The University as a Laboratory for Smart Grid Data Analytics

Duke UNIVERSITY | BASS CONNECTIONS

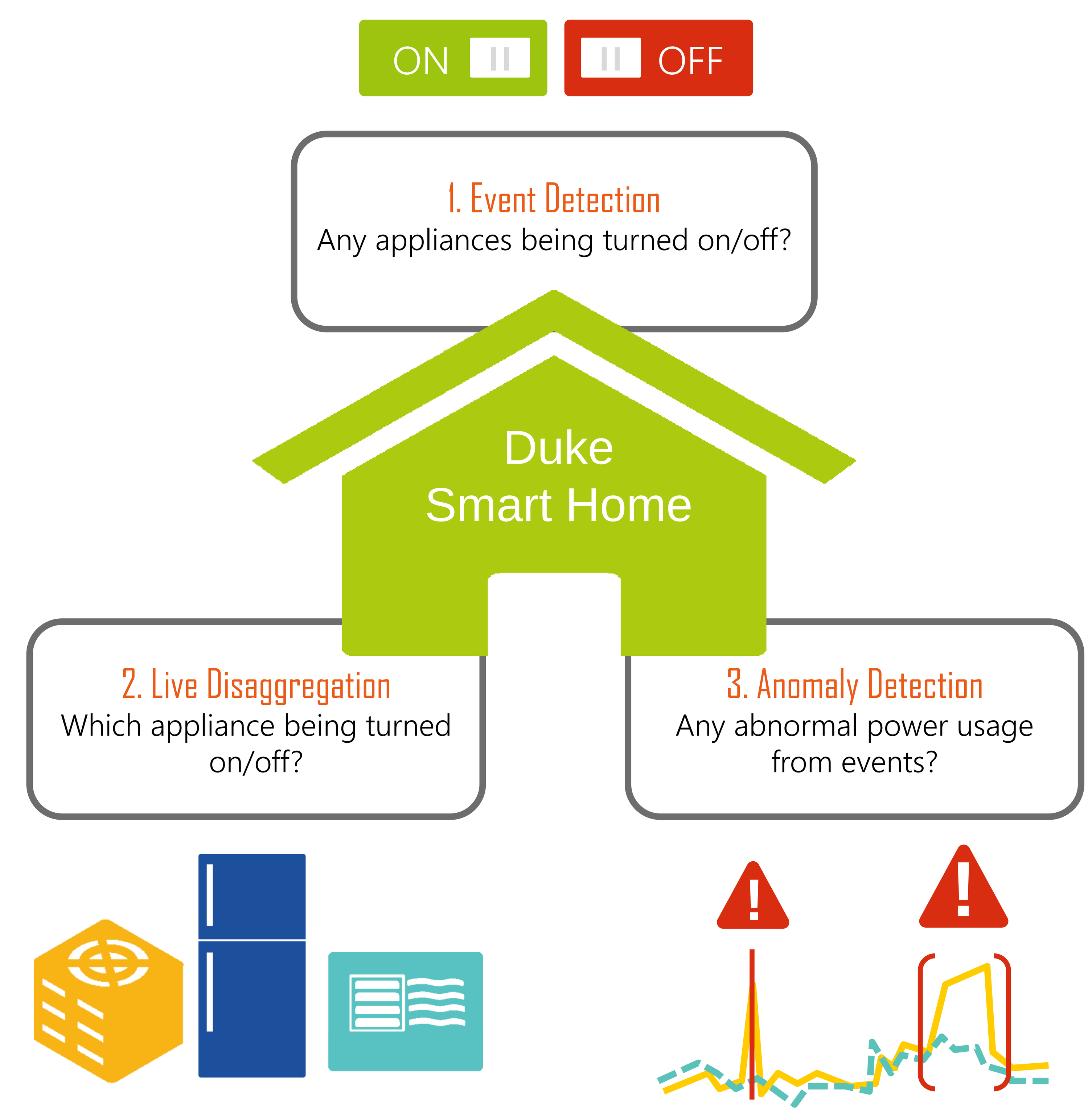
Students: Abhishek Balakrishnan (P'15), Zhenzhen Chen (N'15), Austin Liu (BSE'17), Sharrin Manor (BSE'16), Mindy McTeigue (BSE'17), Raghav Saboo (G'16), Jennifer Cole (N'16)

Co-leaders: K. Bradbury (Energy Initiative), L. Collins (Pratt), T. Johnson (Nicholas), R. Newell (Nicholas, Energy Initiative), K. Raimi (Vanderbilt), S. Palumbo (Facilities)

