High-Frequency, Multiclass Nonintrusive Load Monitoring for Grid-Interactive Residential Buildings

Preprint

Blake Lundstrom,[1] Govind Saraswat,[2] and Murti V. Salapaka[2]

[1] *National Renewable Energy Laboratory*
[2] *University of Minnesota, Twin Cities*

Presented at the 2020 IEEE Conference on Innovative Smart Grid Technologies (IEEE ISGT)
Washington, D.C.
February 17–20, 2020

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5D00-74701
November 2019

High-Frequency, Multiclass Nonintrusive Load Monitoring for Grid-Interactive Residential Buildings

Preprint

Blake Lundstrom,[1] Govind Saraswat,[2] and Murti V. Salapaka[2]

[1] *National Renewable Energy Laboratory*
[2] *University of Minnesota, Twin Cities*

Suggested Citation
Lundstrom, Blake, Govind Saraswat, and Murti V. Salapaka. 2019. *High-Frequency, Multiclass Nonintrusive Load Monitoring for Grid-Interactive Residential Buildings: Preprint*. Golden, CO: National Renewable Energy Laboratory. NREL/CP-5D00-74701. https://www.nrel.gov/docs/fy20osti/74701.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5D00-74701
November 2019

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308 Funding provided by Advanced Research Projects Agency-Energy (ARPA-E) under Grant Nos. DE-AR000701 and DE-AR0001016. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

High-Frequency, Multiclass Nonintrusive Load Monitoring for Grid-Interactive Residential Buildings

Blake Lundstrom[1,2], Govind Saraswat[2], and Murti V. Salapaka[2]
[1]National Renewable Energy Laboratory, Golden, CO, USA
[2]Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA

***Abstract*—Smart buildings with net-load metering and control capabilities can provide valuable flexibility to grid operators. This article develops a novel approach for high-frequency, multiclass nonintrusive load monitoring (NILM) that enables effective net-load monitoring capabilities with minimal additional equipment and cost. Relative to existing NILM work, the proposed solution operates at a faster timescale, providing accurate multiclass state predictions for each 60-Hz ac cycle without relying on event-detection techniques. The approach is validated using a test bed with residential appliances and shown to have high accuracy, good generalization properties, and sufficient response time to support building grid-interactive control at fast timescales relevant to the provision of grid frequency support services.**

***Keywords*—Nonintrusive load monitoring (NILM), multiclass classification, feature extraction, smart buildings, grid-interactive.**

I. Introduction

Modernization of the electric grid with increasing deployments of renewable and distributed energy resources (DERs), decreasing levels of traditional synchronous generation, and evolving load types and capabilities is presenting new challenges and opportunities. These changes, combined with emerging power system operating platforms and concepts, are making behind-the-meter net-load (i.e., DERs and flexible loads) an increasingly important component of system operations. Smart buildings that can manage their net loads to provide additional flexibility to grid operators while maximizing customer preferences and energy cost savings offer a valuable tool for a modernized electric grid. Residential buildings use 38.5% of the total electrical energy produced in the United States [1]; thus, if residential buildings can be converted into smart buildings, significant flexibility to support the grid will be unlocked.

It is crucial to enable circuit-level net-load metering at these residential buildings to gain a better understanding of building performance and customer usage patterns and preferences while enabling the critical monitoring of information needed to perform meaningful grid-interactive control. The simplest approach for such metering is having a sensor at each load, but this is invasive and very costly. Nonintrusive load monitoring (NILM) can extend many benefits of smart buildings to homes using minimal additional equipment and cost. With this method, a single electrical measurement point at the building's point of common coupling (PCC) with the grid is used to derive information (e.g., on/off status, electricity consumption) about the constituent net-load devices in the home.

