Fraunhofer Center for Sustainable Energy Systems

An approximate probabilistic approach for event-based disaggregation

Michael Zeifman

1st International Workshop on
Non-Intrusive Load Monitoring
Pittsburgh
7 May 2012

Fraunhofer USA

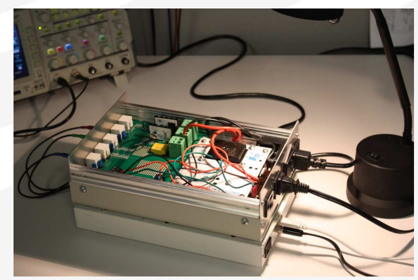
Fraunhofer CSE

- ❑ CSE = Center for Sustainable Energy Systems
- ❑ Located in Cambridge, MA (moving soon to Innovation District in Boston)
- ❑ One of seven Fraunhofer USA centers

Prospective CSE building in Boston

<u>NILM at CSE</u>

- ❑ Interdisciplinary research team (2 behavioral scientists, DAQ engineer, modeling and simulation scientist, interns)
- ❑ High-frequency and low-frequency approaches
- ❑ High-Frequency: combination of several classifiers, need for special hardware (Custom-built DAQ system, National Instruments USB-6251)
- ❑ Low-Frequency: approximate probabilistic approach, OTS sensor (Home Energy Display, e.g., TED – The Energy Detective)



Custom-built DAQ for high frequency sampling

Fraunhofer USA

NILM for HED: User Requirements

- ❑ Features: compatible with 1 Hz, real power only
- ❑ Accuracy: 80-90%
- ❑ No training
- ❑ Real-time capabilities
- ❑ Scalability (up to 20-30 appliances)
- ❑ Various appliance types

