerator in phase A has the most number of events out of all appliances. It also exhibits special behavior in that it switches on and off automatically and frequently, so we let it have a class by itself. All the lights are grouped together into a single class for both phase A and phase B. For the rest of the appliances, we divide the classes based on the average power consumption of each appliance mentioned in the BLUED dataset [14]. Appliances which have less than 150 watts average power consumption are classified into the low-power use class while appliances which have more than 150

234

watts average power use are classified into the high-power use class. For phase A data, we observe that there are actually no appliances in the 150-800 watt range, thus, all the high-power use appliances indeed use very high power. Appliances which had zero events associated with them were not used in our experiments. Events which were labeled as unknown (origin appliance not labeled) were also not used. Also, the microwave and washing machine appliances could not be used in our experiments since the BLUED dataset does not identify labels for events arising from these appliances. All other appliances were grouped into classes as described in the table.

TABLE I. CATEGORIZATION OF APPLIANCES INTO CLASSES

Phase A appliance classes	
Class	Examples
1. Refrigerator	refrigerator
2. Lights	backyard lights, washroom light, bedroom light
3. High-Power ($> 150W$)	hair dryer, air compressor, kitchen chopper
Phase B appliance classes	
Class	Examples
1. Lights	desktop lamp, basement light, closet lights
2. High-Power ($> 150W$)	printer, iron, garage door
3. Low-Power ($< 150W$)	computer, LCD monitor, DVR/blu-ray player

C. Detecting events and non-events

We present the results of using the Fast Shapelets algorithm for classifying event data segments from non-event data segments here. The class label 1 denotes event instances and the class label -1 denotes non-event instances. We report results on both the balanced and imbalanced sampling of the non-events. We use the shapelet-based decision tree learned from the balanced dataset to classify both the balanced and the imbalanced datasets. We observe that this performs better than using the tree learned from the imbalanced dataset to classify itself.

The results for phase A are presented in Table II and for phase B in Table III. Consider the following baseline to compare against: assign a test instance the class label of the majority class in the training data. The accuracy with this baseline for the balanced data is 50% and for the 1:4 imbalanced data is 75%. Our proposed approach performs significantly better than this baseline in all cases.

From the confusion matrices in Table II and Table III we can compute the precision and recall of detecting appliance events from the time series data. (Precision is defined as the fraction of all positives labeled by the classifier that are true positives; recall is defined as the fraction of all positives in the dataset that are identified as true positives by the classifier.) The precision is approximately 98% for detecting events in phase A and phase B and is unchanged when the ratio of event to non-event instances is varied. However, the recall varies with the proportion of training data. It is approximately 99% for phase A events (98% for phase B events) when the ratio of event to non-event instances is $1 : 1$ but reduces approximately 95% for phase A events (92% for phase B events) when the ratio of event to non-event instances is $1 : 4$.

One shapelet was obtained in the experiment on phase A and two shapelets on the experiment on phase B while training. These shapelets are shown in Figure 3 (phase A) and Figure 4 (phase B).

TABLE II. RESULTS FOR EVENT DETECTION ON PHASE A APPLIANCES. BALANCE DENOTES THE RATIO OF EVENT TO NON-EVENT INSTANCES.

Balance	Accuracy	test instances	Confusion Matrix			
			real		event	non-event
1:1	**98.42%**	886	pred	event	432	3
			pred	non-event	11	440
			real		event	non-event
1:4	**98.56%**	2215	pred	event	432	21
			pred	non-event	11	1751

TABLE III. RESULTS FOR EVENT DETECTION ON PHASE B APPLIANCES. BALANCE DENOTES THE RATIO OF EVENT TO NON-EVENT INSTANCES.

Balance	Accuracy	test instances	Confusion Matrix			
			real		event	non-event
1:1	**98.32%**	1072	pred	event	527	9
			pred	non-event	9	527
			real		event	non-event
1:4	**97.91%**	2680	pred	event	527	47
			pred	non-event	9	2097

D. Identifying individual appliances associated with detected events

We now use a multiclass classifier (again, Fast Shapelets) to identify the class of appliances given an unknown event data segment. The division of appliances from phase A and phase B into classes is delineated in Table I. For both phase A and phase B, we have 3 classes. We use all the events from BLUED in this experiment (the train-test split is 50%-50% as described earlier). We report the accuracy of classification of appliance classes from power data on phase A and phase B. We use the same baseline reported earlier: assign to a new test instance the class label of the majority class in the training dataset. The results, including confusion matrices, are presented in Table IV for phase A and Table V for phase B.

