Markovian Models for Electrical Load Prediction in Smart Buildings

Dr. Ihsan Ayyub Qazi

Muhammad Kumail Haider

Asad Khalid Ismail

Outline

- Motivation
- Goal
- Approach
- Evaluation
- Conclusion

Motivation

- Increasing trends towards the Smart Grid
- Developing energy consumption models for smart building important for
  - Demand response
  - Home energy management
  - Distribution network simulation
  - Load prediction/forecasting

Motivation

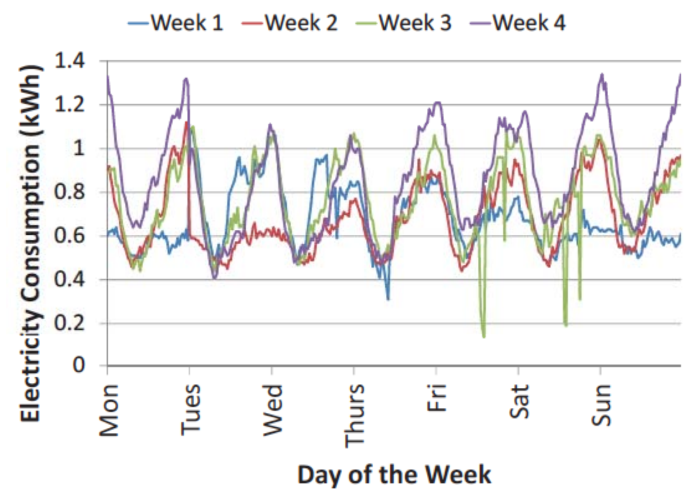
- Prior models lack parsimony
  - Large number of models inputs
  - e.g., set of appliances, the number of occupants, appliance load models, and the occupancy behavior
    - Obtaining them in any realistic situation is challenging
- Due to the need for accuracy and parsimony, this focuses on building Markovian Models
  - Electricity consumption arises from the superposition of a finite set of on-off loads from individual appliances
    - Such superposition have been shown to be well-modeled using Markovian models
  - Markov models have been extensively used to model sequential events (e.g., in human speech) and have been shown to combine both parsimony and descriptive power

Goal of this Work

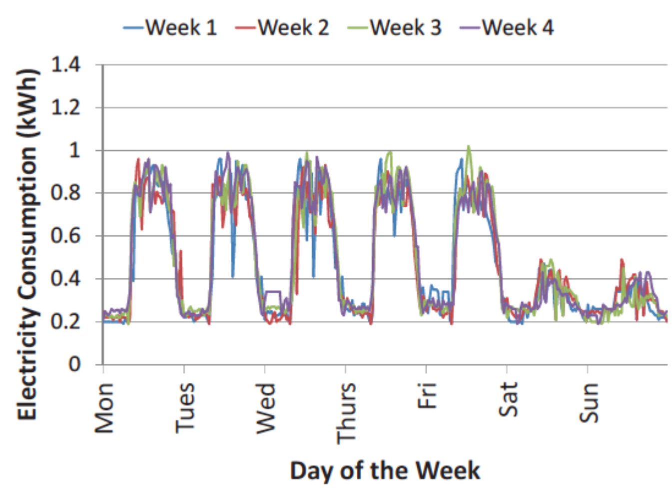
- Given prior load data, derive Markovian models that allow load prediction for different periods of a day as well as for the entire day
  - To this end, we propose analytical models based on Continuous Time Markov Chains (CTMC) and Discrete Time Markov Chains (DTMC)
  - We collect real load measurements for two buildings with widely varying load profiles

Datasets

- Dataset-1
  - Hostel building with 360 occupants at a large University in Pakistan
  - Duration: 7 months
- Dataset-2
  - Large software company
  - Duration: 1 year
- Measurement granularity: 30 mins



Monthly Consumption (dataset-1)



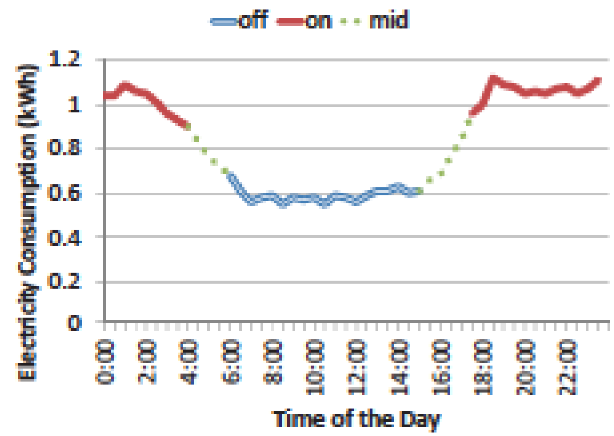
Monthly Consumption (dataset-2)

Key Questions

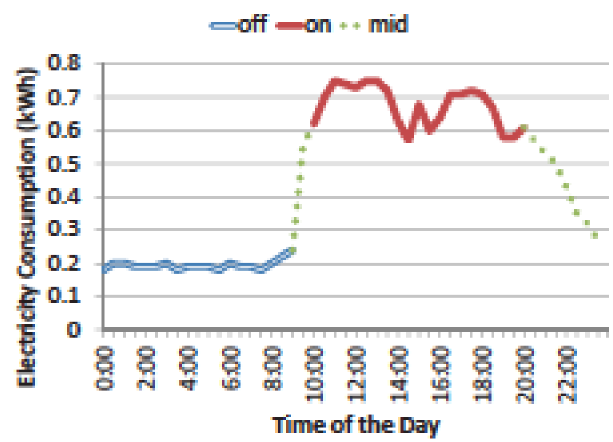
- How many Markov models do we need?
- How are the Markov states chosen?
- How many states do we need in each model?
- How do we merge models to allow prediction for the entire day?