This work was authored in part by Alliance for Sustainable Energy, LLC, the Manager and Operator of the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the Advanced Research Projects Agency-Energy (ARPA-E) under Grant Nos. DE-AR000701 and DE-AR0001016. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Existing work in NILM can be categorized based on the type of measurement inputs used. Generally, NILM approaches use one or both of: 1) steady-state measurement quantities, such as active and reactive power or root mean square (RMS) current, on a macro timescale (generally 1–60 min [2]); and 2) transient features derived from micro-timescale data (generally > 1 Hz [2]). In the first category, [2]–[7] used steady-state power measurements with a variety of contemporary machine learning methods, including multilabel k-nearest neighbor [2], [3], multilabel support vector machines [8], binary relevance [2], [8], hidden Markov models [2], [7], and recurrent neural networks [6]. In the second category, micro-timescale data have been used to derive features based on a frequency-domain transformation [9], [10], instantaneous wave shape [9], switching transients [11], or wavelet transformations [11], [12]. References [13], [14] fit into both categories, using micro timescale data to detect events and macro-timescale data to capture steady-state data throughout the event and build a library of events that can be used for later classification.

Steady-state measurement NILM methods often require a large sample (days to months) of training data and suffer a slow response time following events. Further, the use of steady-state measurements can easily result in misclassification between two different combinations of loads that have nearly identical steady-state measurements but differing instantaneous waveforms. For example, Fig. 1 shows the instantaneous current waveform (top plot) and corresponding RMS value (bottom plot) associated with two different combinations of home loads.

NILM methods based on steady-state measurements, such as the mean RMS current (-rmsMN in Fig. 1) in this example, will not be able to differentiate between these two combinations successfully; whereas NILM methods using

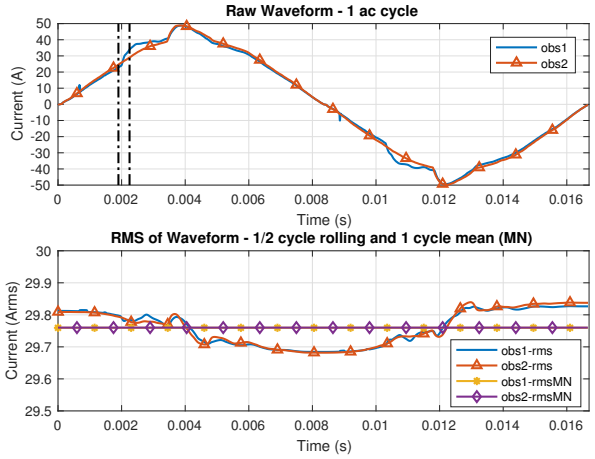

Fig. 1. Comparison of waveform (top plot) and RMS values (bottom plot) from two observations (one 60-Hz ac cycle sampled at 200 kHz) representing two different load configurations in a residential building

micro features, such as harmonics or wave shape, can be more successful. This point is further emphasized in Fig. 2, which shows the distribution of RMS current values (a steady-state quantity) for the case study presented later in this article, where it can be seen that the same steady-state value could exist in many different system load configurations.

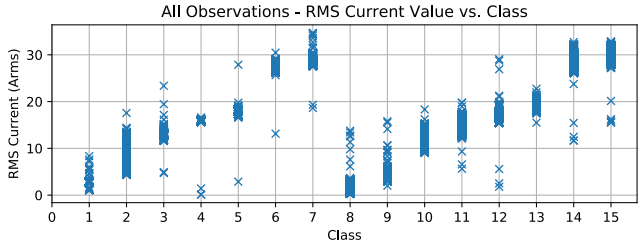

Fig. 2. All observations (all training and test data) with their RMS current feature plotted vs. class for a residential building load classification case study

In both steady-state and transient NILM approaches, it is common to use methods based on event detection; however, such approaches often assume that multiple load transients do not occur at the same time, might suffer confusion between transient events occurring within a single load (e.g., same load switching to a higher power level), and cannot make accurate state predictions until after an entire transient event has occurred. In most existing NILM approaches, system state predictions are often made in no faster than 1 s to 15 min. This might be sufficiently fast for some functions, such as estimating energy usage or customer load usage patterns, but it is too slow to enable a smart building to provide grid services (e.g., fast frequency response using building net load [15]).

This work presents a novel NILM approach relevant to fast smart building control that provides high-frequency (60-Hz) multiload state classification using a combination of frequency transformation, wavelet transformation, and wave shape features that allows for good generalization. The proposed method does not rely on event detection and provides a fast (>60-Hz) response to transient events. The approach is validated using data from a set of actual residential building appliances and is shown to be highly accurate (99%) when tested among all possible multiload state combinations.