NILM: Types of Appliances

- ❑ Permanent

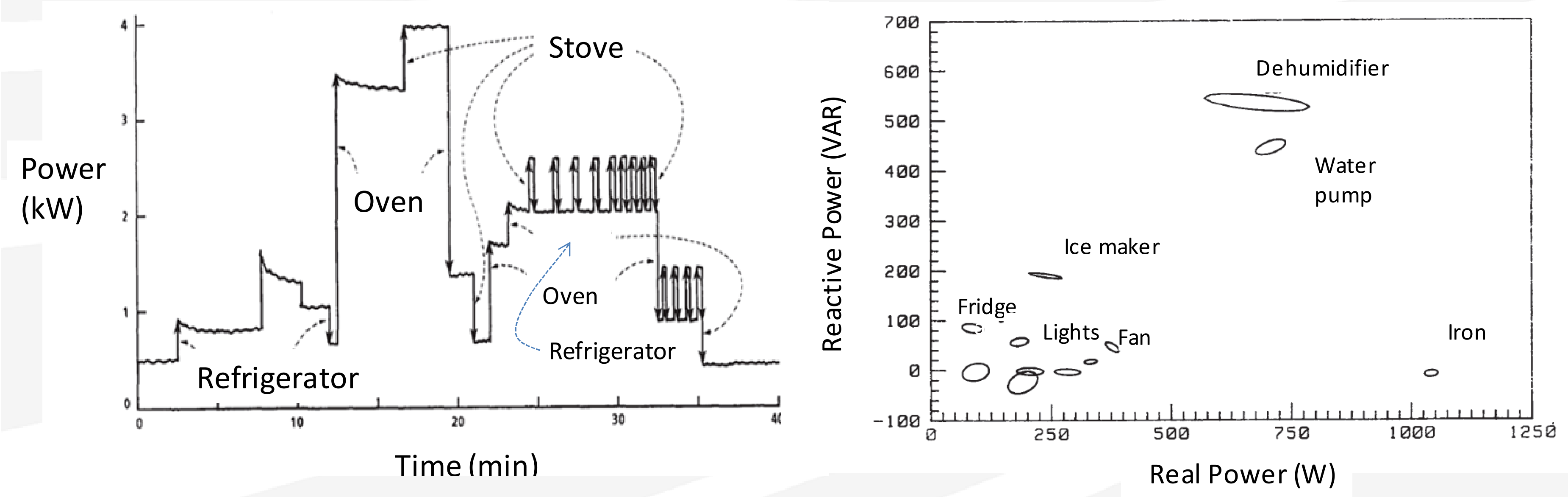
- ❑ On/off

- ❑ Multi State

- ❑ Variable

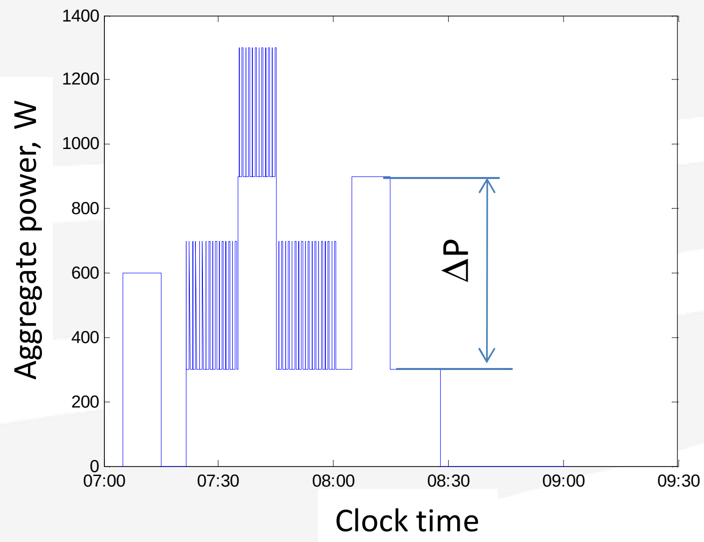
NILM Classic

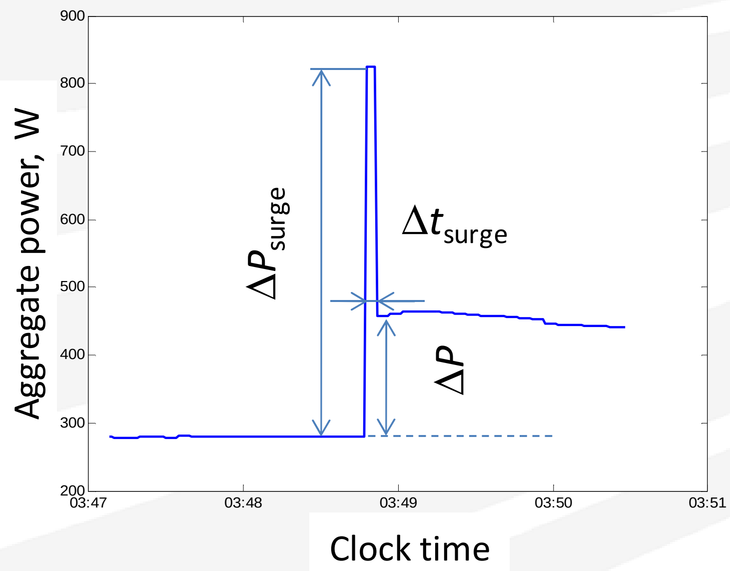
Hart, 1992 (MIT method): What is the problem?