TABLE IV. RESULTS FOR CLASSIFICATION OF APPLIANCES FROM AGGREGATED POWER CONSUMPTION DATA ON PHASE A. THE CLASS LABELS ARE DESCRIBED IN TABLE I. THE BASELINE ACCURACY IS 77.25%.

Accuracy	test instances	Confusion Matrix				
		real	class label	1	2	3
83.75%	400	pred	1	295	36	9
		pred	2	9	26	5
		pred	3	5	1	14

TABLE V. RESULTS FOR CLASSIFICATION OF APPLIANCES FROM AGGREGATED POWER CONSUMPTION DATA ON PHASE B. THE CLASS LABELS ARE DESCRIBED IN TABLE I. THE BASELINE ACCURACY IS 41.32%.

Accuracy	test instances	Confusion Matrix				
		real	class label	1	2	3
77.92%	317	pred	1	95	6	12
		pred	2	5	50	0
		pred	3	31	16	102

The confusion matrices in Table IV and Table V can be converted to precision and recall metrics for each appliance class. These are reported in Table VI for phase A and Table VII for phase B.

We obtained 7 shapelets from the appliance identification experiments on phase A and 14 shapelets from the experiment on phase B. As expected, the number of shapelets is higher than the previous event detection experiment, since there are 3 appliance classes to be distinguished from the data as
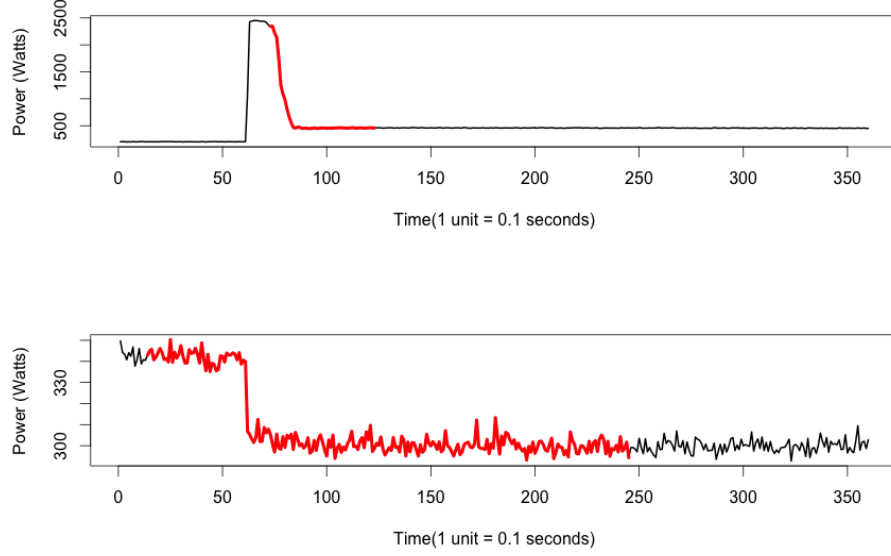
Fig. 4. Two shapelets found from the event detection experiment on phase B power data. The duration of the shapelets is marked in red.
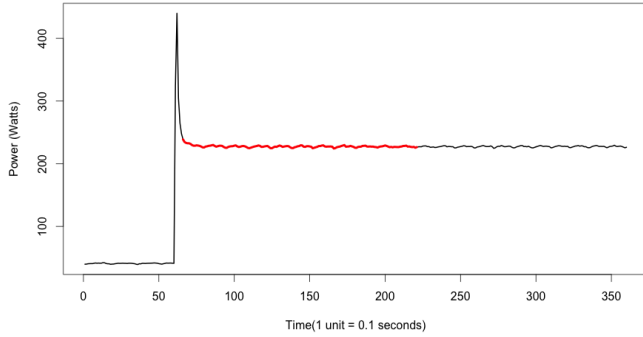
TABLE VII. PRECISION AND RECALL FOR CLASSIFICATION OF APPLIANCE TYPES ON PHASE B FROM AGGREGATED POWER CONSUMPTION DATA. THE CLASS LABELS ARE DESCRIBED IN TABLE I.

Class	Precision (%)	Recall (%)
1	72.5	84.1
2	69.4	90.9
3	89.5	68.5

Fig. 3. The only shapelet found from the event detection experiment on phase A power data. The duration of the shapelet is marked in red.

TABLE VI. PRECISION AND RECALL FOR CLASSIFICATION OF APPLIANCE TYPES ON PHASE A FROM AGGREGATED POWER CONSUMPTION DATA. THE CLASS LABELS ARE DESCRIBED IN TABLE I.

