Resource Provisioning in the Electrical Grid

by

Omid Ardakanian

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2011

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Transformers and storage systems in the electrical grid must be provisioned or sized just as routers and buffers must be sized in the Internet. We prove the formal equivalence between these two systems and use this insight to apply teletraffic theory to sizing the electrical grid, obtaining the capacity region corresponding to a given transformer and storage size. We conduct a fine-grained measurement study of household electrical load. These measurements are essential for two reasons. First, we use them to construct reference models for home loads; these models are used to find the capacity region using the teletraffic theory. Second, these measurements are used in numerical simulations that are done to validate our analysis. More specifically, we compare results of numerical simulations with the results from teletraffic theory. We show not only that teletraffic theory agrees well with numerical simulations but also that it closely matches with the heuristics used in current practice. Moreover, our analysis permits us to develop sizing rules for battery storage electrical grid, advancing the state of the art.

Acknowledgements

I would like to thank my supervisor, Prof. Keshav, and also Prof. Rosenberg for their support and feedback during the preparation of my thesis. I would also like to express my sincere appreciation to those who voluntarily participated in our measurement study. Without their participation this work would never have happened.

Dedication

This is dedicated to my parents for their constant support and encouragement during my studies.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Just as Internet Service Providers size links and routers in the Internet, electric utilities size line and transformer capacities in the electrical grid to be large enough to meet expected peak loads, but not so large as to be too expensive [11]. In both networks, operators use rules of thumb to roughly estimate resource sizing, upgrading capacity piecemeal as dictated by demand growth.

Two trends motivate us to re-examine current design rules for sizing the electrical grid. First, there is expected to be a worldwide surge in grid deployment in the next decade. Smart meters are being installed rapidly in houses. Many coal-fired or nuclear power plants are to be closed in the future. On the contrary, penetration of renewable sources of energy such wind and solar, is increasing constantly. Different types of energy storage including substations' storage systems, battery-operated electric vehicles, and plug-in hybrid electric vehicles will constitute important parts of the grid. Moreover, in the developed world, infrastructure put into place during the rapid postwar growth phase of the 1950's and 1960's is reaching the end of its operational life and must be replaced in the next 10 to 20 years. This is a good time, therefore, to re-examine sizing guidelines.

Second, with the incorporation of renewable energy sources and battery-operated electric vehicles, it is expected that the future grid would have non-trivial amounts of storage [17]. The classical grid has had little storage and provisioning storage is poorly understood. As a result, there is a need to understand how to provision storage in the future grid.

We believe that we can use the techniques from the teletraffic theory to size the electrical grid. We make five specific contributions:

- We conduct a measurement study of homes to determine the fine-grained structure of their electrical loads
- We propose parsimonious reference models of home loads; these models could be used in other studies pertaining to dynamic pricing and demand response
- We prove a formal equivalence between transformers and storage in the grid and routers and buffers in a network, allowing us to use teletraffic theory to analyze the grid
- We provide design rules for provisioning storage in the grid and study the insights gained from these rules
- We show that sizing decisions made using our design rules compares well with the 'ground truth' sizing obtained by directly measuring loads

1.1 Background

The electrical grid consists of three subsystems: generation, transmission, and distribution [19]. Electrical power generators use energy from sources such as coal, natural gas, or falling water to generate alternating currents. These currents flow into a transmission system that moves electric power to distribution networks. The transmission network, like the Internet core, has a mesh structure to meet reliability requirements of the grid. To minimize resistive losses, it operates at very high voltages of 150-500kV. Power from the transmission network is stepped down using transformers before entering the tree-like distribution network, which delivers power from distribution substations to end customers. This structure is analogous to the delivery of video content from content servers in centralized data centers over the Internet core and access networks to end-systems.

Step-down transformers are necessary for distribution networks to interface with the long-distance transmission system. A transformer's capacity or 'size' is the sustained power that it can deliver, measured in kilo Volt Amperes or kVA. Although this rating can be exceeded on rare occasions, grid design rules require that no transformer exceed its rating for more than short time intervals.

Transformers can be expensive. A small 'pole-top' 167kVA single-phase distribution transformer that serves about 10 homes in North America costs around $3,000 [2]. A typical small utility serving a customer base of 30,000 homes would therefore need to

spend $9,000,000 on pole-top distribution transformers alone. High-voltage transformers at substations, which serve thousands of customers, can cost up to $1,500,000.

Sizing a transformer is a critical design decision. A utility could potentially save millions of dollars by choosing smaller transformer sizes when replacing ageing equipment. On the other hand, underestimating the size of a transformer might lead to overloading that would shorten its life, and in the worst case, lead to transformer failure and power outage.

Several issues make transformer-sizing non-trivial. First, like MPEG-encoded video, electrical loads (i.e., the active power consumption, measured in Watts) are highly variable, making it unrepresentative to describe them by their mean values alone. On the other hand, sizing a transformer for peak load may be both overly conservative and expensive. Second, transformers are deployed for twenty to fifty years with only periodic maintenance. Because the load may change over this time, accurate load forecasting must be done. Third, the electrical grid has strict reliability criteria, which, if not met, can lead to dangerous overheating of transformers. To deal with these constraints, utilities use a conservative approach to size transformers. This approach typically results in oversized transformers and lightly used, expensive infrastructure. Our work is the first step in coming up with better sizing guidelines that can help utilities to optimize their infrastructural expenditure without reducing system reliability.

The transformer sizing problem is exacerbated by the imminent widespread availability of energy storage, particularly in the form of battery-electric vehicles. By storing energy during non-peak hours and releasing it to meet peak load, it will be possible to use a smaller transformer in the presence of energy storage. However, the relationship between storage and transformer sizing is currently an open problem. There is, therefore, an urgent need for design rules for distribution systems that incorporate storage.

In choosing transformer size, it is important to ensure system reliability. Reliability is measured by the *loss-of-load probability (LOLP)* [19], which is the probability that the system-wide generation resources fall short of demand. The "one-day-in-ten-years" criterion ($LOLP = 2.74 \times 10^{-4}$) is a benchmark value widely used among utilities in the United States.

This existing definition of reliability has one shortfall: it accounts for generation resources but not storage. The introduction of storage into the grid changes the classical picture of grid reliability because even if the transformer cannot meet the instantaneous aggregated demand, it is possible that the residual demand (i.e., the demand minus the transformer rating) can be met by storage. Therefore, in our work, we say that the reliability criterion is not met when demand cannot be met by a transformer even in the presence of storage because the store is currently empty, so that the demand results in a

store underflow. Then, given a transformer rating and a storage size, we use the probability of storage underflow for a particular set of demands as a measure of the loss-of-load probability.

1.2 Related Work

Transformer sizing in the electrical grid is usually studied in the context of overall distribution system planning. The standard approach to solve the problem is to use linear optimization [4, 11]. However, this approach necessarily models loads using only their peak values, ignoring temporal variations. These models also do not take storage into account.

Storage can be used both to even out variations in demand, as we study, as well as variations in supply, especially in the context of variable-rate generation by wind turbines and photovoltaic cells: see Divya and Ostergaard [10] and Deshmukh et al [9] for further details and a survey of current work in this area. To the best of our knowledge, most prior work on the effect of storage in the power grid has been on the supply side, and has not used concepts from teletraffic theory. For example, Lee and Gushee [17] compute the amount of storage needed across the entire United States to smooth out variations in wind energy generation. Similarly, Roy et al have studied the optimal sizing of batteries to even out variations in a single wind turbine [25]. Although they use the same equations to model the evolution of battery state as we do, their analysis is based on numerical simulations or the simplifying assumption that the generation process follows a Weibull distribution.

The use of storage to deal with variations on the demand side was proposed by Lachs as early as 1995 [16]. However, the lines of work closest to ours are by Ponnambalam et al and Kempton et al. Ponnambalam, et al use a novel moment-based method to study the battery storage process [22]. This approach is complex and relies on stochastic programming rather than teletraffic models. Kempton et al have studied the use of electrical vehicles for supplying energy to the grid, which they term "Vehicle-to-Grid" [15]. Their analysis focuses more on the details of vehicle usage and charging rates than the battery storage process.

1.3 System Model

We study sizing a transformer shared by a set of homes. These transformers, in the North American context, could be either pole-top transformers that are shared by 10-25 homes,

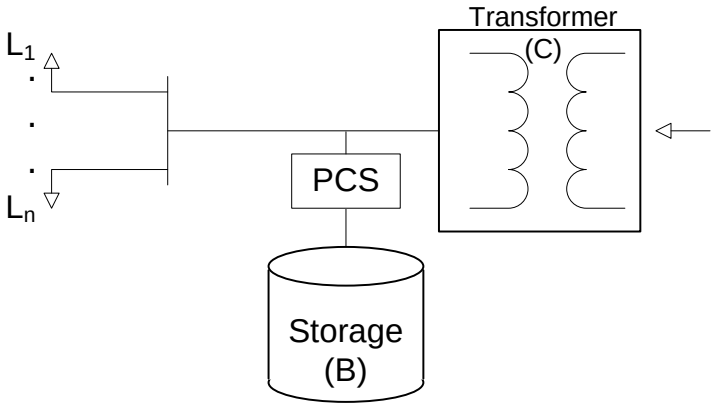


Figure 1.1: A branch of the electrical grid supplied by a transformer and a storage system.

or larger pad-mounted transformers that are shared by up to several thousand homes [19]. Going beyond current practice, we assume that the pole-top or substation may also contain storage to offset peak loads. Our goal is to jointly size the transformer and the storage to make sure that system reliability constraints are met[1].

Figure 1.1 depicts a single distribution branch of the electrical grid associated with a transformer with rating C Volt Amperes and a battery or store with capacity B Watt-hours. These are shared by a set of n homes, indexed by i. Each home places a load of $L_i(t)$ watts on the system at time t. We call the sum of the home loads at any time as the *aggregated* load at that time. We assume that each home's load can be categorized as belonging to one of the N load classes described in Chapter 2, with n_j homes in the j^{th} load class. We assume conservatively that the power generation capacity is larger than C so that it is not a bottlenecked resource. We also assume the presence of a power conversion system, marked 'PCS', that charges the store whenever the aggregate load is smaller than C and meets demand from the store whenever the aggregate load exceeds C.

[1]Note that we do not study systems with electric vehicles in the home, although our analysis can be extended to cover this scenario.

1.4 Outline

The rest of the thesis is laid out as follows. We describe our testbed, define the modeling requirements, and present reference models for home load in Chapter 2. In Chapter 3, we first demonstrate similarities between the electrical grid and a queueing system. We give a formal proof that there exists an equivalent queuing system for the gird. This suggests that the buffer occupancy is exactly similar but inverse of the storage occupancy. Then, we present some of the main results of Teletraffic Engineering obtained over the past 30 years; we introduce the notion of effective bandwidth and review the main results of the asymptotic analysis in buffered and bufferless systems. In Chapter 4, we use the approximations described in pervious chapter to size a buffer in the equivalent queuing system of the electrical grid such that the probability of overflow is less than a threshold. Finally, Chapter 5 sums up the thesis by highlighting the key contributions of our work.