Proposed Method: Features

❑ Power-related



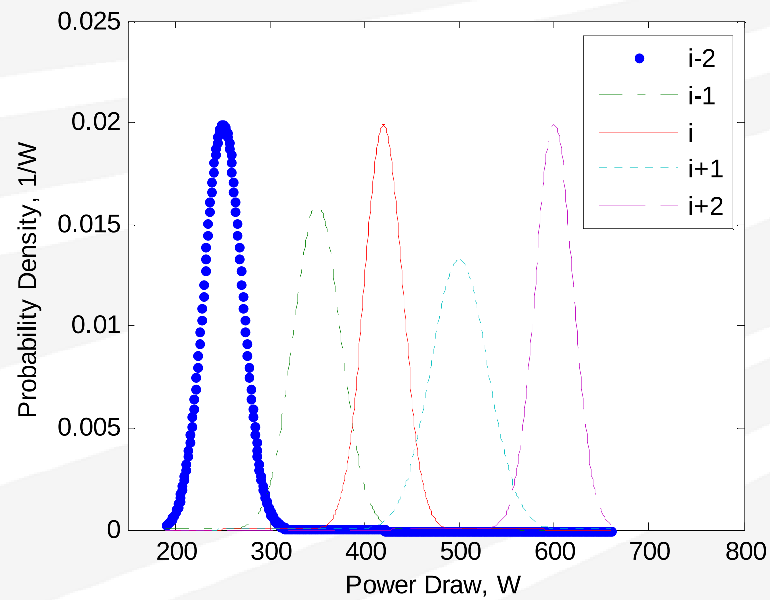


❑ Time-related

- Time-on, time-off statistics
- Can be conditional on time of day

Proposed Method: Key Idea

- ❑ Order appliances by power draw
- ❑ Consider appliance tuplets, e.g., i ±1
- ❑ Track tuplets separately in time
- ❑ Tuplet size – 3, or even 2 (triplet split in two)
- ❑ Foreign and missing data



Transitions within a Tuplet

Tuplet size: 2

System states

STATE	Appliance *i*	Appliance (*i* + 1)
1	OFF	OFF
2	On	Off
3	Off	On
4	On	On

Negative power change, system transitions

TRANSITION BETWEEN STATES	UNDERLYING EVENT(S)
1,1	Is foreign datum
1,2	Was missing datum from state 2 and is foreign datum
1,3	Was missing datum from state 3 and is foreign datum
1,4	Were missing data from states 2 and 3 and is foreign datum
2,1	Appliance *i* is turning off but appliance (*i* + 1) has not turned on
2,2	Is foreign datum
2,3	Was missing datum from state 4 and is 4,3 transition
2,4	Was missing datum from state 4 and is external datum
3,1	Appliance (*i* + 1) is turning off but appliance *i* has not turned on
3,2	Was missing datum from state 4 and is 4,2 transition
3,3	Is foreign datum
3,4	Was missing datum from state 4 and is foreign datum
4,1	Was missing datum from either state 2 or 3 and is 2,1 or 3,1 transition
4,2	Appliance (*i* + 1) is turning off but appliance *i* has not turned off
4,3	Appliance *i* is turning off but appliance (*i* + 1) has not turned off
4,4	Is foreign datum

- Example of transition probability:

$$P_{4,2} \propto p_{i+1}(\Delta P)\, f_{i+1}(\tau_{i+1,on})[1 - F_i(\tau_{i,on})]$$

Modified Viterbi Algorithm

$$\{\hat{s}_t\} = \underset{\{s(t)\}}{\arg\max}[\{s_t\} \mid \{\omega_t)\}]$$

$\{\hat{s}_t\}$ - maximum likelihood estimation of state sequence

$\{s_t\}$ - state sequence

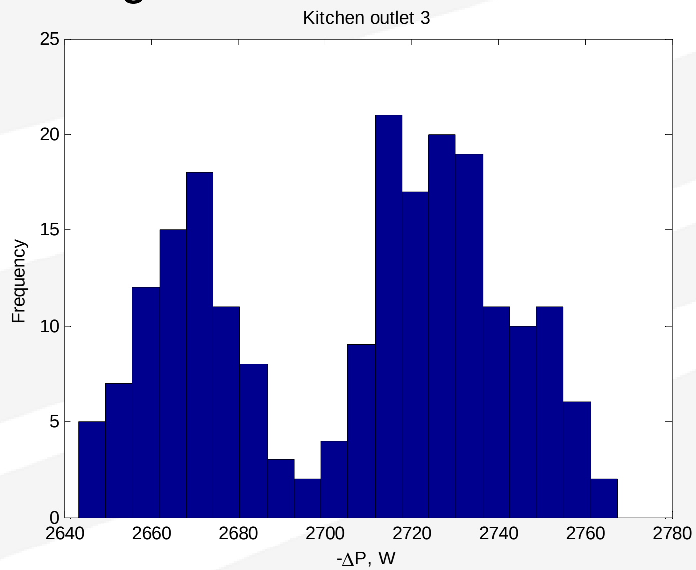
$\{w_t)\}$ - transition observations' sequence

