Utility-Privacy Trade-Offs of Data Manipulation Techniques for Smart Metering

Liang Cheng, Ph.D.

Department of Computer Science and Engineering
Lehigh University
Bethlehem, Pennsylvania 18015

cheng@lehigh.edu

Smart Grids & Smart Meters

Smart meters collecting, processing, storing, and reporting users' energy consumption data with high fidelity

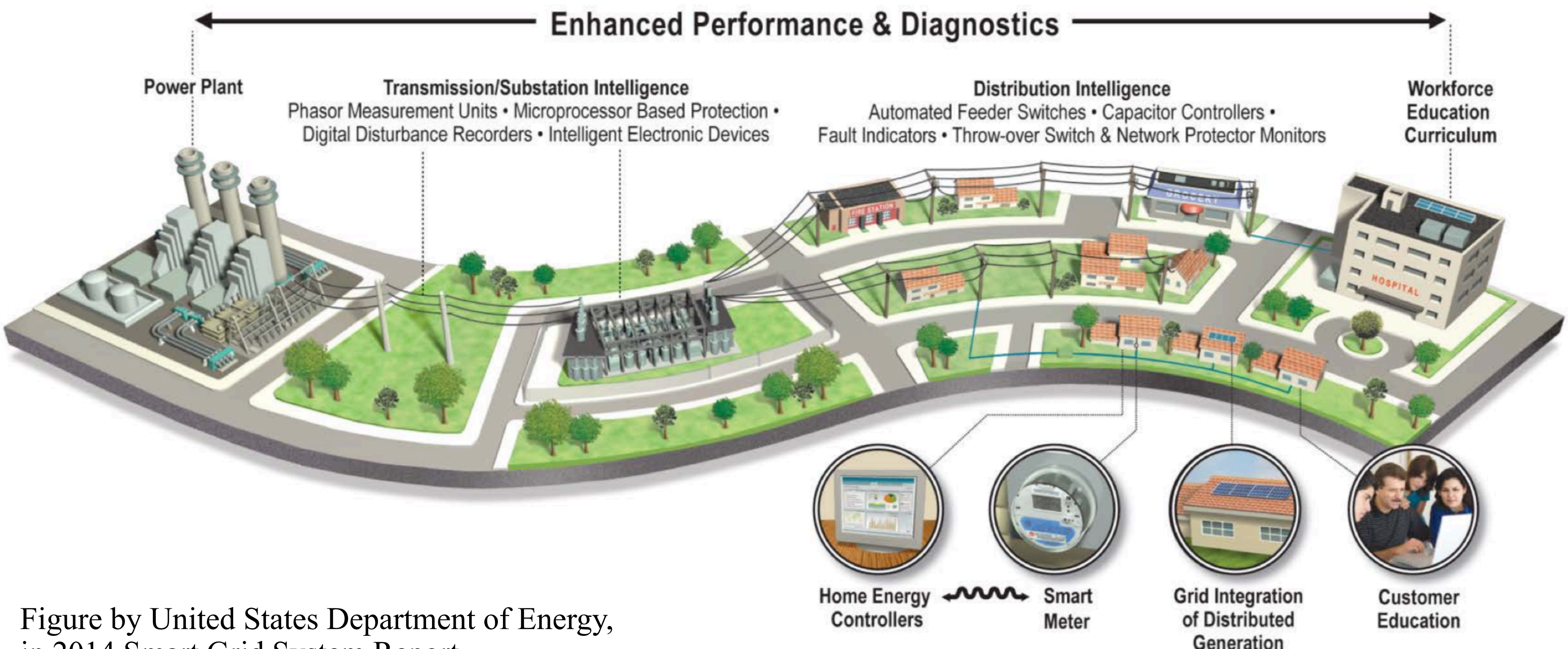


Figure by United States Department of Energy,
in 2014 Smart Grid System Report,
https://www.smartgrid.gov/files/2014-Smart-Grid-System-Report.pdf, accessed on 7/23/2019

Green Button

Allow utility customers to easily and securely access their usage information in a **consumer-friendly** and **computer-friendly** format and control data disclosure