Chapter 2

Measurement & Data Modeling

With the rapid ongoing deployment of the smart grid, there is an increasing need for accurate and parsimonious models for home electrical loads. Accurate home load models are the basis for research in smart appliance design, distribution network sizing, modeling of demand response, smart charging of electric vehicles, and home energy management. Most attempts at home load models in the past have been constrained by the limited measurement infrastructure. Smart meters measure home load every five minutes, at best, which is too slow to capture the on-off behavior of appliances. In path-breaking work, Richardson et al at the University of Loughborough measured 22 homes for a year at a one-minute granularity to construct a home-load generator [23]. However, their generator is not parsimonious, in that its input is the set of appliances in each home, the appliance load model, the number of occupants, and the occupancy behavior. It is difficult to obtain these inputs in any realistic situation.

We believe that it is important for a home electricity consumption model to be parsimonious, that is, have only a handful of parameters. While this is useful from the perspective of analysis and simulation, where a plethora of parameters makes it difficult to cover the model space, parsimony comes at a cost. This is because many factors determine home load such as the time of day, environmental and geographical factors, the type of appliances in the home, the usage pattern of these appliances, the number of occupants of the house, the occupancy pattern, and the size and perhaps even the floor plan of the house. One can easily argue that every one of these factors can significantly affect home electricity consumption. But it is impossible to gather this information for every house! Instead, because *every* model is likely to have a non-trivial error, we believe that it is reasonable in practice to construct a model with only a few parameters. Such a model is likely to be

just as accurate (or, perhaps, inaccurate) as a model with hundreds of parameters, but is far easier to construct and use.

A critical modeling decision is to use finite-state Markov chains. We are motivated by three reasons. First, Markov models are the foundation for many types of mathematical analysis, especially queueing theory and stochastic optimization. Therefore, we need to use Markov models to study the grid sizing problem using teletraffic analysis; we elaborate on this point in Chapter 3. Second, Markov models have been extensively used in the past to study univariate time-varying phenomena. In all these cases, it has been shown that Markov models combine both parsimony with descriptive power. Finally, home electricity usage 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 [24].

Over the years, electrical utilities have built classifications of home loads into a small number of representative classes. We use one of these classifications to try to derive a reference load model per class for different periods of the day. We propose an analytical model based on a k-state Markovian model. This kind of model will be very suitable for use in mathematical analysis. We use even finer-grained measurements than Richardson et al (every six seconds) of nine homes over two months to develop these reference models. We show that each class in each period can be modeled accurately by a Markovian model with less than 5 states. We provide the transition probability matrix, P, and the power consumption matrix, R, for each of these reference models. We believe that these models could be very useful to the research community interested in studying demand-response, transformer sizing, and distribution network simulation. We validate our results with the data that we have.

In the past, detailed models for residential loads have been presented in the power engineering, environmental studies, and civil engineering literature [23, 6, 7, 20]. However, these models suffer from three problems. First, these models tend to be highly parameterized, rather than parsimonious. Second, the data sets on which these models are based are not publicly available. Therefore, we cannot use them to create Markovian models. Finally, to the best of our knowledge, existing models group all homes into a single class. Our measurements show significant differences in demand behavior at different homes.

Therefore, in this chapter we describe our approach to:

- Collect fine-grain load measurements from real usage using inexpensive off-the-shelf components in nine homes. These nine homes can be partitioned into 3 classes using an existing classification.

- Create parsimonious Markovian reference models for these 3 classes for different periods of the day.
- Validate that our reference models generate a realistic trace of home electricity usage by comparing them with ground truth.

Note that we use the same modeling approach to create a model for any home load trace. This trace could be the home load measurement of a particular day or the *concatenated cumulative busy hour trace* described in Chapter 4[1].

2.1 Testbed

Our first step is to obtain real measurements of electrical load. To obtain our own load data set, we built a testbed to measure aggregate loads at nine homes. We deployed measurement nodes (Figure 2.1) at eight houses and one home-based small business covering a range of living area sizes, occupants, appliances, and energy consumption patterns. For the purpose of our study, we used a convenience sample rather than a stratified random sample. Our methodology generalizes to samples chosen using standard population sampling techniques [8].

Each measurement node consists of a Current Cost Envi device [1] and a netbook. The Envi device measures the active power consumption (in Watts) of a house every six seconds and stores it locally in flash memory[2]. A script on the netbook queries the device every six seconds to obtain an XML file that it stores on disk. This is uploaded using a secure SSL connection to a server in our lab once a day. To preserve privacy of the participants in our study, logs files are anonymized before being stored in a secure directory on the file server.

2.2 Classification

Given the high variability of home loads, choosing a classification for them is a challenging task. Fortunately, standard rules based on decades of field experience allow an electric utility to both predict and classify a home load based on a few simple parameters. We

[1]Modeling busy hour traces is fundamental to the teletraffic analysis.

[2]Consequently, the device does not capture load transients that last shorter than this time.

Figure 2.1: A measurement node

obtained such a parametrization, specifically used for transformer sizing, from a major utility in our area (Table 2.1). The key sizing parameters are the house size and the nature of the heating and cooling systems, which constitute the major loads in our geographical area. These are used to compute a 'unit value' that represents the load expected from that home.

To minimally impact participant privacy, we asked each participant to tell us their home's unit value computed using this table. We then placed homes with the same unit value in the same class. Table 2.2 shows the three classes so obtained (note that in our area homes are in general equipped with gas/oil heater).

2.3 Data Overview

Typical loads from three houses in different classes for one week (in winter) are shown in Figure 2.2. The main differences between classes are the amount of peak and base loads, the width of the peak period, and variability of the load (e.g., the average number and the average height of spikes in a given period).

Table 2.3 represents the load characteristics of each class. The base load of each class is the smallest element of the R matrix, described in Section 2.4. Similarly, the peak load is the largest element of the R matrix.

We expect to see that power consumption (either average, peak, or base power consumption) of a home is proportional to its unit size. However, the peak load of class 2 is

Type of Heating	House Size			
	$100m^2$	$200m^2$	$300m^2$	$400m^2$
Baseboard electric heat	3.0	4.0	5.0	6.0
Central electric heat	4.0	5.0	6.0	7.0
Gas/oil heat, no central A/C	1.0	1.5	2.0	2.5
Gas/oil heat, central A/C	1.5	2.5	3.5	4.5
For town or row houses, multiply the unit value by 0.8.				

Table 2.1: 'Unit values' assigned to customer homes by a major utility.

Class	Unit value	Number of houses
1	1.2	3
2	2.5	4
3	4.5	2

Table 2.2: Number of homes in our experiment within each class.

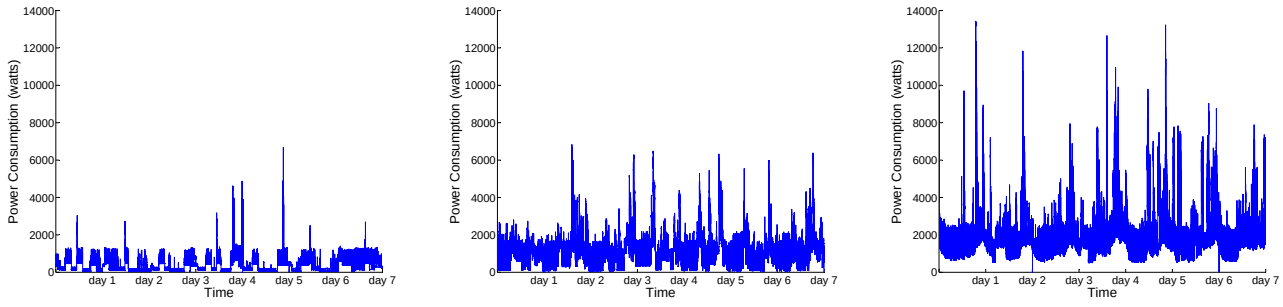


Figure 2.2: Load measurements from houses in three classes for one week.

lower than the peak load of class 1. This is definitely an outlier. We believe that it is due to the limited number of houses that we have in each class. In other words, if we had more homes in each class, the peak load (that is largest element of the R matrix) of that class would be closer to the peak load of a typical home in that class.

	Class 1	Class 2	Class 3
Peak load (W)	7051	6240	10294
Base load (W)	124	226	718

Table 2.3: Load characteristics of each class.

2.4 Markov Modeling

This section describes our approach to building a reference model for each class of homes introduced in Section 2.2. We answer the following questions:

- How many reference models do we need?
- How can we choose the Markov states?
- What metric determines the goodness of a model?
- How many Markov states are needed in each model?

In this work, we assume that the homes selected for measurement in our study are a representative random sample of their assigned class. This assumption is admittedly strong, but can be removed if homes chosen for measurement were chosen from a stratified random sample, which we defer to future work.

2.4.1 Definition

In this section, we introduce both discrete-time and continuous time Markov models. Since the nature of our data is discrete (we measure power consumption every 6 seconds), a finite state discrete-time Markov chain is merely used for the modeling purpose in this chapter. However, in teletraffic theory we usually deal with sources that generate traffic continuously. Therefore, to reuse the rich literature of teletraffic engineering, we assume

that the power consumption of a home is constant during the 6-second intervals. In effect, each home is modeled by a continuous-time Markov process[3] in Chapter 4.

Discrete-time Markov Model

A finite state discrete-time Markov chain is described by its transition probability matrix, P. If a Markov chain has k states then we have a $k \times k$ transition probability matrix. Because of the Markov property, the next state only depends on the current state, and P_{ij} is the the probability of going from state i to state j in the next time step. Accordingly, P_{ii} is the probability of staying in state i in the next time step.

Similar to the approach used in [26] for modeling VBR video sources, we assign a value, R_i, to state i of a Markov chain; this value represents the amount of power consumed in this state. We model the electric load of a home using a k-state Markov chain defined by the $< P, R >$ tuple.

Continuous-time Markov Process

A continuous-time Markov process is the continuous version of a Markov chain. In a Markov chain transitions happen at specific time steps; however, in a continuous-time Markov process the system remains in the previous state for a period of time before it transitions to a new state; these time periods are exponentially distributed. Similar to a Markov chain, dynamics of a k-states Markov process is represented by a $k \times k$ transition rate matrix (also known as the intensity matrix), Q, where q_{ij} is the rate of departing from state i to arrive at state j. Since the transition rates of each state should sum to zero, q_{ii} is defined as[4]:

$$q_{ii} = -\sum_{j \neq i} q_{ij}$$

Like before, we assign a value, R_i, to state i of a Markov process; this value represents the amount of power consumed in this state. We model the electric load of a home using a k-state continuous-time Markov process defined by the $< Q, R >$ tuple.

[3] The key assumption here is that the time that a source resides in each state of the Markov chain is exponentially distributed.

[4] An ergodic continuous-time Markov process has a stationary probability distribution, π, that can be easily computed from its transition rate matrix.

