Computational Sustainability: Smart Buildings

CS 325: Topics in Computational Sustainability, Spring 2016

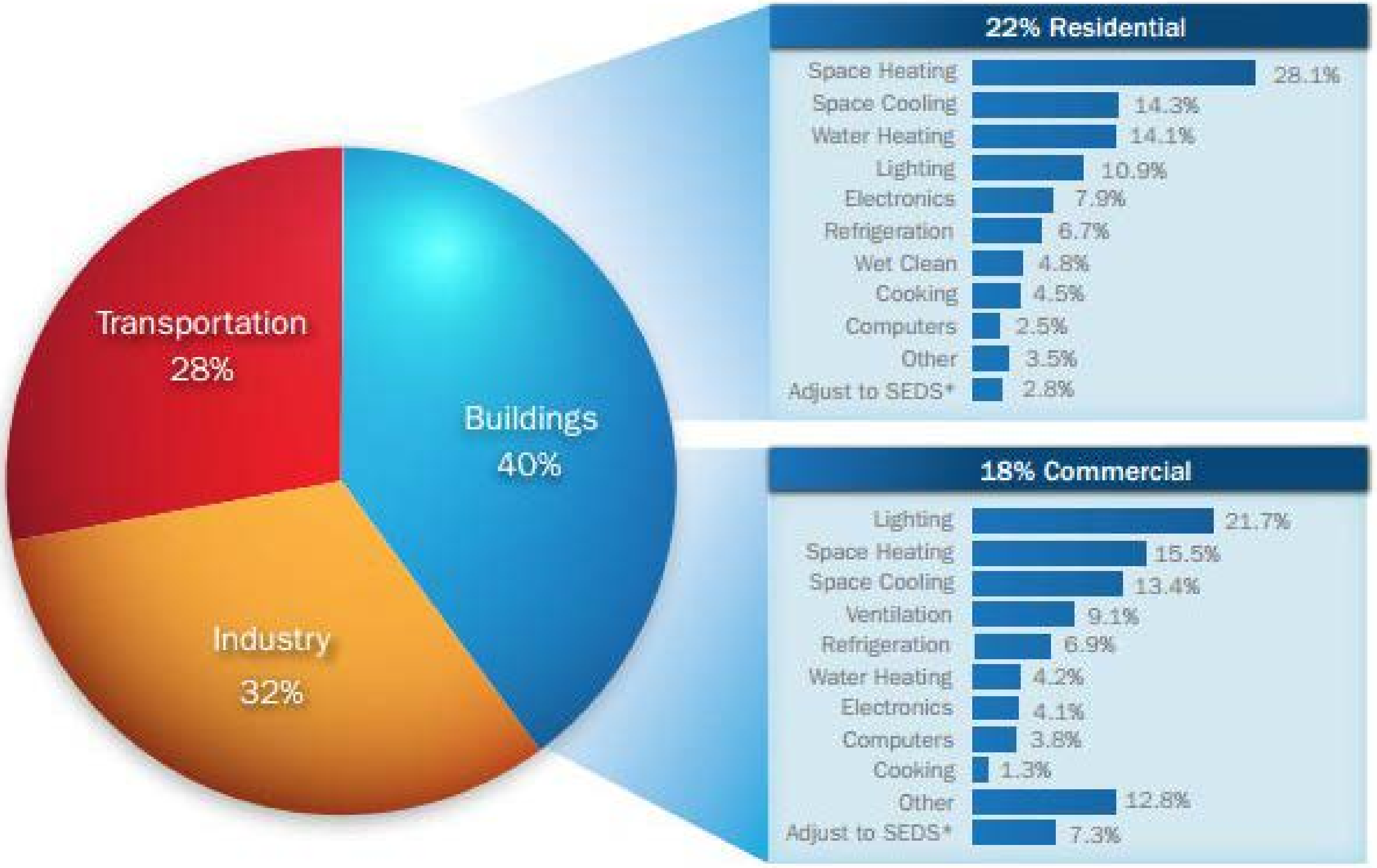
Manish Marwah

Senior Research Scientist

Hewlett Packard Labs

manish.marwah@hpe.com

Building Energy Use



* Energy adjustment EIA uses to relieve discrepencies between data sources. Energy attributes to the commercial buildings sector, but not directly to specific end-users.

Buildings consume 40% of primary energy. Of that, 22% is consumed in residential buildings (dominated by space heating) and 18% in commercial buildings (dominated by lighting).[145]

http://energy.gov/sites/prod/files/ReportOnTheFirstQTR.pdf

Building Energy Management

Buildings consume a lot of energy

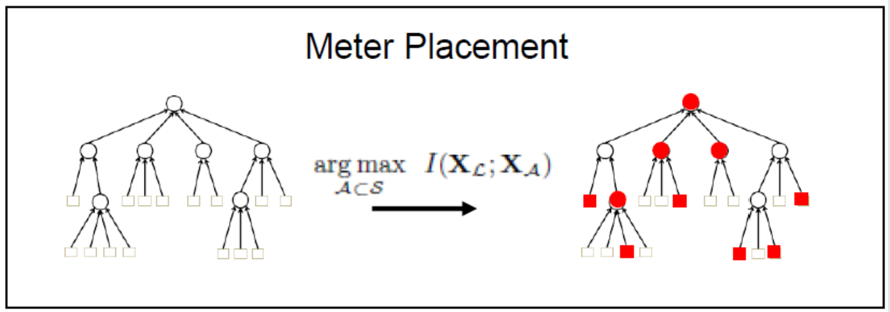
- Commercial buildings
  - 1.3 trillion kWh electricity annually ➔ 1/3 of total US electricity generation
- Annual energy costs > $100 billion

Poorly maintained, degraded, and improperly controlled equipment wastes 15-30% energy in commercial buildings

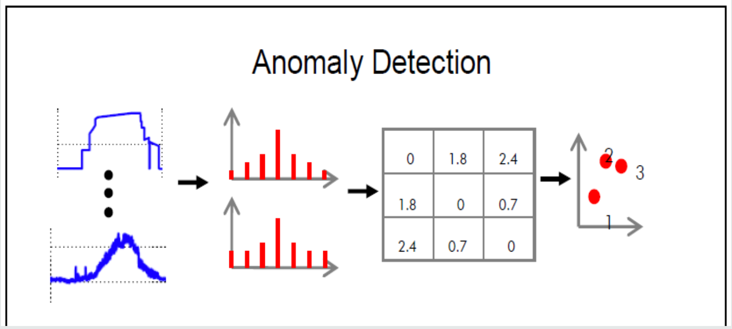
Outline

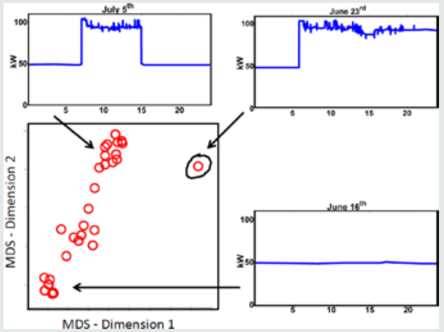
- Meter placement
- Anomaly detection
- Occupancy Modeling
- Energy Disaggregation

Where should meters be installed?

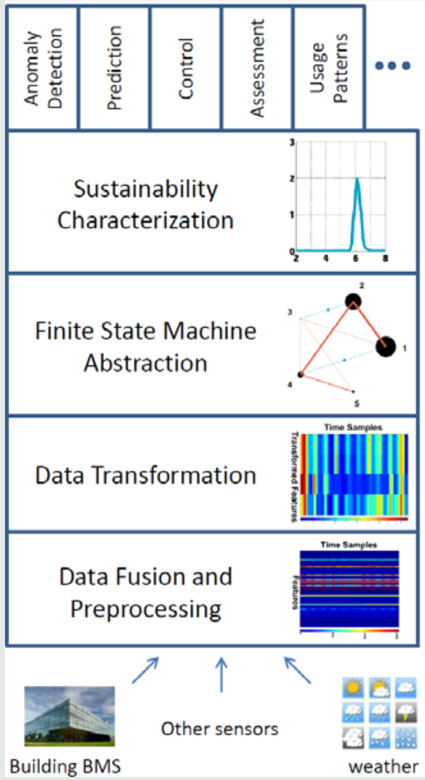


How can we detect anomalous power consumption behavior?

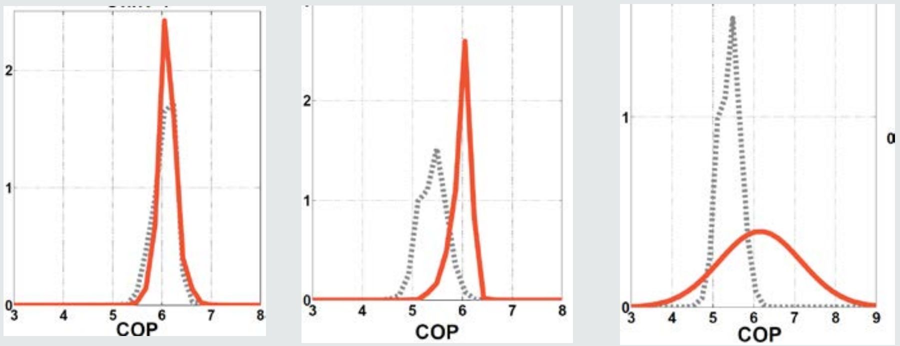




How can we detect degraded performance of equipment/devices in a buildings?



How can we cheaply measure building occupancy?

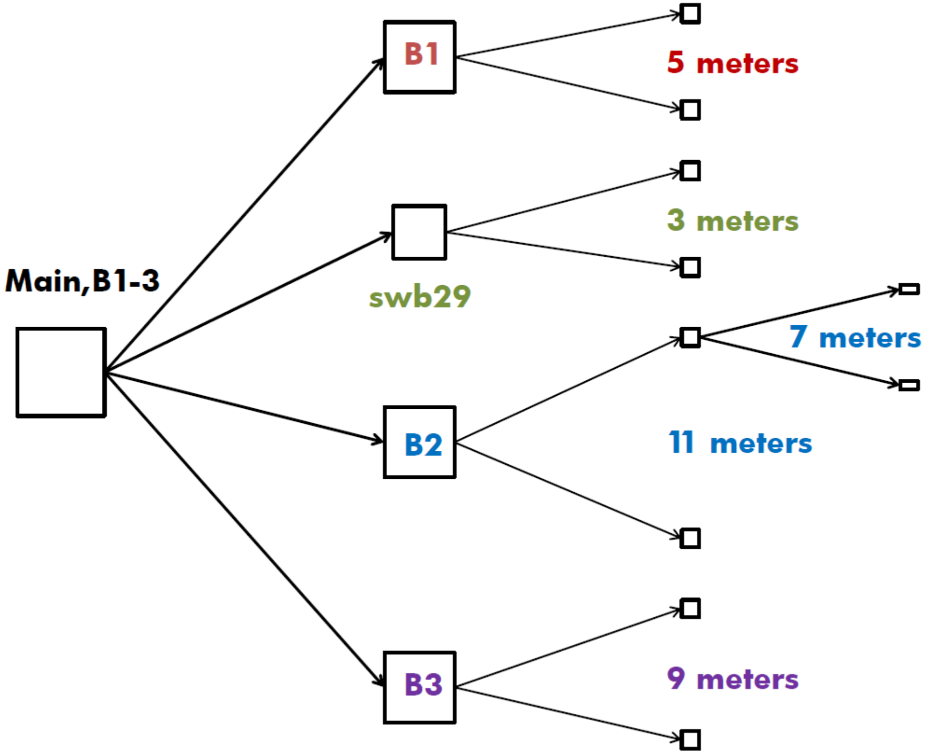


Normal Abnormal Abnormal

Ref.: KDD 2012, ACM BuildSys 2011, 2012

Test Bed

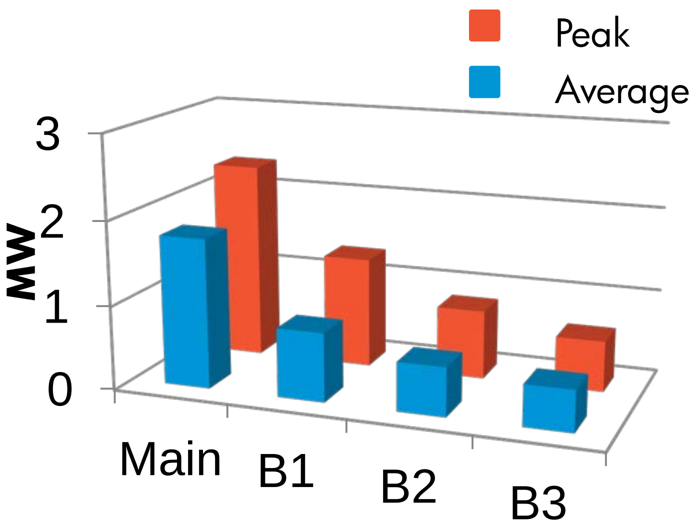
- HP Labs, Palo Alto, CA campus
- Three buildings instrumented with ~40 power meters

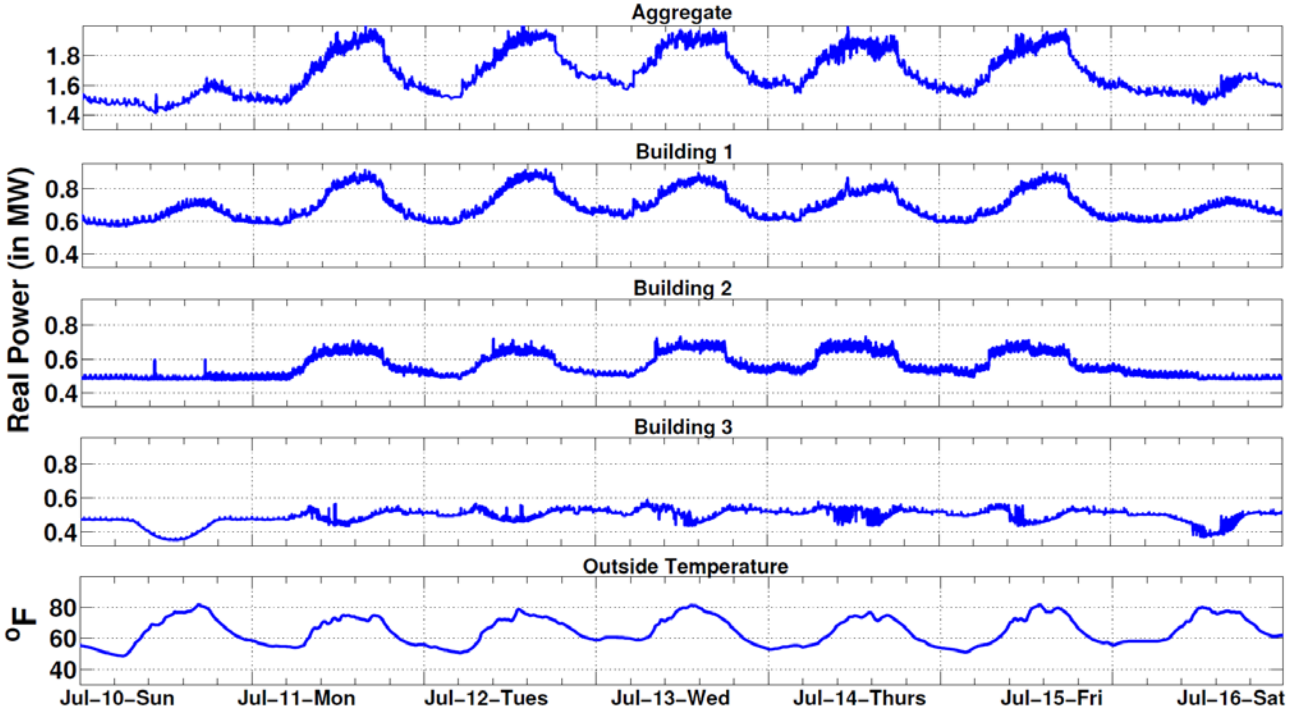


Electrical Infrastructure Topology

Campus Power Use

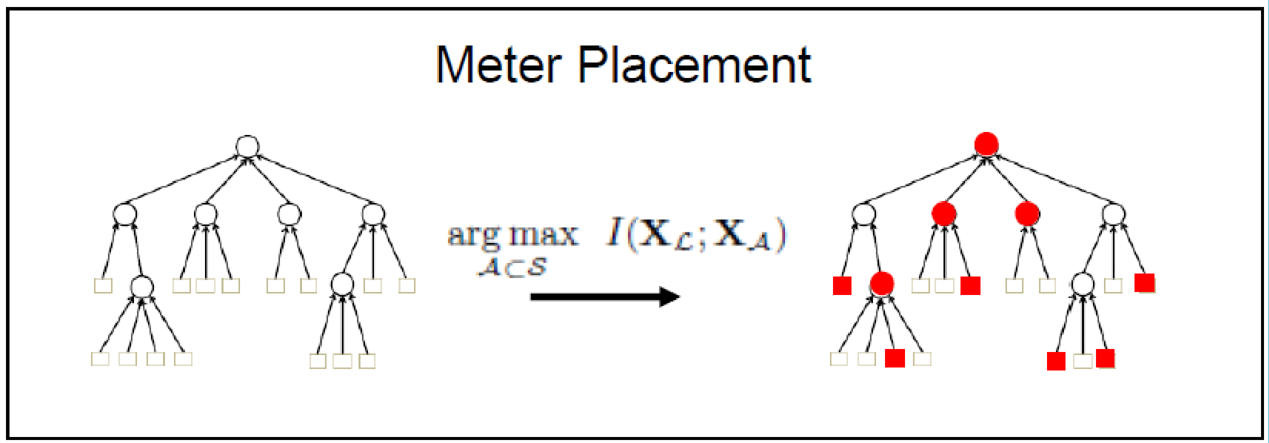
- Power consumption characteristics of Buildings 1, 2 and 3
- Building 3 has a 135kW PV array





Building Power Instrumentation

Where do I place the meters?



Building Power Instrumentation

Motivation: Obtain per-panel power consumption

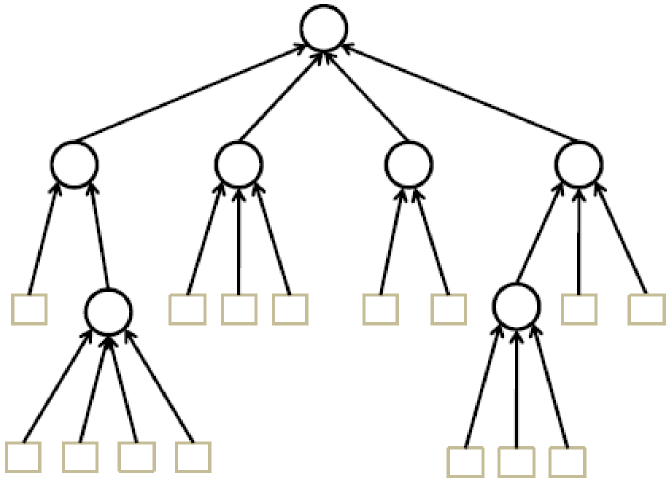
Challenge: Large number of panels, each power meter: $1K-$3K

Goal: Select optimal locations for meter deployment

Approach: Formulate as an optimization problem over panel hierarchy (a tree structure)

Panel Topology & Problem Formulation

Panel Topology



Panel feeding multiple sub-panels

Panel feeding load(s)

Problem Formulation

Select *k* meters:

$$\arg\max_{\mathcal{A}\subset\mathcal{S}} \quad I(X_{\mathcal{L}}; X_{\mathcal{A}})$$
$$\text{s.t.} \quad |\mathcal{A}| \leq k$$

X : Set of meters

$\mathcal{S}$: Set of all possible locations &

$\mathcal{L}$: Set of all leaf locations

$\mathcal{A}$: Selected locations

Greedy, Near-optimal Solution

- Optimal solution is **NP-hard**

- **Greedy optimization:** Select panels sequentially

$$j^* = \arg\max_{j \notin \mathcal{A}} I(\mathbf{X}_{\mathcal{L}}; \mathbf{X}_{\mathcal{A} \cup j}) - I(\mathbf{X}_{\mathcal{L}}; \mathbf{X}_{\mathcal{A}}).$$

- We show objective function is **submodular** [KDD 2012]