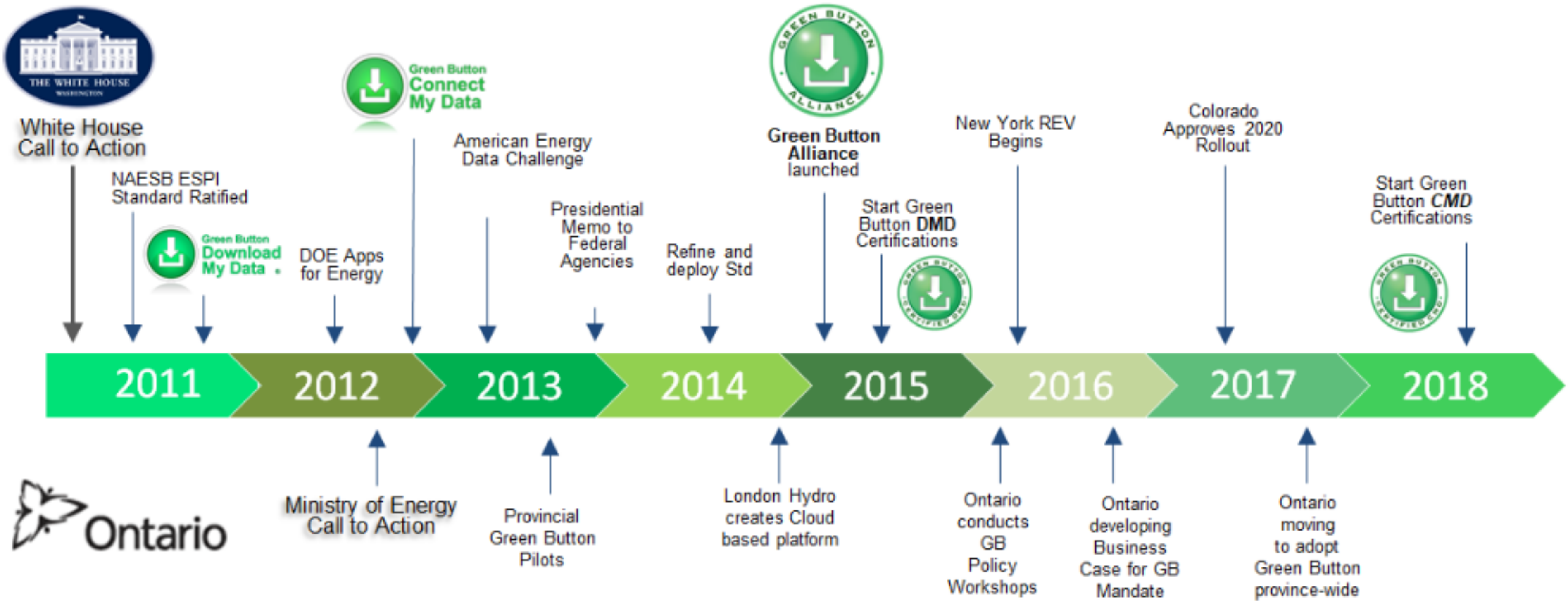
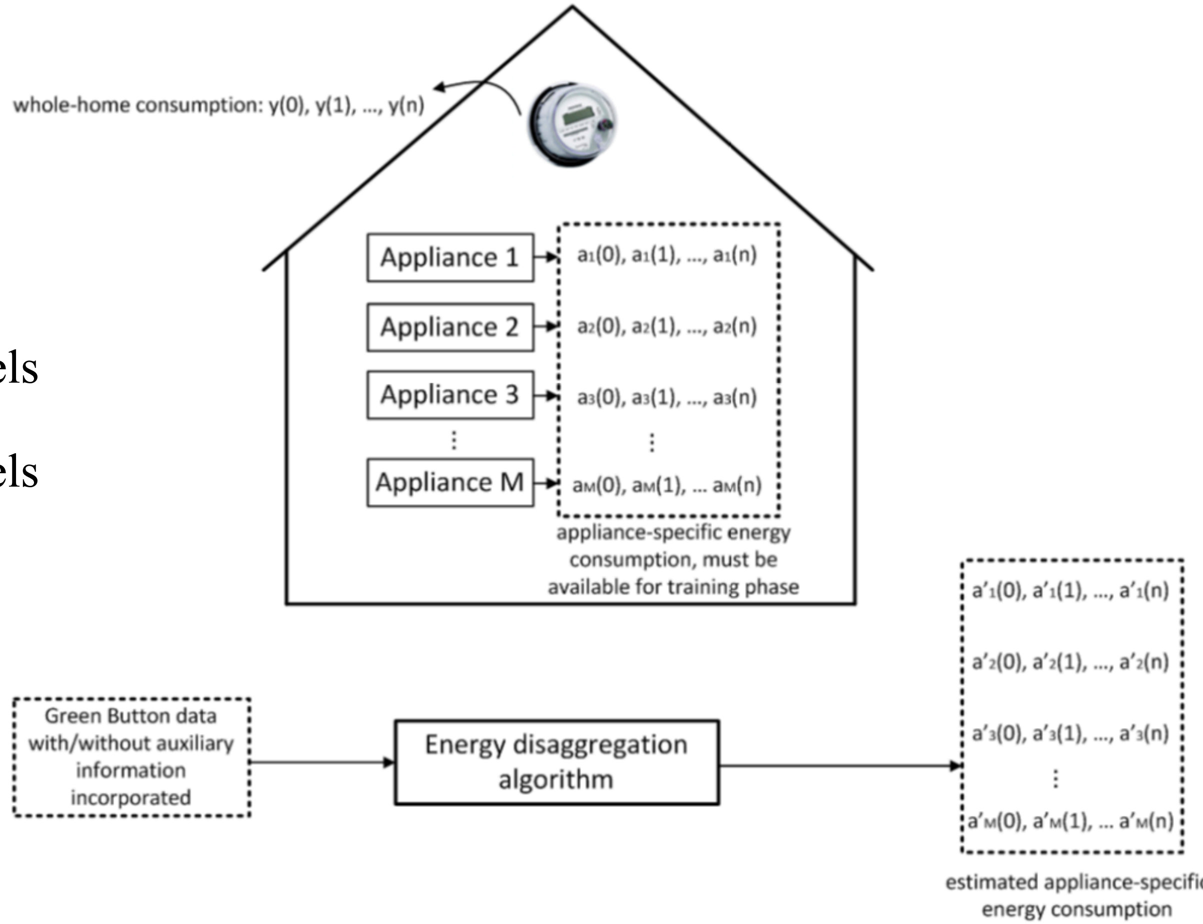


