Enhancing Machine Learning Classification for Electrical Time Series Applications

Mark Valovage
Computer Science and Engineering
University of Minnesota, Minneapolis, MN
valovage@cs.umn.edu

The Benefits Hiding in Electrical Data

Machine learning applications to electrical time series data will have wide-ranging impacts in the near future. Electricity disaggregation holds the promise of reducing billions of dollars of electrical waste every year. In the power grid, automatic classification of disturbance events detected by phasor measurement units could prevent cascading blackouts before they occur. Additional applications include better market segmentation by utility companies, improved design of appliances, and reliable incorporation of renewable energy resources into the power grid. However, existing machine learning methods remain unimplemented in the real world because of limiting assumptions that hinder performance.

My research contributions are summarized as follows: In electricity disaggregation, I introduced the first label correction approach for supervised training samples. For unsupervised disaggregation, I introduced event detection that does not require parameter tuning and appliance discovery that makes no assumptions on appliance types. These improvements produce better accuracy, faster computation, and more scalability than any previously introduced method and can be applied to natural gas disaggregation, water disaggregation, and other source separation domains. My current work challenges long-held assumptions in time series shapelets, a classification tool with applicability in electrical time series and dozens of additional domains.

Electricity Disaggregation

Electricity disaggregation identifies individual appliances from one or more aggregate data streams. By inexpensively collecting and reporting individual appliance power consumption to consumers, disaggregation has the potential to reduce billions of dollars of annual waste. Such data could also be fed into an automated system that could turn off unused appliances or shift their operation to a later time.

Label Correction for Supervised Disaggregation

Supervised learning methods first train on appliance samples recorded in isolation. Afterwards these methods can then identify appliances that are operating simultaneously while being recorded by a single smart meter. However, existing approaches assume error-free labels in training data, an unrealistic assumption for data labeled by naïve consumers. In (Valovage and Gini 2016), I introduced the first method to automatically correct labels in consumer-labeled training samples, enabling realistic application of supervised methods to a single house. I improved this method in (Valovage and Gini 2017) to use a parameter-free model, making it scalable to millions of homes. While these improvements overcome limiting assumptions, supervised learning still requires hours of work by consumers to meticulously label individual appliance samples. To enable a system that requires no consumer setup, unsupervised learning is required.

Unsupervised Electricity Disaggregation

It is more challenging for unsupervised learning methods to accurately identify appliances since they lack training samples and must be able to identify a wide range of appliances. Unsupervised disaggregation requires two distinct steps. First, during *event detection*, the aggregate power data stream is segmented into significant events that represent a state change in one or more appliances. Second, *appliance discovery* reconstructs appliances from these events. Existing methods for both of these steps have their own shortcomings that limit real-world deployment, detailed below.

Parameter-Free Event Detection: Previously introduced event detection methods depend on parameters optimized for a single appliance or dataset, limiting scalability to millions of buildings. In (Valovage and Gini 2017), I introduced the first event detection method that does not require parameter tuning using a modified version of Bayesian change detection. Tests on 2 publicly available datasets containing 7 different houses showed Bayesian change detection performed on par with or better than existing state-of-the-art event detection methods without the need to tune parameters, making it scalable to millions of homes. Furthermore, my modifications to Bayesian change detection reduced its space and time complexity from $O(n^2)$ to $O(n)$, enabling it to run in real-time on inexpensive hardware.

Model-Free Iterative Appliance Discovery: Following event detection, events must be recombined into their respective appliances. Doing this with no previous assumptions is challenging since appliances can operate for different amounts of time and often overlap in operation. In addition, while simple appliances consistently generate the same power signatures, more complex appliances produce a different signature every time they are operated.

To overcome limitations of existing methods, in (Valovage, Shekhawat, and Gini 2018) I introduce *iterative appliance discovery*, an algorithm built around the concept of identifying the simplest appliances first. Once the simplest appliances have been identified, events associated with them are marked so they are not included in subsequent searches for more complex appliances. By iteratively reducing the search space, iterative discovery has the unprecedented ability to discover complex appliances other methods cannot.

Work in Progress: Time Series Shapelets

Shapelets are small subsequences of time series that can be used for fast, accurate classification of unlabeled time series (see figure). Electrical applications of shapelets include electricity disaggregation (Patri et al. 2014) and classification of phasor measurement unit disturbance data (Biswal et al. 2016). However, the accuracy of a shapelet during classification relies solely on its distance from a tested sample, and existing approaches are limited by long-held assumptions.

