Markovian Models for Electrical Load Prediction in Smart Buildings

Muhammad Kumail Haider, Asad Khalid Ismail, and Ihsan Ayyub Qazi

LUMS School of Science and Engineering, Lahore, Pakistan
{13100183,13100004,ihsan.qazi}@lums.edu.pk

Abstract. Developing energy consumption models for smart buildings is important for studying demand response, home energy management, and distribution network simulation. In this work, we develop parsimonious Markovian models of smart buildings for different periods in a day for predicting electricity consumption. To develop these models, we collect two data sets with widely different load profiles over a period of seven months and one year, respectively. We validate the accuracy of our models for load prediction and compare our results with neural networks.

Keywords: Smart Grid, Load Prediction, Markov Processes

1 Introduction

The rapid ongoing deployment of the smart grid has led to an increasing need for accurate and parsimonious models for electrical loads in smart buildings, such as smart homes and offices [1]. Building accurate models is important for several smart grid components and applications. These include modeling of demand response, home energy management, and distribution network sizing. Most attempts at characterizing loads of smart buildings has led to models that lack parsimony as their inputs include the set of appliances, the number of occupants, appliance load models, and the occupancy behavior. Obtaining these inputs in any realistic situation is challenging [2]. Motivated by the need for accuracy and parsimony, in this work, we develop models for load prediction in smart buildings and focus on Markovian models.

Given prior load data, we derive models that allow prediction for different periods of a day. We then analyze the *combined* behavior of these models in predicting the load for the entire day. To this end, we propose analytical models based on Continuous Time Markov Chains (CTMC) and Discrete Time Markov Chains (DTMC). Further, we compare the accuracy of our models with an Artificial Neural Network (ANN). To develop these models, we collect measurements from two buildings with widely different load profiles over a period of seven months and one year, respectively. Our results show that each period of a day can be modeled accurately by a Markov model with only few states. We validate our approach by comparing the cumulative distribution function (CDF) of the load generated by numerical simulation of the models with the actual load.

We make four contributions: (i) we collect real load measurements at a granularity of 30 minutes for buildings with widely different load profiles, (ii) we develop parsimonious Markov models for different periods of the day, (iii) we perform comparison with an ANN, (iv) we validate that our models provide good accuracy for predicting electrical loads.

The rest of the paper is organized as follows: we discuss related work in Section 2, describe our data sets in Section 3, and present models in Section 4. We evaluate the models in Section 5 and offer concluding remarks in Section 6.

2   Background and Related Work

Many deterministic and probabilistic models have been proposed for load forecasting in the past [3, 4]. It is important for models to be parsimonious as highly parameterized models make it difficult to cover model space when doing simulations or analysis. However, parsimony comes at a cost as a parsimonious model has to necessarily ignore one or more of the many factors that determine the load of smart buildings. These factors include the type of appliances, time of day, size of the house, etc. In this work, we focus on Markov models for two reasons. First, electricity consumption arises from the superposition of a finite set of on-off loads from individual appliances. Such superpositions have been shown in the past to be well-modeled using Markovian models. Second, 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 [5].

ANNs are commonly used for load forecasting but can require large number of parameters for achieving high accuracy [4, 6]. Omid et al. proposed CTMC models for transformer sizing which relates to prediction of peak load [7]. Detailed models for smart buildings, such as residential buildings, have also been presented in the civil engineering, power engineering, and environmental studies [2]. However, these models lack parsimony and tend to be highly parameterized.
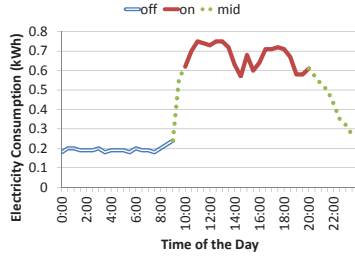
3   Datasets