II. High-Frequency, Multiclass Classification of Residential Appliance Loads

A. Problem Formulation

A total of N_L loads are connected to the main load center panel of a residential building, and each consume an instantaneous current $I_i(t) = q_i(t)\overline{I_i}(t)$, where $\overline{I_i}$ is the load signature, $q_i = \{0,1\}$ is the on/off status (where $q_i = 1$ denotes that the load is on), and I_i is the resulting current consumption of the i^{th} load. The total current consumption of the residential building is $I_{TOT}(t) = \sum_{i=1}^{N_L} q_i(t)\overline{I_i}(t)$. With these N_L loads, there are 2^{N_L} possible system load on/off states, $Y = [q_1, q_2, \ldots, q_i, \ldots, q_{N_L}]$, for the building. The objective is to train a classifier that uses an input vector, X, of features calculated for each 60-Hz ac cycle observation of $I_{TOT}(t)$ to predict the correct class label, Y, representing the building's complete load on/off state. The classifier should provide a state label prediction for every 60-Hz ac cycle of $I_{TOT}(t)$. For the application of high-frequency smart building control, the entire process of splitting the instantaneous $I_{TOT}(t)$ into single-cycle observations, calculating features, and using the classifier to predict an output label should occur within $\tau = 10 / 60 = 0.167\ s$. This response time will ensure state awareness for the fastest grid services, such as fast frequency response (generally <0.33–0.5 s total response time required). To be of practical value, the classifier should be able to generalize well to any $I_{TOT}(t)$ observations from the same group of appliances it is trained on, regardless of whether a similar magnitude $I_{TOT}(t)$ observation is seen during training.

B. Approach

A summary of the high-frequency NILM approach implemented is shown in Fig. 3. This approach extends previous NILM work by using a wider combination of features and by achieving much higher frequency output while maintaining exceptional multiclass performance.

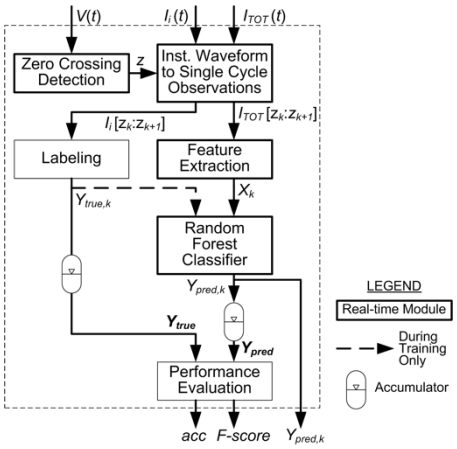

Fig. 3. High-frequency NILM approach

1) Input Data Processing

As shown in Fig. 3, high-speed instantaneous voltage and current waveforms are the inputs to this approach. During the training phase, current measurements from each circuit are used. Voltage measurement data are not strictly required, but if they are available, they are used to provide more accurate detection of zero crossings. The current data are divided at each 60-Hz ac cycle using zero-crossing detection, and then each cycle of the aggregate current measurement becomes an observation for which a label and features are derived.

2) Labeling (Training Phase)

Each aggregate current cycle observation is labeled by examining the instantaneous current data at each circuit, using level detection to derive the on/off state of that circuit, and finally building the multiclass label. This multiclass label could be a multilabel vector or a single integer label with the class encoded using a binary encoding scheme (e.g., for the four-load case, Label 11 corresponds to [1,0,1,1]). Both forms of multiclass labels are used in this work because both single-label and multilabel classifiers were experimented with.