Shapelets: To Normalize or Not To Normalize

Z-normalizing data prior to calculating distance has been accepted as a necessary step in all previous shapelet research. Z-normalization is intended to capture variation in scale and offset in local time series features. However, initial results on datasets from dozens of domains show that Z-normalization actually decreases accuracy more often than it increases it.

Shapelets typically capture a well-defined local feature from a single class. Z-normalization tends to improve accuracy when this feature has low variability and high similarity between samples from that class. However, rescaling caused by Z-normalization can also produce better fits to similar features from *other* classes. This undesirable side effect can diminish the classification ability of a shapelet by reducing (or even eliminating) the distance separation between samples of different classes.

This is especially true for electrical data. In all 6 electrical datasets studied, Z-normalization significantly reduced accuracy. In the most extreme case, shapelet classification on one dataset was 42% using Z-normalized distances, while using raw distance improved shapelet accuracy to over 81%.

In addition to challenging the assumption that Z-normalization always improves accuracy, I am exploring the ability to reliably predict the impact of Z-normalization prior to learning by using cross-validation on the training data.

Learning Distance Metrics for Shapelets

Euclidean distance is the only distance metric that has been attempted with shapelets. While this has worked in the past, deeper analysis of the data shows that different distance metrics have the potential to significantly improve the accuracy of shapelets discovered. I am currently exploring the impact of alternative distance metrics on the accuracy of discovered shapelets and with learning distance metrics in the space specific to a shapelet, enabling higher classification accuracy through finer-tuned measurements.

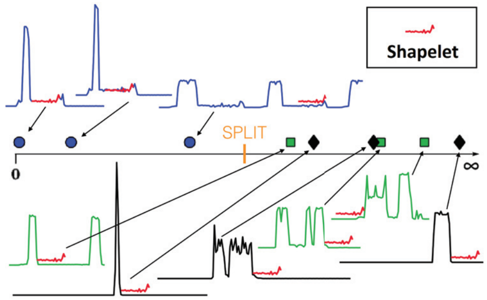

Figure 1: Shapelet example from the LargeKitchenAppliances dataset.

Introducing Virtual Shapelets

By definition, a shapelet must be a subsequence of a training sample in the training data provided. While this limits the search for shapelet candidates to a finite space, any noise present in the sample will also be present in the shapelet.

I am exploring a new structure called virtual shapelets. A *virtual shapelet* is a short time series that can be used in classification the same way a shapelet can, but a virtual shapelet does not have to exist in the training data. Instead, a virtual shapelet can be built from multiple similar shapelets and can better represent a locally discriminative feature.

Virtual shapelets will remove the dependency on a single training sample. By doing so, virtual shapelets will be more robust to noise in any given sample and will have the ability to better capture locally discriminative features. However, since virtual shapelets live in an infinite space, care must be given when crafting any algorithm that produces them to avoid overfitting the training data and ensure the algorithm is robust enough to find virtual shapelets for different domains. I am currently experimenting with multiple algorithms that have the potential to produce virtual shapelets that better represent locally discriminative features and produce higher classifying power.

References

Biswal, M.; Hao, Y.; Chen, P.; Brahma, S.; Cao, H.; and De Leon, P. 2016. Signal features for classification of power system disturbances using PMU data. In *Power Systems Computation Conference (PSCC), 2016*, 1–7. IEEE.

Patri, O. P.; Panangadan, A. V.; Chelmis, C.; and Prasanna, V. K. 2014. Extracting discriminative features for event-based electricity disaggregation. In *Technologies for Sustainability (SusTech), 2014 IEEE Conference on*, 232–238.

Valovage, M., and Gini, M. 2016. Automatic label correction and device prioritization in single household electricity disaggregation. *International Workshop on Artificial Intelligence for Smart Grids and Smart Buildings, at AAAI.*

Valovage, M., and Gini, M. 2017. Label correction and event detection for electricity disaggregation. In *Proc. 16th Conf. on Autonomous Agents and MultiAgent Systems*, 990–998.

Valovage, M.; Shekhawat, A.; and Gini, M. 2018. Model-free iterative temporal appliance discovery for unsupervised electricity disaggregation. In *Proc. AAAI 2018.*