Figure by Green Button Alliance, History of Green Button and the Alliance, https://www.greenbuttonalliance.org/history, accessed on 7/23/2019

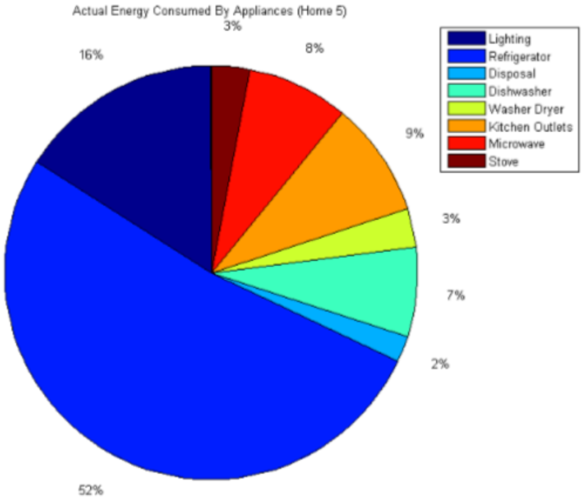
Energy Disaggregation Using Green Button Data

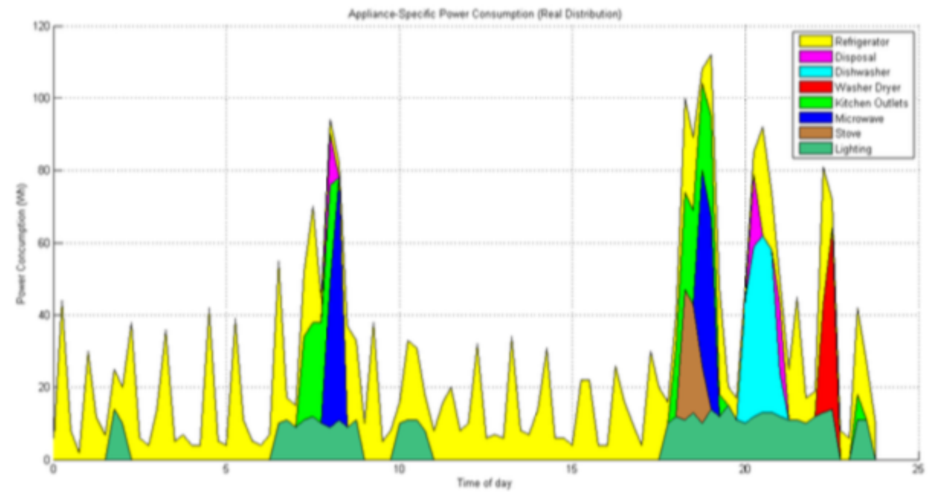
- **Markov Model algorithms**
  - Factorial Hidden Markov Models
  - Conditional Factorial Hidden Markov Models
  - Conditional Factorial Hidden Markov Models Conditional Factorial Hidden Semi-Markov Models
- **K-Nearest Neighbor (KNN) algorithm**
- **Support Vector Machine (SVM) algorithm**