- Developed for 1-st order Hidden Markov models
- Our model can be of higher order
- Keep in memory previous state for each appliance in tuplet

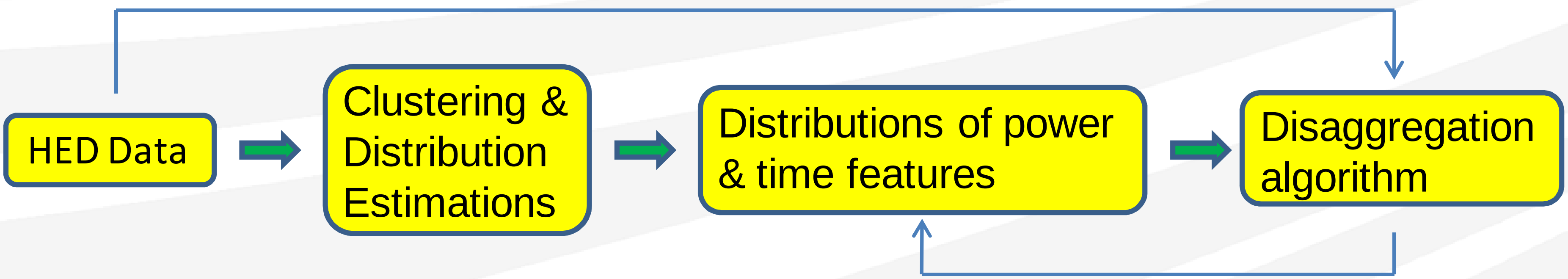
Fraunhofer USA

Estimation of Power and Time Distributions

- ❑ Historical data (~ two weeks)
- ❑ Clustering negative power changes (ISODATA)
- ❑ Matching negative and positive changes, power surges
- ❑ Estimation of time-on and time-off durations
- ❑ Statistical modeling of all distributions