Class	Precision (%)	Recall (%)
1	95.5	86.8
2	41.3	65.0
3	50.0	70.0

compared to the event/non-event classification. Figures 5 and 6 show selected shapelets computed in the experiments from phase A and phase B load data respectively.

V. RELATED WORK

Zeifman et. el [3] and Zoha et. al. [15] survey several NILM approaches for energy disaggregation. Though several techniques for NILM disaggregation have been proposed, we frame the problem in the context of time series classification.

Anderson et. al. [4] survey different approaches for event detection, broadly classified as based on expert heuristics (such as [16]), probabilistic models (such as [17]) or matched filters (such as [18]). However, none of methods take a time series-based approach. Shao et al. [5] mine for temporal motifs from energy consumption time series; however, they do not work with labeled data or extract discriminative features. Motifs are frequent patterns but not discriminative like shapelets, so it may not be possible to differentiate between instances of different classes, which is what we aim to do. A related approach for energy time series forecasting based on pattern similarity is proposed in [19] but it does not deal with disaggregation.

ElectriSense [6] uses electromagnetic interference (EMI) signals during appliance operation to identify and classify individual appliance use. In this work, we have used the concept of *events* as the period of switching on/off an appliance as described in Section IV-A. Instead, we can add rich structure or semantics to these events, as described in the Active Events Model [20] or the Process-oriented Event Model [21] for complex event processing [22]. Initial efforts towards realizing such semantic event models in middleware for a smart grid scenario, specifically via the use of enterprise integration patterns [23], are provided in [24].

VI. DISCUSSION

From both the tables, we can observe that we achieve very high accuracy (close to 100%) for detecting events from
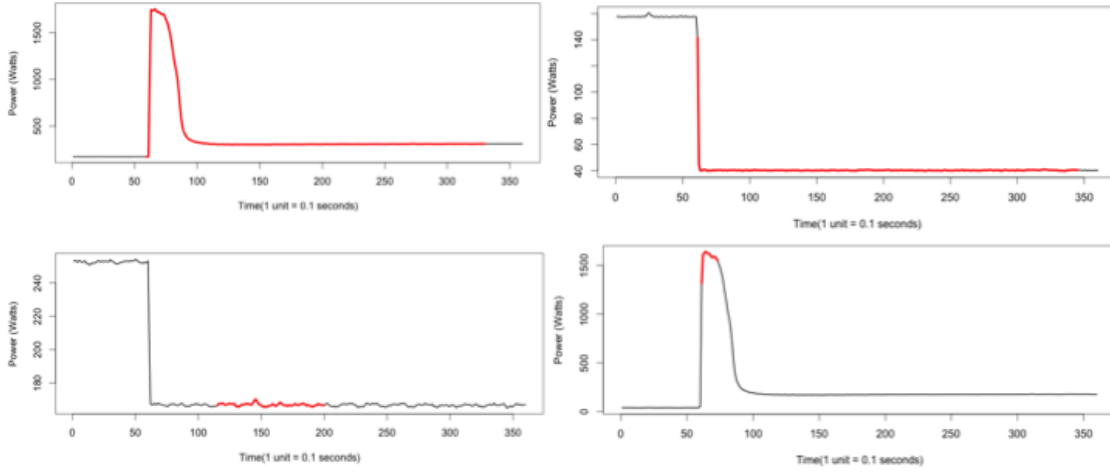
Fig. 5. Selected shapelets computed from the appliance identification experiment on phase A power data. The shapelets are the segments of the original time series marked in red.
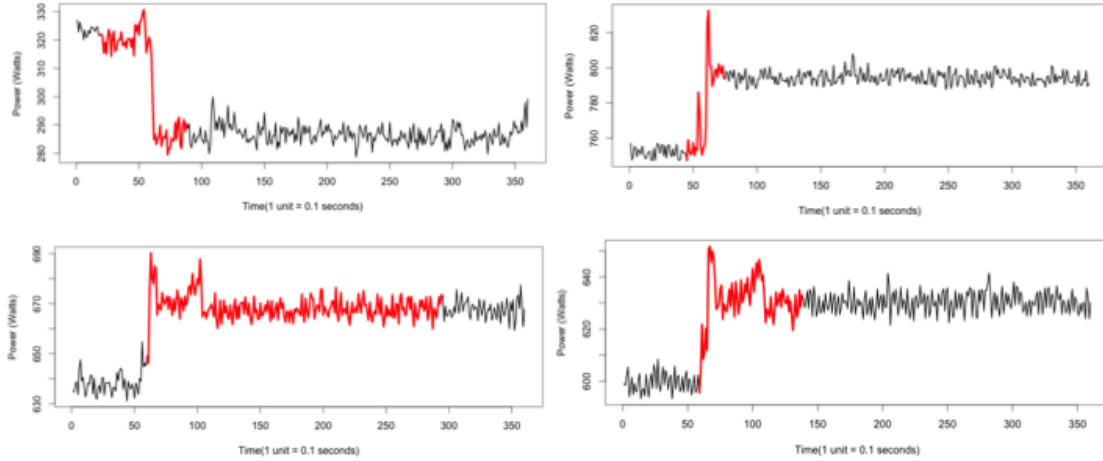