We consider two data sets with different load profiles and measure the real-time energy consumption (in kWh) at a granularity of 30 minutes, which is standard in modern smart meters. The first data set we collected is of a male hostel building with 360 occupants at a large University in Lahore, Pakistan. Figure 1 shows the load variations during a single day as well as across weeks during the month of November in 2011. Observe that the consumption tends to remain high from 6 pm till 2 am and is much lower between 6:30 am and 3 pm even though variations in load exist during these times. During other times, the load generally remains moderate. Moreover, these trends exhibit a great degree of periodicity (see Figure 1(b)) and seem to match the typical academic activities in a University. For instance, during daytime students are on campus and consumption is expected to be low. The second data set we consider is of a large software company, which exhibits the load characteristics of a typical office (see Figures 1(c) and 1(d));
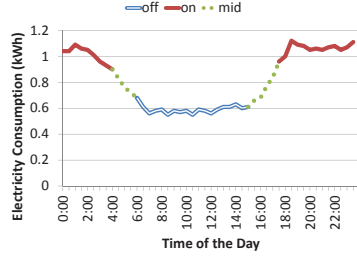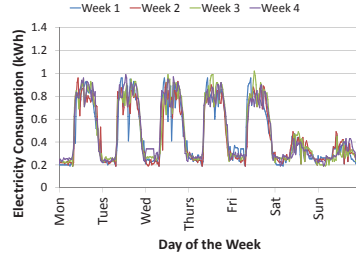
(a) Single-day Consumption (dataset-1).
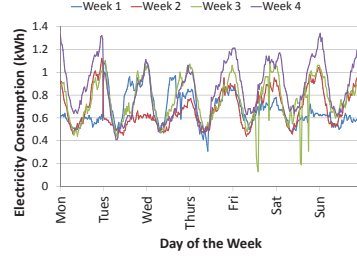
(b) Monthly Consumption (dataset-1).

(c) Single-day consumption (dataset-2).

(d) Monthly Consumption (dataset-2).

Fig. 1. Electric load variations in dataset-1 and dataset-2.

with high consumption from 9 am to 5 pm, low consumption between 12 am-9 am (which represents the *base load*), and medium consumption during other times, suggesting that people tend to work beyond their office hours. Unlike in dataset-1, we observe that weekends exhibit very different load profile compared to weekdays as shown in Figure 1(d).

4   Electricity Consumption Models

We now describe our approach for building Markov models by answering the following questions: (a) *how many Markov models do we need?*, (b) *how are the Markov states chosen?*, (c) *how many states do we need in each model?*, and (d) *how do we merge models to allow prediction for the entire day?*

A Markov chain is a stochastic process which satisfies the Markovian property i.e., *the conditional distribution of any future state, given past states, depends only on the present state*. In DTMCs, state transitions happen at specific time steps whereas in CTMCs they can happen at any time. The dynamics of a k-state DTMC and CTMC are represented by a $k \times k$ transition probability matrix and a $k \times k$ transition rate matrix (denoted by Q), respectively, whose elements q_{ij} represent the rate/probability of departing from state i to arrive at state j. In our Markov models, we assign a value R_i to each state i, which represents

the amount of energy consumed in this state. Consequently, the load is modeled using a k-state Markov process defined by the $\{Q, R\}$ tuple.

The number of Markov models we need depends on the load profile e.g., in our data sets, load is highly sensitive to the time of day; the probability that the minimum energy is consumed at 10 pm in dataset-1 is far lower than at 10 am. To deal with this issue, we divide the load profile of a day into three periods: *on*, *mid* and *off* peak such that the consumption levels remain almost stationary during these periods. Therefore, for each data set, we need three models. In general, however, these periods might span different hours of the day depending on geographical location or seasons. Also, we observed that in case of dataset-2, we need separate models for weekdays and weekends due to different load characteristics.

The number of states in a Markov model depends on the variability in load during the defined periods. For this, we need a goodness-of-fit metric, which we discuss in Section 5. For making a k-state Markov model for a given data set, we first use the k-means clustering algorithm to find k centroids of the data set; these centroids are the required k states of our model and their values act as quantization levels for electricity consumption, giving us a quantized data set. This quantized data set is then used for deriving transition rate and transition probability matrices for CTMCs and DTMCs, respectively. For a k-state CTMC,

$$q_{ij} = \frac{no.\,of\,transitions\,from\,R(i)\,to\,R(j)\,in\,clustered\,load}{total\,time\,spent\,in\,state\,i\,before\,a\,transition\,to\,state\,j} \tag{1}$$

where $i \neq j$ and $q_{ii} = -\sum_{j=1,j\neq i}^{k} q_{ij}$ such that each row of the Q matrix sums to zero. The entries of the Q matrix for a DTMC are defined as:

$$q_{ij} = \frac{no.\,of\,transitions\,from\,R(i)\,to\,R(j)\,in\,clustered\,load}{total\,no.\,of\,transitions\,from\,state\,i\,to\,any\,other\,state\,(including\,i)} \tag{2}$$

Once the Q matrices are derived, we perform numerical simulation of the Markov models to predict load. To allow prediction of at least 24 hours, we need to merge models for different periods. For enabling smooth transitions, the Markov process transitions to the nearest state in the new period and uses the corresponding Q matrix from there on.