High-Level Block Scheme



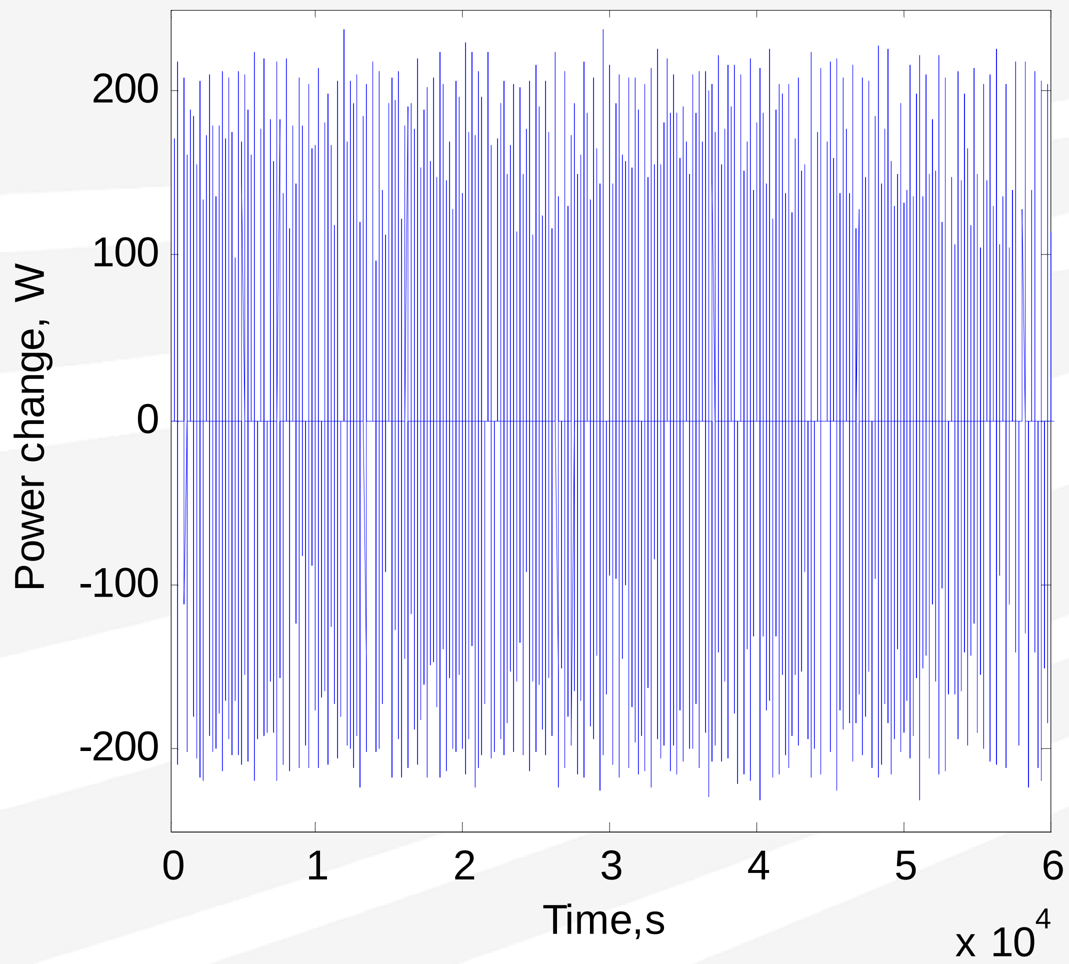
Simulation Example

- 5 on-off overlapping appliances
- Normal (Gaussian) distributions of power changes
- Uniform distributions of time-on and time-off

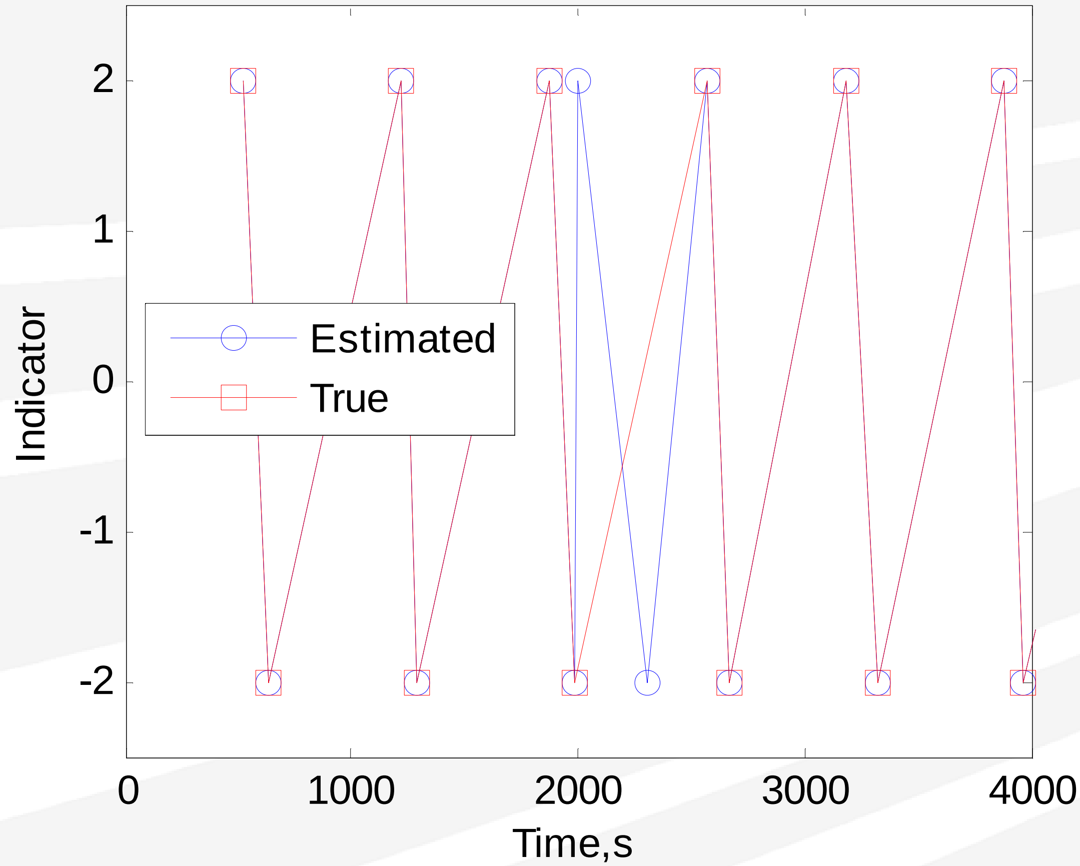
SIMULATED APPLIANCE	MEAN Μ, POSITIVE CHANGE, W	MEAN Μ, NEGATIVE CHANGE (ABSOLUTE VALUE), W	STANDARD DEVIATION Σ, W
1	110	105	10
2	130	135	13.5
3	150	160	10
4	180	190	13.5
5	210	210	10

