An Active Learning Framework for Non-Intrusive Load Monitoring

Xin Jin
Buildings and Thermal Systems Center
National Renewable Energy Laboratory
Golden, Colorado, USA
xin.jin@nrel.gov

***Abstract*—Non-Intrusive Load Monitoring (NILM) is a set of techniques that estimates the electricity usage of individual appliances from power measurements taken at a limited number of locations in a building. One of the key challenges in NILM is having too many data lacking class labels, but being unable to label the data manually for cost or time constraints. This paper presents an active learning framework that helps existing NILM techniques to overcome this challenge. Active learning is an advanced machine learning method that interactively queries a user for the class label information. Unlike most existing NILM systems that heuristically request user inputs, the proposed method only needs minimally sufficient information from a user to build a compact yet highly representative load signature library. Initial results indicate the proposed method can reduce the user inputs by up to 90% while still achieving similar disaggregation performance compared to a heuristic method. Thus, the proposed method can substantially reduce the burden on the user, improve the performance of a NILM system with limited user inputs, and overcome the key market barriers to the wide adoption of NILM technologies.**

Keywords—Electric Load Disaggregation; Active Learning; Machine Learning; Event Classification; BLUED Dataset

I. Introduction

Non-intrusive load monitoring (NILM) is an emerging class of load disaggregation techniques that estimate the electrical consumption of individual appliances from power measurements taken at a limited number of locations in a building. NILM promises to replace expensive submetering and provide actionable information to a variety of customers such as homeowners, building operators, service companies, and utilities [1]. Applications of NILM include energy monitoring, fault detection, and load shed verification [2].

Many new techniques have emerged since the inception of the first NILM technique more than two decades ago [3]. Machine learning techniques have been widely adopted to improve the performance of the NILM systems. Training is a process of learning each appliance's characteristic signatures from the power data with known class labels. A class label is a variable or a string that represents an appliance's type. A major challenge of NILM is having too many data without class labels, but being unable to label the data manually for cost or time constraints. Depending on the level of required effort that users exert in labeling, the training process can be divided into three categories: manual training, sensor-assisted training, and cloud-based training [4]. Although current NILM research is moving toward automatic data labeling, users are still involved in the labeling process to some extent. Ideally, the most informative data should be labeled first, where informativeness is defined as the expected improvement in disaggregation accuracy. Existing techniques lack the intelligence to select the most informative unlabeled data and then obtain the class label from a user; instead, they often require extensive user inputs which result in inconvenience and a low rate of adoption.

This paper presents a novel method for efficiently engaging a user to provide the class label information. The proposed method is built on the framework of active learning [5], a machine learning method that interactively queries the user for the class label information. Unlike existing NILM methods that heuristically request user inputs, the proposed method only needs the minimally sufficient data from a user to build a compact yet highly representative load signature library.

The proposed method does not seek to replace any existing NILM techniques; on the contrary, it augments existing techniques by significantly reducing the burden on a user and improving the disaggregation accuracy. There will be little incremental cost since the proposed method can leverage the existing infrastructure and perform data analytics in the cloud, on embedded systems, or on existing mobile platforms.

II. Background

A. *Categorization of NILM Systems*

NILM systems can be categorized into event-based and model-based (i.e., non-event-based) approaches. Event-based NILM systems rely on the detection and classification of events whereas model-based NILM techniques usually first generate models of each appliance and then use optimization techniques to identify the appliance usage. Event-based methods are able to provide accurate disaggregation results, but they require a fully labeled training set.

NILM systems can also be categorized into supervised and unsupervised approaches depending on whether or not they require a training process to obtain labeled data [6]. A majority of existing NILM techniques follow the supervised learning approach while unsupervised methods are emerging to reduce the effort of acquiring the training data.

The proposed active learning framework is an event-based supervised learning method because it interacts with a user to obtain class label information. The active learning method can identify the optimal threshold for event detection and build a library that contains the most distinctive load signatures for event classification.