3) Feature Engineering

The individual data points of the instantaneous aggregate current cycle observation are not used directly as the input to the classifier (this approach does not result in good prediction). Instead, the instantaneous data are processed into features to reduce the dimensionality of the problem while retaining the important characteristics of the waveform. Four key categories of features are used:

a) Harmonic features: A 1-D discrete Fourier transform is used to compute the magnitudes of the Fourier series coefficients at the fundamental and 3rd, 5th, 7th, 9th, 11th, and 13th harmonics of the 60-Hz nominal aggregate current

waveform. These coefficients are used in both their unscaled and normalized (using the L_2 norm) form, giving 14 possible features.

b) Steady-state RMS feature: A single feature based on calculating the RMS of all data points in the single-cycle aggregate current observation is used.

c) Wavelet features: An 8-level (the maximum decomposition level possible for the wavelet and sampling rate used) 1-D discrete wavelet transform using the order 7 Daubechies (db7) wavelet [16] is used to derive discrete wavelet transform coefficients at multiple levels of decomposition. The db7 wavelet is used because it has been shown to have good performance for power system applications [17]. The level 8 detail coefficients were studied as potential features, and coefficients #12–23 for a total of 12 out of the 25 total coefficients are selected as the final features.

d) Wave shape features: Here, wave shape features refer to features derived using individual data points in the raw instantaneous current waveform that represent unique characteristics of the waveform's shape. Three such features are used. The first is the maximum value in the middle third of the waveform's positive half cycle. The last two are ratios of two specific points in the ac waveform that capture local minima present in a waveform when particular loads are present. For example, in Fig. 1 (top plot), Obs1 and Obs2 contain the same three loads, but Obs1 also includes a fourth, relatively low-magnitude load. Overall, the two observations' waveforms are very similar, making it difficult to differentiate between these two classes using traditional features. In the region between the plot's two vertical black dotted lines, however, a difference in wave shape can be observed. A feature based on the ratio of the waveform's values at the first and second lines allows for reliable separability of these two very similar classes and provides a valuable measure of the presence of this additional load is present in any class. Two such ratios are used, both derived based on local maxima present in two different loads' characteristic waveforms. This approach can be applied generally by deriving a ratio feature for any constituent load that contains local maxima outside of the normal 60-Hz peak in its waveform.

4) Classifier

A random forest classifier (RFC), which comes under so-called ensemble methods, is used in this work. Ensemble methods use multiple uncorrelated base classifiers for better predictive performance. Uncorrelated models usually produce ensemble predictions that are more accurate than any individual predictions [18]. RFCs use decision trees as their base classifier. Each tree provides a class prediction, and the class with the most votes becomes the RFC's predicted class. Decision trees separate a data set so that the resulting subgroups are as different from each other as possible at every split and the observations in the same subgroup are as similar to each other as possible. Decision trees depend heavily on data they are trained on and can be very different if those training data change. RFCs consist of multiple decision trees where each tree is formed by randomly sampling data from the data set (with replacement). Further, each tree in a random forest can use only a random subset of the feature set. This leads to the individual trees being highly uncorrelated, leading to the ensemble's prediction being more accurate than that of any individual tree. Because multiple loads and any number of loads can be turned on, multiclass classification is required for this problem. Simple binary classifiers trained to detect the presence of a single load ignore the inherent correlation of different labels and are not suited for NILM. Instead, the "label power-set method," which takes each combination of possible load states as one class, is used here. This method takes the correlation of different labels into account and is able to provide accurate predictions, as shown in Section IV. The drawback to this approach is that the total number of classes in this method increases exponentially with the number of loads. For this reason, a direct multilabel approach, wherein the classifier uses the multilabel formulation of the load state label, is also considered. The two approaches are found to produce nearly identical results for this problem, but for larger problem sizes, the direct multilabel approach will scale better.

III. Experimental Design

A. Experiment Configuration

A residential-scale demonstration using four household appliances in the Energy Systems Integration Facility at the National Renewable Energy Laboratory is completed. As shown in Fig. 4, this experimental configuration includes four residential appliances, including a combination refrigerator/freezer (General Electric Profile PSQS6YGY), combination oven/range (Maytag MER8674), space heater, and a bank of (10) compact fluorescent light bulbs.