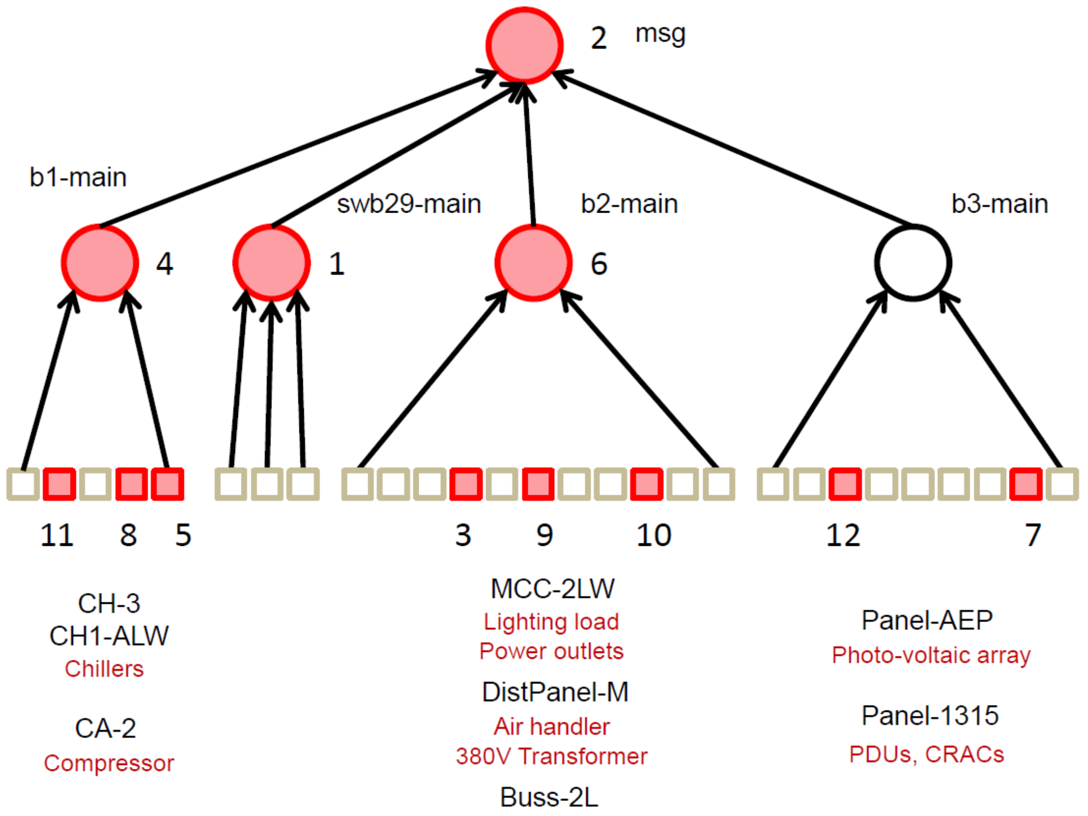
- Thus, solution is guaranteed to be **near-optimal** [Krause *et al. 2006]*

$$I(\mathbf{X}_{\mathcal{L}}; \mathbf{X}_{\mathcal{A}_{\text{greedy}}}) \geq \left(1 - \frac{1}{e}\right) I(\mathbf{X}_{\mathcal{L}}; \mathbf{X}_{\mathcal{A}_{\text{opt}}})$$

~63%

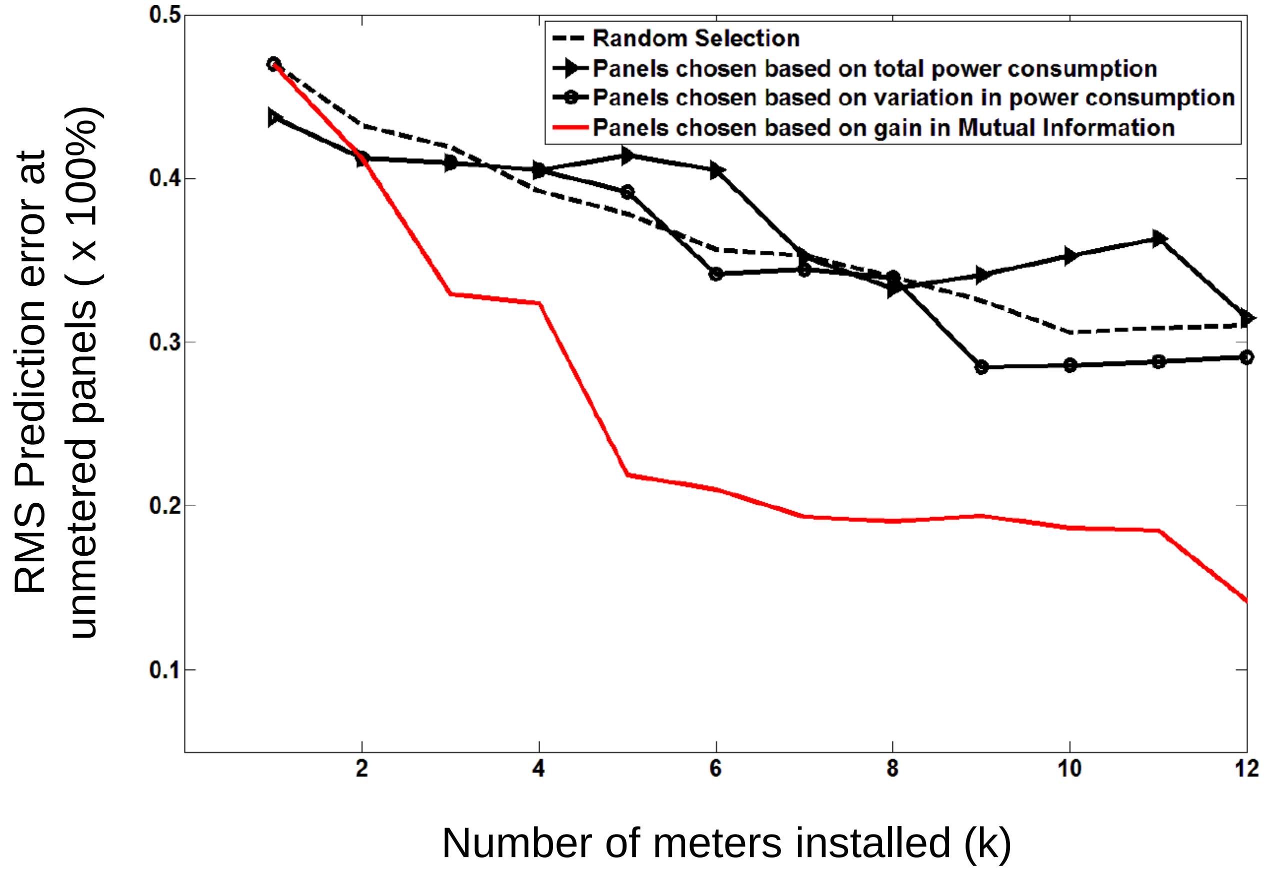
Experimental Results

Panels Selected for k = 12



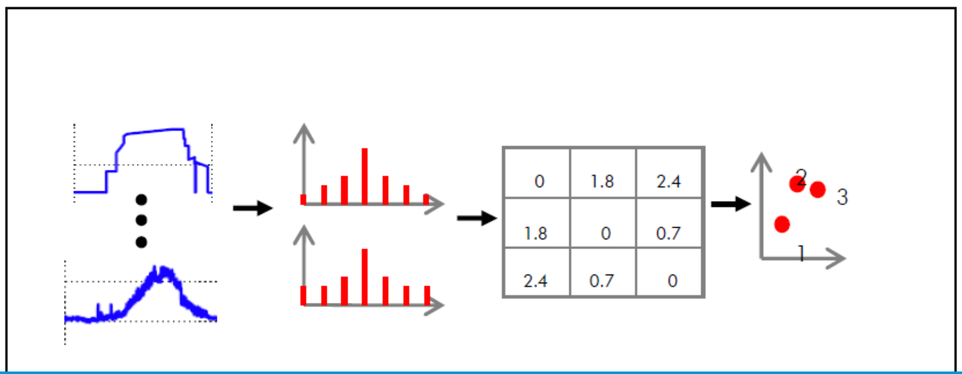
Experiment Results

Prediction ability of the panels selected using the proposed approach



Building Power Management

Meter Anomaly Detection



Anomaly Detection

Motivation:

- Abnormal power usage may indicate:
  - wasted power
  - Failed or faulty equipment

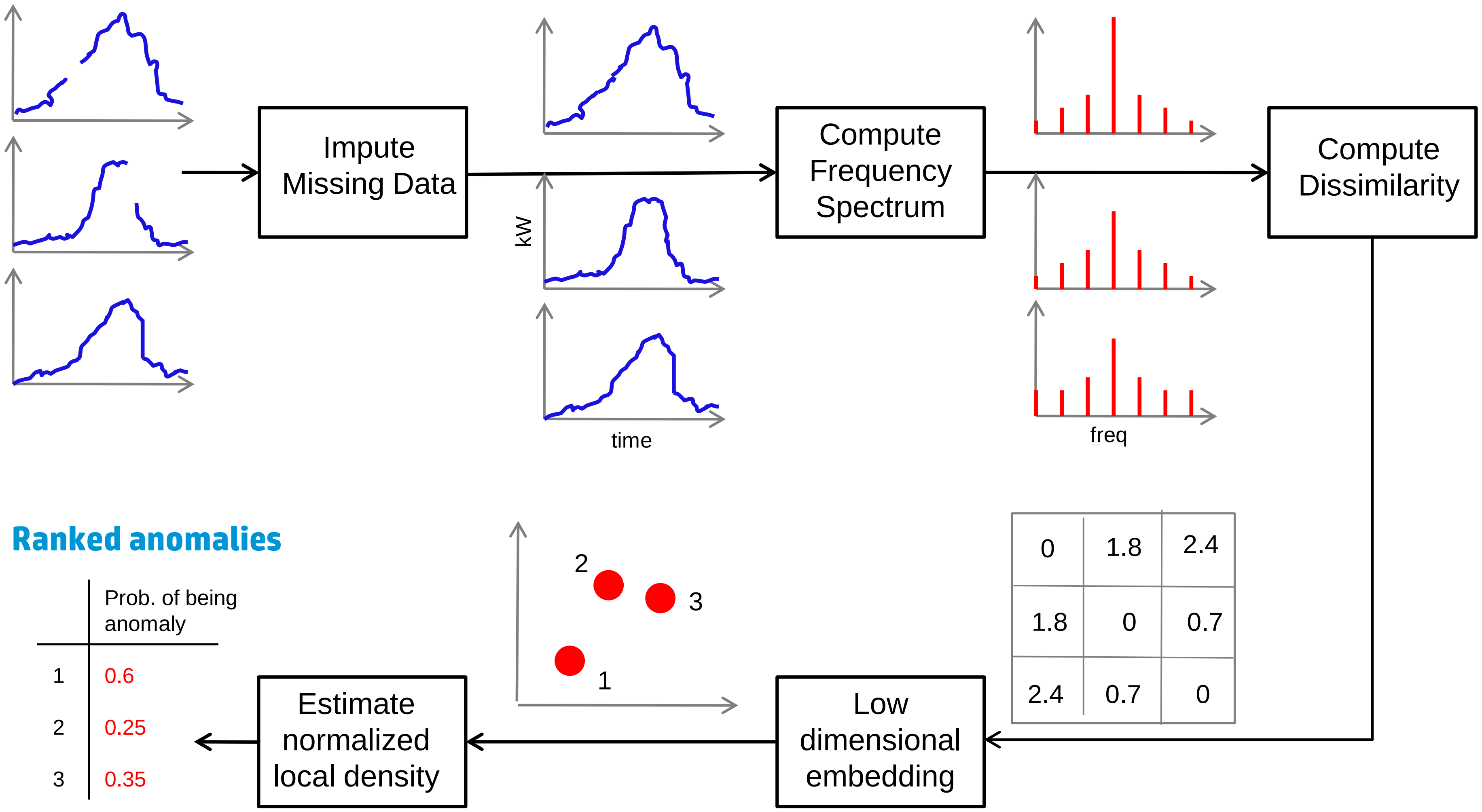
Challenge: Obtaining labeled data is expensive

- requires a lot of manual effort

Goal: Systematically detect abnormal power usage

Approach: Use an unsupervised approach

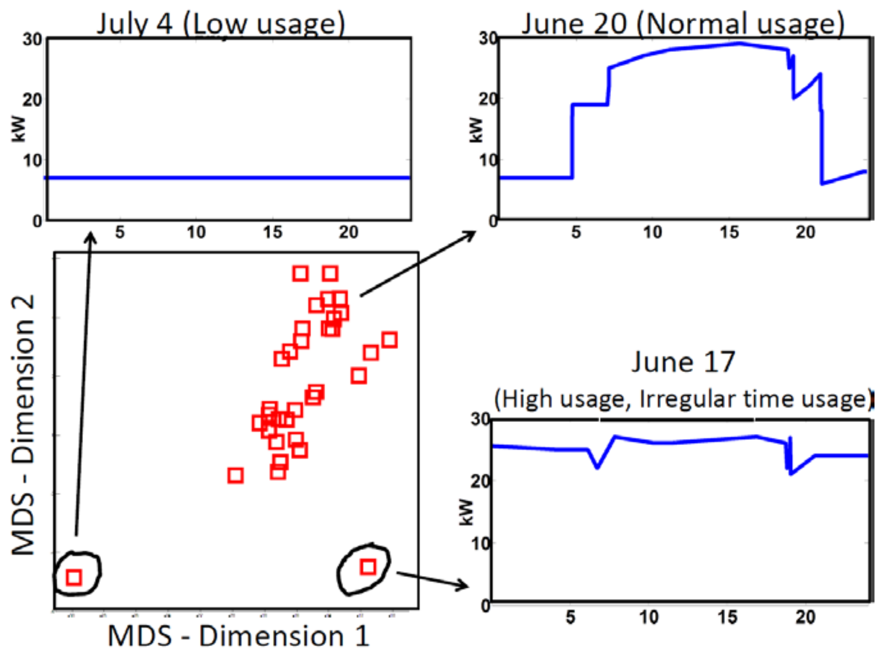
Algorithm


Impute
Missing Data
kW
time
Compute
Frequency
Spectrum
freq
Compute
Dissimilarity
0 1.8 2.4
1.8 0 0.7
2.4 0.7 0
Low
dimensional
embedding
2
3
1
Estimate
normalized
local density
Ranked anomalies
Prob. of being anomaly
1 0.6
2 0.25
3 0.35

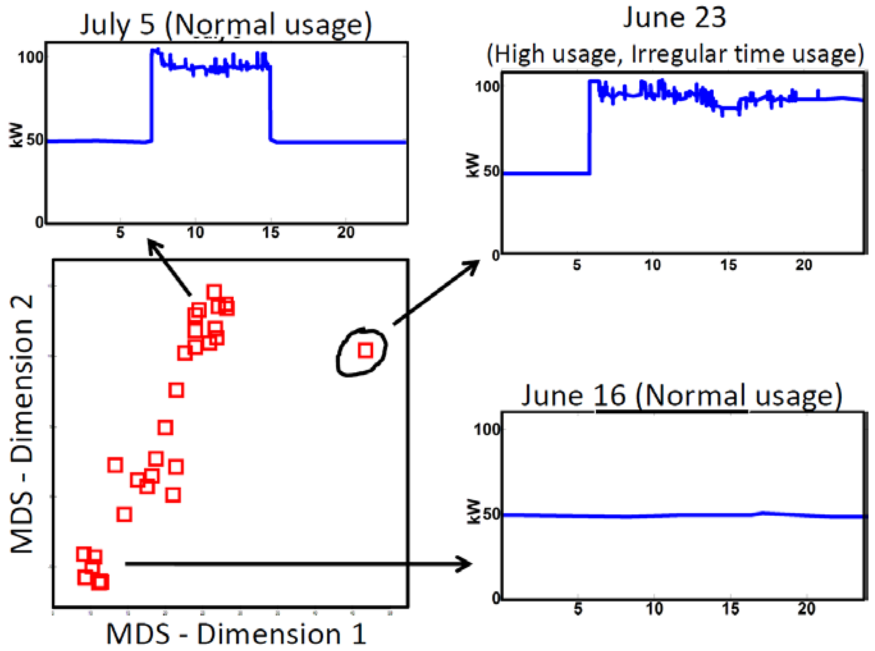
Anomaly Examples

Power Saving Opportunities

Load: Overhead Lighting in Building 1
Anomalies:
- Abnormal low usage (holiday)
- Abnormal time usage; Potential savings ~180 kWh

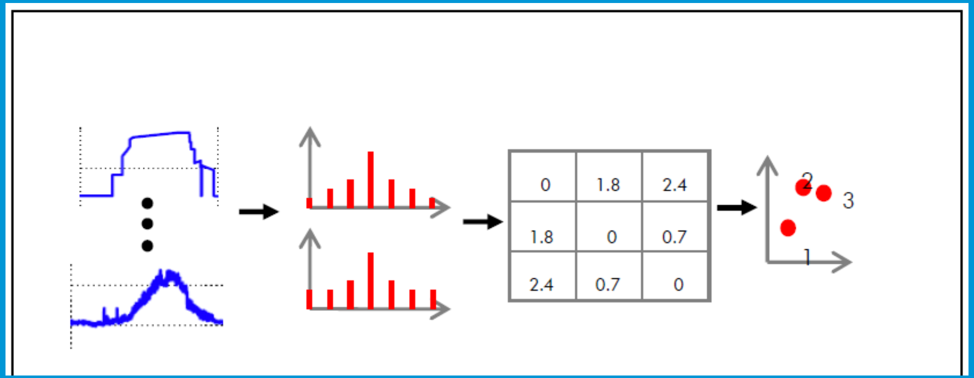


Load: Air Handling Units in Building 2
Anomaly:
- Abnormal time usage; Potential savings ~450 kWh



Building Power Management

Occupancy Modelling



Building Occupancy Estimation

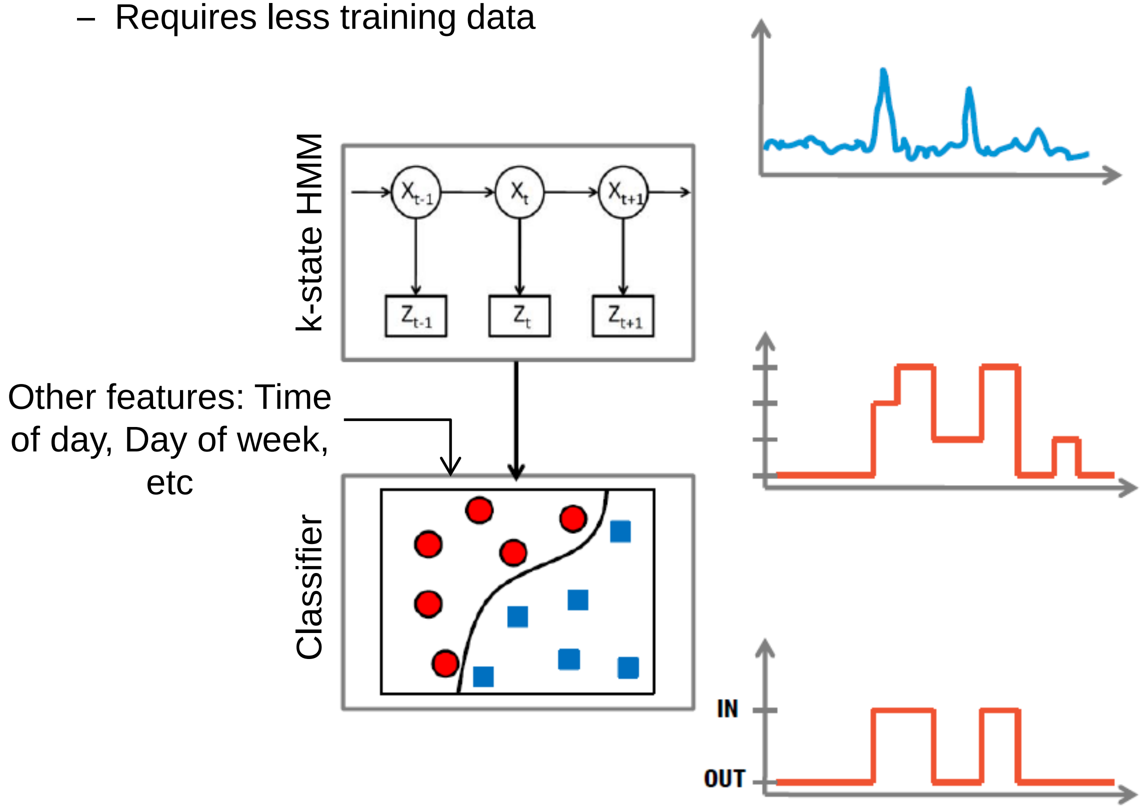
For optimal resource provisioning

- **Motivation:** Save energy via occupancy-based lighting/air conditioning (HVAC) scheduling
- **Challenge:** Fine-grained occupancy information is not available, and requires additional sensors
  - Expensive
  - Intrusive
- **Goal:** Accurately estimate occupancy of a zone using readily available data
- **Approach**
  - Use L2 port-level network statistics as a proxy
  - Semi-supervised method with minimal training data

Methodology[2]

Two stage Semi-supervised Approach

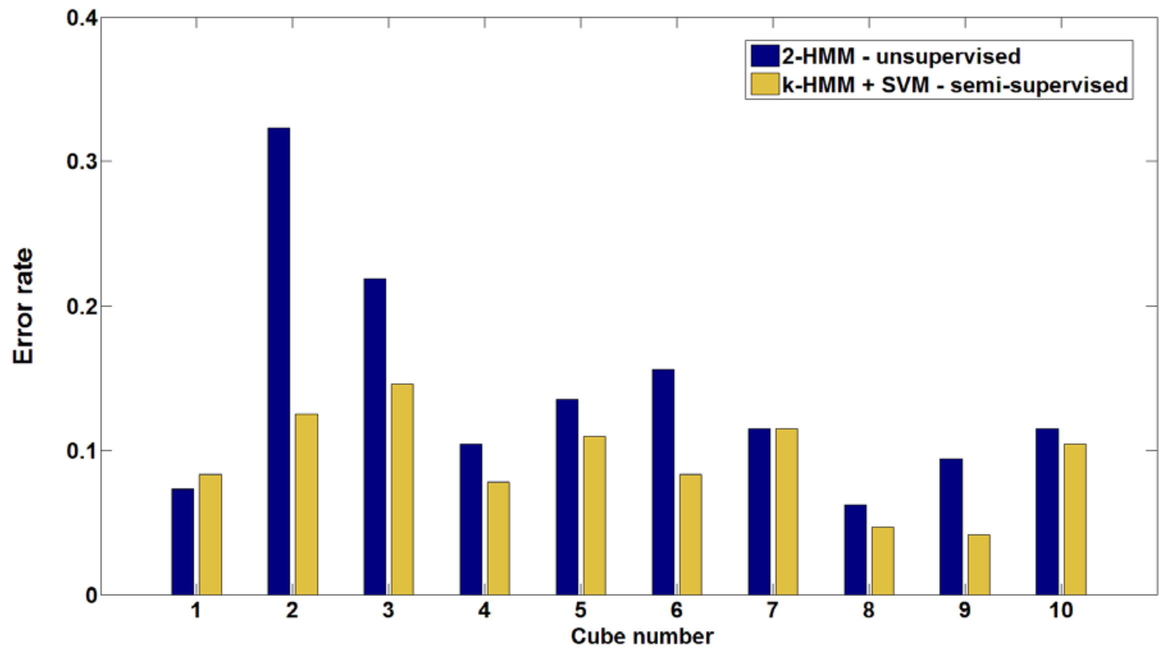
- Can efficiently incorporate external parameters
- Requires less training data