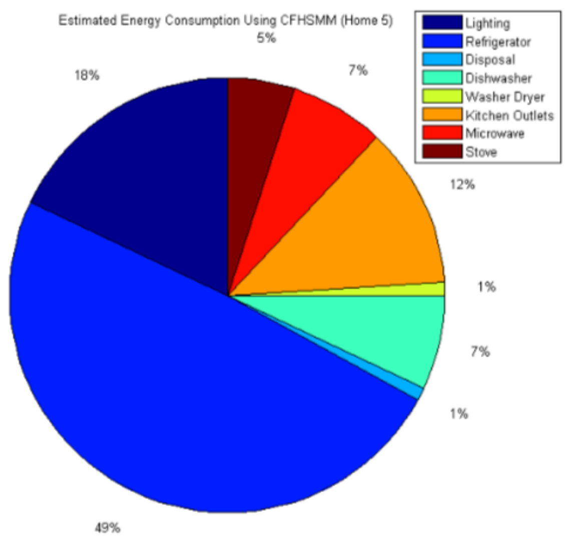


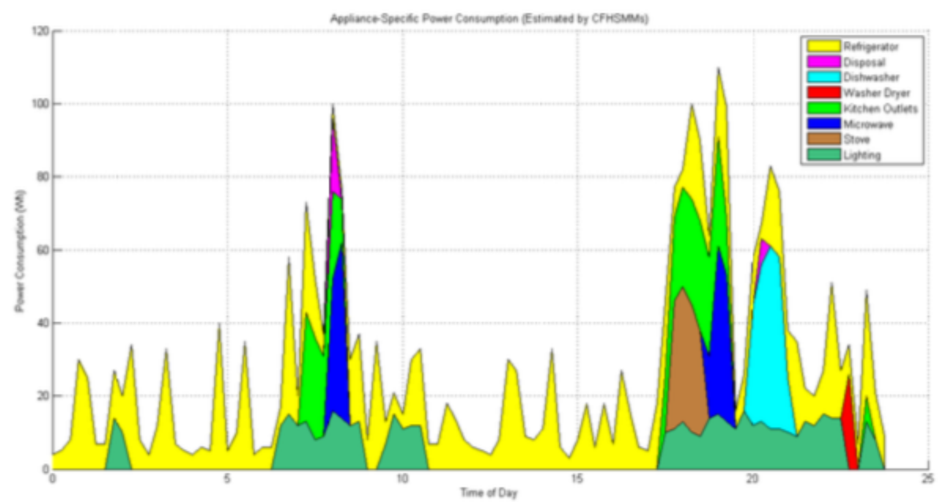
Figures by Huan Yang, *Residential Energy Data Analysis Using Green Button Data*, Technical Report LU-CSE-14-003, Department of Computer Science and Engineering, Lehigh University, 2014.