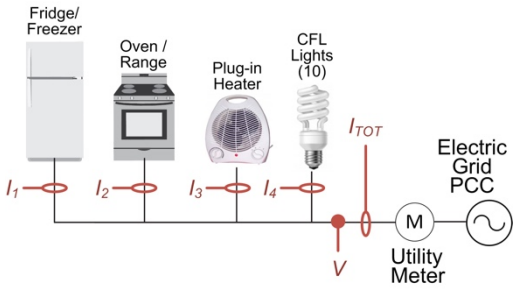

Fig. 4. Experimental configuration

In this four-load configuration, 15 different nontrivial classes that a classifier must predict are possible. Five current probes and one voltage probe at the locations indicated were used to gather instantaneous measurements sampled at 200 kHz over each 200-s data set period using a digital oscilloscope.

B. Data Sets

For classifier training purposes, seven independent 200-second data sets, including V, I_1, I_2, I_3, and I_4 (but not I_{TOT}), were collected. I_{TOT} was obtained as $I_{TOT} = I_1 + I_2 + I_3 + I_4$. For each data set, the appliance loads were put into a different starting condition, and then throughout the data set period, loads were turned on and off to obtain many instances of the 15 different classes. Throughout all data sets, the fridge, oven, and space heater were manually perturbed (e.g., fridge door opened, oven opened, space heater fan speed changed) to capture the variety of operating conditions of the loads. The lights were considered to be either on or off and were not manually perturbed.

For testing purposes, a further six independent 200-s data sets, including V, I_1, I_2, I_3, I_4, and I_{TOT}, were collected. For testing, only I_{TOT} and V are used as input data to the presented NILM approach, and I_1, I_2, I_3, and I_4 are used to assign the true label to each observation. As with training, appliances were cycled on and off and manually perturbed to obtain instances of all 15 classes as well as significant steady-state load variations across their range of operation. An example of one of the testing data sets is shown in Fig. 8.

C. Classifier Training and Tuning

Multiple RFCs using different subsets of the features described in Section II.B.3 are considered in the search for a classifier that is accurate in differentiating the 15 different classes but also that generalizes well to make correct predictions on observations not previously seen before in training data. To illustrate this point, two classifiers (summarized in Table I), both with the same hyperparameters (given below) but using different features, are compared. Classifier1 is the final classifier for which good accuracy and generalization results are presented. Classifier2 has good prediction accuracy on the test data set but does not generalize as well.

TABLE I. FEATURES USED FOR EACH CLASSIFIER

Classifier	Features Included			
	Harmonic Coefficients	*Normalized Harm. Coeff.*	*Wavelet*	*Wave Shape*
Classifier1	1,3,5,7,9,11,13	5,7	All	All
Classifier2	1,3,5,7,9,11,13	None	None	All

Both classifiers are RFC ensembles using bootstrapping and including 200 decision tree classifiers, each with a max depth of 35 and using entropy (information gain) as the split criterion. These optimal hyperparameters were determined using a grid search wherein for each hyperparameter combination, seven-fold cross-validation was performed to obtain the classifier's accuracy as the scoring metric. Seven-fold cross validation was implemented such that each combination of using 1/7 training data sets for testing and the remaining 6/7 training data sets for training were considered.

D. Performance Metrics

Both accuracy and weighted F1-score are used as measures of classifier performance. Accuracy provides an intuitive evaluation of the ratio of correctly labeled observations to the total number of observations. F1-score, a common metric for NILM implementations, is the harmonic mean of the precision and recall (i.e., $F_1 = \frac{2}{R^{-1}+P^{-1}}$, where R is recall and P is precision). Although an effort was made to get a sizeable number of instances of each class, the distribution of instances across classes is not equal, and hence an F1-score weighted by the number of true instances in a class is used to account for class imbalance (it is essentially a "macro" F1-score that accounts for class imbalance).

IV. EXPERIMENTAL RESULTS

A. Overall Results

Table II presents accuracy and F1-score results as averaged over 10 repeated (there is a small variation in results from one run to another because of the randomization present in the RFC) test runs evaluating the trained classifiers using the 6 independent test data sets.