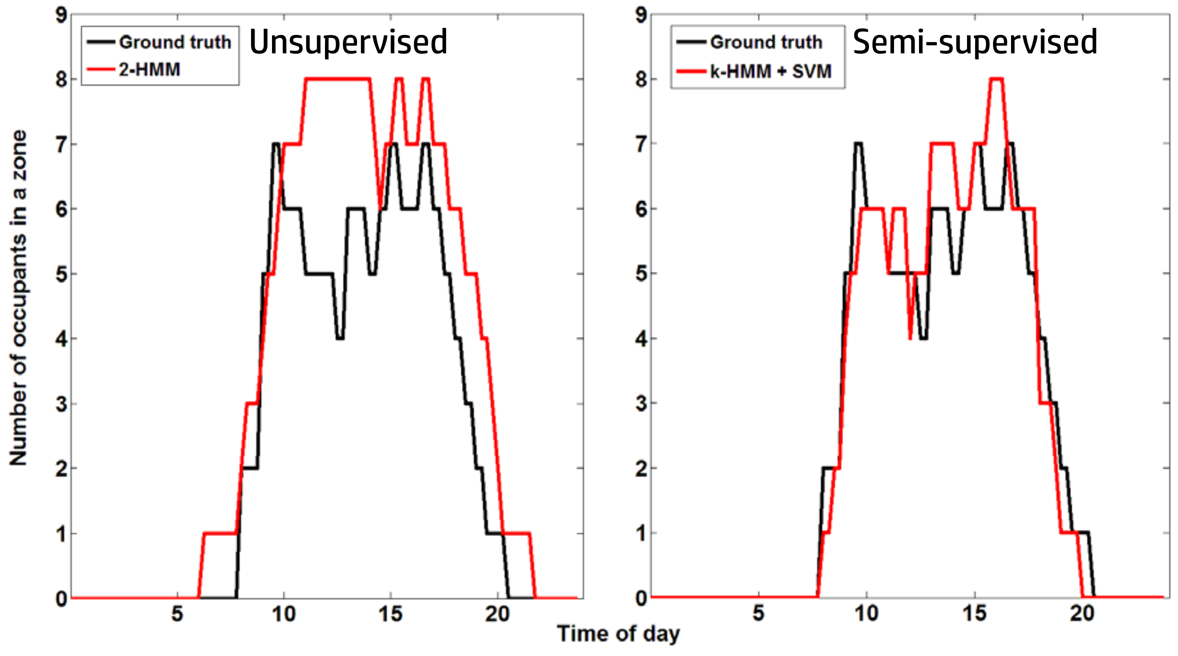
[2] Bellala et.al., "Towards an understanding of campus-scale power consumption," ACM BuildSys 2011.

Experimental Results

Cube/Office-level Occupancy Estimation



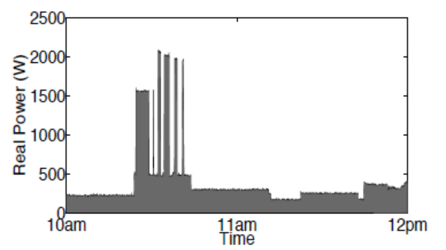
Zone-level Occupancy Estimation

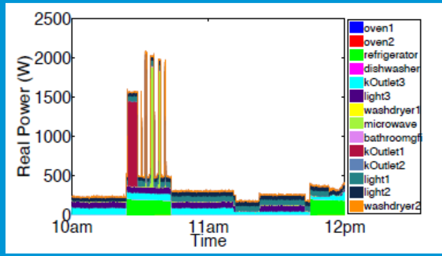


- Occupancy is estimated at cube level (accuracy varied from 85% to 95%)
- This information is aggregated at zone level (8-12 cubes)
- Zone level estimated occupancy is then used to schedule lighting for each zone
- Estimated energy savings using this approach ~ 9.5%

Building Power Management

Energy Disaggregation





Residential Energy Consumption

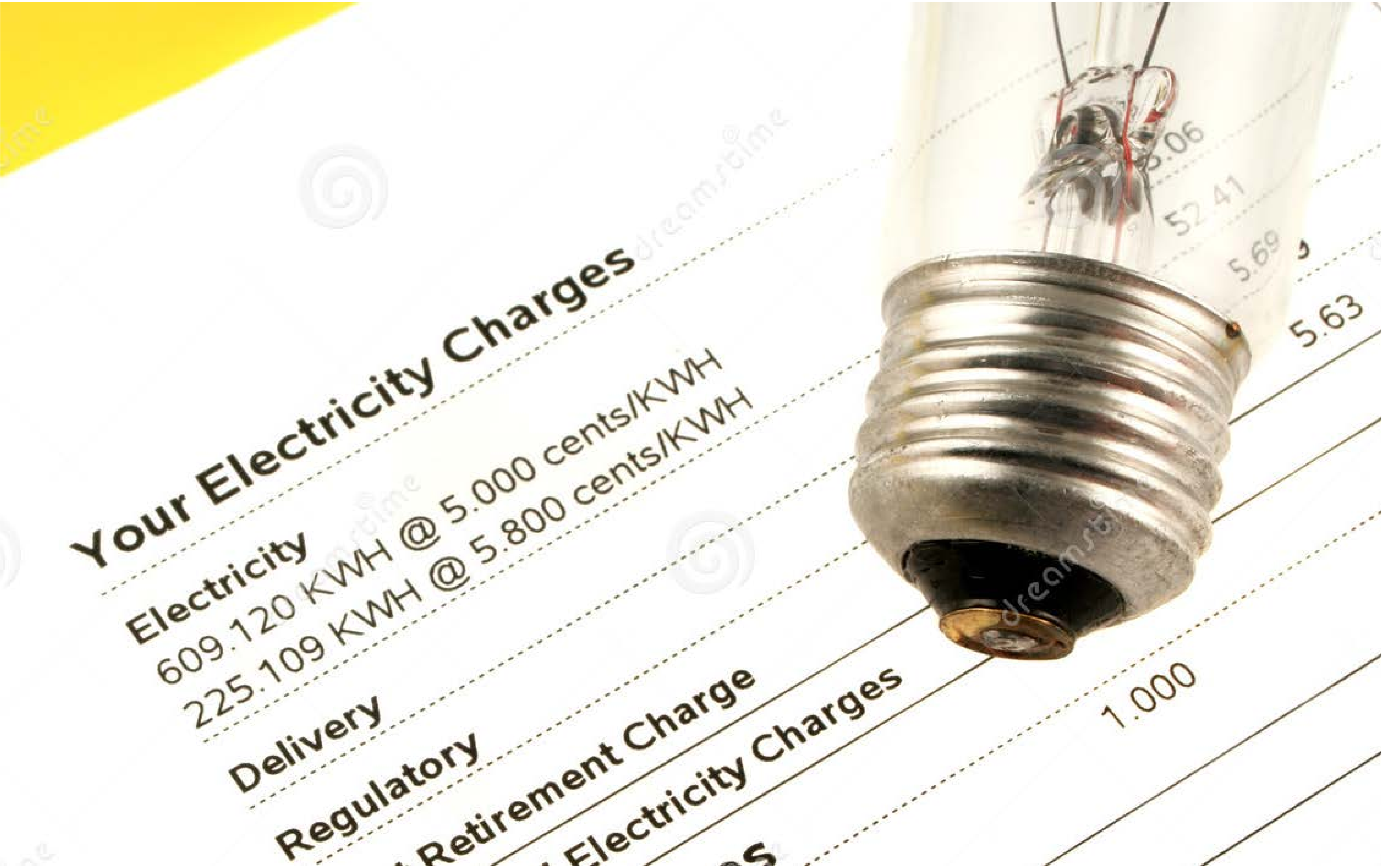
"… the typical American household … is also likely to **use 20 percent to 30 percent more energy** than necessary…"

ACEEE, a non-profit advocacy group

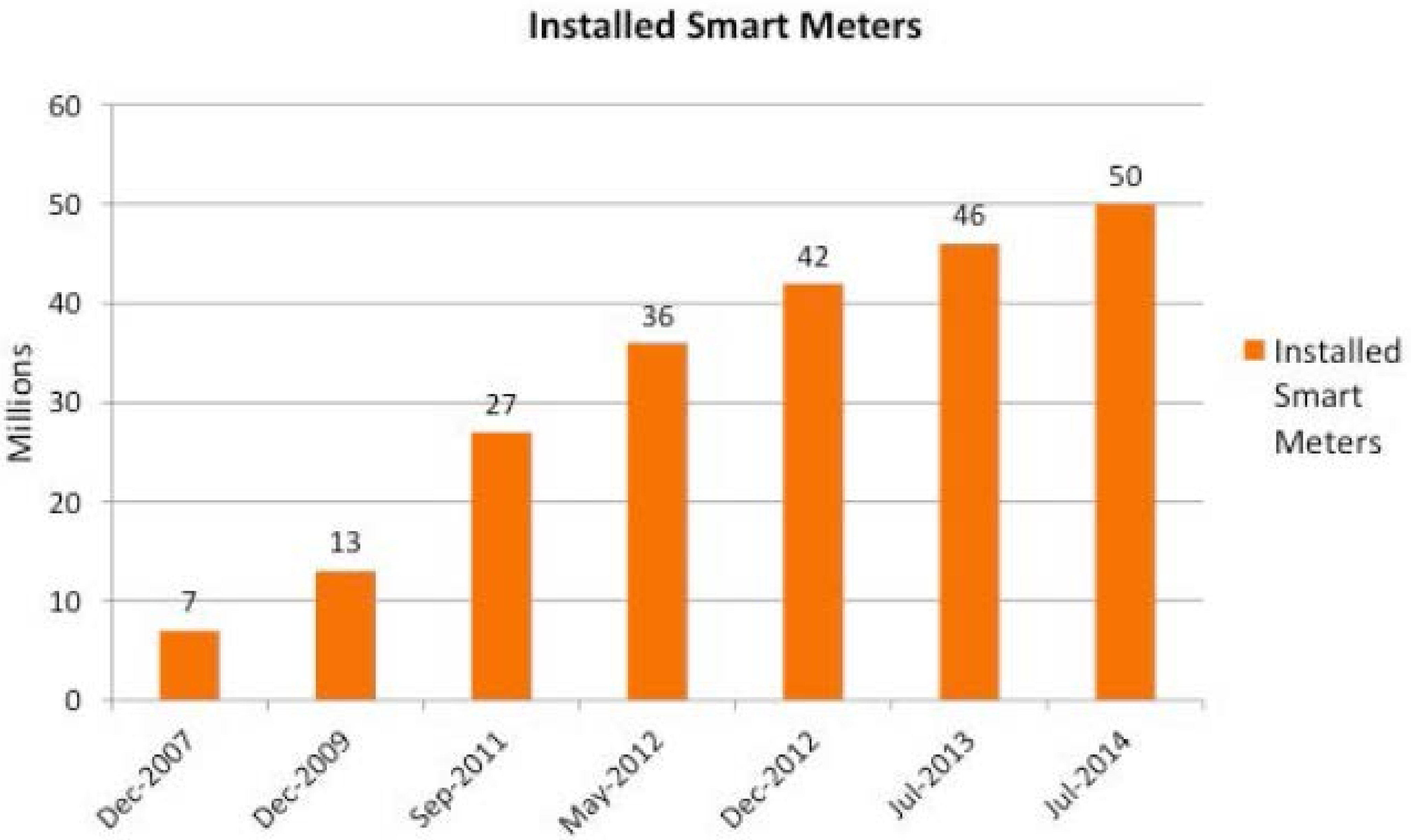
"… Americans could cut their electricity consumption by 12 percent and **save at least $35 billion** over the next 20 years"

ACEEE, a non-profit advocacy group

http://www.withoutagym.net/wp-content/uploads/2014/02/LOWEST-GROCERY-BILL-EVER1.jpg



http://thumbs.dreamstime.com/z/electricity-bill-1565154.jpg

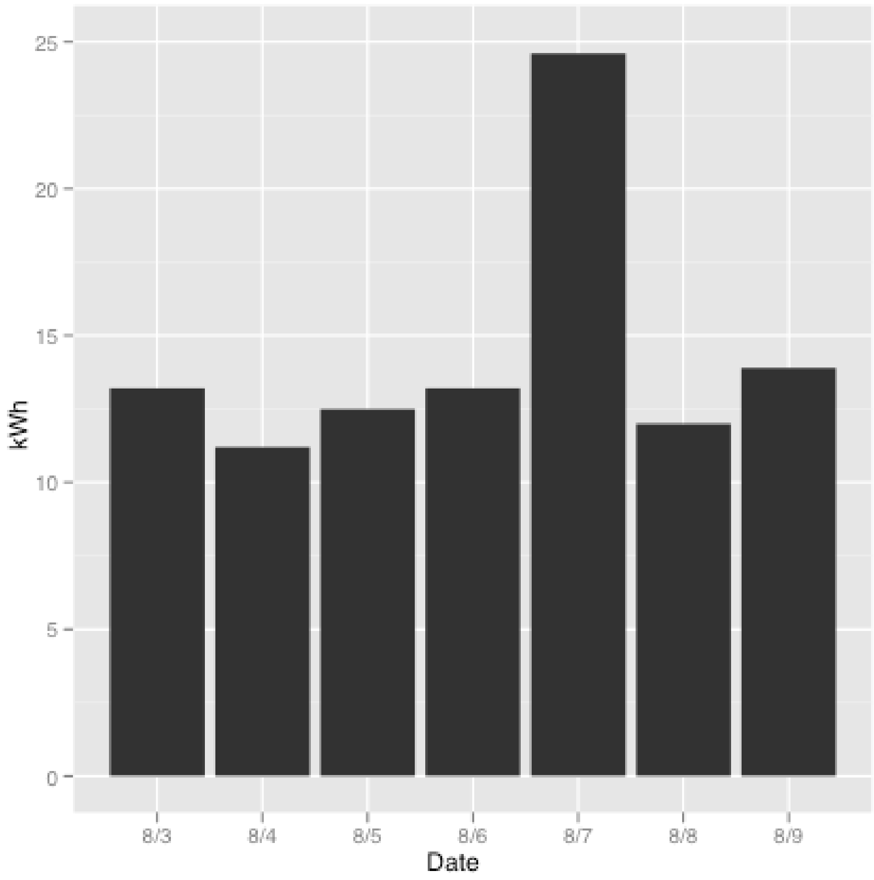


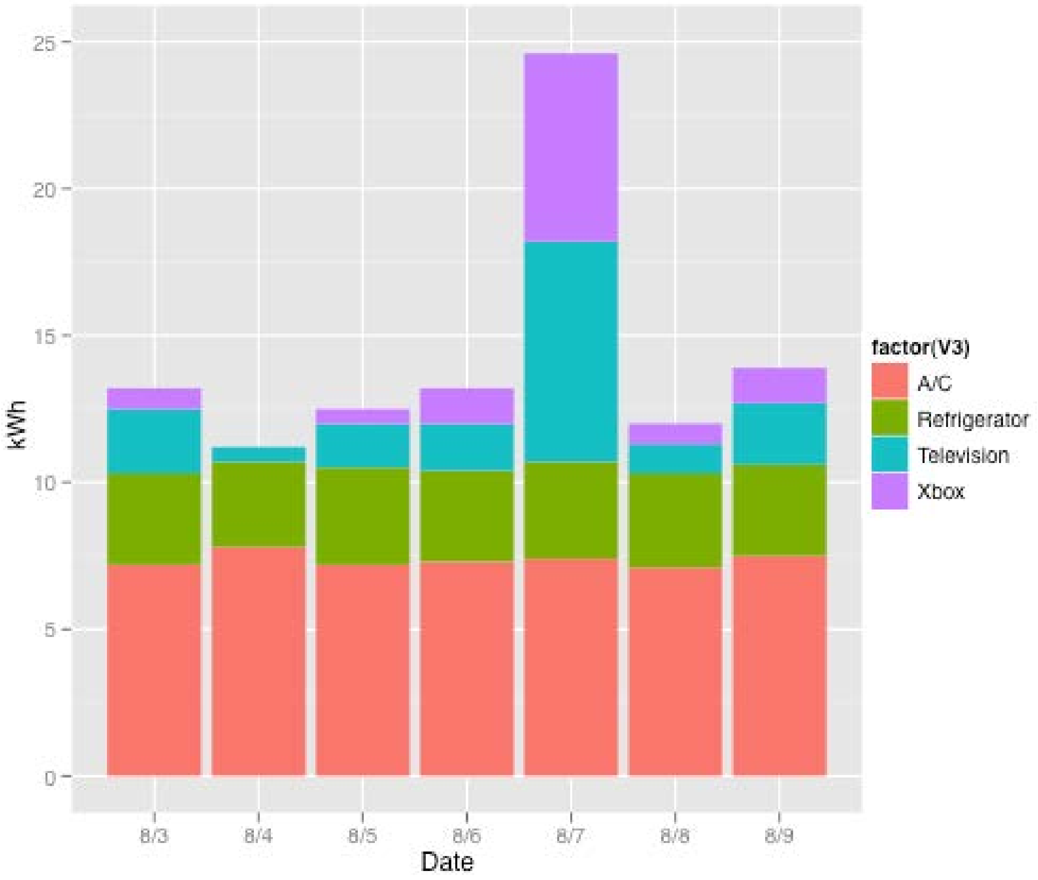
http://www.edisonfoundation.net/

GO BEYOND SMART METERS

– Give customers breakdown of consumption

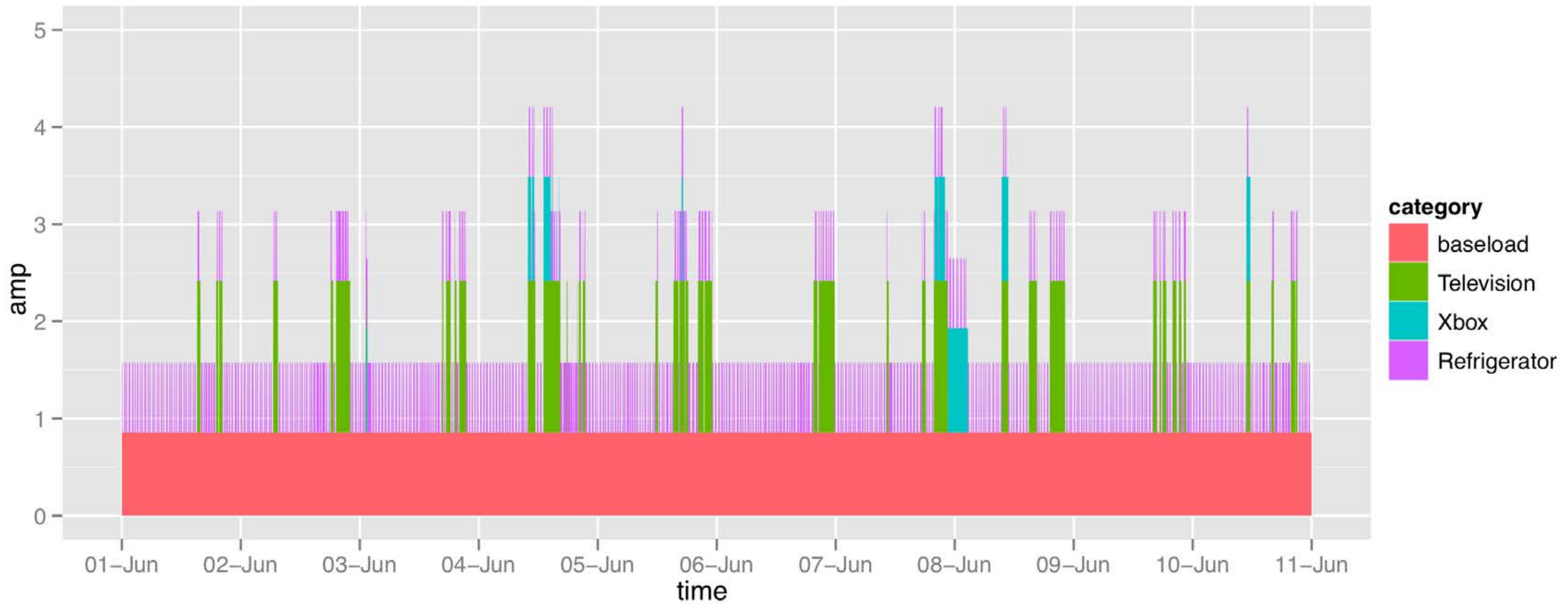
Energy Disaggregation





http://blog.lr.org/wp-content/uploads/2013/08/LordKelvin.jpg

ENERGY DISAGGREGATION



SOLUTION

– Install a meter on every appliance
  - Too intrusive
  - Too expensive

– Non-intrusive load monitoring (NILM) [George Hart, 1984]
  - Figure out appliance usage from the whole house measurement