Abstract

Energy wasted is money wasted. By tightly monitoring and analyzing energy use one becomes more conscious in their energy decisions and in turn can cut down immensely on this waste. Advancements in smart energy meter technology and the rise of extreme data storage capabilities have resulted in household energy monitoring and disaggregation systems that make this level of scrutiny possible. It is not uncommon for these systems to monitor energy use at a rate as fast as one energy reading per second. With this high frequency data, our Bass Connections group has developed and built upon advanced pattern-recognition algorithms that can identify when specific energy appliances are active, how much power these appliances are consuming, and if there is an anomaly in the energy use. Specifically, our system is based on one-second frequency data gathered from the Duke Smart Home. With this automated live process Duke Facilities, which oversees the Smart Home's energy bill, can save money that otherwise would have been lost to wasted energy use or malfunctioning equipment.

System Overview



Our Bass Connections team spent this year constructing an energy disaggregation and anomaly detection system with the intent of automating anomaly detection in energy usage. Specifically, we focused on energy usage of the appliances that use 80% of energy in the Duke Smart Home. With energy use recorded to a SQL database via eGuage units from the Smart Home, we wrote algorithms to determine which appliances are being used at different times and how much energy use each appliances draws. After a brief period of training the algorithms various submetered appliances, all that is required for the algorithm's effectiveness is simply the aggregate total energy usage in the Smart Home rather than independent appliance level data, implying only one affordable monitor needs to be put in place. Furthermore, we can identify when the energy usage is atypical in essentially real time by using a neural net prediction-based model. Our system is a MATLAB-based approach that incorporates some previously developed tools such as a pattern recognition toolbox and interacts with the Smart Home's SQL energy database via Python scripts.

1. Event Detection

The first step to any classification problem is to identify where the actual events to be classified are occurring. In our context, this means clearly labeling on a time series when appliances are turning on and off, as in Figure 1. Our team has implemented a GLR (Generalized Likelihood Ratio)-based event detection scheme. This algorithm first constructs two windows - one globally shifting window (GLR window) and one local window of fixed size around the point of interest. For each point in the GLR window, the GLR is computed as the ratio of the probability distributions of the after local window to the before local window. Next, a test statistic is computed for each point as the sum of the GLR values from that point until the end of the GLR window. Finally, the point with the highest test statistic is given a vote, and the process continues by shifting the GLR window over by one. Our event detection algorithm takes in various statistical variables that allow us to clearly tune the process for desired power levels. These inputs allow us to customize event detection for any appliance in order to obtain the most accurate training data as well as generalize it for a target power consumption threshold with the aggregate power signal.

