Machine Learning and Energy Disaggregation

A Project by Mac Finnie '16, Advised by Professors J. Albrecht and A. Danyluk

Problem

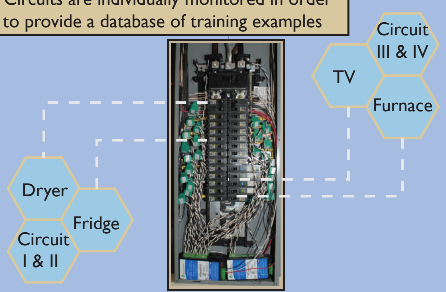
- Capturing circuit-level data is vital to improving energy efficiency
- Manual monitoring is cumbersome and instrusive

Goal

- Our goal is to develop a non-intrusive, Machine Learning based method to detect the use of appliances given only information about total energy usage, day, and time; we are attempting to disaggregate power use.

1 The Data

- We use data obtained from an actual home
- Circuits are individually monitored in order to provide a database of training examples



[Prof. Albrecht's House - Avail. at : egauge.net]

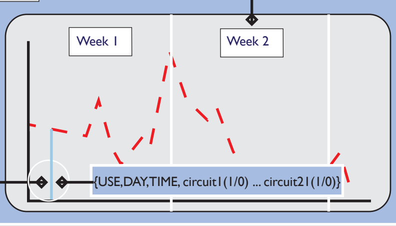
2 Data Formulation

We begin by formulating our data into snapshots (instances) which the learning algorithms can use to build our disaggregation model

Raw wattages from 21 circuits are replaced by binary on/off values

We experiment with 6 weeks of data, grouped by week

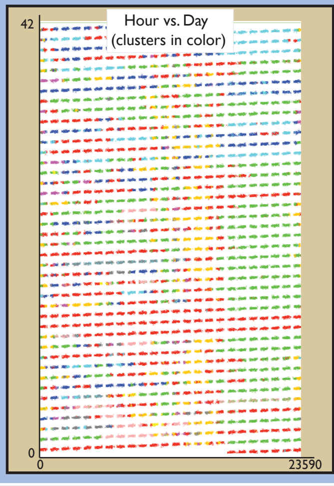
An "Instance"



3 Clustering: *discovering common combinations of appliance usage*

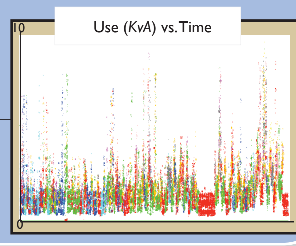
Why?

- With 21 circuits there are 2^21 distinct on/off combinations of circuits
- Some of these combinations are extremely rare or even non-sensical (e.g., A/C and Space Heater both on)



Process

- We use the SimpleKMeans clustering algorithm to find the 20 most common combinations



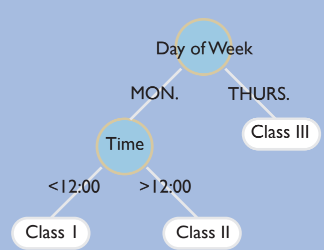
4 Classification: *predicting an appliance combination from aggregate data*

Process

- We build a classifier that will allow us to predict an appliance combination (i.e., a class) based on a snapshot of day, time, and aggregate usage.

- We use a machine learning algorithm (J48 Decision Tree algorithm) to build the classifier from data.

Example Decision Tree



A Hierarchical Classifier Structure

- We developed a hierarchical classifier of 3 decision trees for better accuracy

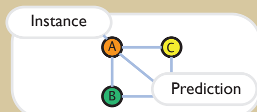
Classifiers:

(A) Train on Week 1, Validated on Week 2

(B) Trained on X errors

(C) Trained on Y errors

[Where X and Y are most often confused classes]



5 Results

Challenges

- A dominant large cluster (no circuits in use) and general class imbalance are a challenge for the learning algorithm.
- Actual usage values for individual circuits make it difficult to identify meaningful clusters.

Phase 1

- Cluster definition improved by replacing circuit usage values with binary on/off calues and using SimpleKMeans to learn clusters.
- Merging similar clusters also led to greater accuracy without the loss of clearly defined clusters.

Phase 2

- The addition of a hierarchical organization for our classifiers yielded the best results; those having good classifier performance *and* clearly defined clusters.

(A) (B) (C)

Correctly Classified Instances
(train week1 / validate week2 / test week3)

(307+5121+2178) = 7606/10080
= 75.5%

6 Looking Forward

An accurate and adaptable energy disaggregation method has wide-ranging applications in energy efficiency:

- optimized background appliance scheduling
- relief of stress on national grid during high-use periods
- personalized energy reports on inefficient usage habits

Acknowledgments:

Many thanks to Jennifer Gossels '13 and her thesis work on the subject of Energy Disaggregation. Additionally, thanks to professors Jeannie Albrecht and Andrea Danyluk for their advice and mentorship.