2.4.2 How many models do we need?

In Section 2.2, we classify the homes being measured into three classes based on the 'unit size' assigned to them. This motivates the need for at least three models. Furthermore, home load is highly sensitive to the time of day. More specifically, it is not stationary over the period of a day, and the probability that the peak power will be consumed at 6am is far lower than it is at 6pm. To deal with this issue, taking a cue from the time periods specified by the electric utility in our region, we divide a day into three periods, namely *on-peak* (7am-11am and 5pm-9pm), *mid-peak* (11am-5pm), and *off-peak* periods (12am-7am and 9pm-12am). Visual observation verifies that the home load is almost stationary in each period. We caution that the definition of these periods might be different in other geographic regions or seasons. Nevertheless, we believe that these three periods can be identified in every region and season. Therefore, we construct three reference models (one for each period) for each class, for a total of nine reference models.

2.4.3 Choosing Markov states

To find a k-state Markov model for each class in each period, we first need to choose a representative load from our measurements. This representative load must include electric loads of all homes within the corresponding class in that period; therefore, we construct it by concatenating these electric loads. For example, the representative on-peak load of a class is the concatenation of on-peak periods of all homes selected for measurements within that class for the entire measurement period. We use the k-means clustering algorithm to find k centroids of the representative load. Values of these centroids constitute the R matrix. To increase the chance of finding the global optimum in the k-means clustering, we run 500 replicates of the k-means algorithm with random start points. Substituting values of the points in each cluster with the value of its centroid, we obtain a clustered home load. Then we use the clustered home load to compute the P matrix from the following

	On-peak	Mid-peak	Off-peak
Class 1	5	4	5
Class 2	5	4	2
Class 3	4	4	4

Table 2.4: Number of states that is enough for representing the home load of a class in a period.

expression [5]:

$$p_{ij} = \frac{no.\ of\ transitions\ from\ R(i)\ to\ R(j)\ in\ clustered\ load}{total\ no.\ of\ times\ we\ have\ R(i)\ in\ clustered\ load}$$

2.4.4 How many Markov states should a model have?

An obvious question is that how many states are needed in each model. To answer this question we need to define a goodness-of-fit metric. We use the area between the cumulative distribution function (CDF) of the measured home load and the CDF obtained from modeling the load of the class it belongs to as this metric. This represents how far two probability distributions are from each other: the greater the metric the higher the modeling error. Therefore, to find a sufficient number of states to model a class, we sum the areas for all homes within that class and study this as a function of the number of states in the model (Figures 2.3, 2.4, and 2.5). The knee or corner point of this curve is the minimal number of states of the corresponding Markov model. Table 2.4 summarizes how many states should be used for each class and for each period.

Alternatively, to create Markovian models of concatenated cumulative busy hour traces for the teletraffic-based sizing of the electrical grid (discussed in Chapter 4), we use the standard Q-Q goodness-of-fit test to determine the number of states that are sufficient for creating a model (Figure 2.6). The only reason that we advocate the use of the Q-Q plot

[5] As mentioned earlier, in Chapter 4, we need to compute the transition rate matrix of a continuous time Markov Process instead of the transition probability matrix of a Markov chain. We know that a non-diagonal element of the transition rate matrix, q_{ij}, is the inverse of the mean of the underlying exponential distribution of waiting times in state i before a transition to state j. Thus, we can compute the Q matrix using the following expression:

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

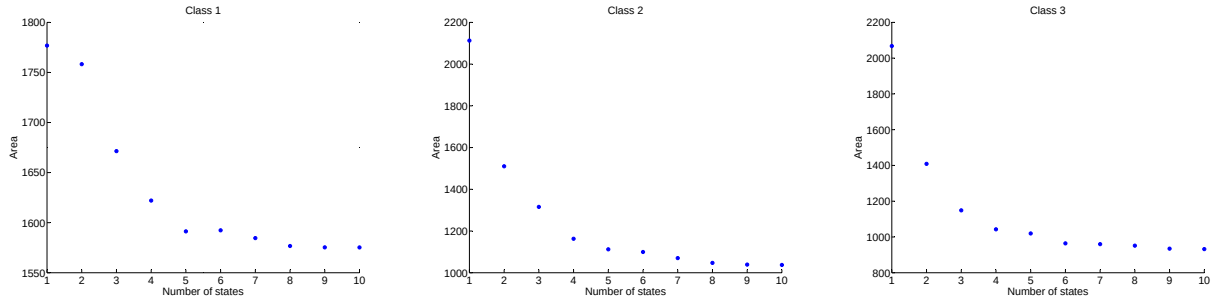


Figure 2.3: The goodness of fit metric versus the number of states for the on-peak period of three classes. Note that the Y axis is exaggerated for emphasis.

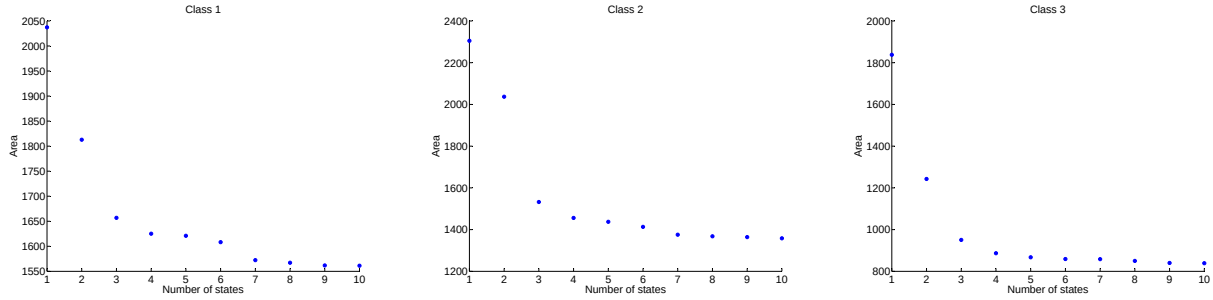


Figure 2.4: The goodness of fit metric versus the number of states for the mid-peak period of three classes. Note that the Y axis is exaggerated for emphasis.

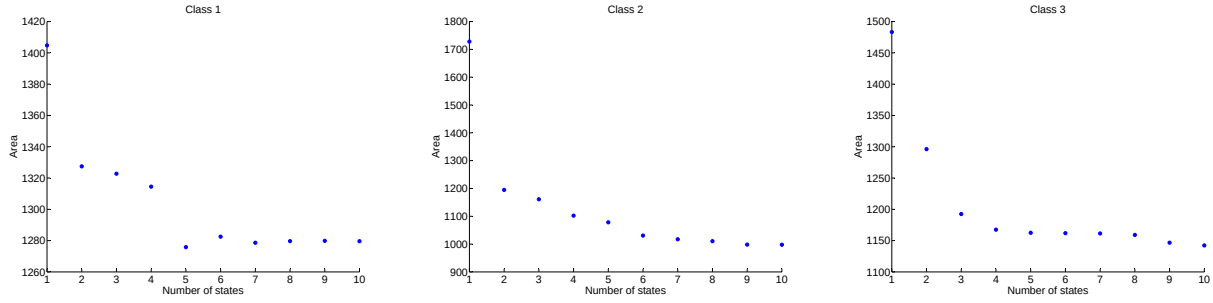


Figure 2.5: The goodness of fit metric versus the number of states for the off-peak period of three classes. Note that the Y axis is exaggerated for emphasis.

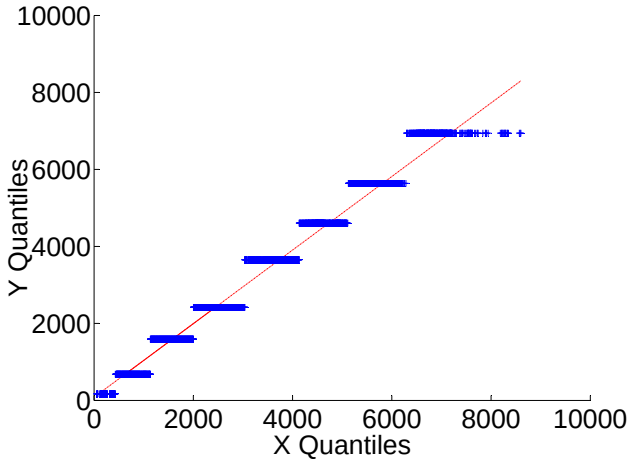


Figure 2.6: Q-Q goodness-of-fit test shows whether the 8-state demand model is sufficient.

in this case is that capturing the highest and lowest states is more important in teletraffic-based sizing and the Q-Q plot allows us to simply verify whether these states are captured by visual inspection. However, the area between CDFs does not show whether the highest and lowest states are captured.

When the highest and lowest states of the load are captured, the skewness of the distribution of the clustered load would be less than the other cases. Comparing the Q-Q plots for different values of k, we can see how the skewness of the distribution changes as a result of increasing the number of states which enables us to determine the sufficient number of states (we elaborate on that in Section 4.1.3).

Reference Models

To give an example, a 5-state Markov model is a good representative of the on-peak load of the second class. It is defined by the following P and R matrices[6]:

$$P = \begin{bmatrix} 0.90264 & 0.00192 & 0.01080 & 0.00834 & 0.07630 \\ 0.00103 & 0.97198 & 0.00006 & 0.02210 & 0.00481 \\ 0.06325 & 0.00110 & 0.91737 & 0.00183 & 0.01645 \\ 0.00336 & 0.01929 & 0.00028 & 0.94352 & 0.03355 \\ 0.02448 & 0.00166 & 0.00108 & 0.03038 & 0.94239 \end{bmatrix}$$

[6]The matrices for the other 8 models are available on `http://blizzard.cs.uwaterloo.ca/~oardakan/models`.

$$R = \begin{bmatrix} 2252 & 500 & 4355 & 1077 & 1614 \end{bmatrix}$$

To sum up, we have shown how to derive a Markovian reference model per class and per time period. We have also provided the P and R matrices for each reference model. We now study how well our models compare to the ground truth obtained from our measurements.

2.5 Validation

In this section our goal is to validate that the nine reference models (3 reference models per class, one for each time period) described previously would generate load traces that are representative of the measured load of the corresponding class and time period. To this end, we divide the measured home loads into two disjoint data sets, training and test. A subset of the training data set that is a good representative of home loads is used to construct the reference models, whereas the test data set contains the home loads that are used in the validation process. Our test data set consists of home loads measured from nine houses over a week. Thus, we validate that traces that are generated from a reference model are representative of the corresponding home loads in the test data set.

Figure 2.7 shows on-peak load of a home in class 1 obtained from the test data set. A quick visual inspection shows that it is similar to the on-peak home load generated from the on-peak model of class 1 (Figure 2.8). This is also true for all the 27 home loads (9 homes and 3 periods).

Moreover, we generate two sample traces (the length of each is the duration of the corresponding period) from each reference model and compute their means and standard deviations. If these traces are similar to the actual home load obtained from the test data set, the first two moments of them would match. Comparing these numbers with the mean and standard deviation of home loads obtained from the test data set, it can be concluded that there is not any significant difference between these numbers (Figure 2.9).