B. Appliance Signatures for Load Disaggregation

In event-based NILM techniques, load signatures are extracted from the power signals of detected events to be used in the subsequent event classification step. Steady-state and transient event-based feature extraction methods are widely used in many NILM techniques [6]. Steady-state methods extract information based on variations in the steady-state signatures, such as variations in real and reactive power, higher order harmonics, and electromagnetic interference signatures. Each appliance's transient behavior also contains distinctive signatures. Information such as transient power can help differentiate appliances that have similar steady-state characteristics.

C. Active Learning

Active learning is a subfield of machine learning and has been widely adopted in scenarios where labeled samples are difficult, time-consuming or expensive to obtain [7]. Active learning overcomes the labeling bottleneck by asking *queries* in the form of requesting class labels of unlabeled samples from a user. In this way, active learning-based systems can achieve high accuracy using as few labeled samples as possible, thereby minimizing the cost of obtaining class labels. Active learning is a perfect fit for the NILM problems where data are abundant but labels are expensive to obtain.

Depending on how a query is generated, active learning can be categorized into three types: membership query synthesis [8], sequential sampling [9], and pool-based sampling [10]. Among these methods, pool-based sampling is well suited for NILM problems because it tackles the scenarios where large collections of unlabeled samples can be collected at once and queries are selectively drawn from the pool. Sequential sampling is very similar to pool-based sampling except it does not evaluate and rank the entire collection before selecting the best query. The membership query synthesis is based on some unrealistic assumptions and not suitable for solving the NILM problems. Therefore, pool-based sampling is used in this paper to optimize the query strategy for handling a large number of unlabeled samples in the NILM problem.

III. Active Learning-Based NILM

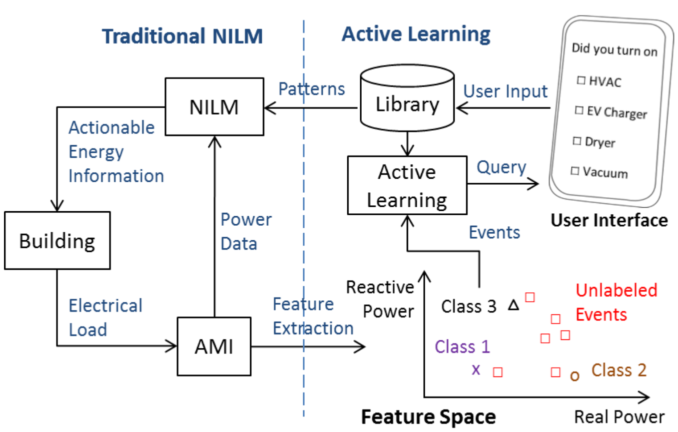

Fig. 1. An active learning framework for augmenting existing NILM systems

Figure 1 shows an overview of the active learning framework that augments the traditional NILM techniques. The power signals obtained from advanced metering infrastructure (AMI) or similar devices are transformed into a feature space to distinguish events generated by different appliances. Each point in the feature space represents the time series of an event. Data analytics is performed at each step to identify the most informative unlabeled samples in terms of the average distances from labeled samples. It is more advantageous to obtain the class label of an unlabeled sample that is far away from all labeled samples. A query is then sent to a user to obtain the class label of the selected sample. The active learning module may also perform statistical inference to help the user to provide accurate labels, which will be stored in a library with the extracted features. The same process repeats until a stopping criterion is met. A formal formulation of the active learning problem is presented below.

A. Problem Formulation

Let $\boldsymbol{X}_K = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\} \in \mathbb{R}^d$ be a set of labeled samples in the feature space with known class labels and $\boldsymbol{X}_U = \{\boldsymbol{x}_{n+1}, \boldsymbol{x}_{n+2}, \cdots, \boldsymbol{x}_{n+m}\}$ be a pool of unlabeled samples. Let $d(\boldsymbol{a}, \boldsymbol{b})$ be a distance measure between samples $\boldsymbol{a}$ and $\boldsymbol{b}$. Active learning aims to find an unlabeled sample $\boldsymbol{x}_q$ such that