Algorithm Comparisons

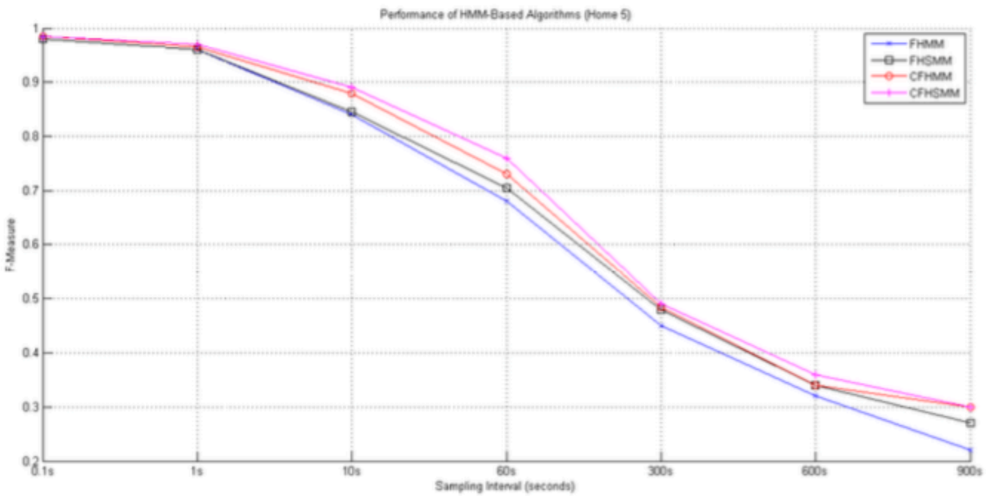






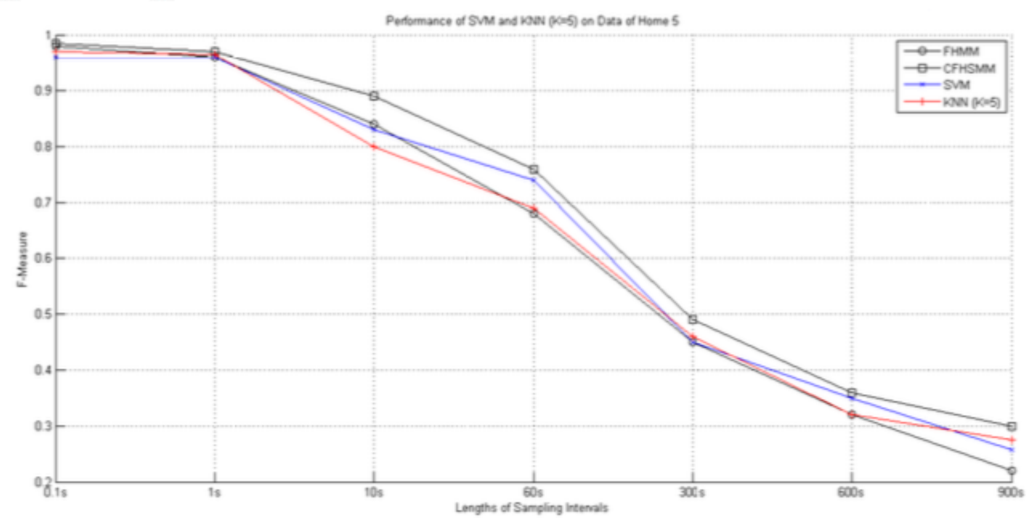


Algorithm Comparisons



- **Precision**
  - TP/(TP+FP)
- **Recall**
  - TP/(TP+FN)

F-measure = 2*Precision*Recall / (Precision+Recall)



Non-intrusive Load Monitoring

Real-Time Itemized Electricity Consumption Intelligence for Military Bases by Omid Jahromi and Alan Meier, NILM Workshop 2018

- **Recommendations:** Install CO2 sensor to control ventilation (estimated saving of 40% ventilation), Install LED lighting & motion sensors (estimated saving of 20% lighting), power-manage office equipment (e.g. disable screensavers, estimated saving of 0-20% office equipment)