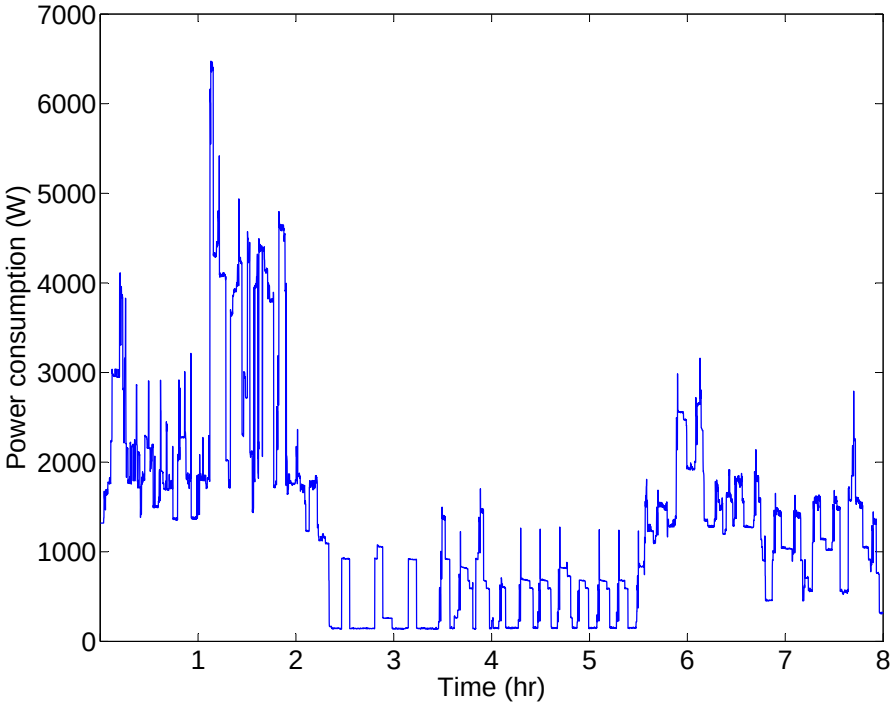


Figure 2.7: On-peak load measured from a house in class 1

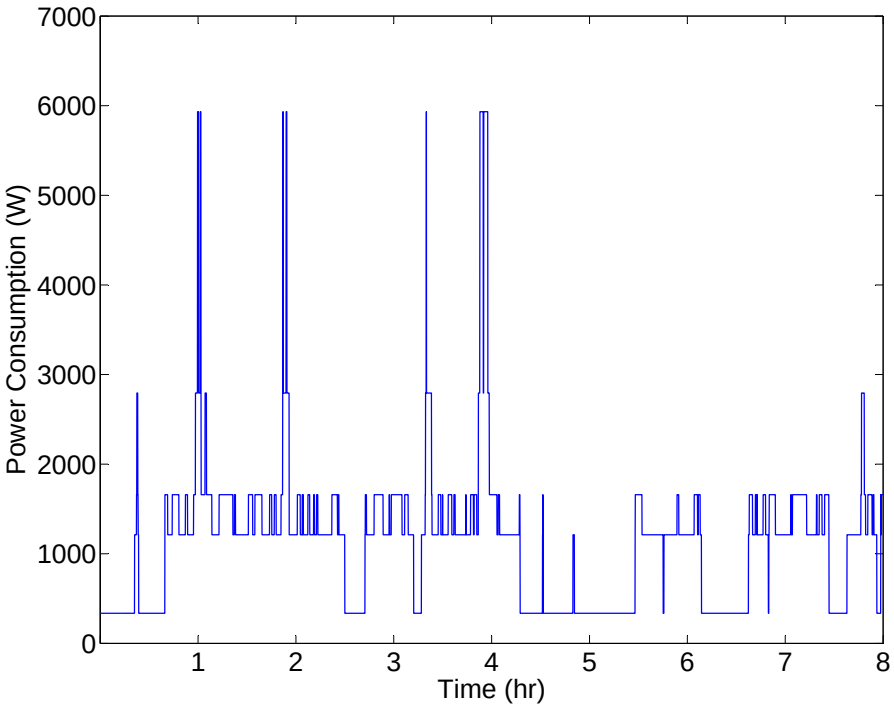


Figure 2.8: On-peak load trace generated from the on-peak model of class 1

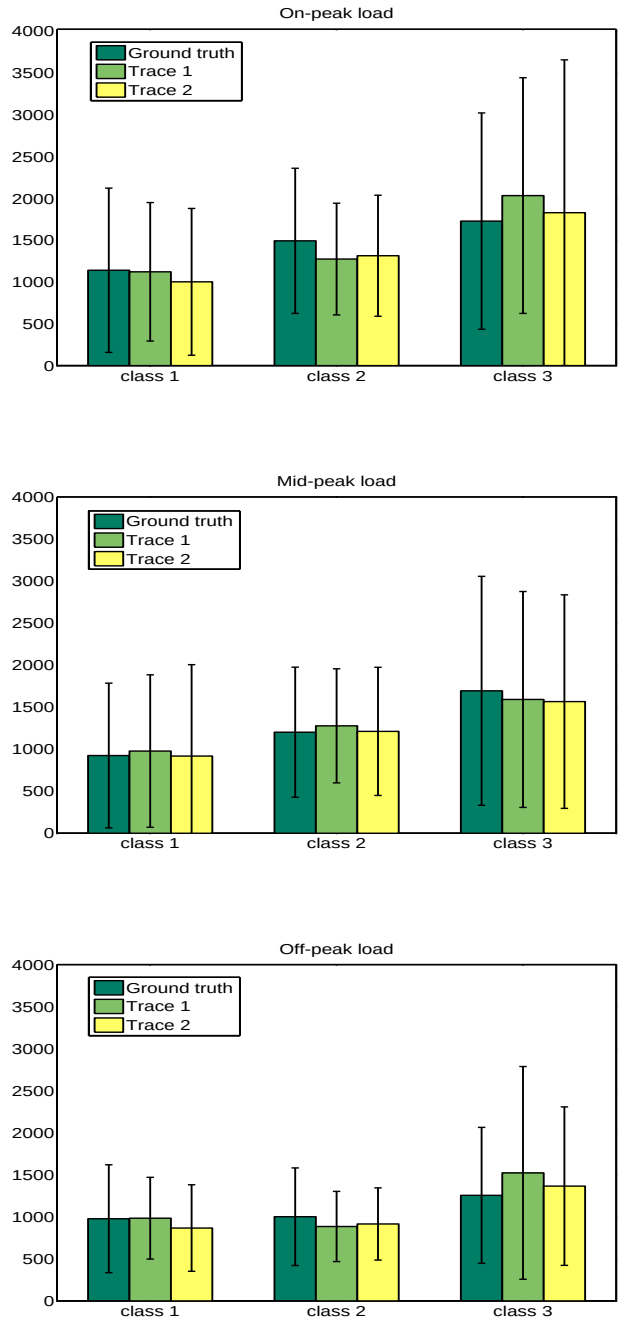


Figure 2.9: Comparison between the mean and standard deviation of the actual load and the synthetic traces. Error bars represent the sample standard deviation.

Chapter 3

Teletraffic Theory

Asynchronous Transfer Mode (ATM) emerged in the late 1980s with the goal of transferring different types of traffic using a shared medium. The main assumption was that network resources such as link bandwidth and buffer capacity of switches can be shared among statistically independent sources generating various types of traffic; this is usually known as statistical multiplexing [18]. The statistical variation of sources makes it possible to allocate fewer resources than sum of the maximum traffic generated by sources, achieving a considerable multiplexing gain. However, there is always a non-zero probability of maximum traffic outputs of all sources occurring simultaneously; thus, if the system can tolerate some packet loss, then we can design the system in a way that is much more cost-efficient and in which the resources are better utilized. To quantify the multiplexing gain, the notion of *effective bandwidth* is introduced in Section 3.2. Effective bandwidth represents how much bandwidth should be assigned to a source at a given time and for a specific system's population (i.e., the number of sources that share the resources).

In ATM, the network design problem is related to the *Quality of Service* (QoS). Usually, QoS is expressed in terms of average throughput, delay, and packet loss for each type of traffic. For example, the packet (cell) loss probability in ATM is about 10^{-6} which shows that the loss event happens rarely. These requirements must be taken into account when designing a network. Therefore, assuring that a resource allocation scheme satisfies these rigid requirements is quite challenging. We need to use some techniques to approximate small magnitude parameters such as packet loss. These approximations will be discussed at length in Section 3.3.

We start this chapter by presenting our key contribution, which is showing an equivalence between a branch of the electrical grid and a queueing system in Section 3.1. It

readily follows from this equivalence that the problem of jointly sizing a transformer and a storage system such that the storage underflow probability is limited (i.e., the reliability criterion is met) is the same as the problem of sizing a server along with a buffer in a queuing system such that the buffer overflow probability or the packet loss probability is limited (i.e., a QoS requirement is satisfied). This equivalence is essential as it allows us to use the same approximations derived for the buffer overflow probability to compute the storage underflow probability and to size the electrical grid accordingly.

3.1 Equivalence

To achieve our top level goal of determining sizing rules for storage in the electrical grid, we begin by constructing a queueing system to model a distribution branch in the grid. We proceed in two steps. First, we develop an intuitively appealing equivalence between a branch of the distribution grid and a simple computer network in Section 3.1.1. Then, in Section 3.1.2, we formalize our intuition by showing that an electrical grid with storage can be modeled as a non-traditional $D/G/1/B$ queueing system that can, nevertheless, be analyzed as a standard $G/D/1/B$ queue.

3.1.1 State Evolution Equivalence

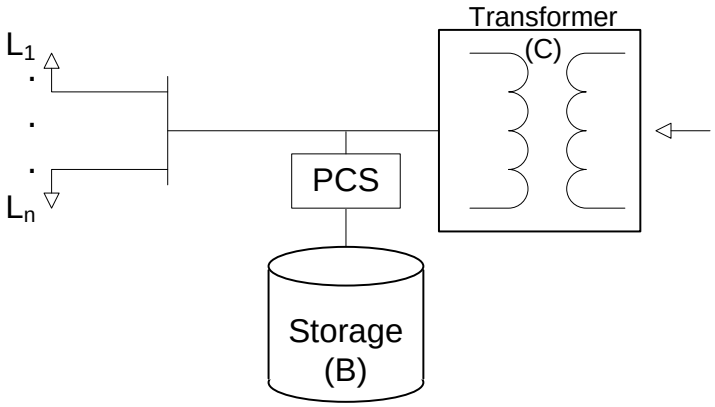
We compare the branch of the electrical grid shown in Figure 3.1a with a shared buffer of size B bytes accessed by a communication channel of capacity C bits/second shared by a set of sources, indexed by i, and with a transmission rate of $L_i(t)$ bits/second in Figure 3.1b.