TABLE II. OVERALL TEST RESULTS (10-RUN AVERAGE)

Test Data Set #	Test Accuracy [%]		F1-Score (Weighted) [%]	
	Clf. 1	Clf. 2	Clf. 1	Clf. 2
All	**99.44**	**98.97**	**99.45**	**98.94**
1	99.02	97.95	99.44	98.95
2	99.07	99.87	99.06	99.90
3	99.57	97.15	99.68	97.23
4	99.99	99.99	99.99	99.99
5	99.97	99.96	99.97	99.97
6	98.87	99.06	98.90	99.05

Fig. 5 compares the classifiers' accuracy in predicting each class in the full test data set (all 6 data sets). These results show that both classifiers perform well in terms of accuracy and F1-score on all data sets and among all classes.

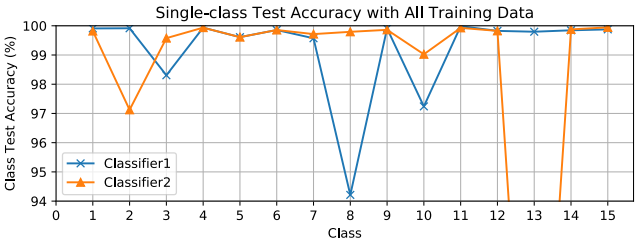

Fig. 5. Test accuracy (full test data set) by class when using all training data. Note: outlier on Classifier2 for Class 13 test accuracy is 80.92%.

B. Classifier Generalization

Generalization is important because practical NILM implementations should have good multiclass prediction performance across a variety of load operating conditions and it is not practical to train the classifier with every possible condition. To evaluate generalization performance, the training data set is first segmented by class, then each segment is sorted by cycle observation RMS value, and the middle 95% of that class segment's training data is discarded before proceeding to training and testing. An example of this for Class 6 is shown in Fig. 6, where the difference between the green triangular and red diamond line plots of the training data show that all training instances with RMS current values between ~26.5 and ~28.75 have been removed in the 5% training data set.

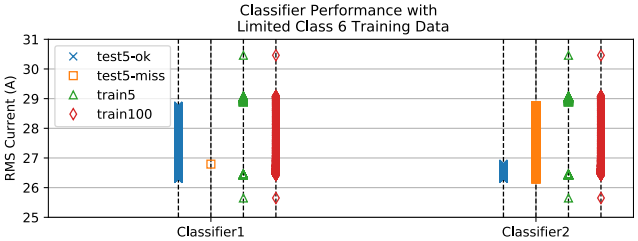

Fig. 6. Comparison of classifier performance on Class 6 (Y=[0,1,1,0]) test data when the middle (by RMS value) 95% of Class 6 instances in the training data set are removed before training

The blue cross and orange square line plots in Fig. 6 show how the range of instances in the test data set, which includes many values of RMS current in the range 26.5–28.75 where no similar training instances were used, were predicted. Classifier1, which has additional normalized harmonics and wavelet features, generalizes significantly better (97% accuracy vs. 22% accuracy) than Classifier2, which does not have these features. As shown in Fig. 7, this trend holds among most classes, further emphasizing the importance of these features in obtaining a classifier with good generalization performance.

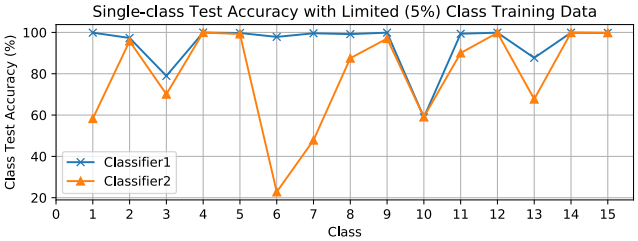

Fig. 7. Test accuracy by class when the middle (by RMS value) 95% of each class' instances in the training data set are removed before training

C. Transient Performance

Fig. 8 shows that label predictions match the true values throughout the entire data set (as expected from the 99.97% test accuracy in Table II), even including most transient periods. Fig. 9 shows an example of one such transient period (near $t = 133\ s$) where the Load 4 (CFL lights) turn-on event transient

lasts nearly 14 ac cycles. During this event, the classifier correctly predicts each cycle, despite the cycles having a variety of wave shapes and magnitudes. This demonstrates the high-frequency and high-accuracy performance of the approach and its advantage over many existing NILM approaches, which would not predict a state change until after the entire 14 ac cycle event signature was detected.