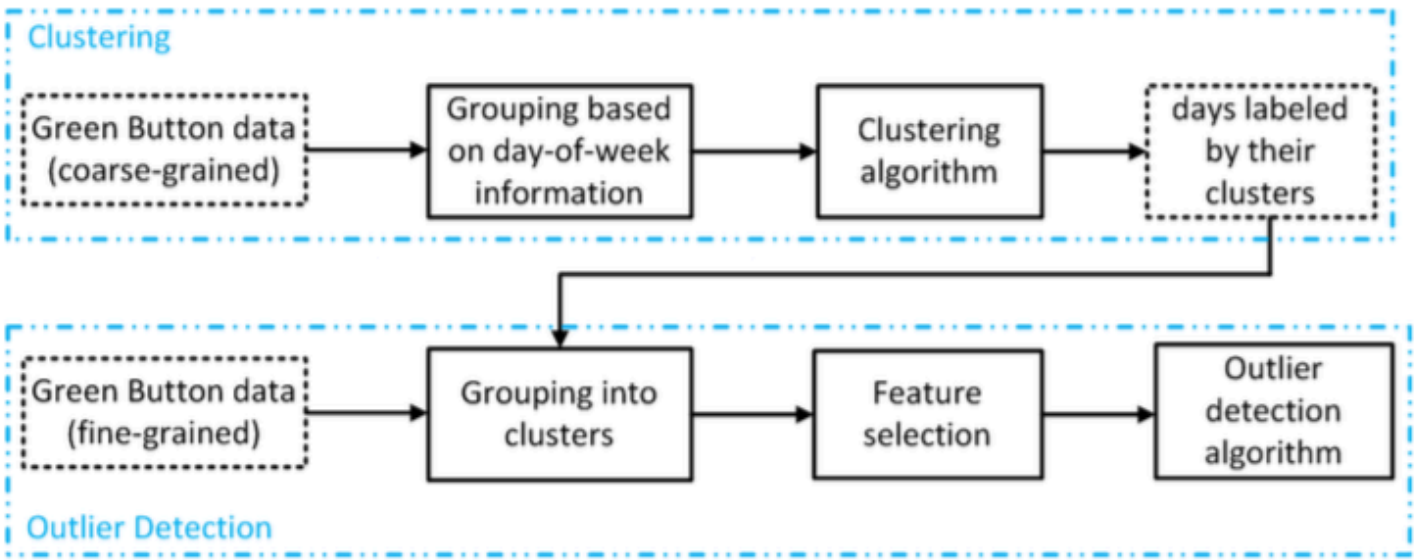
Energy Disaggregation for Commercial Buildings: A Statistical Analysis by Simon Henriet, Umut Simsekli, and Gael Richard, NILM Workshop 2018

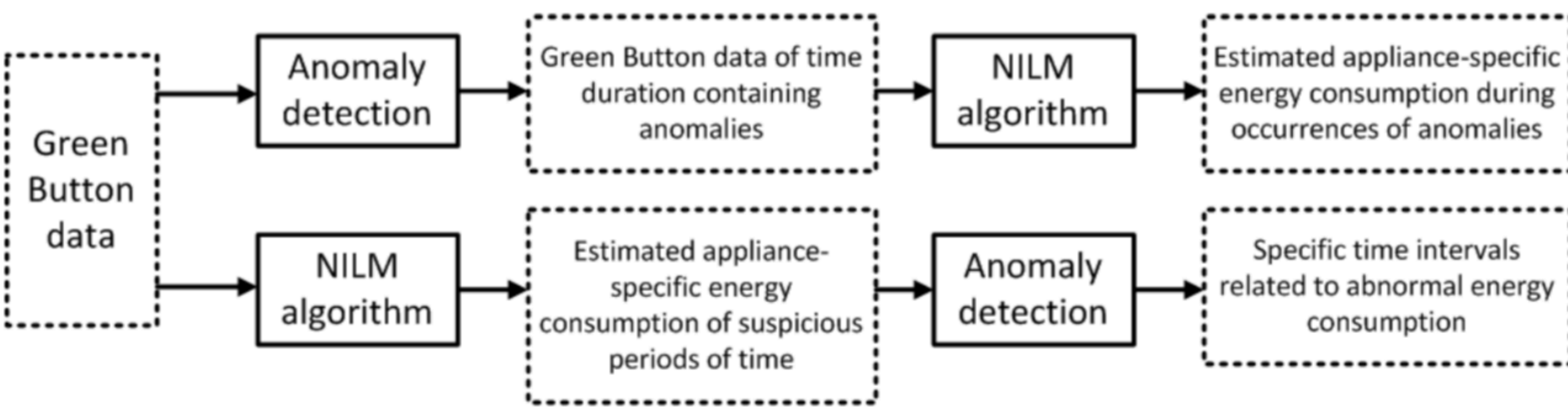
Load Disaggregation of Industrial Machinery Power Consumption Monitoring Using Factorial Hidden Markov Models by Pedro Martins, Pedro Bittencourt, and Raphael Pinto, NILM Workshop 2018

* NILM Workshop: http://nilmworkshop.org
** EU NILM Workshop: http://www.nilm.eu

and many more …

NILM + Anomaly Detection





Privacy Concerns

Sensitive information can be extracted from appliance-specific energy usages.