SIMULATED APPLIANCE	TIME-ON, MINIMUM, S	TIME-ON, MAXIMUM, S	TIME-OFF, MINIMUM, S	TIME-OFF, MAXIMUM, S
1	30	100	400	600
2	70	150	500	600
3	100	180	350	500
4	150	270	200	400
5	40	140	300	700

Simulated Time Series



Simulated Time Series: Disaggregation Results



Indicator = ±2 → Appliance 2 is on(off)

Fraunhofer USA

Simulated Time Series

- Accuracy metric: F-measure

$$F = 2\frac{\frac{TP}{TP+FP}\cdot\frac{TP}{TP+FN}}{\frac{TP}{TP+FP}+\frac{TP}{TP+FN}}$$

SIMULATED APPLIANCE	BAYESIAN CLASSIFIER	OUR ALGORITHM, NO TIME-OFF STATISTICS USED	OUR ALGORITHM
1	0.859	0.910	0.985
2	0.666	0.835	0.865
3	0.749	0.850	0.925
4	0.743	0.865	0.970
5	0.859	0.950	0.965

Fraunhofer USA

Real Household Data: Appliance Characteristics

- ❑ Kolter and Johnson (2011) collected data in six households in Massachusetts
- ❑ Submetered on individual circuit level
- ❑ Selected 1 household with 9 "good" circuits recorded over 26 days

CIRCUIT (APPLIANCE)	POSITIVE, MEAN, W	POSITIVE, STANDARD DEVIATION, W	NEGATIVE, MEAN, W	NEGATIVE, STANDARD DEVIATION, W
1 Refrigerator	204.3	3.9	179.5	2.4
2 Dishwasher	217.3	20.6	212.1	19.7
3 Kitchen outlet 1	1084.8	14.5	1074.6	15.1
4 Microwave	1548.5	52.4	1504.9	43.9
5 Kitchen outlet 2	1543.5	21.1	1533.5	14.5
6 Bathroom GFI	1613.2	18.6	1607.3	18.4
7 Oven 1	1651.4	22.4	1640.7	21.1
8 Oven 2	2474.9	29.5	2448.4	24.2
9 Kitchen outlet 3	2767.9	51.3	2706.3	33.1