5   Evaluation

To compare Markov models with different number of states, we need a metric that captures the efficacy of a model. We use the area between the CDF of the measured load and the CDF obtained from the Markov model, which represents *how far* two probability distributions are from each other [7]. We then normalize this by the area under the CDF of measured load to calculate the CDF error as:

$$CDF_{error} = \frac{Area\,between\,CDF\,of\,measured\,load\,and\,modeled\,load}{CDF\,Area\,of\,the\,measured\,load} \times 100. \tag{3}$$

(a) CDF_{error} for individual models   (b) CDF_{error} for merged models
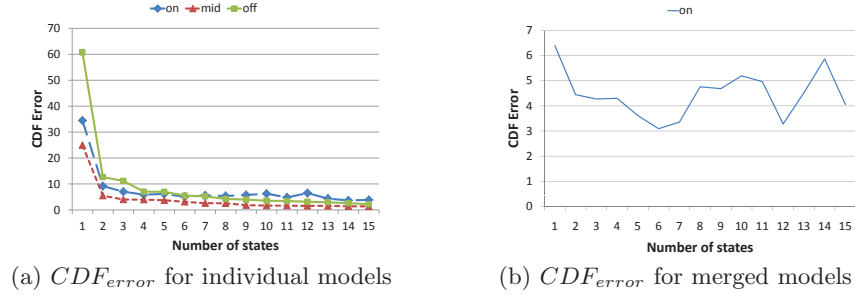
Fig. 2. Testing individual model's suggestion of optimum number of states

The higher the value of this metric, the greater the modeling error. Note that when the modeled load exactly matches the measured load, CDF_{error} is zero.

Due to the measurement granularity of 30 mins in our data sets, we have *only* 48 load samples per day. Consequently, we obtain the model parameters (i.e., the Q and R matrices) by considering data of a month, which gives us a total of 1440 samples. This allows us to effectively capture the load variations in the data sets. In addition, the study of our data sets reveals that consumption patterns do not show much change during a month. With improved sampling granularity, the same models can be used while considering less history [2][7].

We first evaluate the Markov models using CTMC for each period separately. To do this, the load profile is divided into *on*, *mid* and *off* peak periods based on the time of day, and then these smaller data sets are used to obtain the Q and R matrices. Figure 2(a) shows the CDF_{error} as a function of the number of states for each model for dataset-1 for November 2011. Observe that the CDF_{error} decreases with the number of states and is minimum when 15 states are used.

Intuitively, using 15-state Markov models for each period should yield the minimum CDF_{error} if the periods are uncorrelated. To validate this, we analyze the merged model in which all three models are used together to generate a day-long load profile. We use 15 states for *mid* and *off* periods and vary the number of states in the *on*-model. Figure 2(b) shows that we can achieve the minimum CDF_{error} using smaller number of states in the merged model than by using 15-states in each model (Note that similar results were obtained when the *mid* and *off* states were varied). This is due to the fact that the boundaries of the *on*, *mid* and *off* periods, where state transitions occur across period models, are correlated and form a reasonable portion of the sample set (e.g., 5 out of 48 samples in a day for dataset-1 are on these boundaries). This inter-dependence between models increases if we divide a day into more periods. Due to this, running the models independently and then stitching together the results does not correspond well with the results suggested by the merged model.