- **Occupancy states**
  - M. Jin, R. Jia, Z. Kang, I. C. Konstantakopoulos, and C. J. Spanos, "PresenceSense: Zero-Training Algorithm for Individual Presence De- tection Based on Power Monitoring," in 1st ACM Conf. on Embedded Systems for Energy-Efficient Buildings, 2014, pp. 1–10.
- **User activity patterns**
  - J. Alcala, J. Urena, and A. Hernandez, "Activity Supervision Tool Using Non-Intrusive Load Monitoring Systems," in 2015 IEEE Conf. on Emerging Technologies Factory Automation, 2015, pp. 1–4.
- **Multimedia contents being played on a TV set**
  - U. Greveler, P. Glosekotterz, B. Justusy, and D. Loehr, "Multimedia Content Identification through Smart Meter Power Usage Profiles," in Int. Conf. on Information and Knowledge Engineering, 2012.

Privacy Protection Techniques

- **Encryption-based** techniques
  - F. Benhamouda, M. Joye, and B. Libert, "A New Framework for Privacy-Preserving Aggregation of Time-Series Data," ACM Trans. Inf. Syst. Secur., vol. 18, no. 3, pp. 10:1–10:21, 2016.
- **Battery-based load hiding** (BLH) techniques
  - J. Zhao, T. Jung, Y. Wang, and X. Li, "Achieving Differential Privacy of Data Disclosure in the Smart Grid," in IEEE Conf. on Computer Communications, 2014, pp. 504–512.
  - L. Yang, X. Chen, J. Zhang, and H. V. Poor, "Optimal Privacy-Preserving Energy Management for Smart Meters," in IEEE Conf. on Computer Communications, 2014, pp. 513–521.
- **Data manipulation** techniques
  - P. Barbosa, A. Brito, and H. Almeida, "Defending Against Load Monitoring in Smart Metering Data Through Noise Addition," in 30th Annu. ACM Symp. on Applied Computing, 2015, pp. 2218–2224.

Utility-Privacy Tradeoff

- How well can data manipulation techniques prevent leakage of appliance-level energy consumption information?
- When are investments on BLH techniques necessary to protect privacy?

- Adversary model
- Data utility model
- Privacy model

Definition 1 (Data Utility Metric): Given two time series X_i^T and $\hat{X_i^T}$ for appliance i, the distortion between $\hat{X_i^T}$ and X_i^T can be measured by their distance $d(X_i^T, \hat{X_i^T})$. Suppose that there are N samples in X_i^T (and $\hat{X_i^T}$), we use the average distortion $\bar{d} = \frac{d(X_i^T, \hat{X_i^T})}{N}$ as the utility metric for i.

The greater the distortion, the less the utility.

Definition 2 (Privacy Metric): Given two time series X_i^T and $\hat{X_i^T}$ over the same time period T for appliance i, the mutual information $I(X_i^T, \hat{X_i^T})$ between the two series is

$$I(X_i^T, \hat{X_i^T}) = \sum_{x \in X_i^T} \sum_{y \in \hat{X_i^T}} \ln \frac{p(x,y)}{p(x)p(y)},$$

where $p(x)$ and $p(y)$ are the probability density functions of random variables $x \in X_i^T$ and $y \in \hat{X_i^T}$, and $p(x,y)$ is the joint probability density function.

The greater the mutual information, the more the privacy leakage.