Number of Markov Models

- Depends on the correlation between different periods of the day
- For our data sets, we found that a day could be divided into three periods
  - ON, MID, and OFF
  - 3 Models: One Markov model for each period
  - For dataset-2, we need separate models for weekdays and weekends



Single-day Consumption (dataset-1)



Single-day consumption (dataset-2).

Choosing states of Models

- States chosen using the k-mean clustering algorithm
  - Helps in finding k centroids of the data set which form the k states of the model
  - Results in a quantized data set which used for deriving transition probability and rate matrices for DTMCs and CTMCs, respectively
- Number of states depend on the variability in load during a period
  - Need a good-of-fit metric

Merging models

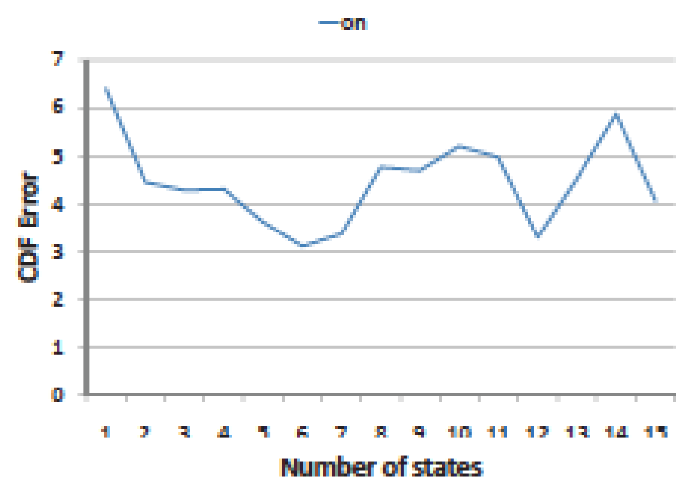
- We derive models for each period in a day
  - Need to merge models to allow prediction for the entire day
  - For enabling smooth transitions, the Markov process transitions to the nearest state in the new period

Evaluation

- Goodness-of-metric for comparing models

$$CDF_{error} = \frac{Area\,between\,CDF\,of\,measured\,load\,and\,modeled\,load}{CDF\,Area\,of\,the\,measured\,load} \times 100.$$

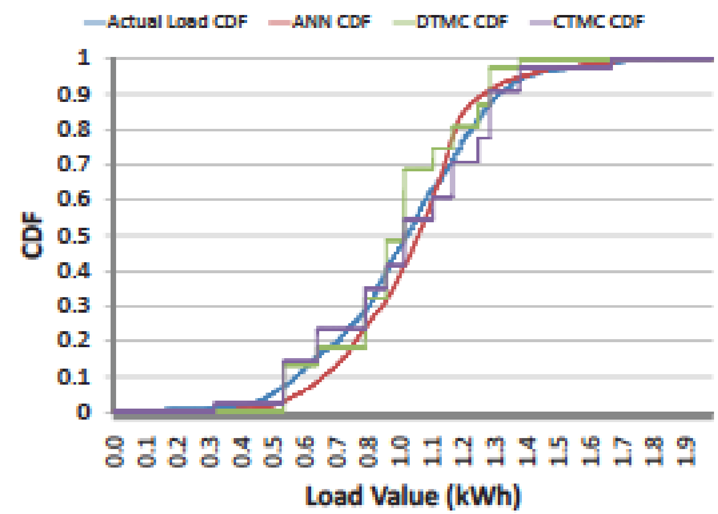
  - represents how far two probability distributions are from each other
- Insight-1
  - Suppose we choose the number of states for each period such that the CDFerror in individual periods is minimized
  - Will it also minimize CDFerror in the merge model
    - Not necessarily!
    - Why? Boundaries of ON, MID, and OFF periods are correlated and boundary points form a reasonable portion of the sample set



(b) CDF_{error} for merged models

Evaluation Results

- CDF for CTMC , DTMC, and a ANN
  - ANN trained using the Levenberg-Marquardt back propagation algorithm
- MSE
  - ~0.08 for ANN
  - ~0.077 for the Markov models



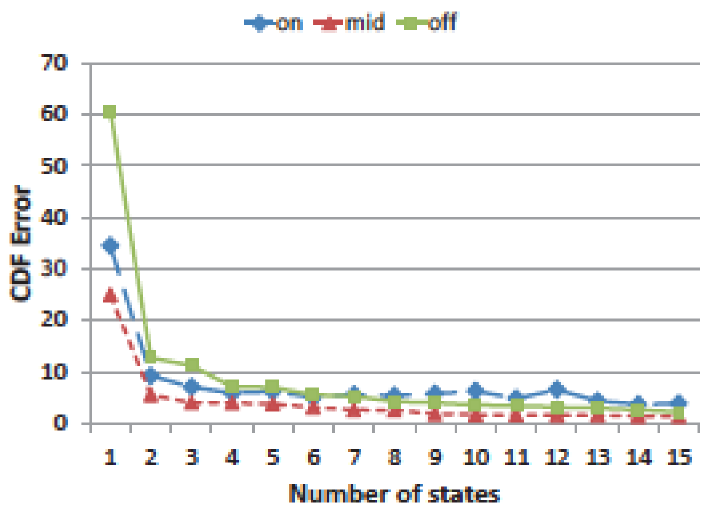
(a) CDFs for three models

Conclusion

- Proposed k-state Markov models for load prediction
- Our results show that the model requires small number of states, are parsimonious, and are accurate enough for load prediction

Thank you

Backup Slides



(a) CDF_{error} for individual models