Comparison between CTMC and DTMC models: We use a 3-tuple (a,b,c) to represent the number of states used in the *on*, *mid* and *off* models, respectively. To compare these models, we find the least number of states which
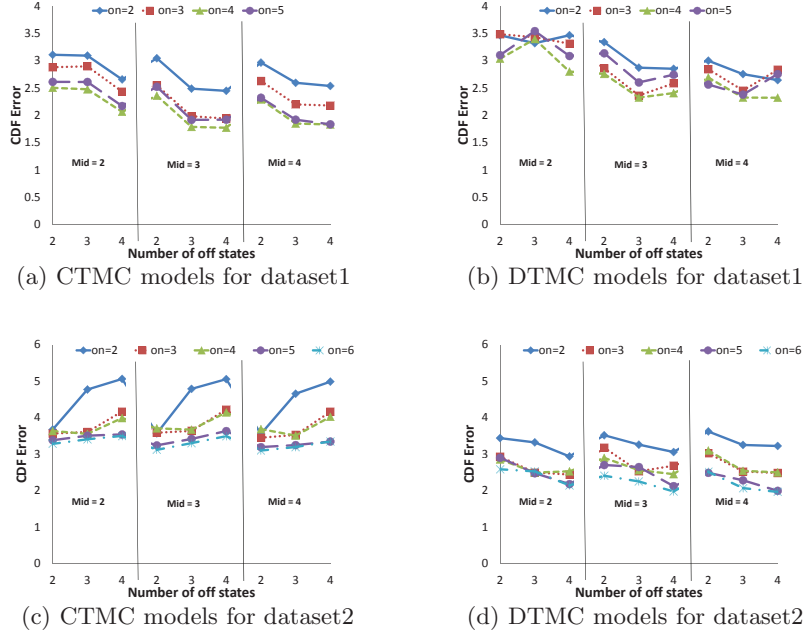
(a) CTMC models for dataset1   (b) DTMC models for dataset1

(c) CTMC models for dataset2   (d) DTMC models for dataset2

Fig. 3. Modeling errors with variations in states

give us a CDF_{error} below a threshold of 3%. Figure 3(b) shows the CDF_{error} of the DTMC for the month of December. Observe that a minimum of 8 states are required to obtain an error lower than 3% corresponding to tuples such as (2,3,3). With CTMC, 7 states are required (e.g., with (3,2,2)) as shown in Figure 3(a). In dataset-2, the DTMC and CTMC models require 7 and 10 states corresponding to tuples (3,2,2) and (5,3,2), respectively (see Figures 3(d) and 3(c)). The CTMC model requires much more states in dataset-2 than the DTMC model because of the larger variability in the *on* period of the data set, which increases the penalty of a misprediction, especially when the process spends more than 30 minutes in a state. Increasing the number of states decreases the likelihood of a large deviation from the true load, thus reducing the penalty of a misprediction. Another trend observed in Figure 3(c) is the increase in CDF_{error} with the number of *off* states. Since the *off*-period captures the base load (which is fairly constant), 2 states are enough to model the load and increasing the number of states results in inaccurate centroids, which increases the error. These results show that modeling error depends on how much time the process spends in a period and how variable the consumption is within that period.

Prediction: We now use CTMC and DTMC based models for load prediction. In particular, we derive the Q and R matrices for the month of November to predict load for the next 15 days of December. Figure 4(a) shows the CDF_{error} of the CTMC model for a 15-day prediction for dataset-1. Observe that the
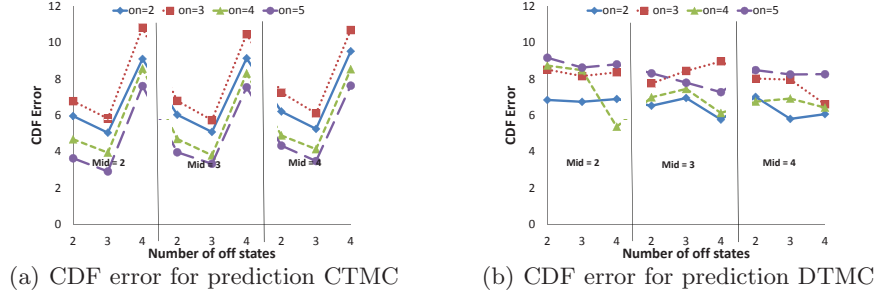
(a) CDF error for prediction CTMC   (b) CDF error for prediction DTMC

Fig. 4. CTMC and DTMC for 15 days prediction

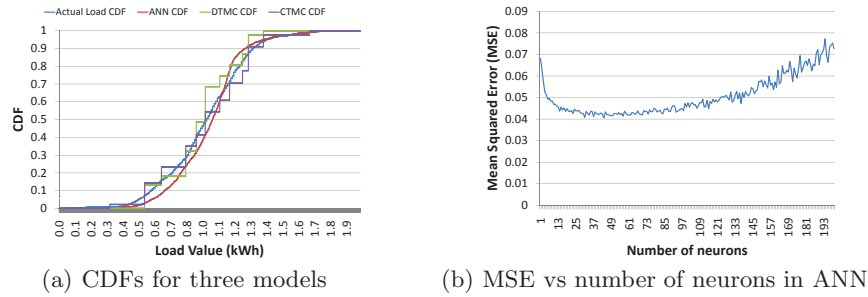(a) CDFs for three models   (b) MSE vs number of neurons in ANN

Fig. 5. Error comparisons for ANN and Markov models

minimum CDF_{error} is 2.92% given by (5,2,3). In case of DTMC, the minimum CDF_{error} is 5.37% given by (4,2,4) as shown in Figure 4(b). These results show that Markov models based on CTMCs and DTMCs can be used for prediction with CTMCs yielding potentially better results for data sets similar to dataset-1.