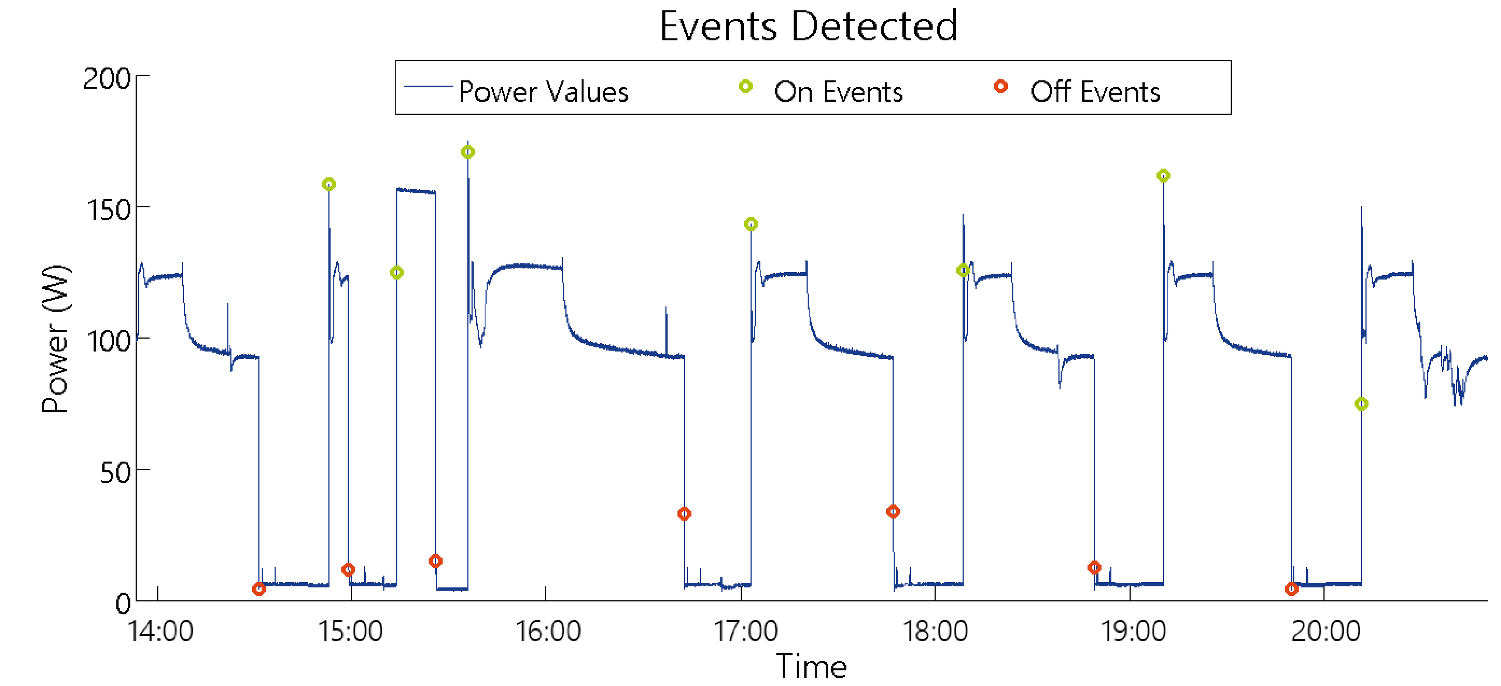


Figure 1: Depiction of event detection algorithm which analyzes an energy signal to detect appliances turning on and off.

2. Live Disaggregation

Once the events have been identified, the next step is to classify these events. We first searched for clear "features" that separate one appliance from another. Since the data is sampled at 1 second intervals, we needed to explore alternatives to methods involving frequency analysis which were heavily referenced in literature. The features we identified to extract from each appliance are the change in power between ON-OFF states and the slope of the curve between ON-OFF states. We ran event detection and feature extraction on hundreds of thousands of submetered time series data for each of our target appliances to construct a training dataset. Then we applied the same event detection and feature extraction to the aggregate signal in real time. Each event's features were compared using a kNN classification algorithm, which classifies an event according to the cluster of training data it is closest to in the feature space. Figure 2 provides a visual of this clustering by features. We used Euclidean distance to construct a radius of high probability in order to keep appliances outside of our target range from being mislabeled. We are able to achieve up to 70% accuracy in classification of ON and OFF appliances, as shown in Figure 3.

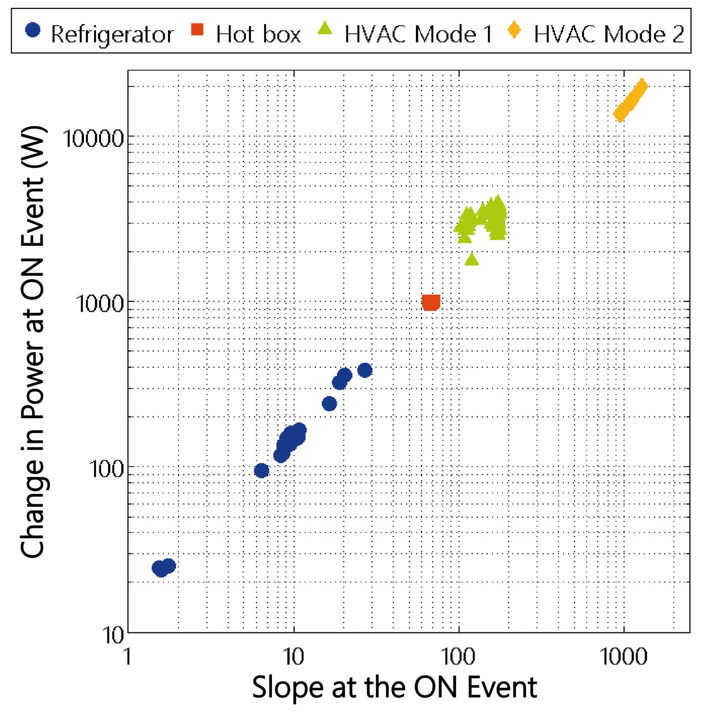


Figure 2: Feature Space of ON events by appliance from submetered data.