We first consider the evolution of the shared store in the grid. If the sum of demands is less than C, then the store charges at the rate $C - \sum_i L_i(t)$, unless it is full. Denoting the amount of energy in the store (i.e., its *workload*) at time t by $W(t)$, we write this as:

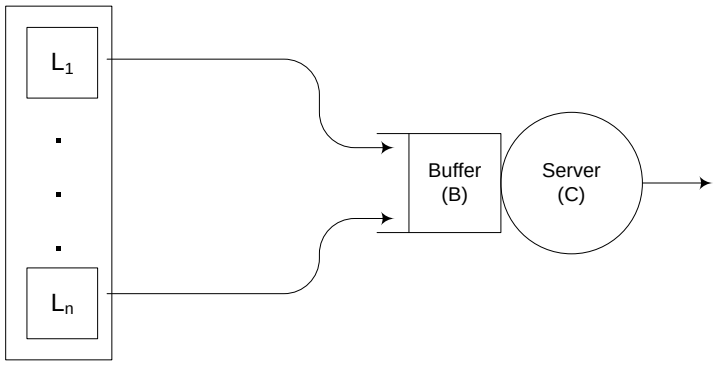
$$dW(t)/dt = \begin{cases} C - \sum_i L_i(t) & \text{if } (W(t) < B), \\ 0 & \text{otherwise} \end{cases}$$

Symmetrically, if the sum of loads exceeds C, then the store can be used to supply energy to the homes, unless the store is empty, in which case the voltage supply received by the homes will drop, which can be viewed as a failure of reliability. We write this as:

$$dW(t)/dt = \begin{cases} C - \sum_i L_i(t) & \text{if } (W(t) > 0), \\ 0 & \text{otherwise} \end{cases}$$



(a) A branch of the electrical grid with n loads L_i where the capacity of the battery is B Watt-hours and the base rating of the transformer is C Volt Amperes. The Power Conversion System (PCS) drains and fills the store depending on load conditions.



(b) A $G/D/1/B$ fluid queue with n sources L_i. The capacity of the buffer is B bytes and the service rate of the server is C bytes/second.

Figure 3.1: The storage system and a small queuing network.

Combining the two, we write

$$dW(t)/dt = \begin{cases} C - \sum_i L_i(t) & \text{if } (0 < W(t) < B), \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

Now, consider the rate at which the network buffer changes over time. Denote the amount of information in the buffer at time t by $\overline{W}(t)$. If the sum of the arrival rates from the sources exceeds C, then the excess arrivals are stored in the buffer if space permits. Therefore, we can write:

$$d\overline{W}(t)/dt = \begin{cases} \sum_i L_i(t) - C & \text{if } (\overline{W}(t) < B), \\ 0 & \text{otherwise} \end{cases}$$

On the other hand, if the sum of arrivals is less than C, then the buffer drains out at the rate $C - \sum_i L_i(t)$ unless it is empty, in which case its drain rate is 0. We write this as

$$d\overline{W}(t)/dt = \begin{cases} \sum_i L_i(t) - C & \text{if } (\overline{W}(t) > 0), \\ 0 & \text{otherwise} \end{cases}$$

Combining the two, we can write:

$$d\overline{W}(t)/dt = \begin{cases} \sum_i L_i(t) - C & \text{if } (0 < \overline{W}(t) < B), \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

Comparing equations 3.2 and 3.1, we see that they are symmetrical. This suggests that it should be possible to model the two queueing systems analogously. We formalize this intuition next.

3.1.2 Equivalent Queueing Models

Observe that the queueing model corresponding to our electrical storage system is a $D/G/1/B$ fluid queue. This is because electrical power generated at a constant rate C is precisely a fluid arrival bringing work to the system at the deterministic rate C. Moreover, the load from home i can be viewed as a fluid service rate, so that the service rate corresponding to the aggregate load $\sum_{i=1}^{n} L_i(t)$ that drains the buffer can be modeled as a G (general) service-rate process[1] The critical aspect of this queueing system that we want

[1]Note that, if $\lambda = E(\sum_i L_i)$ is the average service rate then typically, for this queueing system, $C > \lambda$, i.e., we have a finite queue with a utilization factor $\rho > 1$.

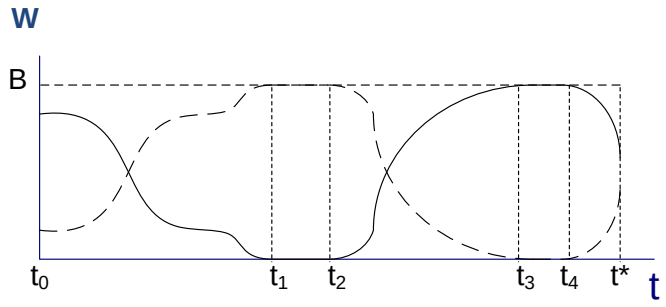


Figure 3.2: Workload of the storage system. The dashed line represents $\overline{W}(t)$ while the solid line represents $W(t)$.

to quantify is its *underflow* probability, i.e., the probability that a service finds the store empty. Unfortunately, standard queueing models do not deal with this question.

However, teletraffic analysis can be used to analyze the overflow probability of the standard $G/D/1/B$ queueing system [14]. Based on the intuition from the previous section, our plan of attack is to show that we can model a $D/G/1/B$ system with an equivalent $G/D/1/B$ system, permitting the use of teletraffic analysis.

Let the *workload trajectory* of a queue denote a specific instance of the function $W(t)$, i.e., the store size at time t. Let $W(\infty)$ denote the stationary workload process [28]. Our main theoretical result is the Equivalence Theorem:

Equivalence Theorem *Every workload trajectory in the $D/G/1/B$ queuing system corresponds to an equivalent trajectory in the $G/D/1/B$ queuing system such that $\forall t, W(t)+\overline{W}(t) = B$.*

Lemma 1 *If at any time t, $W(t) + \overline{W}(t) = B$ then $W(t^*) + \overline{W}(t^*) = B \ \forall \ t^* > t$.*

Proof: Let $t_{2i-1}, \ i = 1, 2, \cdots$, be the i^{th} time in the interval $[t \ \ t^*]$ that the storage system becomes either full or empty and persistently in this state until t_{2i} (Figure 3.2). Similarly, we define $\bar{t}_{2i-1}$ to be the i^{th} time that the buffer in the model becomes either full or empty and persistently in this state until $\bar{t}_{2i}$.

We prove the lemma in two parts. First, we use induction to prove that $t_i = \bar{t}_i$ for all values of i and that if $W(t_{i-1}) + \overline{W}(t_{i-1}) = B$ then $W(t_i) + \overline{W}(t_i) = B$. In the second part, we show that these results hold in the last interval prior to t^*.
Part 1

Base case: Without loss of generality, we assume that $t_1 \leq \bar{t}_1$. Since $W(t)$ and $\overline{W}(t)$ are differentiable on the interval $[t \;\; t_1]$, the Fundamental Theorem of Calculus allows us to write:

$$W(t_1) - W(t) = \int_t^{t_1} dW(s)$$

$$\overline{W}(t_1) - \overline{W}(t) = \int_t^{t_1} d\overline{W}(s)$$

From Section 3.1.1, in this interval we have $dW(t)/dt = -d\overline{W}(t)/dt$. Thus, it can be readily seen that $W(t_1) - W(t) = -(\overline{W}(t_1) - \overline{W}(t))$. Since we have $W(t) + \overline{W}(t) = B$, then $W(t_1) = 0$ (*or* B), implies that $\overline{W}(t_1) = B$ (*or* 0), and it implies that $t_1 = \bar{t}_1$.

Inductive step: Given $W(t_k) + \overline{W}(t_k) = B$, we prove that $t_{k+1} = \bar{t}_{k+1}$, and $W(t_{k+1}) + \overline{W}(t_{k+1}) = B$. Again w.l.g., assume that $t_{k+1} \leq \bar{t}_{k+1}$. Since $W(t)$ and $\overline{W}(t)$ are differentiable on the interval $[t_k \;\; t_{k+1}]$, the Fundamental Theorem of Calculus allows us to write:

$$W(t_{k+1}) - W(t_k) = \int_{t_k}^{t_{k+1}} dW(s)$$

$$\overline{W}(t_{k+1}) - \overline{W}(t_k) = \int_{t_k}^{t_{k+1}} d\overline{W}(s)$$

Since $dW(t)/dt = -d\overline{W}(t)/dt$ in this interval, we have $W(t_{k+1}) - W(t_k) = -(\overline{W}(t_{k+1}) - \overline{W}(t_k))$. Therefore, we conclude that $W(t_{k+1}) + \overline{W}(t_{k+1}) = B$. Clearly, $W(t_{k+1}) = 0$ *or* B implies that $\overline{W}(t_{k+1}) = B$ *or* 0, which in turn, results in $t_{k+1} = \bar{t}_{k+1}$.

Part 2

Now assuming that $t_n \leq t^* \leq t_{n+1}$, the last part of the proof is to show that $W(t^*) + \overline{W}(t^*) = B$ given that $W(t_n) + \overline{W}(t_n) = B$. Again, $W(t)$ and $\overline{W}(t)$ are differentiable on the interval $[t_n \;\; t^*]$, and we can write:

$$W(t^*) - W(t) = \int_{t_n}^{t^*} dW(s)$$

$$\overline{W}(t^*) - \overline{W}(t) = \int_{t_n}^{t^*} d\overline{W}(s)$$

Using the fact that $dW(t)/dt = -d\overline{W}(t)/dt$, we can write $W(t^*) - W(t_n) = -(\overline{W}(t^*) - \overline{W}(t_n)$. By induction, we previously showed $W(t_n) + \overline{W}(t_n) = B$; thus, we conclude that $W(t^*) + \overline{W}(t^*) = B$ and the proof is complete. □

Proof of the Equivalence Theorem
Let the initial workload state in the storage system, that is, the $D/G/1/B$ system be $W(t_0)$. Then, in the $G/D/1/B/$ model, we set $\overline{W}(t_0) = B - W(t_0)$. It follows from Lemma 1 that there is a one-to-one mapping from trajectories of the $D/G/1/B$ queuing system to trajectories of the $G/D/1/B$ queuing system and that $\forall t, W(t) + \overline{W}(t) = B$. □

Corollary 1 *It follows from the above theorem that:*

$$\mathbb{P}(W(\infty) > B) = \mathbb{P}(\overline{W}(\infty) < 0)$$
$$\mathbb{P}(W(\infty) < 0) = \mathbb{P}(\overline{W}(\infty) > B)$$

Where $W(\infty)$ is the stationary workload process.

One consequence of the Equivalence Theorem is that the probability of storage underflow in the storage system is precisely the probability of buffer overflow in the network system (Corollary 1). The latter probability has been thoroughly investigated in teletraffic theory, to which we turn to next.

3.2 Effective Bandwidth

Effective bandwidth is a powerful mathematical concept that is a deterministic quantity associated to a source [14]. It adequately represents the statistical characteristics of a source and provides a sense of how much resources should be reserved for this source to make sure that specific QoS requirements are satisfied. The effective bandwidth of a source depends on two free variables, the space and time scaling. Values of these variables are assigned based on characteristics of the source. It is defined as:

$$\alpha(s,t) = \frac{1}{st} \log E[e^{sX[0,t]}]$$

where $X[0,t]$ is the amount of work that arrives from the source in the interval $[0,t]$.

We are also interested in understanding how a mix of sources can impact the sizing problem. To this end, we have to define the effective bandwidth of all these sources, $\alpha(s)$. Assume that sources belong to N classes where all sources in a class are i.i.d and sources from different classes are mutually independent. Moreover, assume that n_i sources are in class i and the effective bandwidth of each source is denoted by $\alpha_i(s)$. Thus, we have:

$$\alpha(s) = \sum_{i=1}^{N} n_i \alpha_i(s)$$

This is called the additive property of effective bandwidth.