$$\boldsymbol{x}_q = \underset{\boldsymbol{x} \in \boldsymbol{X}_U}{\operatorname{argmax}} \frac{1}{n} \sum_{i=1}^{n} d(\boldsymbol{x}, \boldsymbol{x}_i)$$

where $\boldsymbol{x}_i \in \boldsymbol{X}_K$ and $\frac{1}{n}\sum_{i=1}^{n} d(\boldsymbol{x}, \boldsymbol{x}_i)$ is the *informativeness measure* that can be interpreted as the mean distance between an unlabeled sample $\boldsymbol{x}$ and all labeled samples in set $\boldsymbol{X}_K$. This results in selecting an unlabeled sample that is the farthest from all labeled samples, which brings the maximum amount of information. After obtaining the label from a user, the sample $\boldsymbol{x}_q$ is moved from set $\boldsymbol{X}_U$ to set $\boldsymbol{X}_K$ and the entire process is repeated until a stopping criterion is met.

We use a gradient-based convergence method [11] to stop the active learning process when more labeled samples from the pool do not contribute more information, indicated by when the decrease in the informative measurement becomes less than a fraction of the initial descent. A flow chart is provided in Figure 2 to illustrate the active learning process.

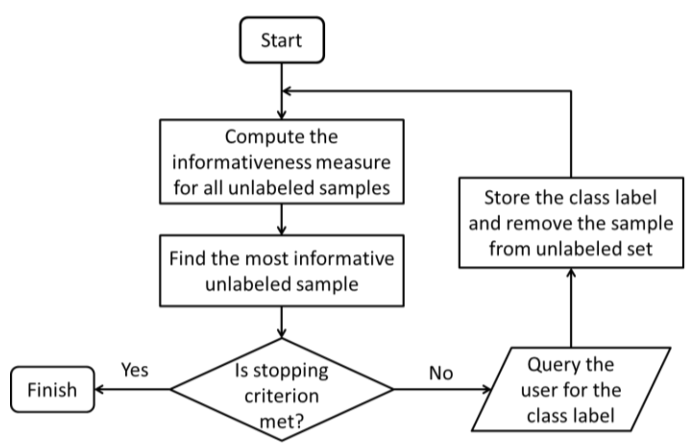

Fig. 2. Flow chart of the proposed active learning framework for NILM

IV. Experiment and Results

A. BLUED Data set

The proposed active learning algorithm is evaluated on the BLUED data set [12], which is a fully-labeled public data set for NILM research. Every state transition of each appliance in the test home is labeled and time-stamped, providing the ground truth for evaluation of event-based NILM algorithms. More than 40 different household appliances are included in the data set, although some of the appliances do not have any recorded events.

This paper focuses on evaluating the performance of the active learning framework, so only the event classification step is considered. Real and reactive power data are generated from the raw voltage and current measurements and stored at 60 Hz. A window of 150 data points (i.e., 2.5 seconds) is taken around an event, simulating the outcome of an ideal event detector.

For simplicity, we extract three features from the power data, namely steady-state variations of real (ΔP) and reactive power (ΔQ) and transient real power overshoot ($P_{overshoot}$). All three features are normalized. Figures 3 and 4 show examples of Phase A and B events in the feature space where each data point represents an event. Overall, the appliances have quite distinct load signatures except for some of the lights and the computers.