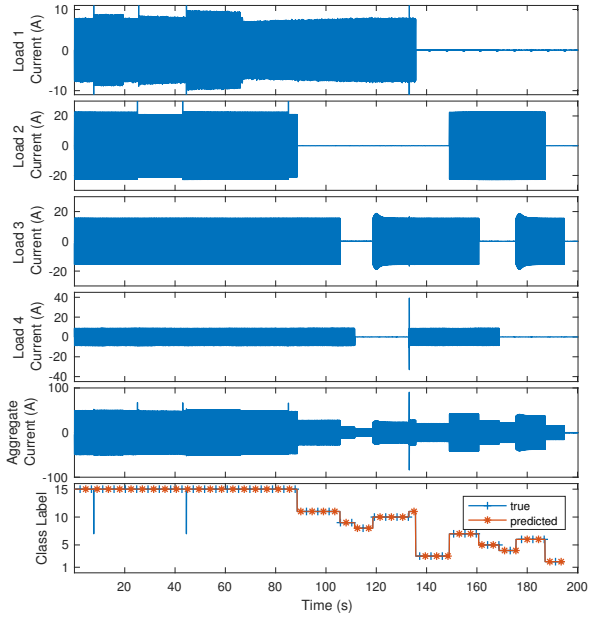

Fig. 8. Entire Test Data Set #5 with predicted versus true labels

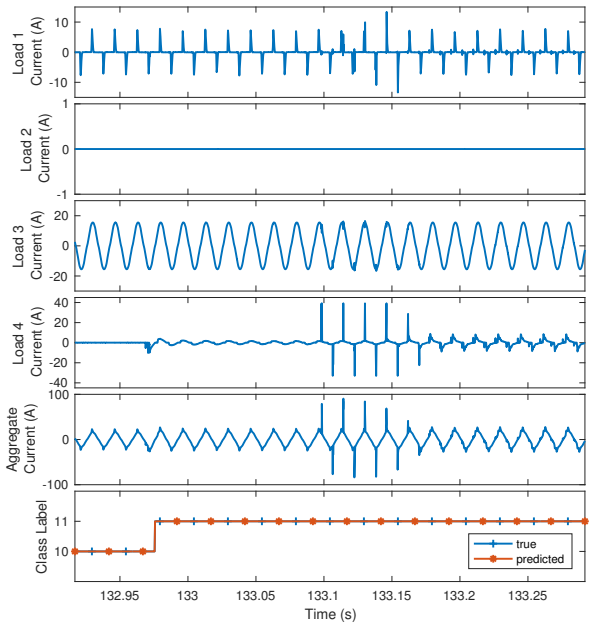

Fig. 9. Zoomed-in view of Test Data Set #5 near $t = 133\ s$

For this proof-of-concept study, the NILM approach was not implemented in real time, but the total processing time—including zero-crossing detection, feature extraction, and classifier prediction—is measured (average over thousands of cycles) to be 14.16 ms, which is fast enough to support high-frequency prediction and certainly well within the desired $\tau = 0.167s$ response time. This is measured on a laptop with a 2.8-GHz Intel i7 processor and assuming no data acquisition time, but with a response time 10 times faster than the desired response time, there is sufficient room for data acquisition and decreased performance if implemented on a slower computational platform while still meeting the response time requirement.

V. Conclusion

This paper developed a novel approach for high-frequency, real-time, multiclass nonintrusive load monitoring. This method was validated using a test bed with four residential appliances and was shown to have high accuracy, good generalization properties, and sufficient response time to support building grid-interactive control at fast timescales relevant to the provision of grid frequency support services. With minimal additional equipment and cost, the solution developed provides high-speed monitoring capabilities to better measure building performance and understand load usage patterns while further enabling smart buildings to support a modernized grid.

References

[1] U.S. EIA, "Electricity Explained: Use of Electricity," 2018. [Online]. Available: https://www.eia.gov/energyexplained.