PROBLEM STATEMENT

– Input

- $Y = <y_1, y_2, \ldots, y_T>$, a sequence of aggregated power consumption
- M, the number of appliances

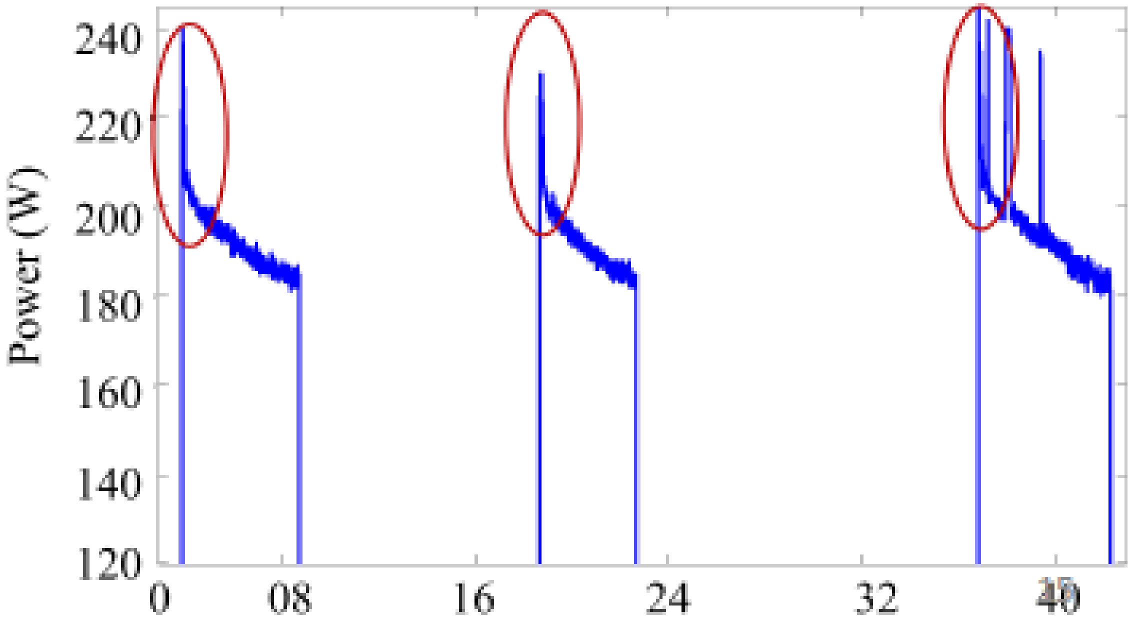
– Output

- $S_1 = <s_1, s_2, \ldots, s_T>$, a sequence of consumption for *Appliance* 1
- $S_2 = <s_1, s_2, \ldots, s_T>$, a sequence of consumption for *Appliance* 2

  …
- $S_M = <s_1, s_2, \ldots, s_T>$, a sequence of consumption for *Appliance* M

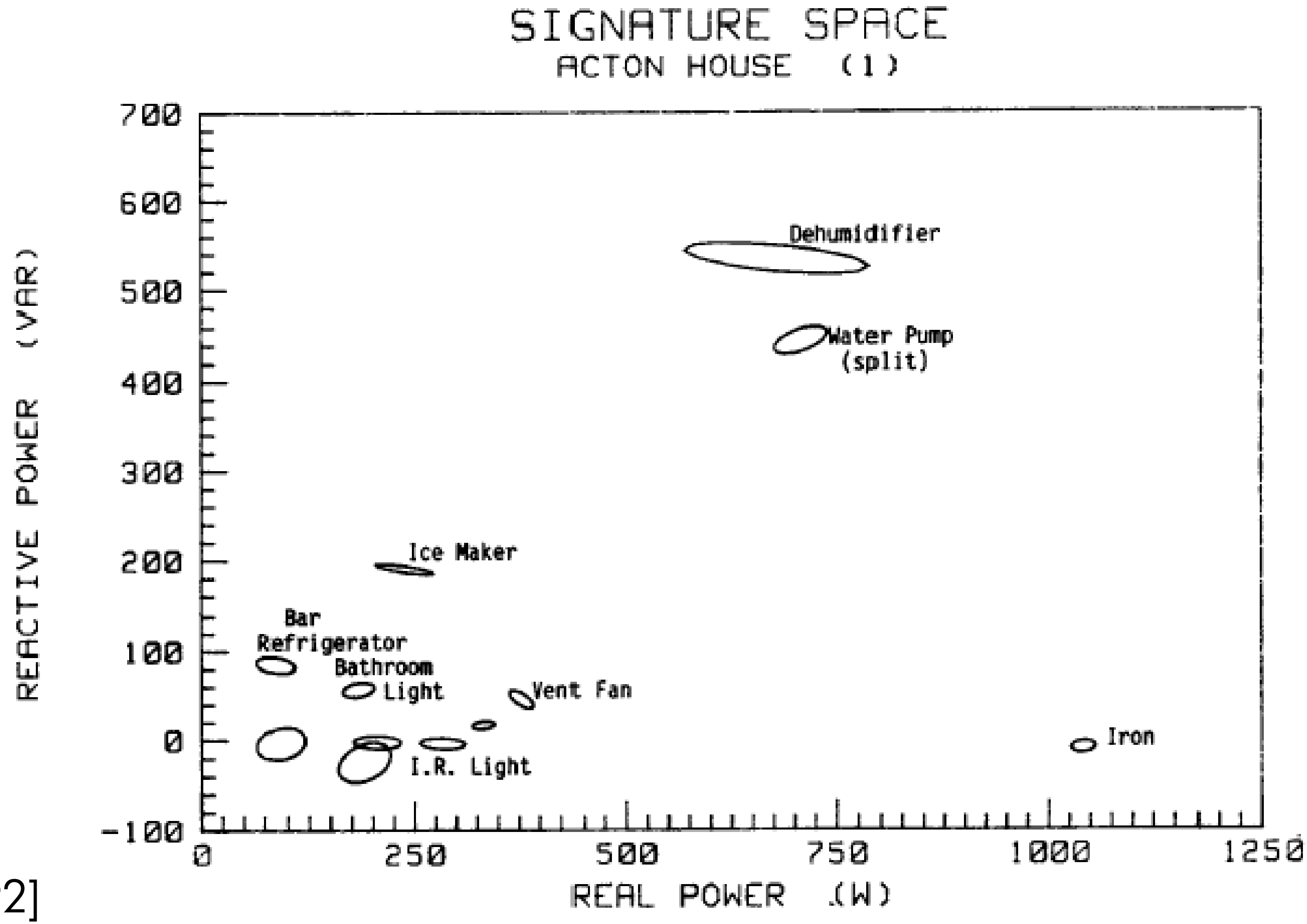
FEATURES

- Sampling frequency
  - Low (minutes to hours)
  - Medium (~ 1Hz)
  - High (in kHz)
- Stable state features
- Transient features
  - Require special HW
- Real and reactive power
- Non-power features
  - Time of day
  - Day of week
  - Weather
  - Sensors
  - State of other appliances



EVENT IDENTIFICATION

- Compute delta in real and reactive power



[Hart 1992]

APPLIANCE STATE MACHINES

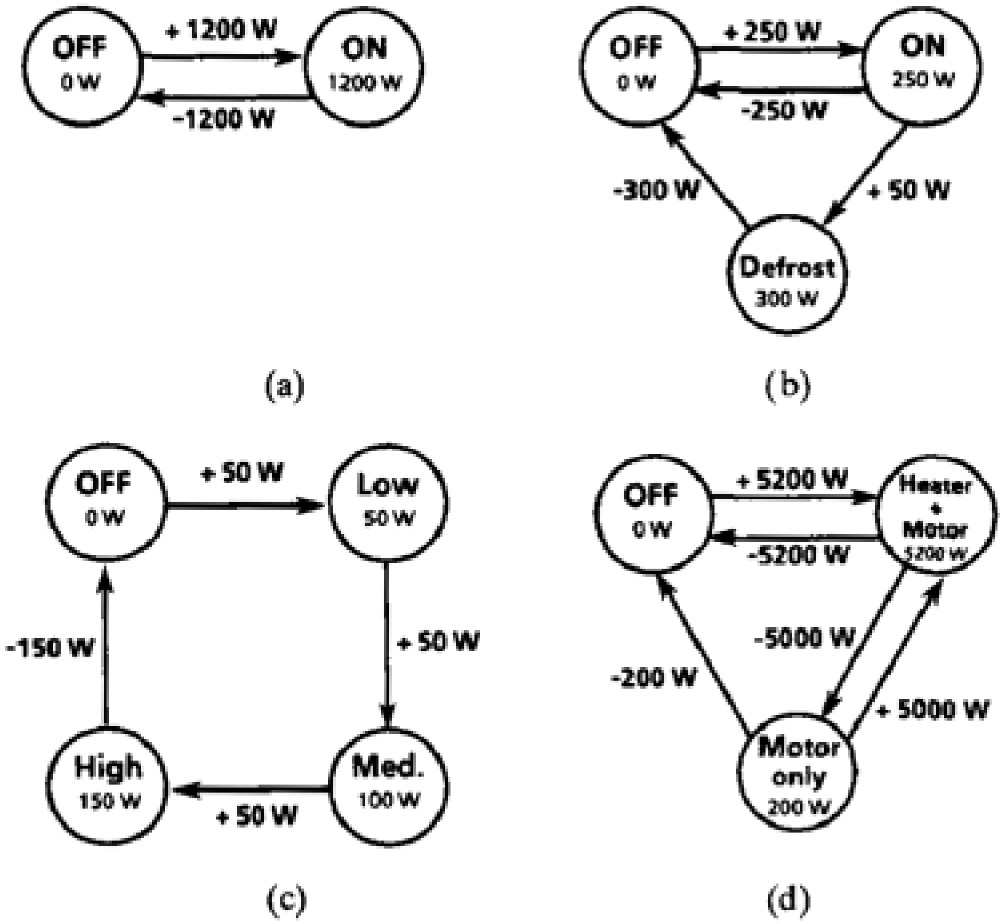


Fig. 4. Finite-state appliance models: (a) generic 1200 W two-state appliance, e.g., toaster; (b) refrigerator with defrost state; (c) "three-Way" lamp; (d) clothes dryer.

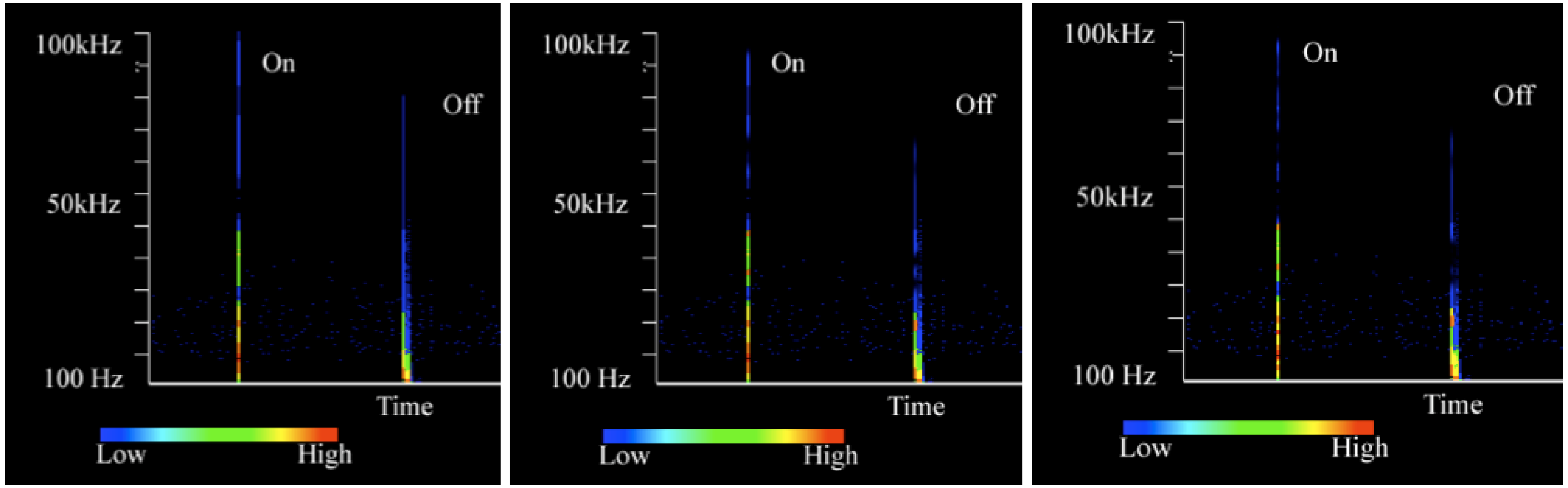
[Hart 1992]

SUPERVISED APPROACHES

– High frequency samples (100KHz)

– Labelled event data

– Train a classifier (e.g. SVM)



S.N. Patel et al. (2007)

DRAWBACKS OF EVENT-BASED METHODS

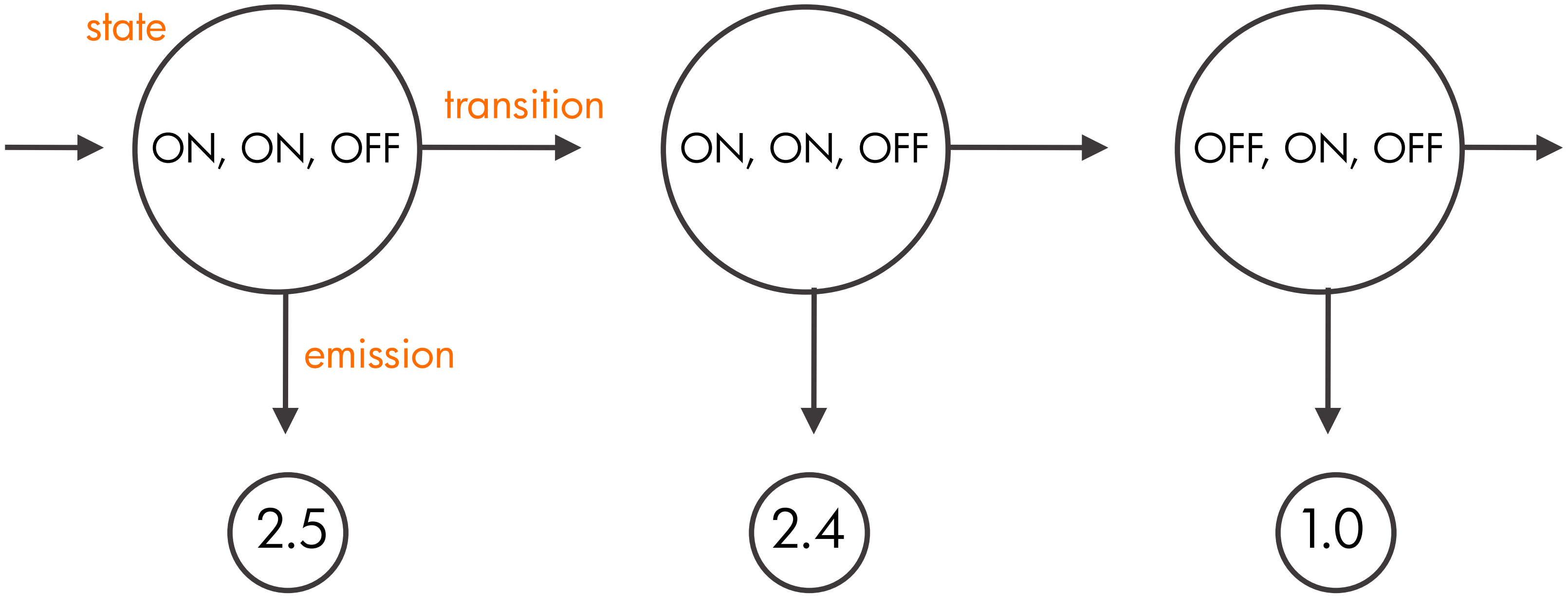
- Require labelled data
- Events considered in isolation
- Most require high frequency data

HMM-BASED MODELS

- **General algorithm outline**
- 1. Define a model
- 2. Learn the parameters in the model from data
- 3. Make predictions (Inference)

HMM

Time	1	2	3	4	5	6	7	8	...
readings	2.5	2.4	1.0	1.1	1.7	1.6	0.8	0.7	...
A	1.4	1.5	0	0	0	0	0	0	...
B	1.1	0.9	1.0	1.1	1.0	0.9	0	0	...
C	0	0	0	0	0.7	0.8	0.8	0.7	...



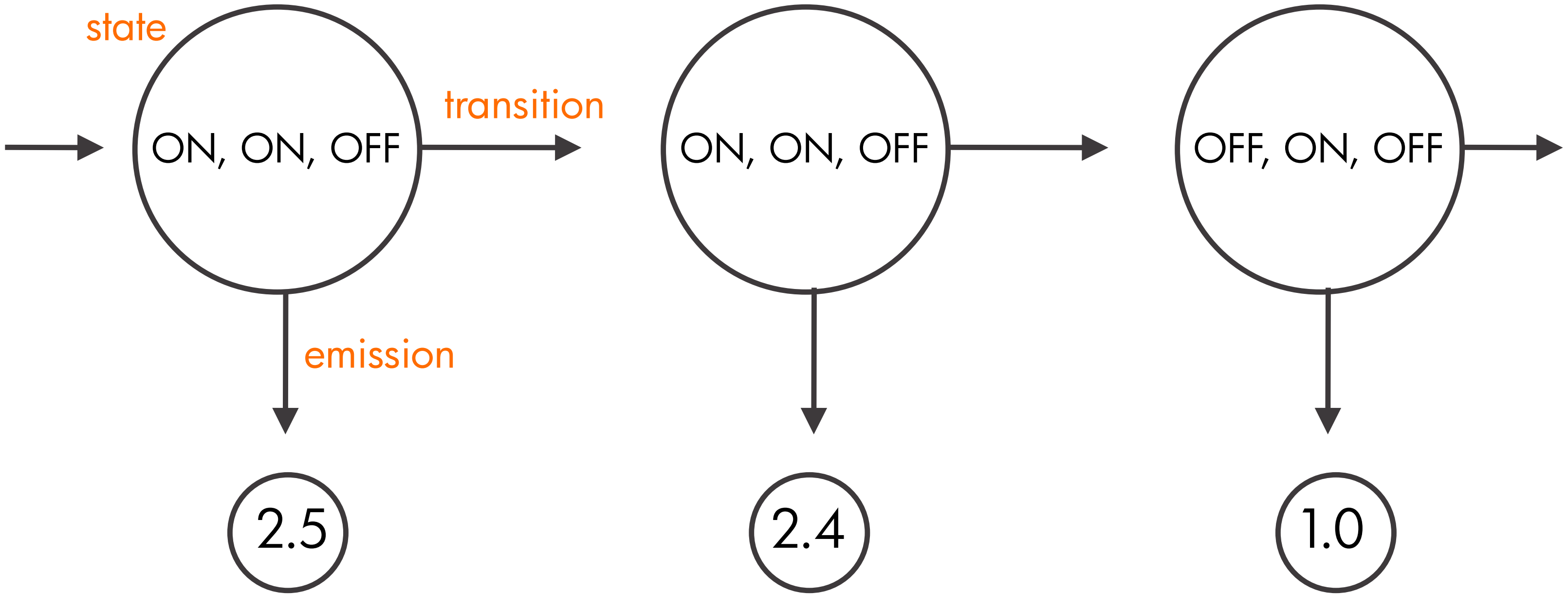
HIDDEN MARKOV MODEL

– Transition probability

$$\Pr(s_{t+1} = i \mid s_t = j) = \pi_{ij}$$

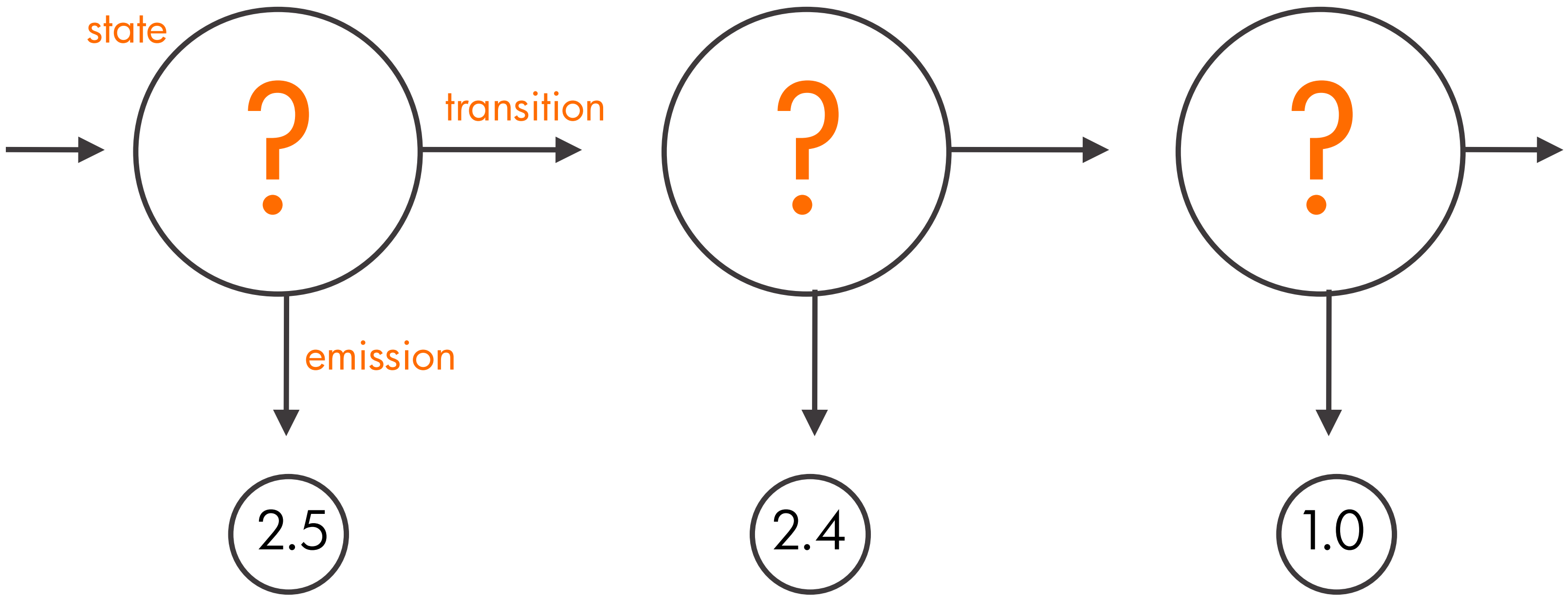
– Emission probability

$$\Pr(y_t = v \mid s_t = i) \sim \mathrm{Normal}(w_i,\ e),$$ where *e* is the noise variance



HIDDEN MARKOV MODEL

– S, the sequence of the internal states, is not observable



HIDDEN MARKOV MODEL

– Transition probability

$$\Pr(s_{t+1} = i \mid s_t = j) = \pi_{ij}$$

– Emission probability

$$\Pr(y_t = v \mid s_t = i) \sim \text{Normal}(w_i, e),$$ where e is the noise variance

– Let $\theta = \{\pi_{ij}\} \cup \{w_i\} \cup \{e\}$, the set of the parameters in HMM

– If both S and Y are observable, we can find the parameters θ by Maximum Likelihood (ML)

– But… S is unknown

– If Y and θ are known, we can perform inference to compute S

– Chicken and egg problem!