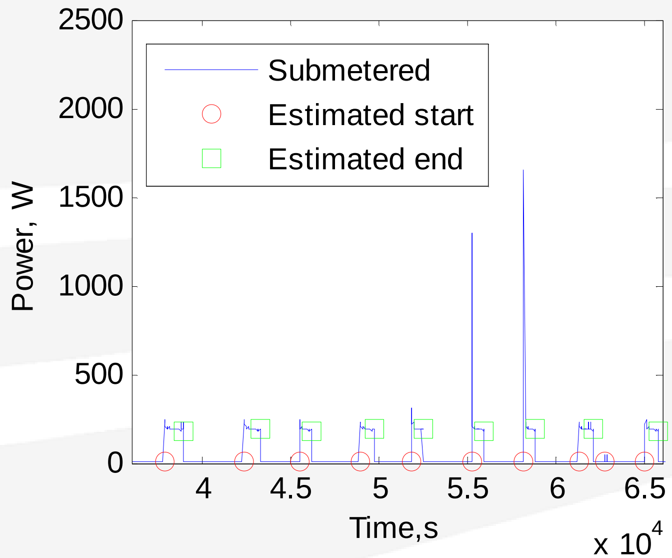
W → V·A (apparent power)

Fraunhofer USA

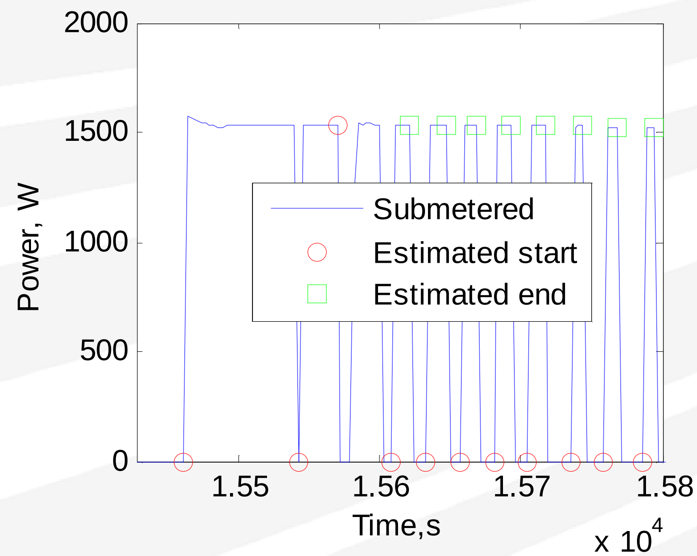
Real Household Data: Disaggregation Results

APPLIANCE FROM TABLE VI	BAYESIAN CLASSIFIER	OUR ALGORITHM
1 Refrigerator	0.859	0.831
2 Dishwasher	0.881	0.846
3 Kitchen outlet 1	0.989	0.936
4 Microwave	0.775	0.899
5 Kitchen outlet 2	0.409	0.840
6 Bathroom GFI	0.753	0.927
7 Oven 1	0.800	0.908
8 Oven 2	1.0	0.962
9 Kitchen outlet 3	1.0	0.971

Real Household Data: Disaggregation Results



1. Refrigerator



5. Kitchen outlet 2

Needs Improvement

- ❑ Fully develop triplet
- ❑ Better pre-processing (filtering, incremental power changes)
- ❑ Better change detection (change-point problem)
- ❑ More advanced clustering procedure
- ❑ Matching between found clusters and real appliances (use of estimated time-on and time-off features)

Conclusions

Use Requirements	Our Method
❑ Features: compatible with 1 Hz, real power only	❑ Yes
❑ Accuracy: 80-90%	❑ Accuracy: 80-90%
❑ No training	❑ Learning from historical data. Matching with real appliances forthcoming
❑ Real-time capabilities	❑ Yes
❑ Scalability (up to 20-30 appliances)	❑ Yes. Complexity linear with number of appliances
❑ Various appliance types	❑ Not yet, but in the process