[2] K. Basu, V. Debusschere, S. Bacha, U. Maulik, and S. Bondyopadhyay, "Nonintrusive Load Monitoring: A Temporal Multilabel Classification Approach," *IEEE Trans. Ind. Inform.*, vol. 11, no. 1, pp. 262–270, Feb. 2015.

[3] S. M. Tabatabaei, S. Dick, and W. Xu, "Toward Non-Intrusive Load Monitoring via Multi-Label Classification," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 26–40, Jan. 2017.

[4] S. Welikala, C. Dinesh, M. P. B. Ekanayake, R. I. Godaliyadda, and J. Ekanayake, "A real-time non-intrusive load monitoring system," in *2016 11th International Conference on Industrial and Information Systems (ICIIS)*, 2016, pp. 850–855.

[5] B. Liu, Y. Yu, W. Luan, and B. Zeng, "An unsupervised electrical appliance modeling framework for non-intrusive load monitoring," in *2017 IEEE Power Energy Society General Meeting*, 2017, pp. 1–5.

[6] M. A. Devlin and B. P. Hayes, "Non-Intrusive Load Monitoring and Classification of Activities of Daily Living Using Residential Smart Meter Data," *IEEE Trans. Consum. Electron.*, vol. 65, no. 3, pp. 339–348, Aug. 2019.

[7] H. He, Z. Liu, R. Jiao, and G. Yan, "A Novel Nonintrusive Load Monitoring Approach based on Linear-Chain Conditional Random Fields," *Energies*, vol. 12, no. 9, p. 1797, May 2019.

[8] D. Li and S. Dick, "Whole-house Non-Intrusive Appliance Load Monitoring via multi-label classification," in *2016 International Joint Conference on Neural Networks (IJCNN)*, Vancouver, BC, Canada, 2016, pp. 2749–2755.

[9] T. Hassan, F. Javed, and N. Arshad, "An Empirical Investigation of V-I Trajectory based Load Signatures for Non-Intrusive Load Monitoring," *IEEE Trans. Smart Grid*, vol. 5, no. 2, p. 10, 2014.

[10] R. A. S. Fernandes, I. N. da Silva, and M. Oleskovicz, "Load Profile Identification Interface for Consumer Online Monitoring Purposes in Smart Grids," *IEEE Trans. Ind. Inform.*, vol. 9, no. 3, pp. 1507–1517, Aug. 2013.

[11] C. Duarte, P. Delmar, K. W. Goossen, K. Barner, and E. Gomez-Luna, "Non-intrusive load monitoring based on switching voltage transients and wavelet transforms," in *2012 Future of Instrumentation International Workshop (FIIW) Proceedings*, 2012, pp. 1–4.

[12] J. M. Gillis, S. M. Alshareef, and W. G. Morsi, "Nonintrusive Load Monitoring Using Wavelet Design and Machine Learning," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 320–328, Jan. 2016.

[13] D. He, W. Lin, N. Liu, R. G. Harley, and T. G. Habetler, "Incorporating Non-Intrusive Load Monitoring Into Building Level Demand Response," *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 1870–1877, Dec. 2013.

[14] M. Nardello, M. Rossi, and D. Brunelli, "An innovative cost-effective smart meter with embedded non intrusive load monitoring," in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, 2017, pp. 1–6.

[15] B. Lundstrom, S. Patel, S. Attree, and M. V. Salapaka, "Fast Primary Frequency Response using Coordinated DER and Flexible Loads: Framework and Residential-scale Demonstration," in *2018 IEEE Power Energy Society General Meeting (PESGM)*, 2018, pp. 1–5.

[16] I. Daubechies, *Ten lectures on wavelets*, vol. 61. SIAM, 1992.

[17] S. Kashyap and A. K. Singh, "Most suitable mother wavelet for measurement of power system harmonics using DWT in view of IEEE Standard 1459-2000," in *2008 13th International Conference on Harmonics and Quality of Power*, 2008, pp. 1–6.

[18] L. I. Kuncheva, "Measures of Diversity in Classifier Ensembles and Their Relationship with the Ensemble Accuracy," *Mach. Learn.*, vol. 51, pp. 181–207, 2003.