Fig. 6. Selected shapelets computed from the appliance identification experiment on phase B power data. The shapelets are the segments of the original time series marked in red.

non-events. This is expected as there is a significant change in power consumption when an appliance is switched on or off. The shapelet algorithm identifies the specific pattern of this change in load (Figure 3, phase A) and uses it to discriminate against non-event instances, which do not exhibit this pattern. Figure 4 shows that both the shapelets from phase B are discriminative patterns which span the spike occurring due to activity of the appliance (switching on/off). Note that the shapelet from phase A did not capture the spike resulting from switching while it was happening; it covers the duration just after this activity ends. Despite this, it provides a high classification accuracy. This experiment justifies our pre-processing method of considering the event segment as 1 second before the switching and 5 seconds after the switching. If we had instead considered only the instantaneous change during the switching, this shapelet would not have been found. This experiment also shows that the shapelets method is robust to noise or scale changes in the electricity load data and can automatically determine the most discriminative patterns.

For the appliance class identification task, the accuracy

achieved by our approach is also significantly better than that of the baselines for events in both phase A and phase B (Tables IV and V). For phase A events, the baseline accuracy is 77.25% while we achieve 83.75%. For phase B events, the baseline accuracy is only 41.32% while we achieve 77.92%. However, the classification precision and recall vary with the appliance class (Tables VI and VII). For phase A events, the precision of identifying Class 1 (refrigerator) appliances is much higher than that of identifying Class 2 (lights) and Class 3 (high power appliances $> 150W$) appliances. For phase B events, the precision of identifying Class 3 (low power appliances $< 150W$) appliances is higher than that of identifying Class 1 (lights) and Class 2 (high power appliances, $> 150W$) appliances but its recall is lower than that of both Class 1 and Class 2 appliances. This illustrates the trade-off between precision and recall in a classifier and indicates that disaggregation methods need to be fine-tuned to achieve the desired balance between precision and recall for specific appliance types.

From the shapelets shown in Figure 5 and Figure 6,

237

we observe that the shapelets computed from the appliance identification experiments capture a wide range of patterns from the load data. These patterns also vary widely in length from each other. The appliance identification task is harder than the event detection task since there are multiple classes to be distinguished as compared to the two (event vs. non-event) in the event identification task. Even in this case, the shapelet-based classification method is able to achieve high accuracy and provide visually interpretable results.

VII. CONCLUSIONS

We presented the use of time series shapelets for energy disaggregation. Using the publicly available BLUED dataset, we showed that this method is able to achieve high accuracy in two key tasks in energy disaggregation: (i) identifying events from non-events, and (ii) identifying which appliance is associated with a specified event. We were able to do large-scale classification fast as the shapelet-based method requires only the discovered discriminative data segments to be stored in memory for classification (instead of the full training data). Furthermore, the method makes no assumptions on the nature of the real-world data. Through the use of shapelets, we were able to identify the exact temporal subsequences which are most discriminative and possess the representative characteristics of the appliance classes. The method is able to discriminate appliance power on/off events from the load data with high accuracy. The accuracy of identifying the specific appliance class that were operated varies with the appliance types comprising a class. These results can be further analyzed and interpreted to discover which appliances are responsible for critical events or durations of power use. Such appliance-level feedback can be sent to the consumers so that they can reduce their usage accordingly, and manage their energy use efficiently. In future work, we will look at more fine-grained events, such as differentiating between the events of switching on and switching off of appliances.

ACKNOWLEDGMENTS

REFERENCES

[1] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart gridthe new and improved power grid: a survey," *Communications Surveys & Tutorials, IEEE*, vol. 14, no. 4, pp. 944–980, 2012.

[2] S. Aman, Y. Simmhan, and V. K. Prasanna, "Energy management systems: state of the art and emerging trends," *Communications Magazine, IEEE*, vol. 51, no. 1, pp. 114–119, 2013.

[3] 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.