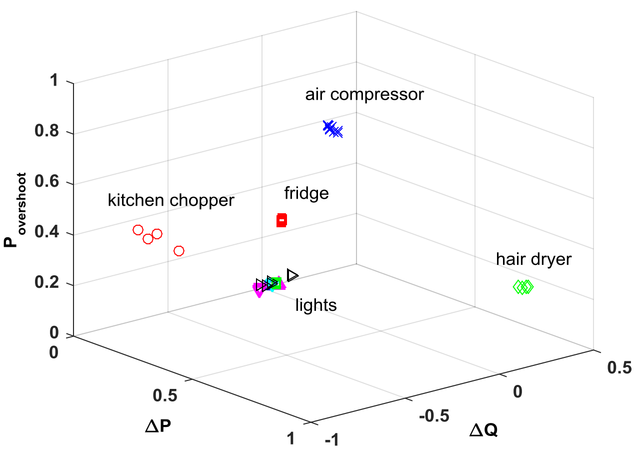

Fig. 3. An example of Phase A events shown in the feature space

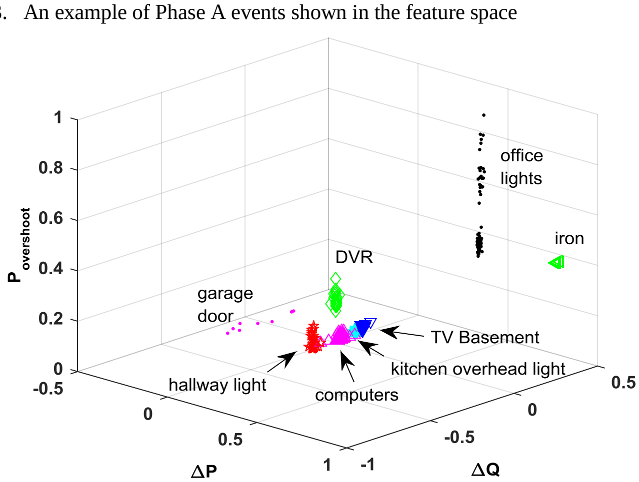

Fig. 4. An example of Phase B events shown in the feature space

Active learning is performed separately on Phases A and B to utilize the phase information as an additional feature. The underlying assumption here is that appliances always connect to the same phase. This is a reasonable assumption because very few appliances are frequently moved around the home and switched to a different phase.

B. Initial Results

A four-fold cross-validation is implemented in the experiment. In other words, 75% of the entire data set is used as the training set whose labels are to be provided by a user and 25% of the data set is used as the testing set where the ground truth is held back for evaluation. We swap the training set and the testing set until all the samples have been included in the testing set.

In the proposed active learning algorithm, a *k*-Nearest Neighbors (*k*-NN) classifier is included in the learning loop to perform event classification every time a new label is provided by the user. For comparison, a random selection algorithm is implemented following the same cross-validation process. In contrast to active learning, which optimizes the query strategies based on the informativeness measure, the random selection algorithm heuristically selects the unlabeled samples and asks the user for the class label.

To mitigate the effects of randomness on the results, both algorithms are implemented for 50 iterations. The results are shown in Figures 5 and 6 for Phases A and B, respectively. The classification accuracies and stopping point are the mean derived from the 50 iterations. The stopping points are determined based on the gradient convergence of the informativeness measure (not shown in the figures) instead of the classification accuracy. For comparison, we continue evaluating the classification accuracy of the active learning algorithm even though the stopping criterion is met.

In both cases, the classification accuracy of the random selection method increases rapidly with the first few user inputs, but slows down afterwards. In contrast, by carefully selecting user inputs, the active learning method can achieve vastly improved accuracy with minimally sufficient user inputs. The active learning method reaches its peak or near-peak performance by automatically selecting about only 10% of the total user-labeled samples. This result is very significant because it indicates that an active learning-based NILM system is able to reduce the amount of required user inputs by 90% while still achieving similar disaggregation performance that can be accomplished in the other heuristic method only by obtaining the labels of all training samples. The results of both algorithms are summarized in Table I.