5.1 Comparison with Artificial Neural Networks

The neural network we use is a two-layer feed-forward network with sigmoid hidden layer neurons and linear output neurons. The network is trained using the *Levenberg-Marquardt* back propagation algorithm using its implementation available in the Neural Network Toolbox in MATLAB [8]. A commonly used error metric for ANNs is the Mean-Squared Error (MSE). In order to compare the results obtained using the neural network and the Markov models, we use both MSE as well as CDF_{error}. Figure 5(a) shows the CDF of the actual load (for dataset-1), ANN predicted CDF, DTMC predicted CDF, and CTMC predicted CDF. Observe that the Markov models show a *stair-case* kind of CDF. This is due to the quantization of load values based on k-means clustering. The CDF_{error} for the CTMC model with 6 *on* states, 3 *mid* states and 3 *off* states was 4.1%, which was the minimum we observed across all the models. The

CDF_{error} for the neural network with eighty hidden layer neurons and 30%, 35%, 35% training, validation and testing sample percentages, respectively was 4.9%. We use eighty neurons because we found that the CDF_{error} stabilized after ≈ 60 neurons. Moreover, we tested various combinations of sample percentages for the three phases and found the MSE and CDF_{error} to be minimum for the above combination. The CDF_{error} for the DTMC Model with 6 *on* states, 3 *mid* states and 3 *off* states was 5.4%.

Figure 5(b) shows the MSE of ANN as a function of the number of neurons. While the MSE remains less than 0.08, it increases when the number of neurons are increased beyond ≈ 80. This trend is likely due to over-fitting of data which reduces ANN's ability to generalize. In case of Markov models, we did not observe any significant variations in MSE as we varied the number of states. However, the mean MSE for the Markov model as a function of the states (not shown due to space limitations) was ≈ 0.077. Note that MSE may not be an ideal metric for evaluating Markov models as it does not capture the distribution of loads.

6   Conclusion and Future Work

We presented Markov models for load prediction and performed comparison with ANNs. We derived k-state models based on CTMCs and DTMCs for different periods of the day. Our results show that the models require small number of states, are parsimonious, and are accurate enough for load forecasting. In the future, we plan to evaluate these models on data sets with finer-granularity.

References

1. Keshav, S., Rosenberg, C.: Direct adaptive control of electricity demand. Technical Report CS-2010-17, University of Waterloo, September (2010)
2. Richardson, I., Thomson, M., Infield, D., Clifford, C.: Domestic electricity use: A high-resolution energy demand model. In: Energy and Buildings, vol. 42, issue 10, pp. 1878-1887 (2010)
3. Asrari, A., Javan, D.S., Javidi, M.H., Monfared, M.: Application of gray-fuzzy-markov chain method for day-ahead electric load forecasting. In: Przeglad Elektrotechniczny-Electrical Review, vol. 2012, no. 3, pp. 228-237 (2012)
4. Hippert, H.S., Pedreira, C.E., Souza, R.C.: Neural networks for short-term load forecasting: A review and evaluation. In: IEEE Transactions on Power Systems., vol. 16, pp. 44-55 (2001)
5. Rabiner, L.R.: A tutorial on hidden markov models and selected applications in speech recognition. In: Proc. IEEE, vol. 77, no. 2, pp. 257-285, Feb. (1989)
6. Hayati, M., Shirvany, Y.: Artificial neural network approach for short term load forecasting for illam region. In: International Journal of Electrical, Computer, and System Engineering, vol. 1, no. 2, pp. 121-125 (2007)
7. Ardakanian, O., Keshav, S., Rosenberg, C.: Markovian models for home electricity consumption. In: Proc. ACM SIGCOMM Green Networking Workshop (2011)
8. Neural Network Toolbox, `http://www.mathworks.com/products/neural-network/`