If X_i is the instantaneous fluid generation rate corresponding to the choice of $\alpha_i(s) = \lim_{t\to 0} \alpha_i(s/t, t)$, we can easily derive that:

$$\alpha_i(s) = \frac{1}{s} \log E[e^{sX_i}]$$

We use the above definition of effective bandwidth in the next section.

Using the definition of effective bandwidth, we can derive a formula for the effective bandwidth of a Markovian source. In [13], the authors prove that the effective bandwidth $(\alpha(s))$ of a Markovian source represented by the $< Q, R >$ tuple is the maximal real eigenvalue of the matrix $R_d - \frac{1}{s}Q$, where $R_d = diag(R)$.

3.3 Asymptotic Analysis of the Loss Probability

This section briefly states standard results from teletraffic theory to compute approximations for the overflow probability in a $G/D/1/B$ system in both bufferless and buffered systems under the assumption that the arrivals are Markovian. We validate our use of teletraffic theory in Chapter 4.

We make the technical assumptions that each individual load $L_i(t)$ is stationary and Markovian. Let Y_i be this stationary distribution. Let Y be the stationary distribution of the aggregate load. Without storage, C has to be dimensioned so as to allow for large variations in the aggregate load (i.e., peaks). By introducing finite storage, we will be able to dimension C less conservatively. If $B = \infty$, then there is no overflow and the system is stable as long as $\lambda < C$. Typically, our requirement is that the overflow probability in the original system is less than a desired small value ϵ, which corresponds to LOLP target, typically 2.7×10^{-4}.

3.3.1 Bufferless Model

We can write our requirement as:

$$\log P(Y \geq C) \leq -\beta = \log \epsilon \tag{3.3}$$

Following Kelly [14], we use Chernoff's bound to obtain:

$$\begin{aligned}\log P(Y \geq C) &\leq \log E[e^{sY}] - sC \\ &\leq \inf_s\{\log E[e^{sY}] - sC\} \\ &\leq \inf_s\{s(\alpha(s) - C)\}\end{aligned}$$

Where $\log E[e^{sY}]$ is the logarithm of the moment-generating function of Y and $\alpha(s)$ is the effective bandwidth of a source with the stationary fluid generation rate Y.

An improved approximation for the loss probability can be derived using the approach of El Walid *et al* [12]:

$$P(Y \geq C) \sim \frac{e^{s^*(\alpha(s^*)-C)}}{s^*(2\pi\sigma^2(s^*))^{\frac{1}{2}}} \quad as \quad C \to \infty \tag{3.4}$$

where s^* is a point where $s(\alpha(s) - C)$ attains its infimum, and $\sigma^2(s)$ is defined as follows:

$$\sigma^2(s) = \frac{\partial^2}{\partial s^2}(s\alpha(s))$$

Therefore, if we approximate $\log P(Y \geq C)$ by $\inf_s\{s(\alpha(s) - C)\}$, the capacity region; i.e., the values of C that satisfy (3.3), will be:

$$Capacity \quad region = \{C | \inf_s\{s(\sum_{i=1}^{N} n_i\alpha_i(s) - C)\} \leq -\beta\} \tag{3.5}$$

Otherwise, using the improved approximation, the capacity region will be:

$$Capacity \quad region = \{C | \log(\frac{e^{s^*(\alpha(s^*)-C)}}{s^*(2\pi\sigma^2(s^*))^{\frac{1}{2}}}) \leq -\beta\} \tag{3.6}$$

Thus, given the aggregate fluid generation rate Y, C can be computed so that the overflow probability is less than ϵ. We note that this is an asymptotic formula, i.e., the formula is valid under the assumption that the total number of sources is large and we are interested in the tail of the distribution.

3.3.2 Buffered Model

Here, our first goal is to compute the overflow probability in a system given C and B. For this we compute

$$\log p(W(\infty) \geq B) \tag{3.7}$$

where $W(\infty)$ is the stationary distribution of the workload. Whitt [28] states several different asymptotic forms for the steady state distribution of the workload of stable queues. Of these, we focus on the exponential tail approximation of the workload for large buffers originally studied by ElWalid *et al* [12], *i.e.*,

$$p(W(\infty) \geq B) \sim e^{-nc_1} e^{-c_2 B} \quad as \quad B \to \infty \tag{3.8}$$

where e^{-nc_1} is the loss probability in the bufferless case (see Equation 3.4). To find the loss probability of the buffered case (i.e., the overflow probability), we only need to compute c_2. It can be shown that $-c_2$ is the dominant eigenvalue of a buffered multiplexing system which determines the tail behaviour of the workload, and for Markovian sources, we can compute it by finding the solution of the following problem [3] [12]:

$$f(z) = \sum_{j=1}^{N} n_j MRE(R_{j_d} - \frac{1}{z} Q_j) - C = 0 \tag{3.9}$$

Where MRE gives the maximal real eigenvalue of a matrix, n_j is the number of sources in class j, and R_{j_d} and Q_j are the diagonal traffic generation rate matrix and intensity matrix of sources in class j respectively. As we defined earlier, $MRE(R_{j_d} - \frac{1}{z} Q_j)$ is the effective bandwidth of a single source (that is represented by the $< Q_j, R_j >$ tuple).

Now, we are interested in the capacity region so that

$$\log p(W(\infty) \geq B) \leq -\beta = log\epsilon \tag{3.10}$$

we have:

$$Capacity \quad region = \{C | \log(\frac{e^{s^*(\alpha(s^*)-C)}}{s^*(2\pi\sigma^2(s^*))^{\frac{1}{2}}}) + zB \leq -\beta\} \tag{3.11}$$

This is an asymptotic formula, i.e., the formula is valid under the assumption that the number of sources is large, the buffer is large, and we are interested in the tail of the distribution..

To sum up, teletraffic analysis allows us to associate an overflow probability (or LOLP) with a particular choice of B and C as $n \to \infty$ and for a given Markovian aggregate

workload $\sum_i L_i(t)$. We view these as our 'design rules' that allow us to size transformer and storage capacities, i.e., a (C, B) tuple, to meet the demands of a given workload with a certain reliability constraint.

Chapter 4

Sizing the Electrical Grid

In this chapter, we pursue two goals. First, we state our approach to validate the use of teletraffic theory to size the electrical grid. Second, we show how to size the grid using the approximations that are mentioned in Chapter 3. We conclude that our sizing approach is in excellent agreement with the heuristics used by utilities. The measurements obtained from the testbed and the home classification introduced in Chapter 2 are used to validate our approach and to size the grid.

4.1 Validation of Our Approach

Our modeling and analysis of storage systems in the electrical grid allows us to use teletraffic theory to determine transformer and storage sizing rules. Teletraffic theory makes a number of strong assumptions about the nature of the workload. Are the results of teletraffic analysis really applicable to the electrical grid? This section describes our approach to answering this critical question.

Our overall approach is to use real measurements of electrical load to empirically determine the storage and transformer sizes needed to serve them. We then compare the sizes so determined with those determined from teletraffic analysis. We show that the results obtained in these two ways are comparable. An overview of our approach is shown in Figure 4.1. We explain the details of this approach in the remainder of this section.

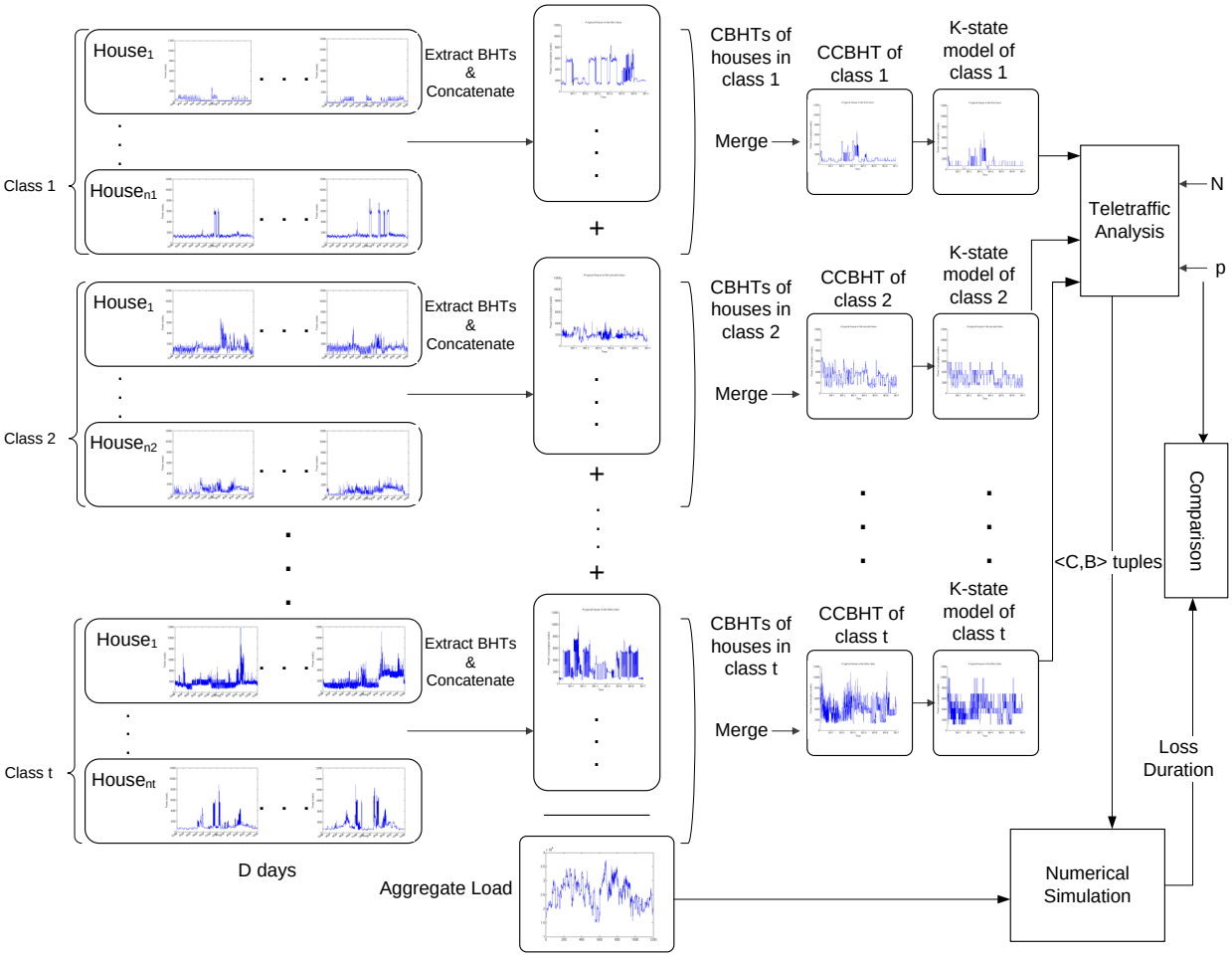


Figure 4.1: An overview of our validation methodology.