[4] K. D. Anderson, M. Berges, A. Ocneanu, D. Benitez, and J. M. Moura, "Event detection for non intrusive load monitoring," in *IECON 2012-38th Annual Conference on IEEE Industrial Electronics Society*. IEEE, 2012, pp. 3312–3317.

[5] H. Shao, M. Marwah, and N. Ramakrishnan, "A temporal motif mining approach to unsupervised energy disaggregation," in *Proceedings of the 1st International Workshop on Non-Intrusive Load Monitoring, Pittsburgh, PA, USA*, vol. 7, 2012.

[6] S. Gupta, M. S. Reynolds, and S. N. Patel, "Electrisense: Single-point sensing using emi for electrical event detection and classification in the home," in *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, ser. Ubicomp '10. New York, NY, USA: ACM, 2010, pp. 139–148. [Online]. Available: http://doi.acm.org/10.1145/1864349.1864375

[7] J. Froehlich, E. Larson, S. Gupta, G. Cohn, M. Reynolds, and S. Patel, "Disaggregated end-use energy sensing for the smart grid," *IEEE Pervasive Computing*, vol. 10, no. 1, pp. 28–39, Jan. 2011. [Online]. Available: http://dx.doi.org/10.1109/MPRV.2010.74

[8] L. Ye and E. Keogh, "Time series shapelets: a new primitive for data mining," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009, pp. 947–956.

[9] J. Shishido, "Smart meter data quality insights," in *ACEEE Summer Study on Energy Efficiency in Buildings*, 2012.

[10] K. Carrie Armel, A. Gupta, G. Shrimali, and A. Albert, "Is disaggregation the holy grail of energy efficiency? the case of electricity," *Energy Policy*, vol. 52, no. C, pp. 213–234, 2013.

[11] T. Rakthanmanon and E. Keogh, "Fast shapelets: A scalable algorithm for discovering time series shapelets," in *Proceedings of the thirteenth SIAM conference on data mining (SDM)*, 2013.

[12] J. Zakaria, A. Mueen, and E. Keogh, "Clustering time series using unsupervised-shapelets," in *Data Mining (ICDM), 2012 IEEE 12th International Conference on*. IEEE, 2012, pp. 785–794.

[13] Z. Xing, J. Pei, S. Y. Philip, and K. Wang, "Extracting interpretable features for early classification on time series." in *SDM*, 2011, pp. 247–258.

[14] 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, Beijing, China*, 2012, pp. 12–16.

[15] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey," *Sensors*, vol. 12, no. 12, pp. 16 838–16 866, 2012.

[16] L. Farinaccio and R. Zmeureanu, "Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses," *Energy and Buildings*, vol. 30, no. 3, pp. 245–259, 1999.

[17] M. Berges, E. Goldman, H. S. Matthews, L. Soibelman, and K. Anderson, "User-centered nonintrusive electricity load monitoring for residential buildings," *Journal of Computing in Civil Engineering*, vol. 25, no. 6, pp. 471–480, 2011.

[18] S. R. Shaw, S. B. Leeb, L. K. Norford, and R. W. Cox, "Nonintrusive load monitoring and diagnostics in power systems," *Instrumentation and Measurement, IEEE Transactions on*, vol. 57, no. 7, pp. 1445–1454, 2008.

[19] F. Martinez Alvarez, A. Troncoso, J. C. Riquelme, and J. S. Aguilar Ruiz, "Energy time series forecasting based on pattern sequence similarity," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 23, no. 8, pp. 1230–1243, 2011.

[20] P. N. Gross, S. Gupta, G. E. Kaiser, G. S. Kc, and J. J. Parekh, "An active events model for systems monitoring," 2001.

[21] O. P. Patri, V. S. Sorathia, A. V. Panangadan, and V. K. Prasanna, "The process-oriented event model (poem): A conceptual model for industrial events," in *Proceedings of the 8th ACM International Conference on Distributed Event-Based Systems*, ser. DEBS '14. New York, NY, USA: ACM, 2014, pp. 154–165.

[22] G. Cugola and A. Margara, "Processing flows of information: From data stream to complex event processing," *ACM Comput. Surv.*, vol. 44, no. 3, pp. 15:1–15:62, Jun. 2012.

[23] G. Hohpe and B. Woolf, *Enterprise integration patterns: Designing, building, and deploying messaging solutions*. Addison-Wesley Professional, 2004.

[24] O. P. Patri, A. V. Panangadan, V. S. Sorathia, and V. K. Prasanna, "Semantic management of enterprise integration patterns: A use case in smart grids," in *Data Engineering Workshops (ICDEW), 2014 IEEE 30th International Conference on*. IEEE, 2014, pp. 50–55.