– Expectation Maximization (EM)

HIDDEN MARKOV MODEL

- The number of states: 2^M
- The number of parameters: $2^M + 2^{2M}$
  - 2^M emission-parameters
  - 2^{2M} transition-parameters
- Exponential increase with number of appliances
- That’s too many parameters!

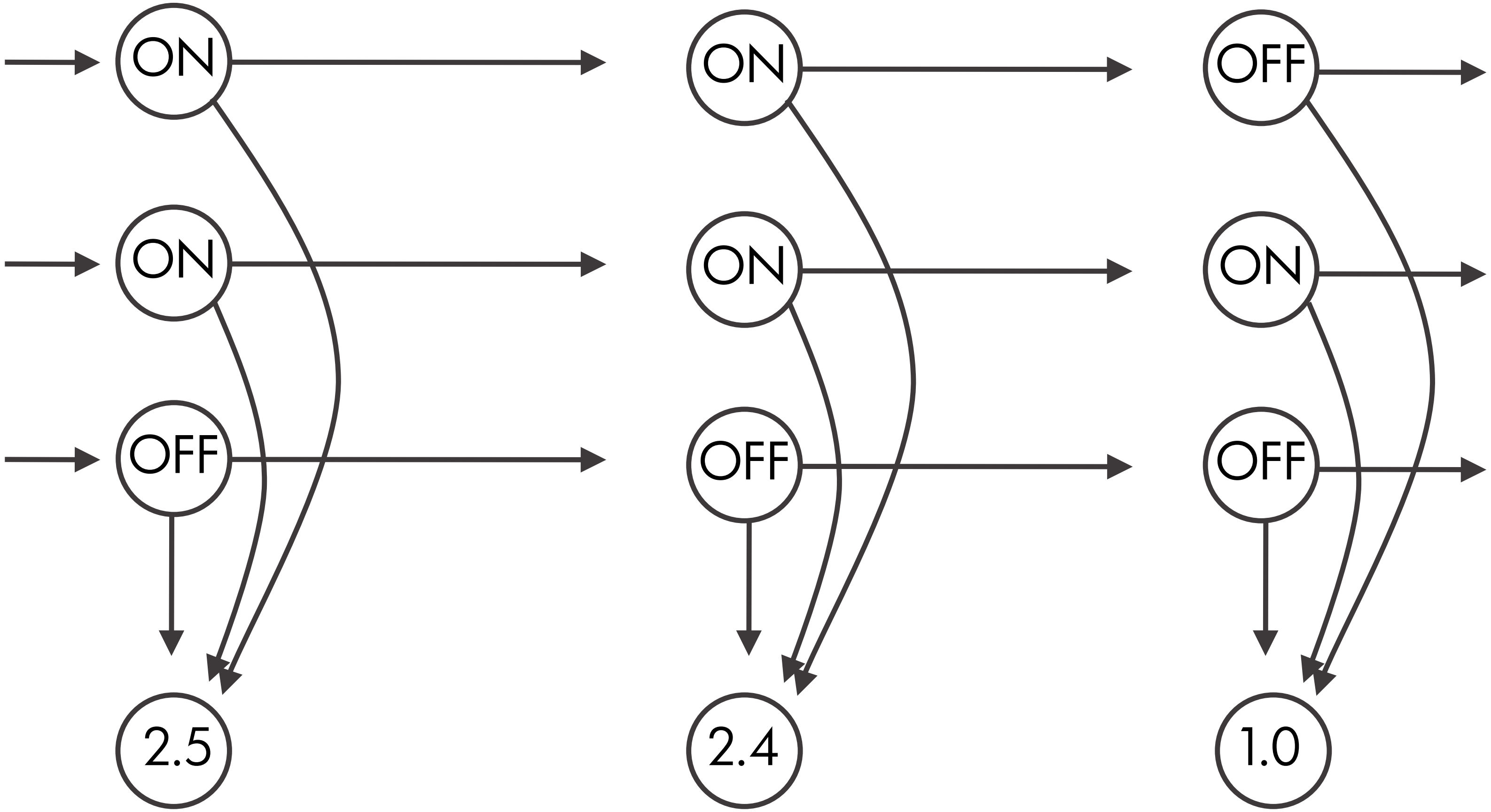
FACTORIAL HIDDEN MARKOV MODEL

– The number of states: 2M
– The number of parameters: 6M
    • 2M emission-parameters
    • 4M transition-parameters

– Much better!

FACTORIAL HIDDEN MARKOV MODEL

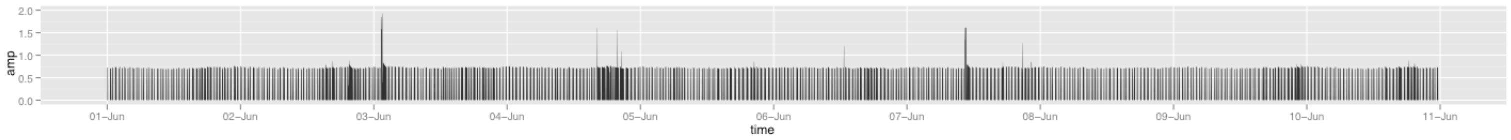
- Assumption: *Appliances are used independently*
- The observation is a linear combination of the emissions of the markov chains