TABLE I. Performance Comparison between Active Learning (AL) and Random Selection (RS)

Phase	Method	Use Stopping Criterion		Peak Performance	
		# of Queries	*Accuracy*	*# of Queries*	*Accuracy*
A	***AL***	23	99.72%	23	99.72%
	RS	N/A	89.99% (w/ 23 queries)	260	99.72%
B	***AL***	23	98.82%	31	99.38%
	RS	N/A	92.79% (w/ 23 queries)	206	99.42%

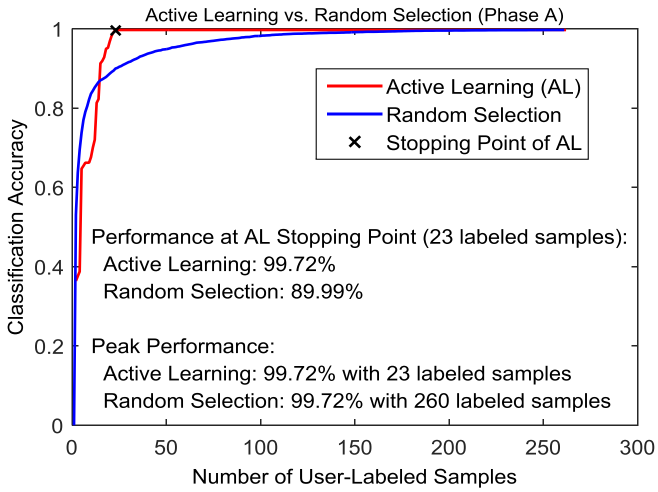

Fig. 5. Performance comparison between active learning and random selection on Phase A. Results shown here are the average of 50 iterations.

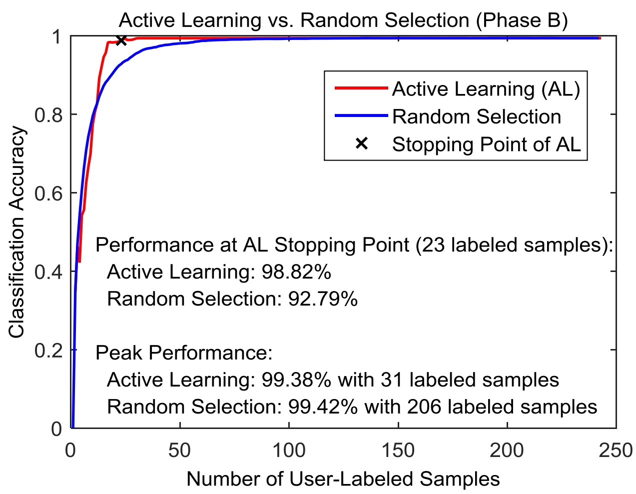

Fig. 6. Performance comparison between active learning and random selection on Phase B. Results shown here are the average of 50 iterations.

C. Future Research

We will explore proactive learning, semi-supervised learning, informativeness measure, labeling assistance, and holistic integration with other NILM systems in future research.

Although the proposed active learning method has shown to significantly reduce the labeling efforts compared to the random selection method, active learning assumes the user is infallible or indefatigable, which may be unrealistic in many real-world situations. A *proactive learning* method is proposed in [13] to bridge the gaps between traditional active learning and many practical problems. The goal of extending active learning to proactive learning is to reach out to the user with the appropriate query at the appropriate cost without assuming the user is infallible or indefatigable.

Unlike active learning methods that exploit the least confident samples, *semi-supervised learning* methods exploit the most confident ones from the unlabeled samples [14]. It may be advantageous to combine these two methods to generate a new learning scheme where only highly uncertain samples are labeled by a user while all others are automatically labeled to further reduce the labeling effort.

The *informativeness measure* used in the paper is based on the Euclidean distance measure. Other measures such as information-theoretic measures should be explored. We will also generalize the proposed method to make it work with other NILM techniques that are not event-based.

The time-of-use and frequency-of-use information can be used to assist the labeling process because thermostatically-controlled appliances (e.g., HVAC, water heater, refrigerator) behave quite differently from the appliances with user-initiated cycles (e.g., dishwasher, clothes dryer). We will develop a *labeling assist* module to infer the class label and further reduce the number of queries for a user.

In addition to the electric load disaggregation, NILM techniques have been implemented to identify the sources of water and gas consumption [15]. Active learning has been used to calibrate the HydroSense system that uses a single pressure sensor on the plumbing line to infer water usage [16]. Integrating these active learning-based NILM techniques will address all of the energy sensing needs in residential buildings in a systematic and efficient way and create a holistic solution for home energy management.