Assigned Label

True Label	Other	Refrige-rator	Hot box	HVAC Mode 1	HVAC Mode 2	
Other	77.3	5.7	0	14.8	2.3	[88]
Refrige-rator	34.8	60.9	0	4.3	0	[23]
Hot box	22.6	0	67.7	9.7	0	[31]
HVAC Mode 1	38.5	23.1	0	38.5	0	[13]
HVAC Mode 2	0	0	0	0	100	[2]

70.06% Accuracy of ON Events

Figure 3: Confusion Matrix showing accuracy of appliance disaggregation; diagonal cells show percentage of events correctly classified.

3. Anomaly Detection

After the successful disaggregation of appliances, the classified data is passed to an anomaly detection system—the vital final component to our system. The system was modeled after the system described in the article "Real-time detection of anomalous power consumption" [3]. Our approach creates a hybrid Auto-regressive Neural Net and Auto-regressing Integrated Moving Average model. This ANN-ARIMA hybrid uses past historical sub-metered energy data (only required once) to decide on various prediction parameters.

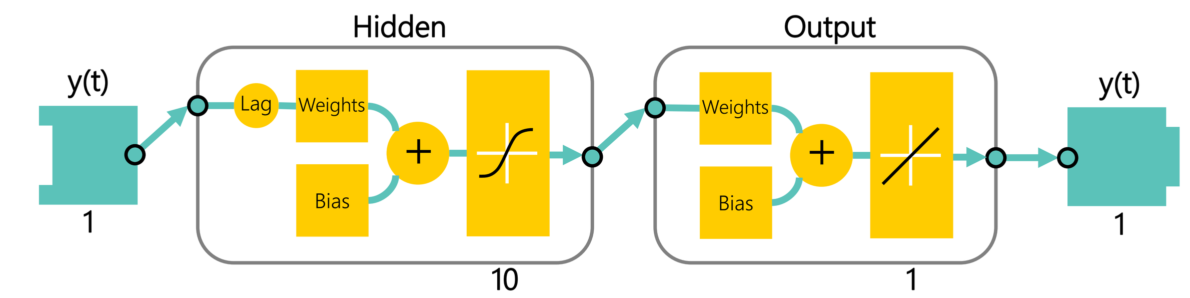


Figure 4: The ARIMA-ANN system model uses historical data to create sophisticated parameters with which to predict future values.

Now the model is ready to predict future values from the recently disaggregated data. In the case of missing data (a substantial issue described in Figure 5), the system interpolates these values. Because so much data was missing prior to January 2015, our prediction system works solely with data beginning in January. As more quality data is gathered, the accuracy of the prediction system will improve and can factor in seasonality.

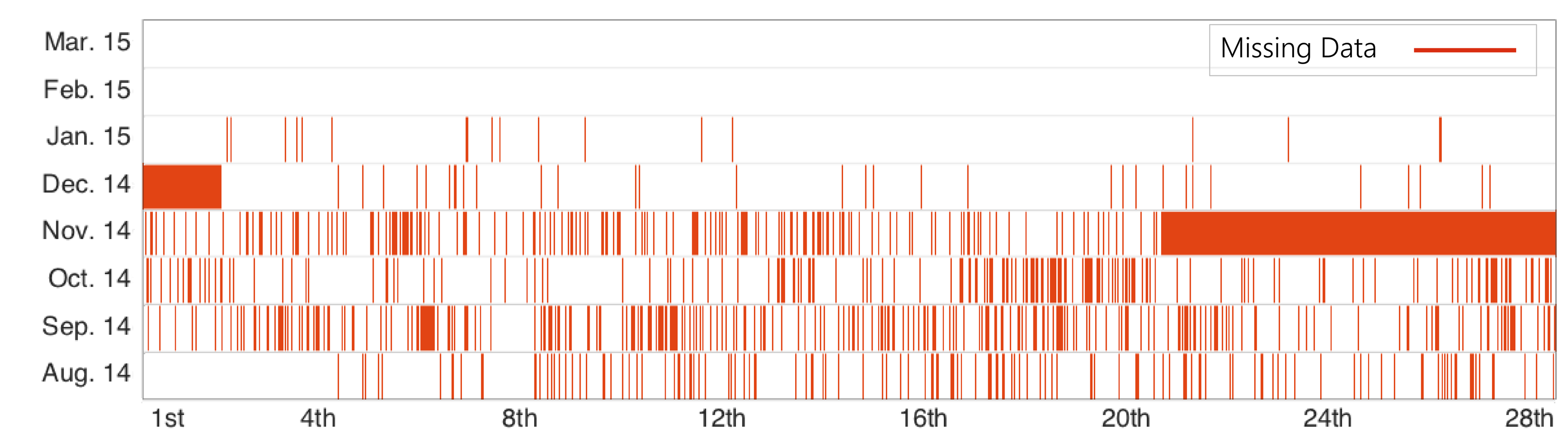


Figure 5: Missing values from the database are illustrated in red. Some instances of missing data are quite severe, but an update to the database hardware in January has nearly eliminated missing data points.