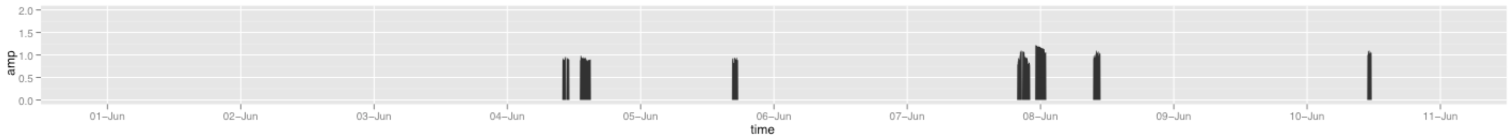
EXAMPLE APPLIANCE DATA

– 3 appliances: Refrigerator, Xbox, TV

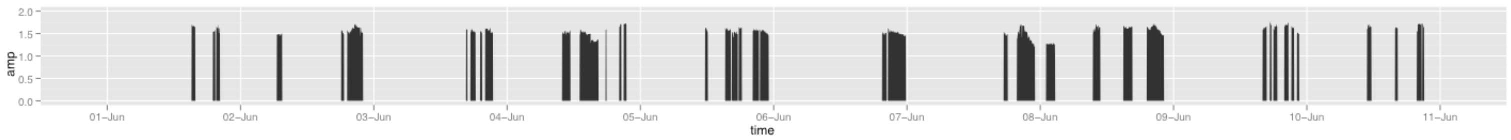
Ref



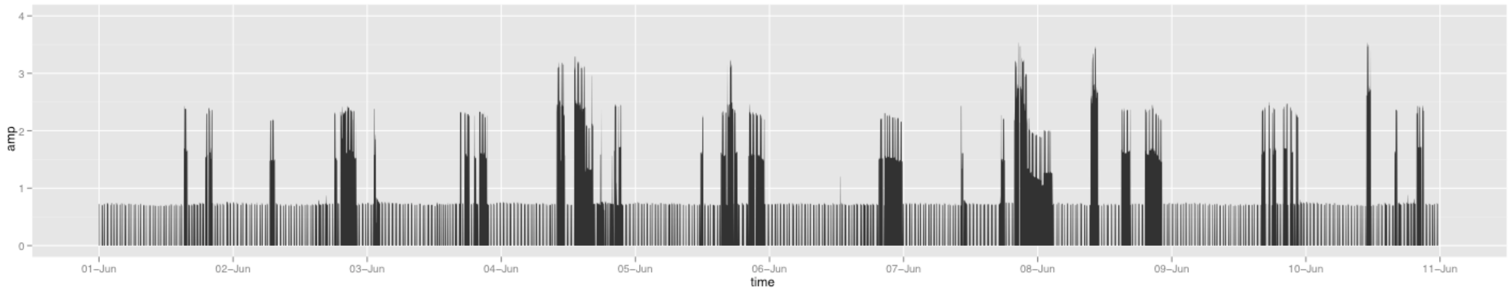
Xbox



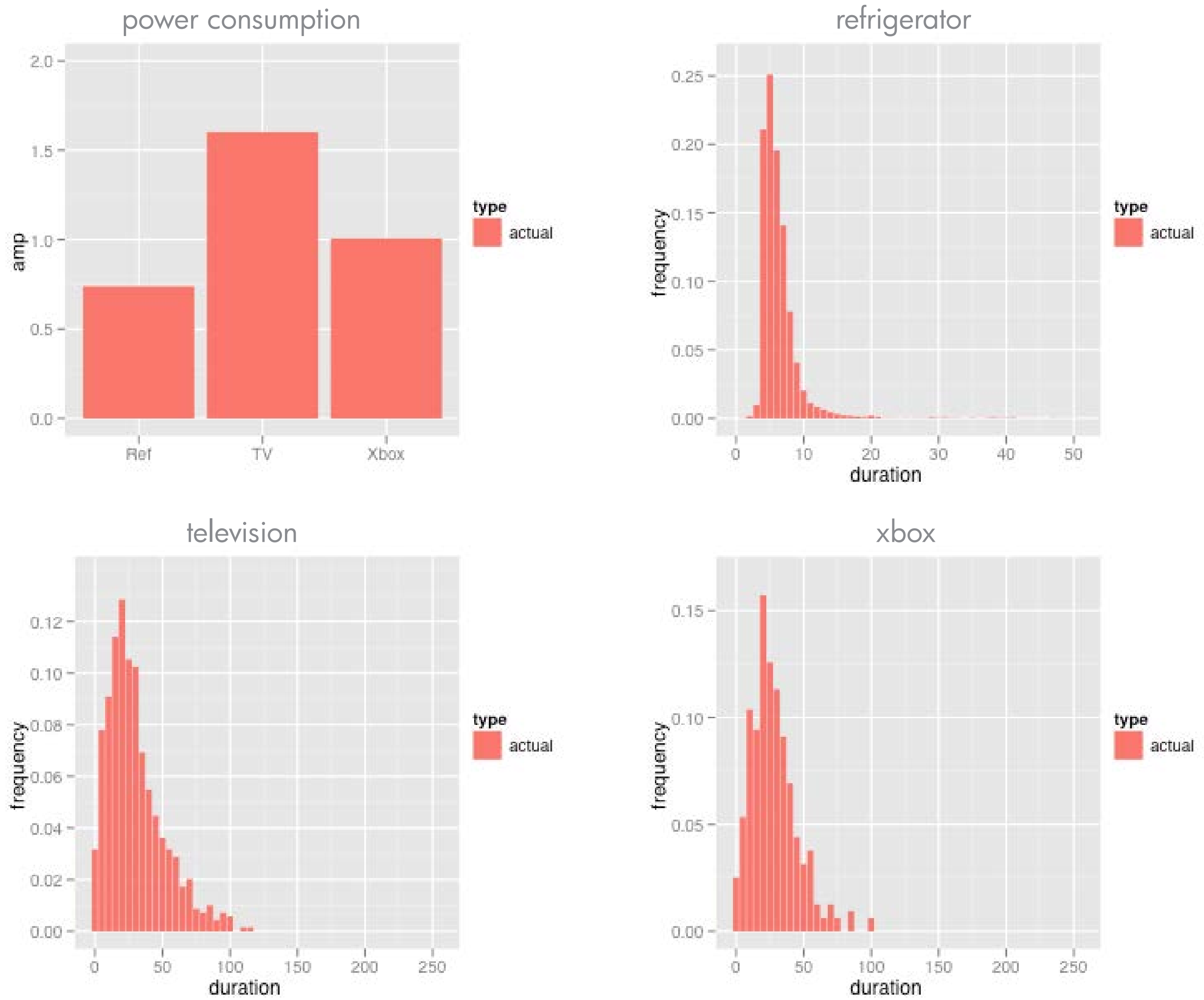
TV



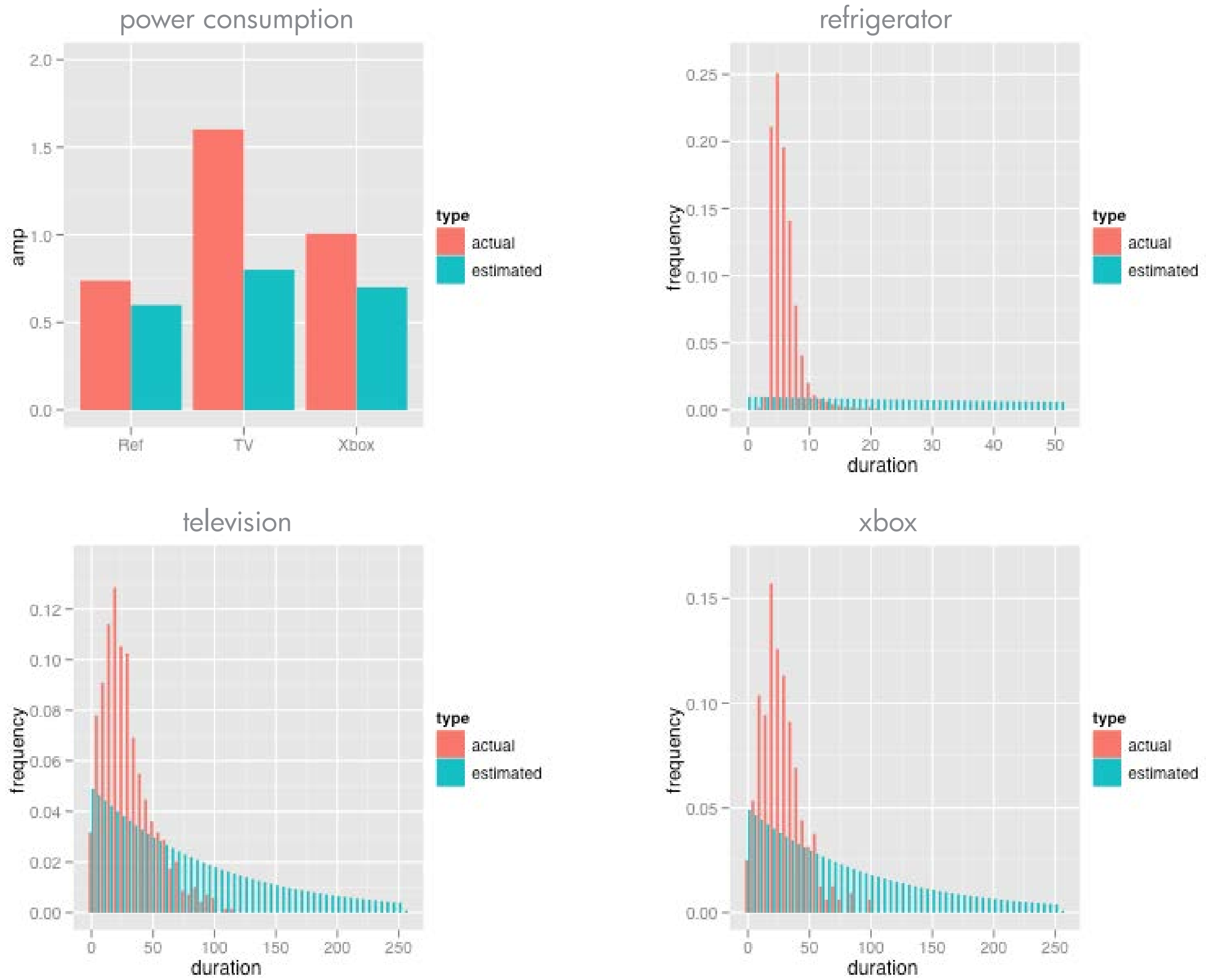
aggregate



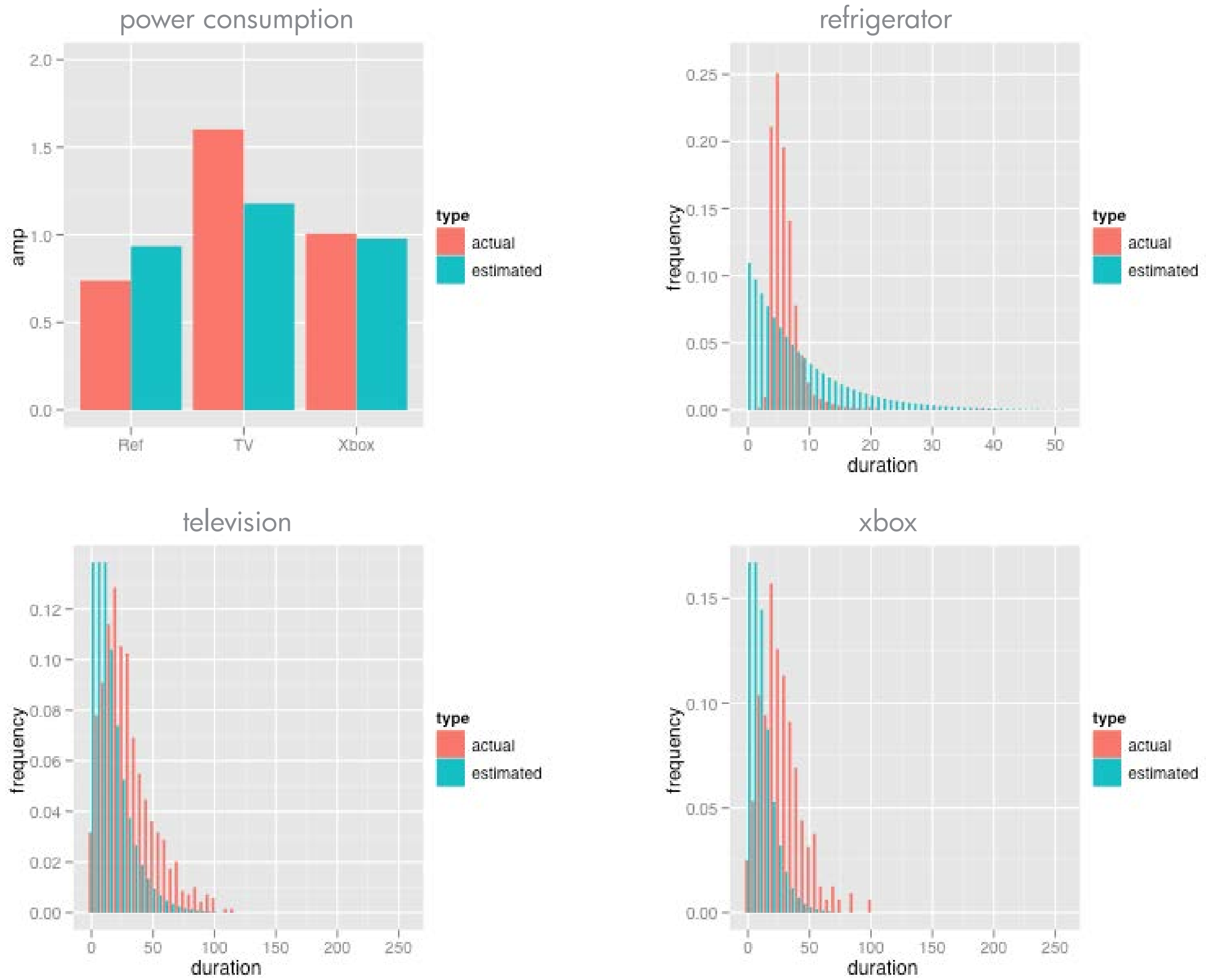
APPLIANCE DISTRIBUTIONS



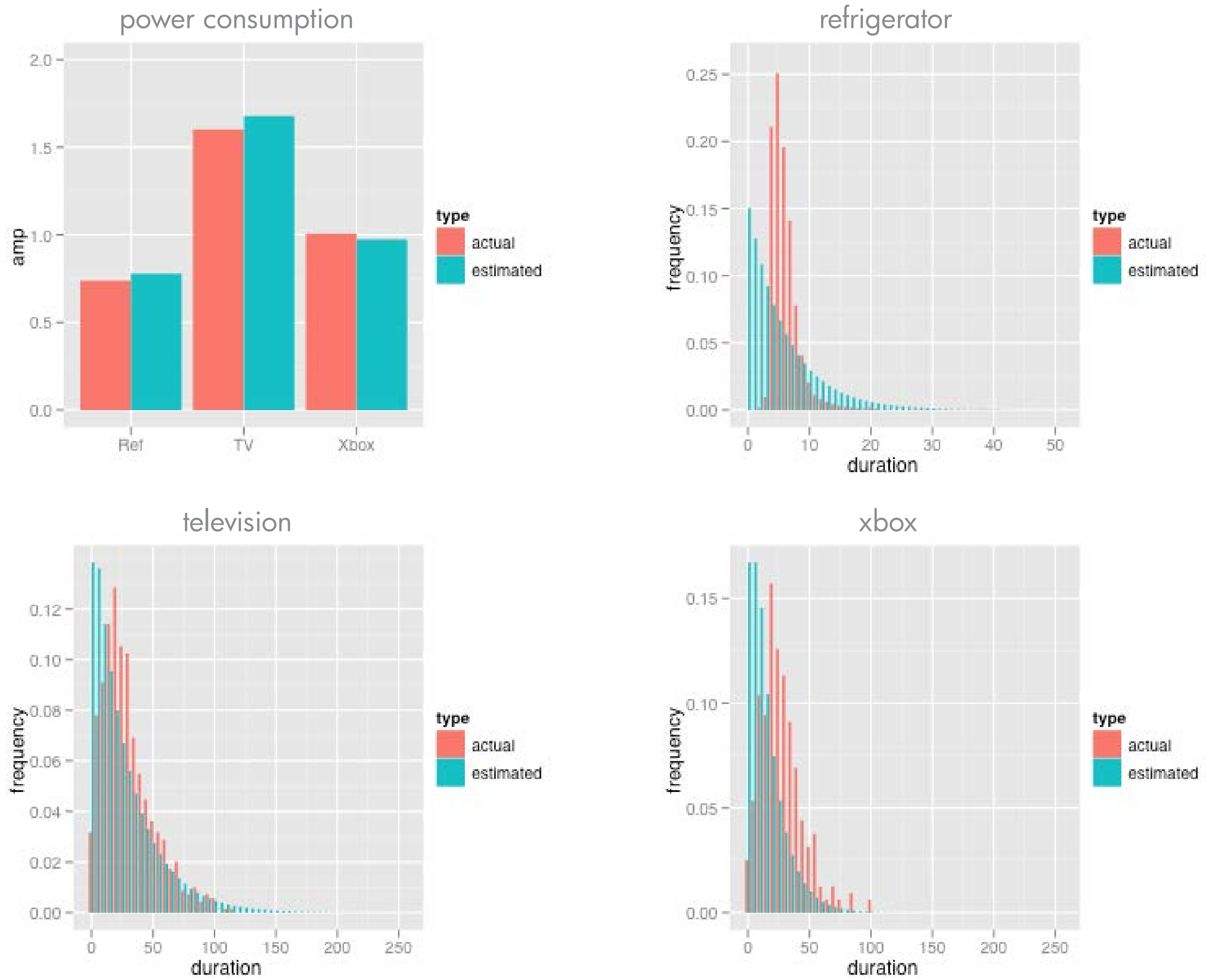
FHMM – EM ITERATION 0



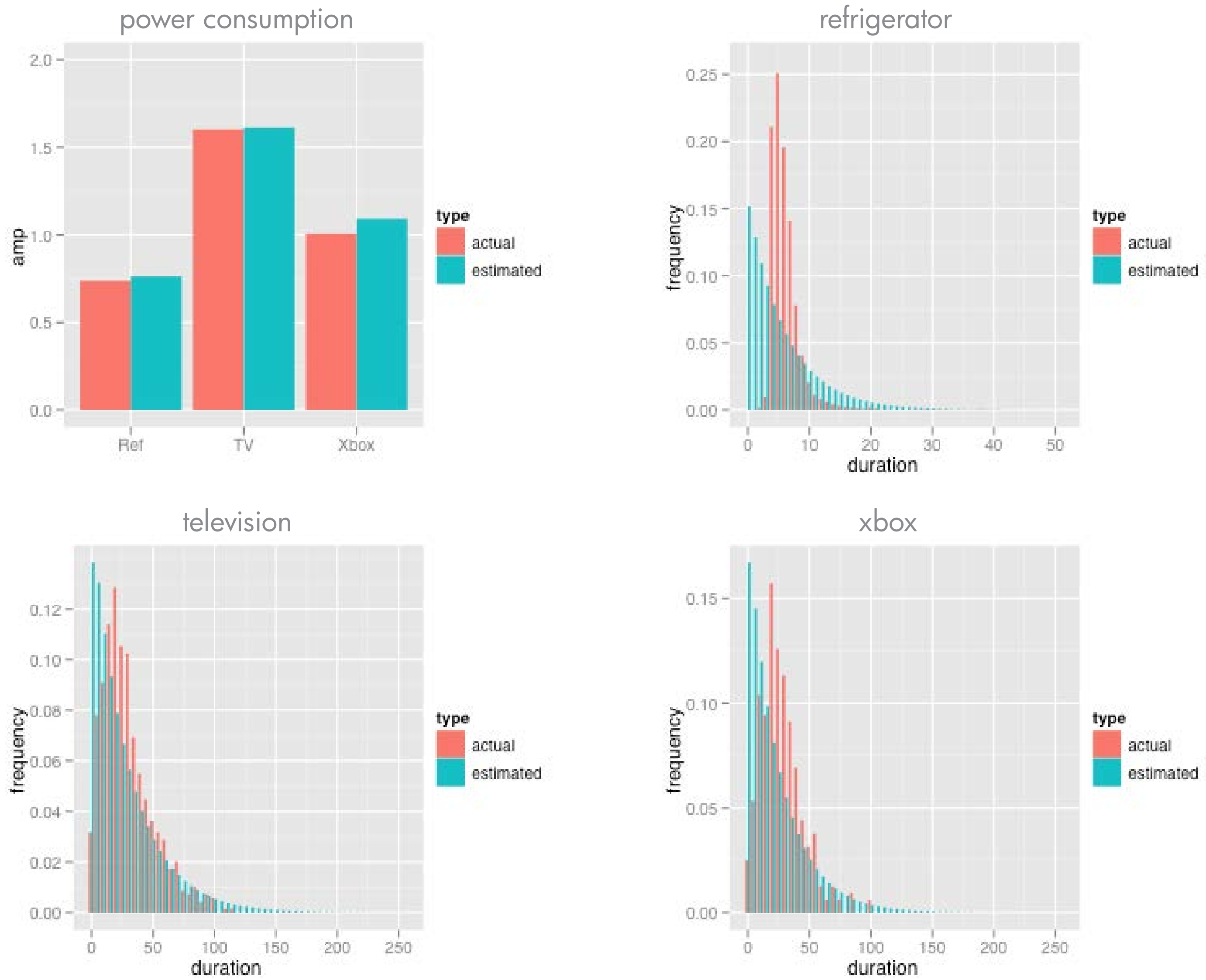
FHMM – EM ITERATION 4



FHMM – EM ITERATION 10



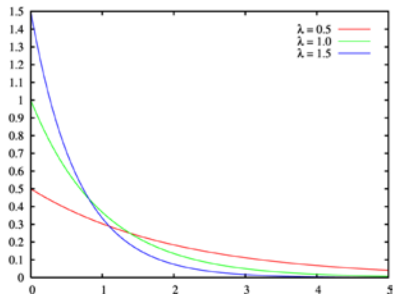
FHMM – EM ITERATION 20



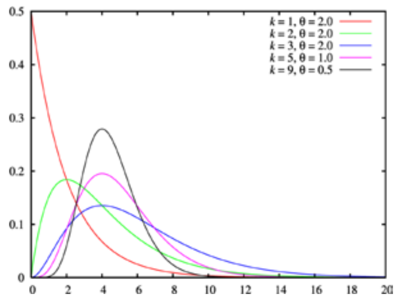
CHALLENGE: STATE-DURATION DISTRIBUTIONS

– In the family of Hidden Markov Model, the state-durations have **exponential distributions**

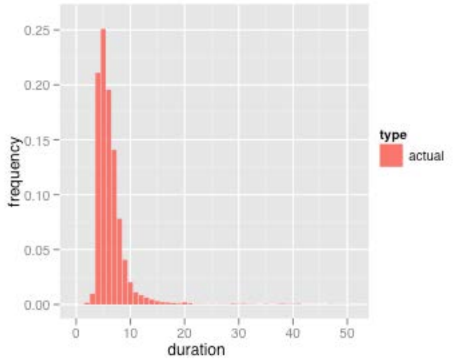
– But, the state-durations for appliances follow **gamma distributions**