Experiment Settings

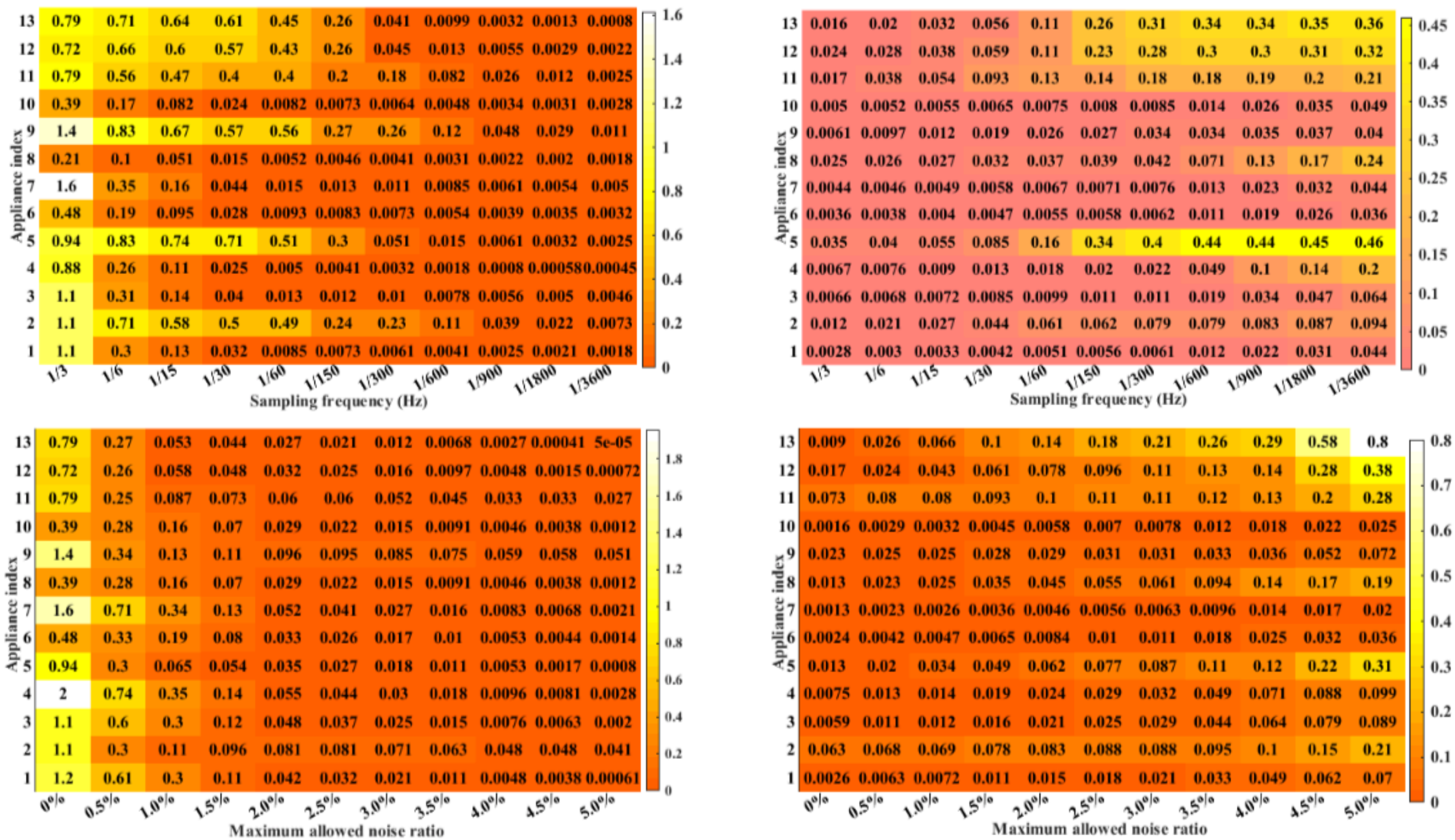
- 1 house, 13 appliances
- Sampling rates (<=1/3 Hz)
- Training set (1 week) and Testing set (11 days)
- Random noise w/ uniform distribution
- FHMM algorithm
- A 50W bin size for computing mutual information

LIST OF APPLIANCES IN THE EXPERIMENTS

Index	Appliance	Index	Appliance
1	oven	2	refrigerator
3	dishwasher	4	kitchen-outlets-1
5	lighting-1	6	washer-dryer
7	microwave	8	bathroom-gfi
9	electric-heat	10	stove
11	kitchen-outlets-2	12	lighting-2
13	lighting-3		

Figures by J. Z. Kolter and M. J. Johnson, “REDD: A Public Data Set for Energy Disaggregation Research,” in SustKDD workshop on Data Mining Applications in Sustainability, 2011.

Evaluation Results



Huan Yang, Liang Cheng, and Mooi Choo Chuah, Evaluation of Utility-Privacy Trade-Offs of Data Manipulation Techniques for Smart Metering, 2016 IEEE Conference on Communications and Network Security (CNS): International Workshop on Cyber-Physical Systems Security (CPS-Sec), Philadelphia, PA, October 19, 2016.

Conclusion

- NILM / Energy disaggregation may enhance energy consumption awareness and enable additional smart grid applications
- NILM / Energy disaggregation may reveal private information of energy consumers
- Users can balance between privacy revelation and data utility by choosing the proper data manipulation techniques
  - Although extra investments on batteries and control infrastructure are not required, the granularity of control supported by these techniques is coarse

This work was supported by PPL Corporation and Lehigh University. Any opinions, findings, and conclusions or recommendations expressed in this talk/paper are those of the author(s) and do not necessarily reflect the views of the sponsors of the research.