The ANN-ARIMA hybrid model we implemented makes heavy use of MATLAB's "narnet" function to recursively predict the future points and incorporates both linear and nonlinear trends in the data into this prediction. Finally we compare the true energy values to the values predicted and use a difference in 2 standard variations to determine if an anomaly is occurring.

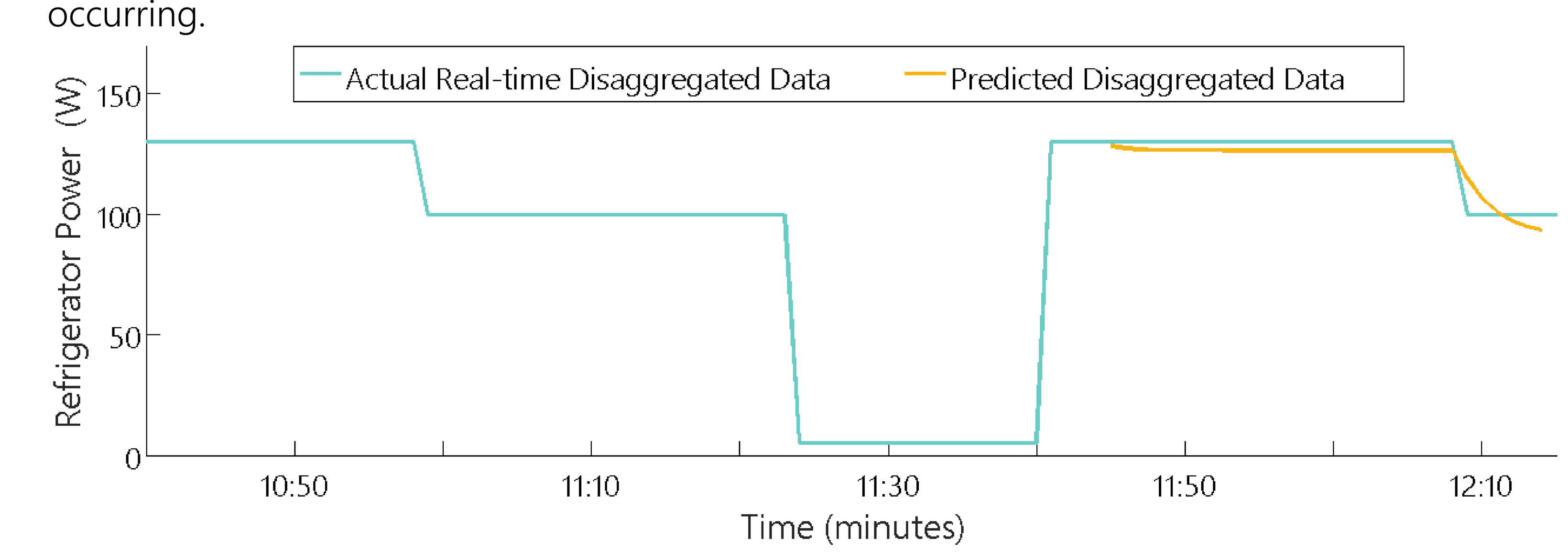


Figure 6: Real-time disaggregated and predicted power values

A log of anomalies is recorded and accessible by the user of the system. When a recent anomaly occurs, the user is notified. With this automated anomaly detection system in place, serious energy use issues can be identified and stopped before they cause serious energy waste.

References

[1] Armel, C., Gupta, A., Shrimali, G., & Albert, A. (2013). Is disaggregation the holy grail of energy efficiency? The case of electricity. *Energy Policy*, 52, 213– 234.

[2] Anderson, K. D., Berges, M. E., Ocneanu, A., Benitez, D., & Moura, J. M. (2012). Event detection for non intrusive load monitoring. *Proceedings of 38th Annual Conference on IEEE Industrial Electronics Society*, 3312–3317.

[3] Chou, J., Telaga, A. S. (2014). Real-time detection of anomalous power consumption. *Renewable and Sustainable Energy Reviews*, 33, 400-411.