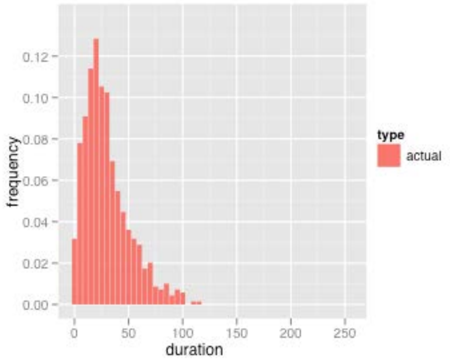


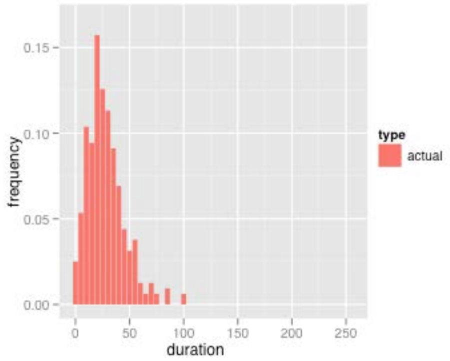
Exponential distribution



Gamma distribution

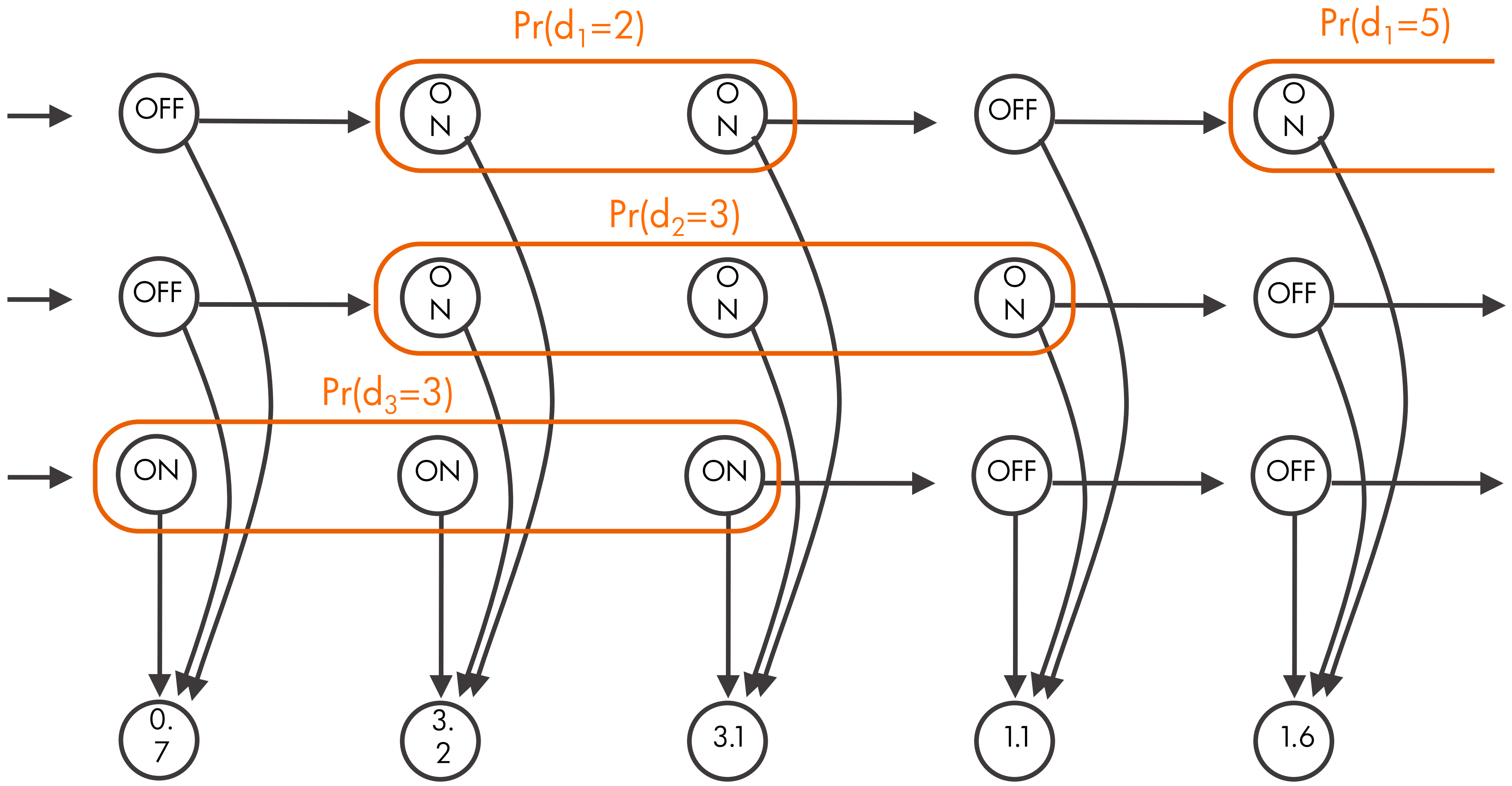




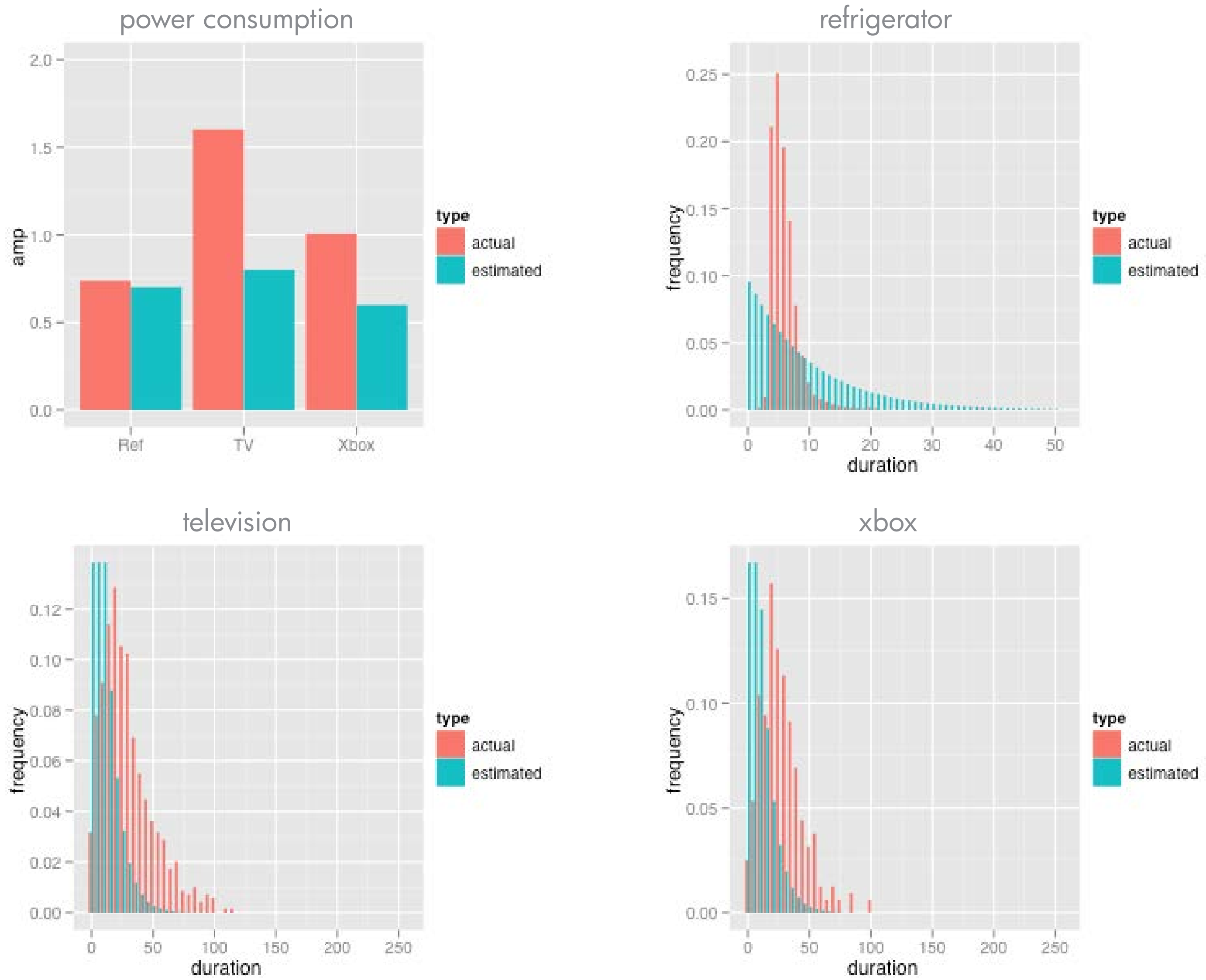


images from http://en.wikipedia.org/

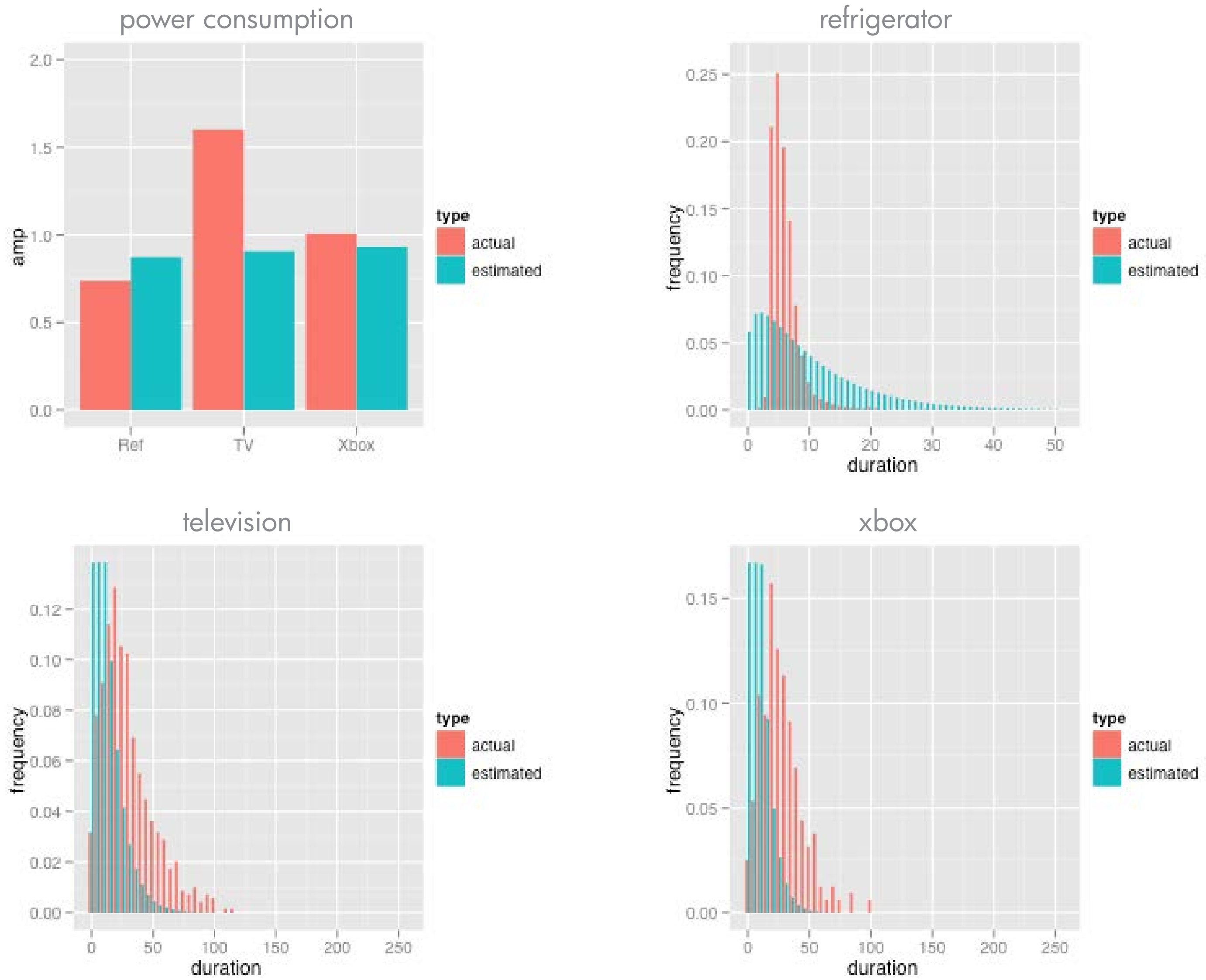
FACTORIAL HIDDEN SEMI-MARKOV MODEL



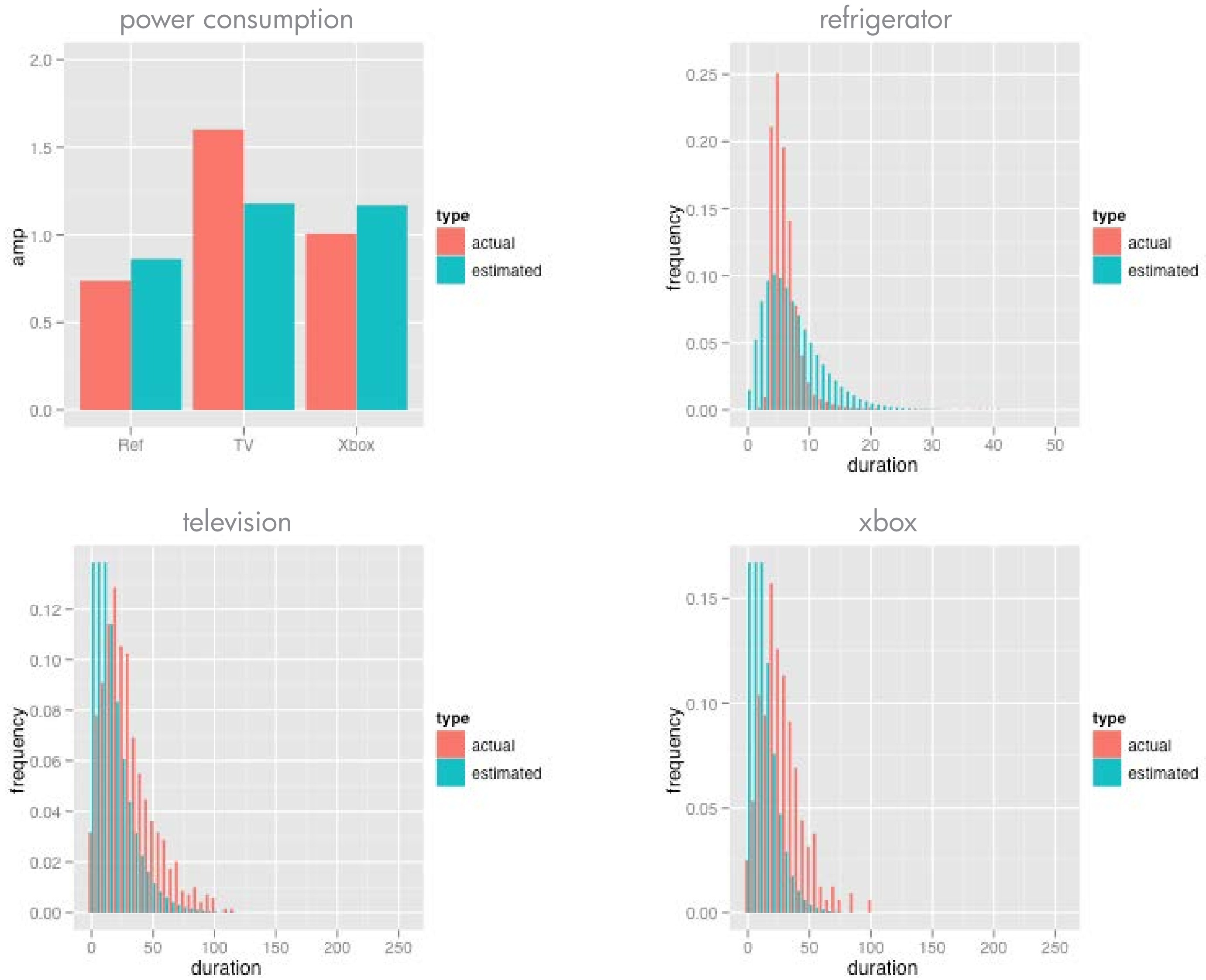
FHSMM – EM ITERATION 0



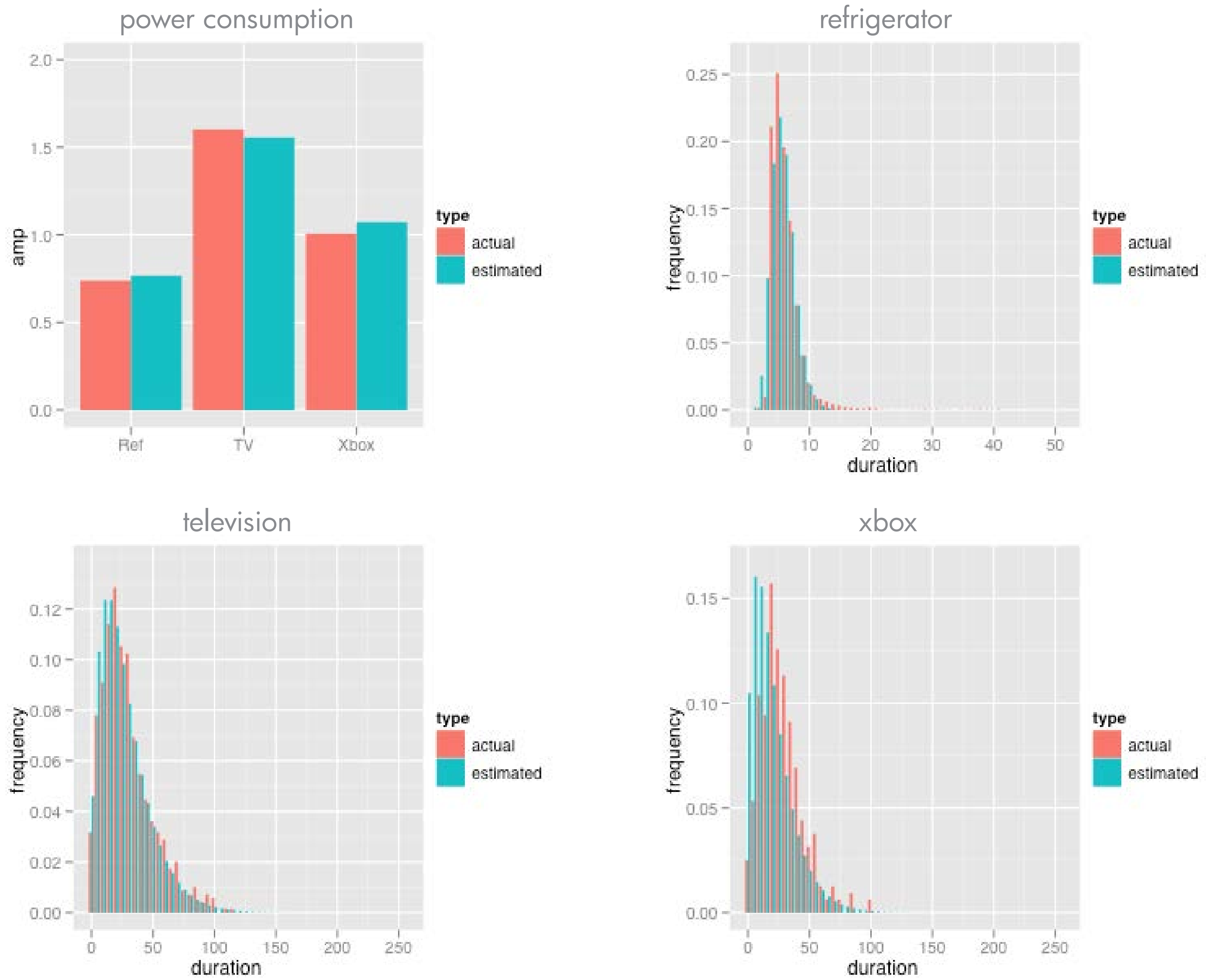
FHSMM – EM ITERATION 1



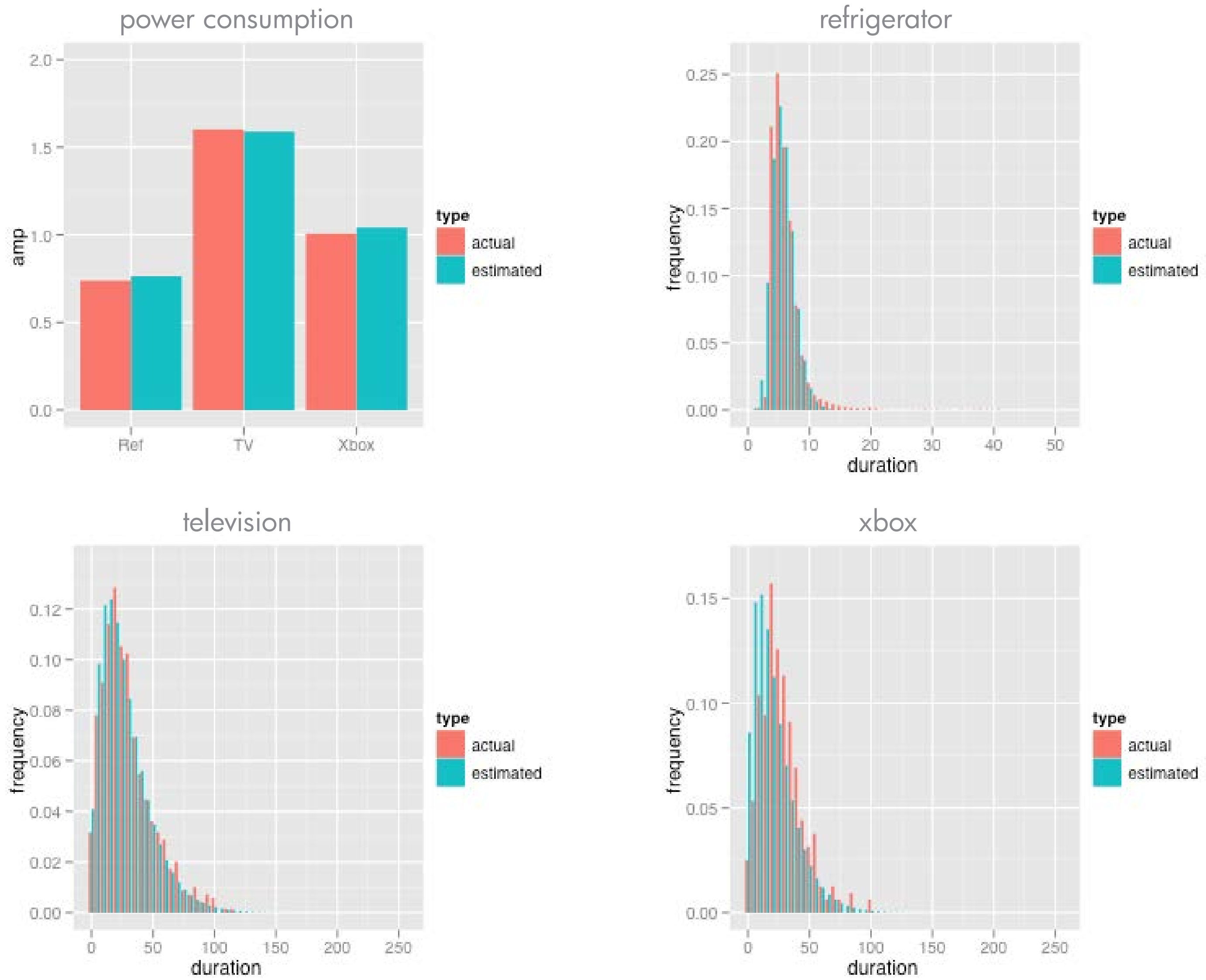
FHSMM – EM ITERATION 5



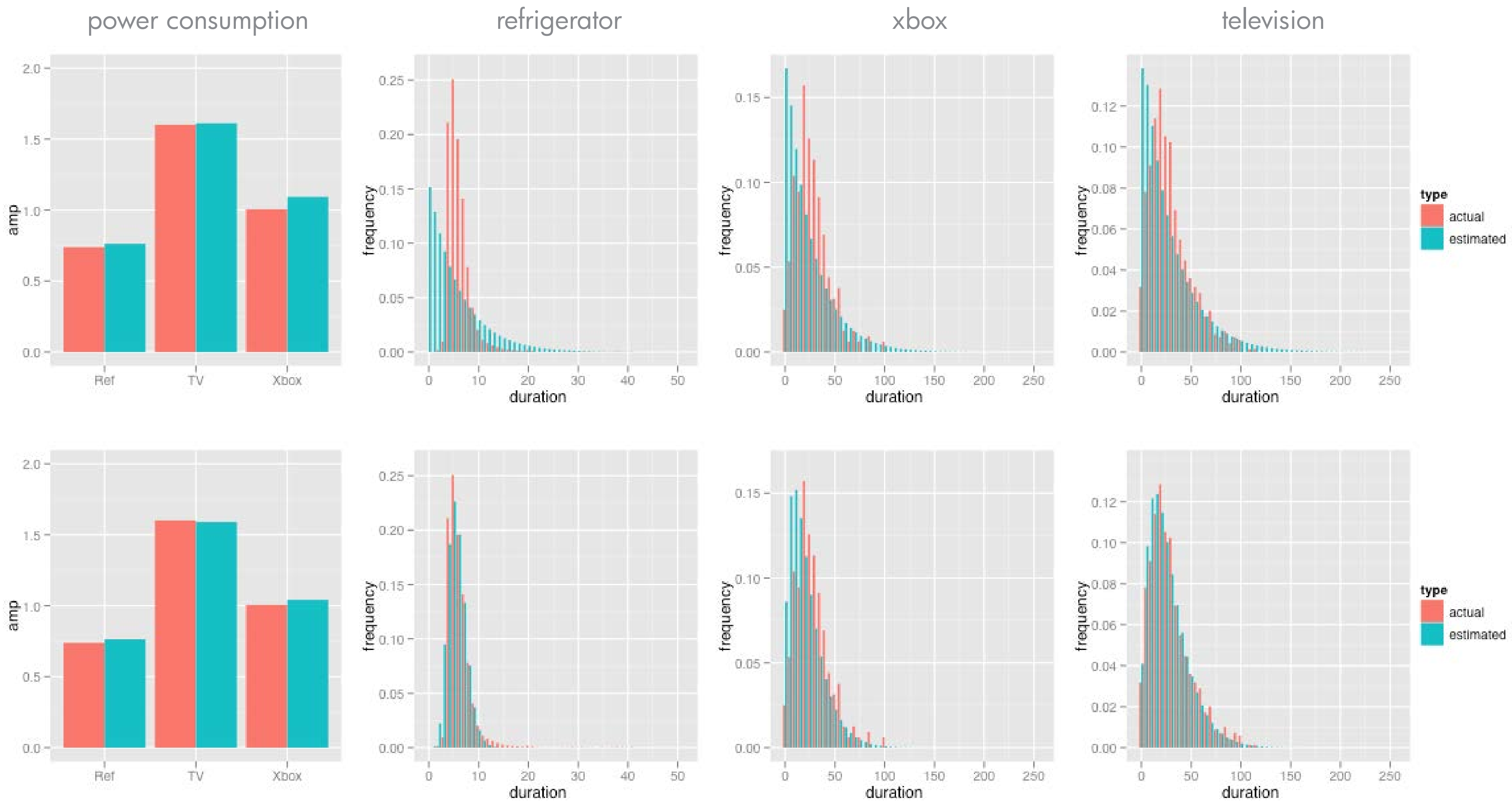
FHSMM – EM ITERATION 10



FHSMM – EM ITERATION 20



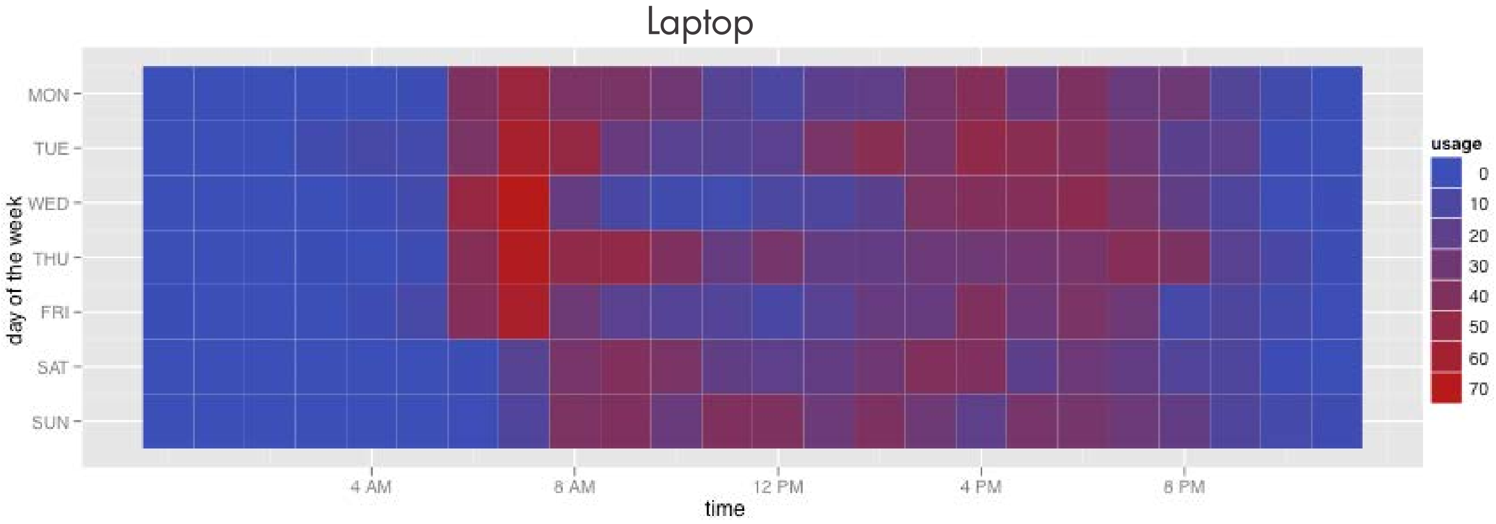
FHMM vs. FHSMM

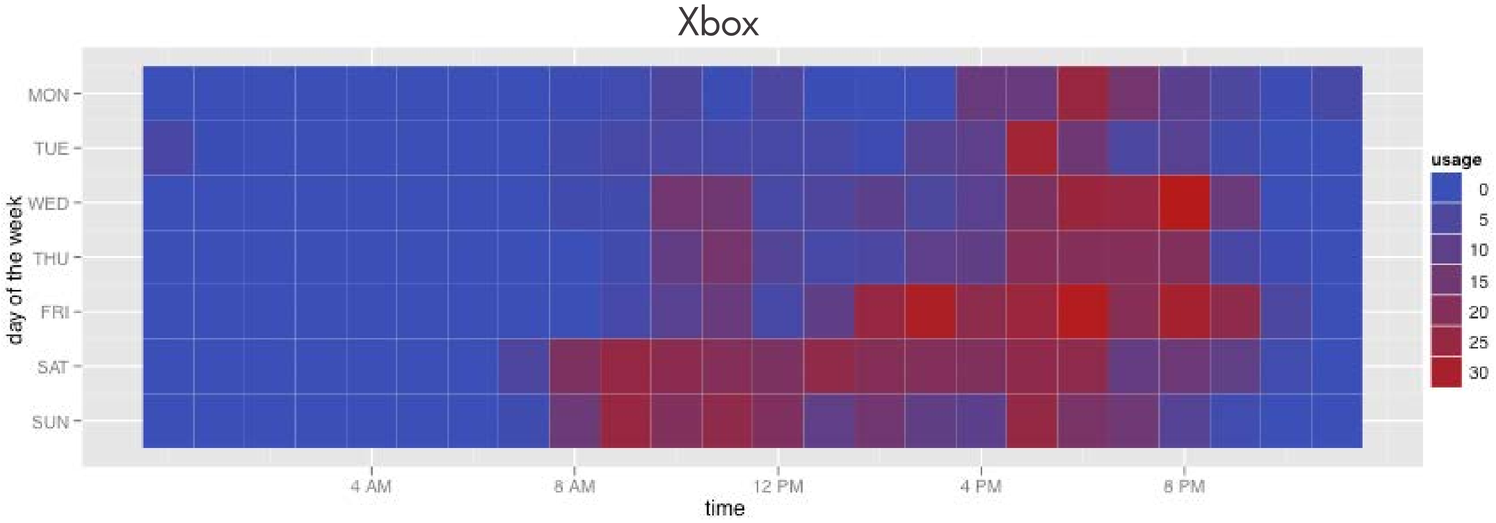


CHALLENGE: MODELING DEPENDENCIES

- There are many factors which affect the usage of appliances
- There can be additional contextual features
- Example
  - Weather (e.g. heater, A/C)
  - Day of the week (e.g. more TV on the weekends)
  - Time of day (e.g. more Xbox in the afternoon)
  - Seasons (e.g. more laundry in summer)
  - User’s schedule (e.g. more laptop use in early morning)
  - Other appliances (e.g. TV is on when Xbox is in use)