V. Conclusion

This paper presents an active learning framework to reduce the labeling effort from a user, which is one of the key challenges in the NILM research. Unlike most existing NILM techniques that heuristically request user inputs, the proposed method only needs the minimally sufficient information from a user to build a compact yet highly representative load signature library by querying the most informative samples first. Initial results on the BLUED data set indicate the proposed method is able to reduce the user inputs by up to 90% while still achieving similar disaggregation performance compared to a heuristic method. Therefore, the proposed method has the potential to substantially reduce the burden on the user, improve the performance of a NILM system with very limited user inputs, and overcome the key market barriers to the wide adoption of NILM technologies. The proposed method can be implemented in the cloud, on embedded systems, or on mobile platforms, which will minimize the incremental cost over existing methods.

Acknowledgment

This work was supported by the U.S. Department of Energy under Contract No. DE-AC36-08-GO28308 with the National Renewable Energy Laboratory.

The author would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper. The author is also grateful to Dane Christensen, Lieko Earle, Kyri Baker, and David Roberts for their contributions to this paper.

REFERENCES

[1] D. Christensen, L. Earle, and B. Sparn, “NILM Applications for the Energy-Efficient Home, ” 1st International Workshop on Non-Intrusive Load Monitoring, Pittsburgh, PA, May 7, 2012, available as NREL Technical Report: NREL/TP-5500-55291.

[2] D.C. Bergman, D. Jin, J.P. Juen, N. Tanaka, and C.A. Gunter, “Nonintrusive Load-Shed Verification,” IEEE Pervasive Computing, 10 (1), pp. 49-57, January-March 2011.

[3] G. Hart, “Nonintrusive Appliance Load Monitoring,” IEEE, 80 (12), December 1992.

[4] Y. Li, M. Trayer, S. Ramakrishna, and P. Lai, “Training Schemes in Non-Intrusive Load Monitoring,” 1st International Workshop on Non-Intrusive Load Monitoring, Pittsburgh, PA, May 7, 2012.

[5] P. Jain, and A. Kapoor, “Active Learning for Large Multi-class Problems,” IEEE CVPR 2009, pp. 762-769, 2009.

[6] A. Zoha, A. Gluhak, M.A. Imran, and S. Rajasegarar, “Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey.” Sensors, 12(12), 16838-16866, 2012.

[7] B. Settles. "Active learning literature survey." University of Wisconsin, Madison 52.55-66 (2010): 11.

[8] D. Angluin. “Queries and concept learning. Machine Learning,” 2:319-342, 1988.

[9] D. Cohn, L. Atlas, and R. Ladner. “Improving generalization with active learning”. Machine Learning, 15(2):201-221, 1994.

[10] D. Lewis and W. Gale. “A sequential algorithm for training text classifiers.” In Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 3-12, 1994.

[11] F. Laws, and H. Schutz, “Stopping Criteria for Active Learning of Named Entity Recognition,” in Proceedings of the 22nd International Conference on Computational Linguistics, 465-472, Manchester, 2008.

[12] 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,” SustKDD’12, Beijing, China, Aug. 2012.

[13] P. Donmez, “Proactive Learning: Towards Learning with Multiple Imperfect Predictors,” Carnegie Mellon University. CMU-LTI-10-002, 2010.

[14] T. Ma, J. Ge, and J. Wang, “Combining Active Learning and Semi-Supervised for Improving Leaning Performance,” ISABEL ’11, Barcelona, Spain, October 26-29, 2011.

[15] J. Froehlich, E. Larson, S. Gupta, G. Cohn, M. Reynolds and S. Patel, "Disaggregated End-Use Energy Sensing for the Smart Grid," in IEEE Pervasive Computing, vol. 10, no. 1, pp. 28-39, Jan.-March 2011.

[16] E. C. Larson, “Consumer Centered Calibration in End-use Water Monitoring,” 2nd International Workshop on Non-Intrusive Load Monitoring, Austin, TX, June 3, 2014.