4.1.1 Assumptions for Empirical Sizing

We now turn our attention to using our load measurements to sizing transformers and storage in the grid. Suppose we had fine-grained load measurements from all the homes in one neighborhood for a period of several years. Then, we could simply add these to create the true aggregate load. Given the aggregate load, a trivial numerical simulation suffices to determine the aggregate duration of service disruption corresponding to a particular transformer sizing and a particular storage size. This simulation uses the discretized version of Equation 3.2 to update the state of the store given a particular demand and transformer size, recording the durations of underflow.

However, it is impractical to measure all the loads from a neighborhood for several years before making a sizing decision. Moreover, even if such a trace were to be obtained, it would be difficult to determine the degree to which the trace would be representative of other neighborhoods or of the same neighborhood two decades hence. Therefore, we have to make the following assumptions even when doing an empirical sizing of transformers and storage:

1. Household energy demands can be categorized into a few distinct classes corresponding to sampling strata, where demands within a class are homogeneous and the classes are mutually exclusive.

2. The homes selected for measurement in our study are a representative random sample of their assigned class.

3. The proportion of homes selected for measurement is representative of the true proportion of homes in each class.

4. The electrical demand during the *busy hour* (defined below) at each home is a conservative upper bound on its demand.

The first three assumptions are rather strong, but can be removed if homes chosen for measurement were chosen from a stratified random sample, which we would advocate in a real-world application of our design rules. In this case, the aggregate load in Figure 4.1 would be a reasonably good representative of the true aggregate load.

We also note that using the busy hour to size the system is the standard approach used in telecommunication systems. This is the one-hour period during which a home uses the most energy (it may or may not include the daily peak power point). It is generally accepted that a sizing that is based on the busy hour alone is more conservative than that

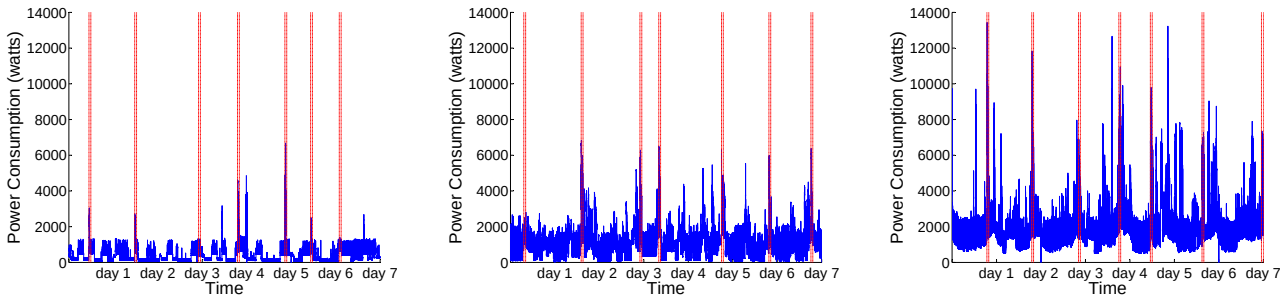


Figure 4.2: Load measurements from houses in three classes for one week with busy hours marked by vertical lines.

using the entire day and therefore provides a sufficient cushion against measurement bias and lack of complete measurement data.

4.1.2 Empirical Sizing

Given the assumptions in Section 4.1.1, we now consider the problem of empirically determining the size of a transformer and storage pair for a residential neighborhood.

We use the following methodology. First, based on Assumption 4, we find the busy hour for each home for each day. This is the one-hour period with the maximum area under the power consumption profile (see Figure 4.2). Usually, the busy hour happens during the peak hours, i.e., 7am-11am and 5pm-9pm during the winter[1]. We call the load during the busy hour for a home as its 'Busy Hour Trace' or BHT. Second, we concatenate the BHTs of each house for a specific number of days to obtain the cumulative busy hour trace (CBHT) for that period. Third, we sum the CBHTs of different homes to find the aggregate power consumption. A typical aggregate load is shown in Figure 4.3. Fourth we use this aggregate power consumption in a numerical simulation to obtain the aggregate duration of load disruption corresponding to a particular transformer and storage sizing. We note that a similar approach can be used by an electric utility to empirically size storage *without* needing to use teletraffic analysis. The results from this numerical simulation are presented in Section 4.2.

[1]All our measurements have been obtained during the winter.

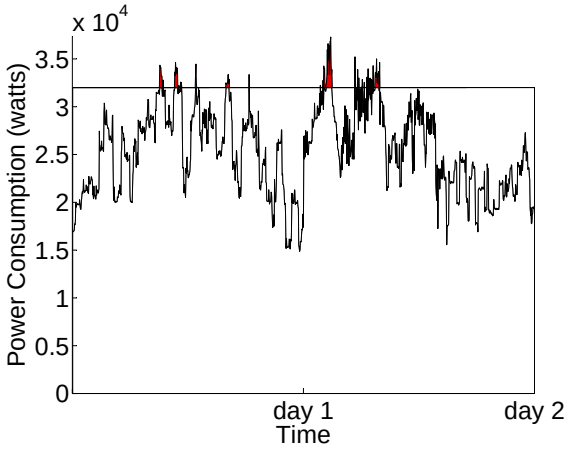


Figure 4.3: An example of the aggregate workload over a period of two days. The shaded areas above the horizontal line represent the times when demand is met from the store for a transformer size of 32.4 kiloWatts.

4.1.3 Teletraffic-based Sizing

As discussed earlier, using teletraffic theory to size transformers and storage has several advantages over an empirical approach. Indeed, applying the theory allows us to readily compute the effect of varying the number of homes, the buffer size, or the proportion of the homes in each class without having to recompute or re-measure the aggregate load and run onerous numerical simulations.

To gain these advantages, however, we need to make some additional assumptions about the nature of electrical demands. These are:

5. Cumulative busy hour traces (CBHTs) from different homes in the same class can be concatenated to represent the aggregate demand from the class. We call the concatenated cumulative busy hour traces the CCBHTs.

6. CCBHTs are independent.

7. CCBHTs are adequately represented by a k-state continuous-time Markov model. This implicitly assumes that busy hour behavior is stationary and ergodic.

8. Asymptotic limits can be used even for the fairly small number of homes and CCBHTs in our study.

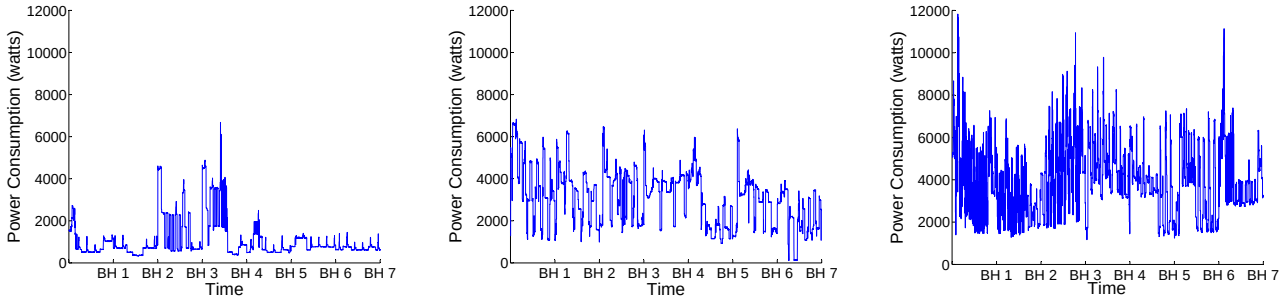


Figure 4.4: CCBHTs for three classes for one week.

Assumption 5 is a representativeness assumption. It allows us to concatenate CBHTs of homes in the same class to get the concatenated CBHT (CCBHT) of that class, which represents its busy hour behavior. Figure 4.4 shows the typical CCBHT of three classes in our measurement study. The last three assumptions are technical assumptions needed for teletraffic analysis.

We first use Assumption 7 to create a Markov model for each class using the approach explained in Chapter 2, then describe the algorithm for sizing using teletraffic theory. We jointly validate assumptions 5-8 by comparing the loss duration predicted by teletraffic analysis to those computed by numerical simulation in Section 4.2.

Creating a Markov Model for a Class

Home loads are due to the superposition of loads from different electrical appliances [23]. Both the literature and our observations suggest that each appliance can be modeled as an ON-OFF source with exponentially distributed ON and OFF periods. An appliance i consumes P_{ON_i} watts when it is ON, and P_{OFF_i} watts when it is OFF (usually, P_{OFF_i} is zero). Therefore, it is plausible that the power consumption of a class can be modeled as a k-state continuous-time Markov process. However, this still leaves the assignment of power levels to Markov states open.

To address this issue, as we discussed in Section 2.4, we use the *k-means* clustering algorithm to cluster the CCBHT for each class into k levels. Using these levels, we construct a modified CCBHT by replacing a measured power consumption value with the value of the center point of the cluster that it belongs to. Since k is an unknown, to determine the appropriate value of k for each class, we run the clustering algorithm with different values of k.

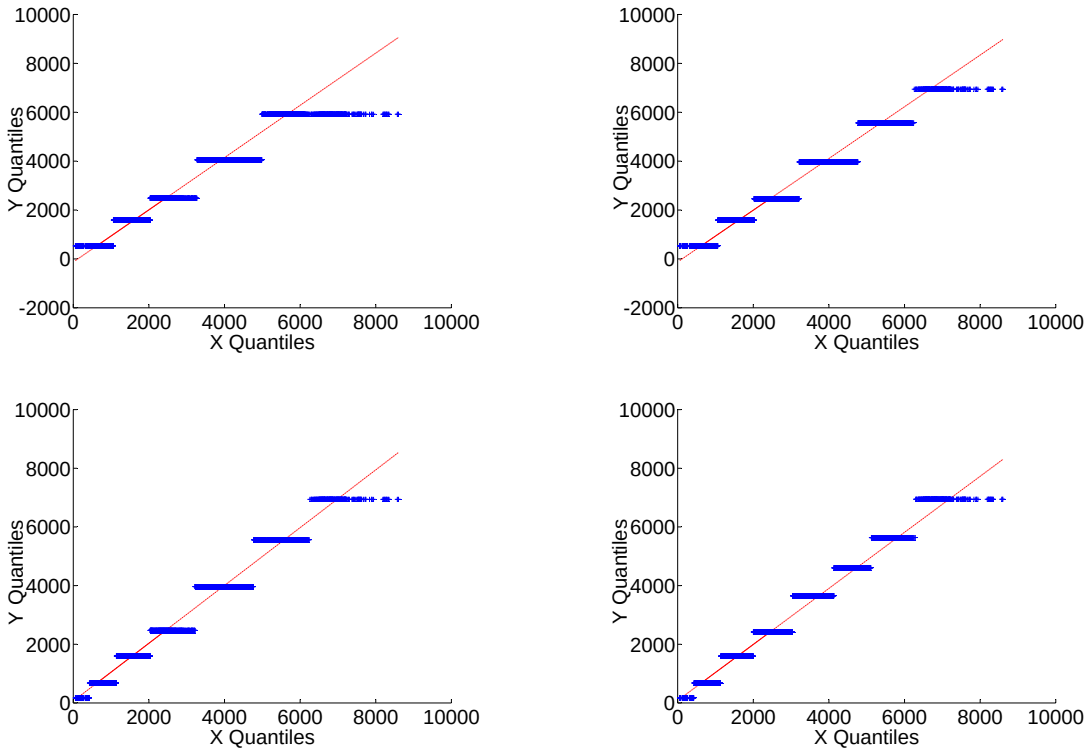


Figure 4.5: Q-Q plots comparing the representativeness of 5-,6-,7-, and 8-state demand models.