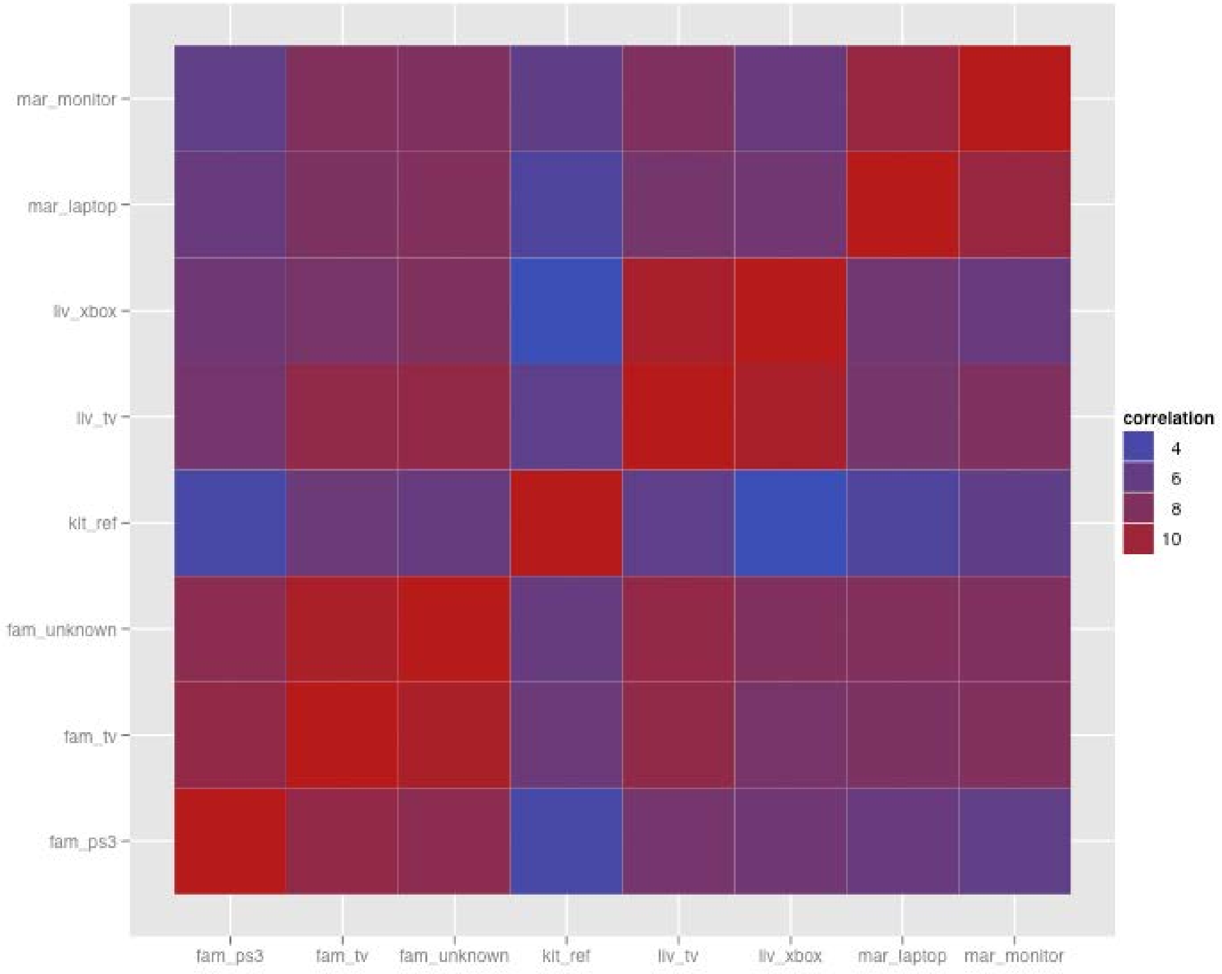
DEPENDENCY 1.
TIME AND DAY OF THE WEEK





DEPENDENCY 2.
OTHER APPLIANCES

Correlations between appliances



T(*x*, *y*) = log of the value of chi-square test for appliance *x* and *y*

CONDITIONAL FACTORIAL HIDDEN MARKOV MODEL

– In FHMM, the transition probability is constant

$$Pr(X_{t+1} = q | X_t = p) = \pi_{pq}$$

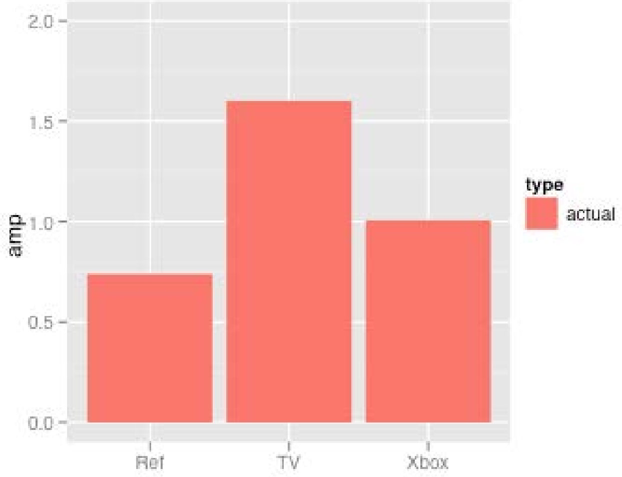
– In CFHMM, the transition probability depends on several conditions, and it is computed by assuming each condition is independent (Naïve Bayes assumption)

$$Pr(X_{t+1} = q | X_t = p, C_1 = c_{1t}, C_2 = c_{2t}, \ldots, C_k = c_{kt}) = \frac{\pi_{pq}}{Z} \prod_{i=1}^{k} Pr(C_i = c_{it} | X_{t+1} = q)$$

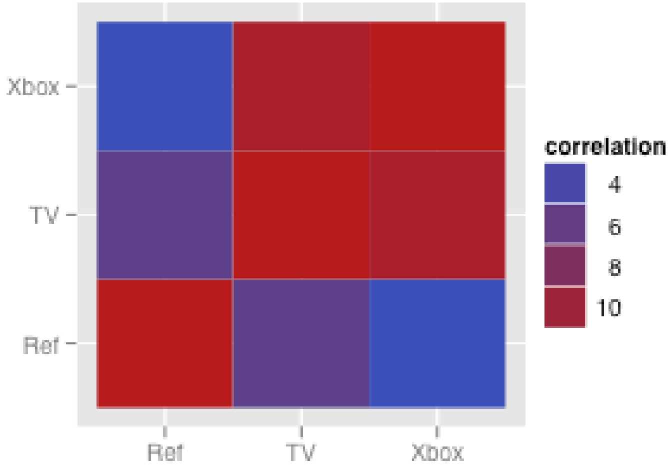
where Z is the normalization factor

APPLIANCE STATISTICS

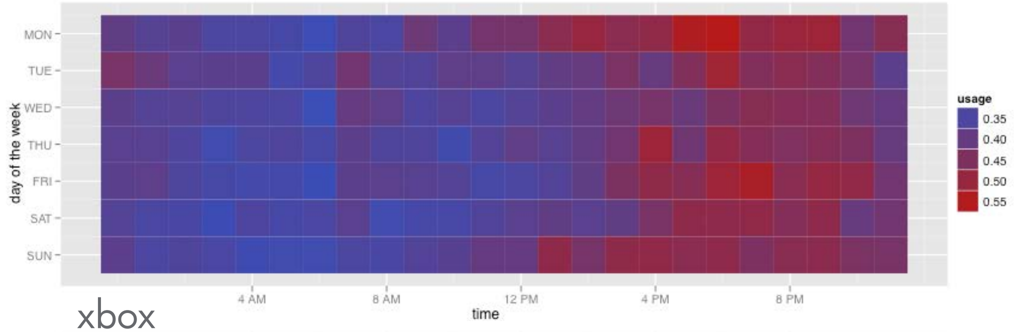
power consumptions



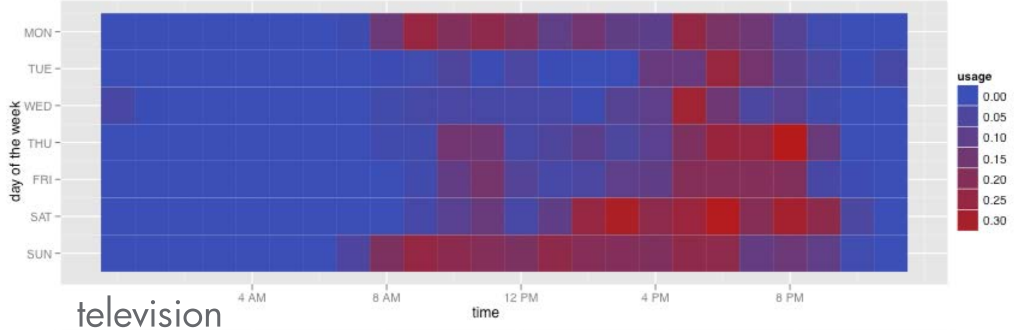
appliance correlations



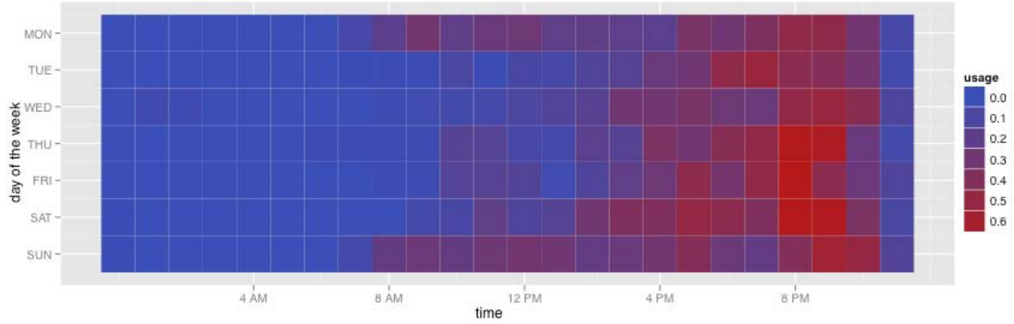
refrigerator



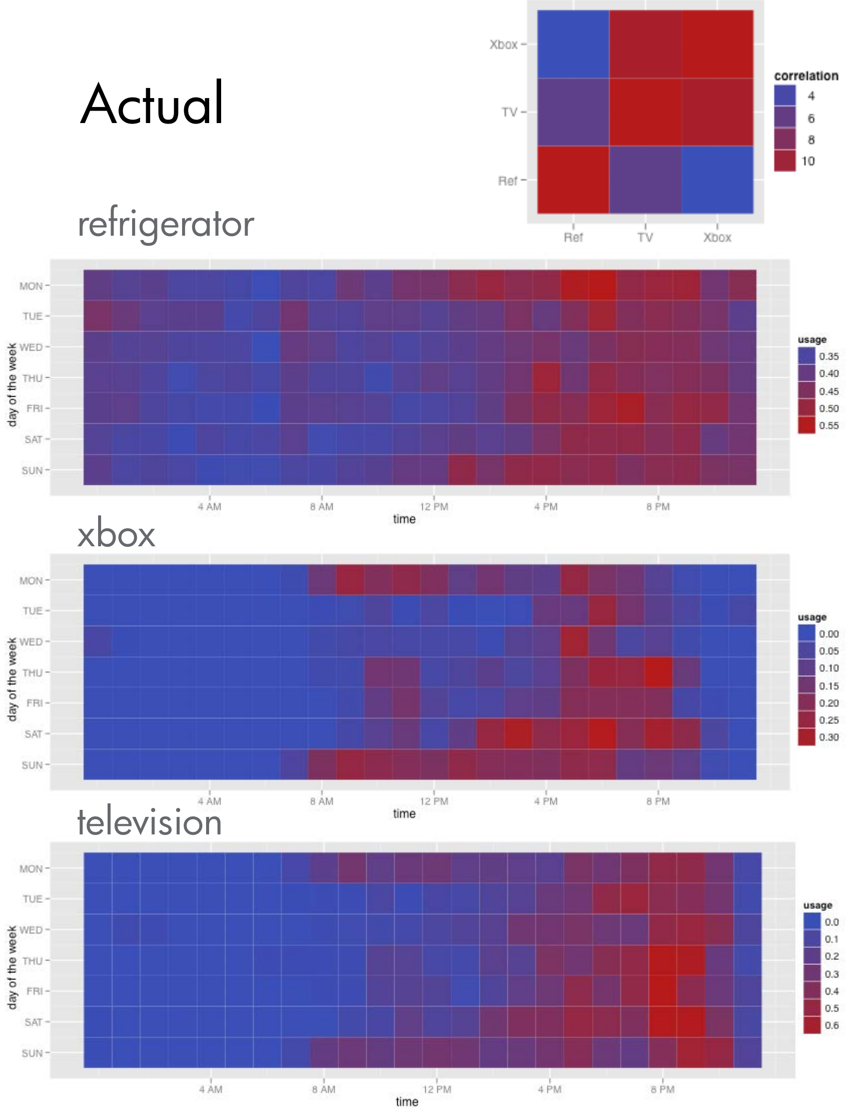
xbox

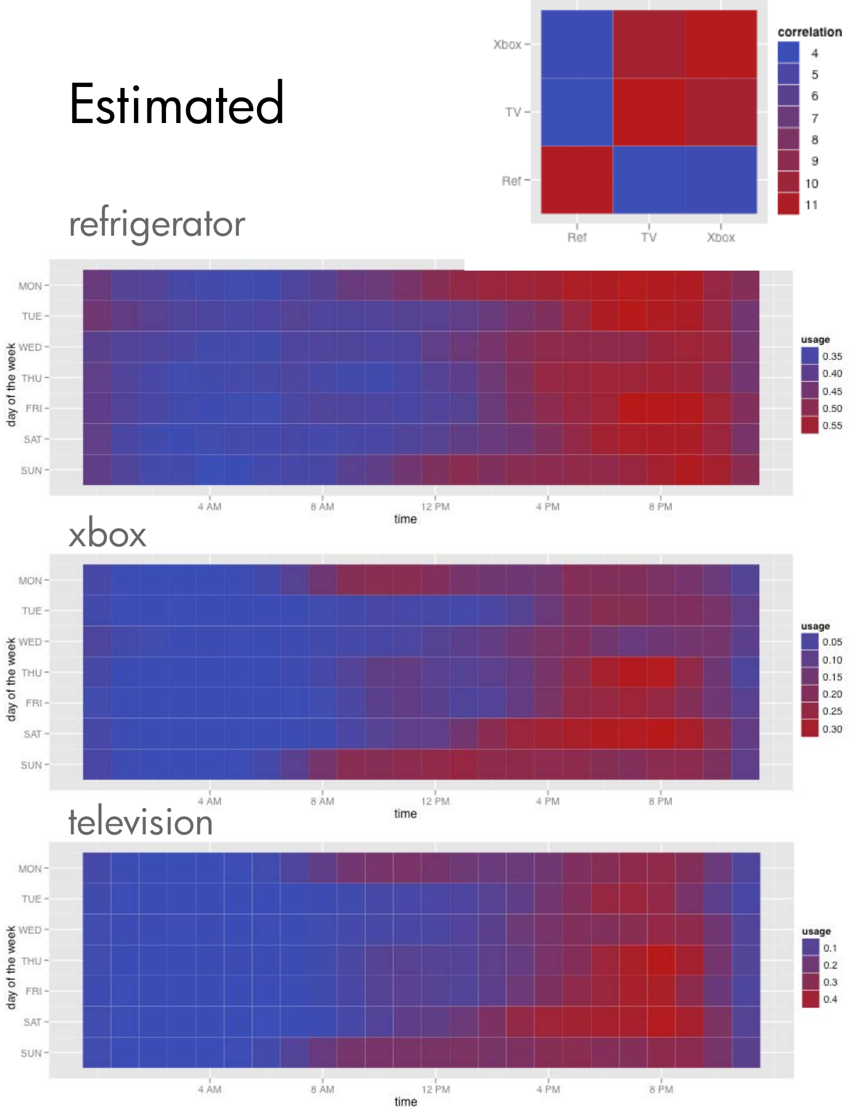


television

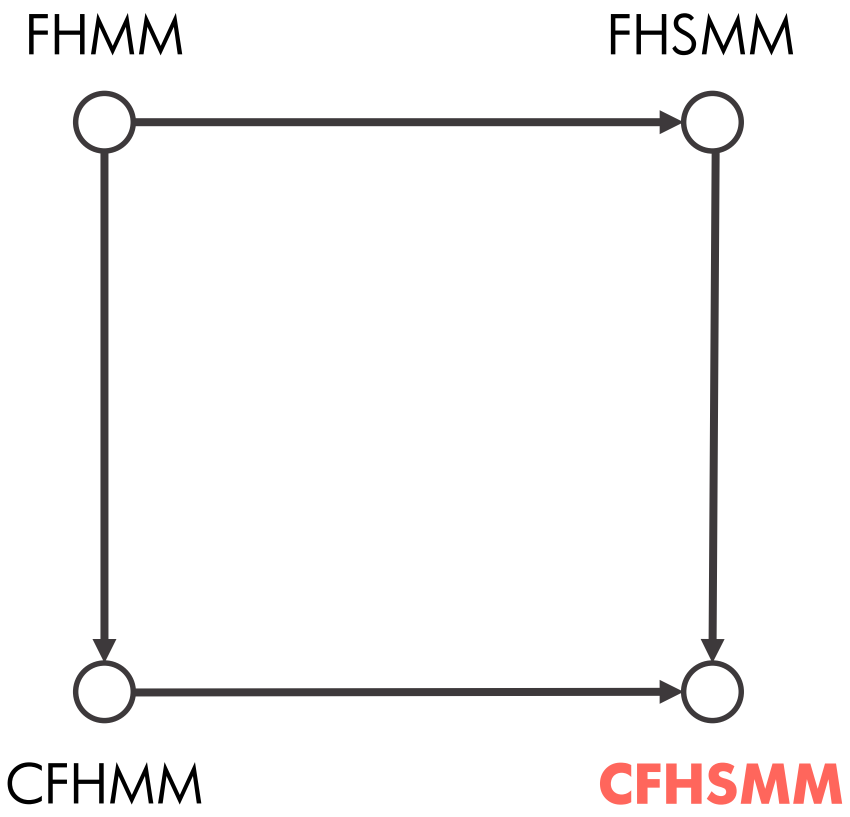


CFHMM – RESULT





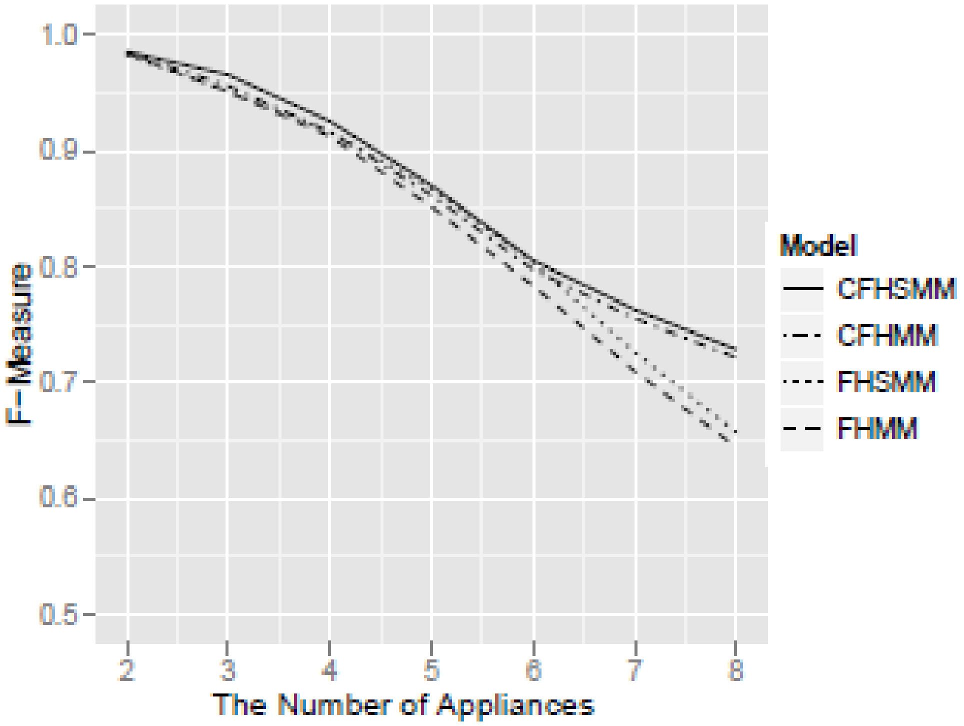
CONDITIONAL FACTORIAL HIDDEN SEMI-MARKOV MODEL



RESULTS

Home ID	Num. of Appliances	FHMM	CFHSMM
Home 1	4	0.983	0.998
Home 2	6	0.899	0.930
Home 3	6	0.859	0.881
Home 4	7	0.625	0.693
Home 5	8	0.713	0.781
Home 6	8	0.641	0.722
Home 7	10	0.796	0.874

Table 4: The evaluations on several homes.



Acknowledgements

Martin Arlitt, Cullen Bash, Gowtham Bellala, Hyungsul Kim, Geoff Lyon, Martha Lyons, Chandrakant Patel

References

- **[SIAM 2011]** Hyungsul Kim, Manish Marwah, Martin Arlitt, Geoff Lyon and Jiawei Han, "Unsupervised Disaggregation of Low Frequency Power Measurements", SIAM International Conference on Data Mining (SDM 11), Mesa, Arizona, April 28-30, 2011.
- **[BuildSys 2011]** Gowtham Bellala, Manish Marwah, Martin Arlitt, Geoff Lyon, Cullen Bash, "Towards an understanding of campus-scale power consumption." In ACM BuildSys, November 1, 2011, Seattle, WA.
- **[BuildSys 2012]** G. Bellala, M. Marwah, A. Shah, M. Arlitt, C. Bash, "A Finite State Machine-based Characterization of Building Entities for Monitoring and Control", Proceedings of ACM Workshop on Building Systems (Buildsys), Nov 2012.
- **[KDD 2012]** Gowtham Bellala, Manish Marwah, Martin Arlitt, Geoff Lyon, Cullen Bash, "Following the Electrons: Methods for Power Management in Commercial Buildings." In KDD, August 12-16, 2012, Beijing, China.
- **[Hart 1992]** Hart, G.W., ``Nonintrusive Appliance Load Monitoring," Proceedings of the IEEE, December 1992, pp. 1870-1891.