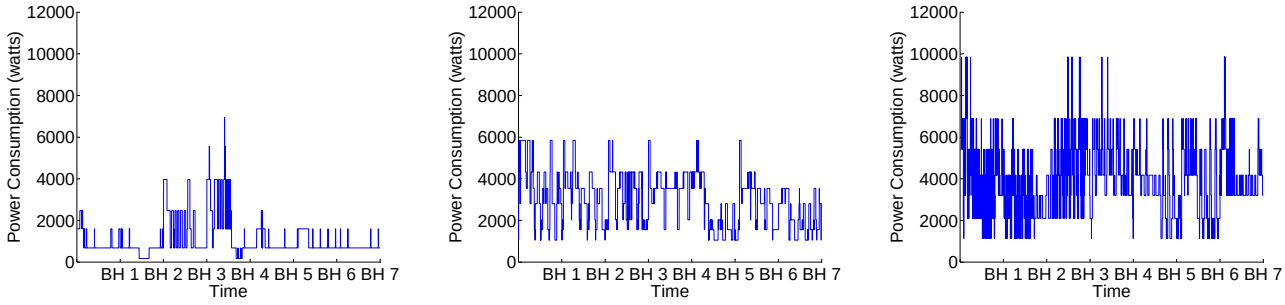


Figure 4.6: Seven-state models corresponding to the three CCBHTs.

We use the standard Q-Q goodness-of-fit test, discussed in Section 2.4.4, to see how close the modified CCBHT (the one that has only k levels) is to the original one. Clearly, adding more states increases both the goodness-of-fit and the computational complexity of our analysis. Therefore, our goal is to find a sufficient number of states that leads to a model that is a good representative of that class. Figure 4.5 shows the Q-Q plot for different values of k. We find that seven states seems to be a good compromise because, like the eight-state model, it captures both the lowest and the highest demand levels, which the six-state model does not. Figure 4.6 represents the three CCBHTs in Figure 4.4 when modeled with seven states.

The Teletraffic-based Sizing Algorithm

Given the set of Markov models, one for each class, using teletraffic theory to compute sizing requires four additional steps. First, we compute the power consumption rate matrix, R, and the intensity matrix, Q, of each class from its modified CCBHT as follows. The rate matrix represents the amount of power consumed by houses in each state. Values of the center points of the clusters (in the clustered CCBHTs) that are found for $k = 7$ are elements of the power consumption rate matrix, R. The intensity matrix specifies how fast the amount of power consumption is changed. We construct the intensity matrix of the Markov models by finding the average time that it takes to transition from the state i to the state j, which gives us $1/q_{ij}$ (q_{ij}s are elements of the intensity matrix, Q).

Second, from the Q matrix we compute the stationary probability distribution of the continuous-time Markov process. Suppose that π_i is the stationary probability of being in state i, we write the moment generating function of the stationary power consumption of a Markovian source

$$M(s) = \sum_i \pi_i e^{sr_i} \tag{4.1}$$

where r_i is the power consumption in state i.

Third, using the moment generating function of a Markovian source we can derive a formula for the effective bandwidth. We elaborated on this process in Section 3.2.

Finally, in the fourth step, we use the approximations in Equations (3.4) and (3.8), to find (C, B) tuples such that for a given number of houses from each class, the loss probability is less than a specific value. To allow us to study distribution branches with different numbers of houses than in our study, we assume that the proportion of houses in class i, say ρ_i, is constant and known. Therefore, if we have n houses, the number of houses in class i is $n\rho_i$, which we can then use in our teletraffic-based design rules.

4.2 Results

We present our results in four parts. First, we describe the methodology by which we placed home loads into one of three classes. Second, we validate our use of teletraffic theory by comparing the aggregate duration of load outage, for a particular sizing, obtained using numerical simulations and teletraffic analysis. Third, we compare the sizing obtained from our model with those used by a major electricity utility in our geographical area. Finally, we use teletraffic models to study the behavior of the electrical grid in response to changes in transformer and storage sizing. This allows us to gain insights into the operation of this complex system.

4.2.1 Classifying Home Loads

As mentioned earlier, standard rules based on decades of field experience allow an electric utility to both predict and classify a home load based on a few simple parameters. In Section 2.2, we used such a parametrization to compute the 'unit value' for each home and categorize homes into a few distinct classes based on these unit values.

Here, we use the same classes obtained for the modeling purpose (Table 2.2). We computed the CCBHTs for the homes in each class as discussed in Section 4.1 to carry out teletraffic analysis and these are shown in Figure 4.4.

4.2.2 Comparing Results from Numerical Simulation and Teletraffic Theory

We used both teletraffic theory and numerical simulations to compare the expected aggregate duration of load disruption for the set of nine homes in our measurement study keeping the LOLP fixed at two values: 10^{-3} and 10^{-5}. Although we collected data from some homes for as long as 60 days, we had reliable data from all nine homes for only seven days. Therefore, our results are from the concatenation of the busy hours from only these seven days.

For a particular LOLP, we computed equivalent pairs of (B, C) values, as shown in Table 4.1. We chose a wide range of B values and computed the corresponding C values. Note that B values are in Watt-TimeUnits, where one TimeUnit is the granularity of our measurement, i.e., six seconds. Therefore, a B value of 10^4 Watt-TimeUnits, for example, is $10^4/10$ Watt-minutes or 16.7 Watt-hours.

LOLP	(B,C) tuples (Watt-TimeUnits, VA)	Teletraffic theory (TimeUnits)	Numerical simulation (TimeUnits)
10^{-3}	(B=0 , C= 46670)	4.2	0
	(B=10^5 , C= 31030)	4.2	0
	(B=10^6 , C= 26894)	4.2	37
	(B=10^7 , C= 25495)	4.2	0
	(B=10^8 , C= 25206)	4.2	0
10^{-5}	(B=0 , C= 54001)	0.04	0
	(B=10^5 , C= 33464)	0.04	0
	(B=10^6 , C= 27437)	0.04	0
	(B=10^7 , C= 25636)	0.04	0
	(B=10^8 , C= 25223)	0.04	0

Table 4.1: Loss duration for seven days of measurements conducted in nine houses. The time unit is six seconds.

Given a (B, C) pair, we used numerical simulation to compute the actual duration of load disruption. Table 4.1 compares the duration of load outage predicted using the two techniques. We see that the predictions from theory closely match simulation results, except for one anomalous result. We attribute this to the fact that a buffer value of $B = 10^6$ is not in the asymptotic regime (also see our discussion below). Therefore, in practice, we advocate using our design rules only for buffer sizes larger than $B = 10^6$.

To compensate for the limited duration of our trace and to additionally validate our approach, we also synthetically generated the electricity demand for 100 homes for 100 days using a 1-minute-grain simulator developed at the University of Loughborough [23]. This workload generator has been shown to closely approximate real domestic demands. To generate this synthetic trace, we chose all the homes to have four occupants and with a randomly selected mix of appliances. All other values were those set by default including the occupancy pattern. The subsequent modeling and analysis of this data set was identical to that used for our own data set. However, it represents both a homogeneous load population as well as a much longer CCBHT for load modeling. Table 4.2 shows the results of this comparison. We see that for buffer values smaller than $B = 10^6$, predictions from teletraffic analysis are not consistent with ground truth. However, beyond this threshold, the agreement is excellent.

To further validate our results, we computed load models from the entire 24-hour trace, rather than just the busy hour both for our data set and the synthetic data set. We

LOLP	(B,C) tuples (Watt-TimeUnits, VA)	Teletraffic theory (TimeUnits)	Numerical simulation (TimeUnits)
10^{-3}	(B=0 , C= 564098)	6	85
	(B=10^5 , C= 525624)	6	382
	(B=10^6 , C= 494944)	6	47
	(B=10^7 , C= 489557)	6	0
	(B=10^8 , C= 488859)	6	0
10^{-5}	(B=0 , C= 591809)	0.06	0
	(B=10^5 , C= 546127)	0.06	40
	(B=10^6 , C= 498406)	0.06	11
	(B=10^7 , C= 489915)	0.06	0
	(B=10^8 , C= 488895)	0.06	0

Table 4.2: Loss duration for 100 days for synthetic data for 100 statistically identical houses. The time unit is one minute.

compared predictions from this load model with numerical simulations over the entire 24-hour trace. We found that for all LOLPs and for all values of B, the predictions were a tight upper bound on the simulation results.

4.2.3 Comparing Our Sizing with Industry Practice

The transformer sizing rules used by a major utility in our geographical area are shown in Table 4.3. We now compare the sizing obtained by using our analysis and these rules.

Total unit value	Transformer size (kVA)
1-3	10
4-9	25
10-24	50
25-36	75
37-50	100
51-88	167

Table 4.3: Transformer sizing rules used by a major utility.

The total unit value of the nine homes in our study was 23.5. Thus, for the industry standard LOLP of 2.74×10^{-4} the transformer size is 50kVA. From Table 4.1, we predict

that for an LOLP of 10^{-3} with no storage, the transformer size should be 46.6 kVA and for an LOLP of 10^{-5}, the size should be 54 kVA. This is in excellent agreement with the heuristics used by the utility. This indicates that a careful load modeling based on measurements matches heuristics developed over decades of field experience, validating our analysis.

Note that had the sum of unit values been even slightly larger (greater than 24), the heuristic would have advocated a size of 75 kVA, which would have been 50% greater than strictly necessary to meet the LOLP. Even greater savings can be achieved by adding storage. For example, our analysis indicates that, for the same set of homes, by adding 10^7 Watt-time units, or 16.7 kWh of storage, keeping LOLP at 10^{-5}, it is possible to reduce the transformer size from 54 kVA to 25.6 kVA, a reduction of 52%.

To further investigate this agreement, we used equivalent pairs of (B, C) values for the industry standard LOLP of 2.74×10^{-4} computed from the synthetic data (electricity loads of 100 statistically identical homes[2] for 100 days generated by the simulator developed at the University of Loughborough). Figure 4.7 illustrates how close our sizing is to the industry practice [3]. We again note that there is no storage sizing guideline in practice; thus, we can only compare our results in the bufferless case with the heuristic approach adopted by electric utilities to size transformers. Since transformer capacity is quantized in practice, we found the closest transformer size that is larger than the C value computed in the bufferless case (Figure 4.8). It can be seen that our sizing is just as conservative as the industry practice, implying that the teletraffic-based sizing works pretty well.

4.2.4 The Effect of Storage on the Electrical Grid

We now use teletraffic analysis to study the insights embodied by our design rules, that is, the inter-relationship between transformer size C, the storage size B, the number of homes n and the loss probability p.

We first study the effect of storage size and loss probability on transformer size for 100 and 1000 homes (Figures 4.9, and 4.10). We choose to study 100 and 1000 homes because these are the typical number of homes that we can expect to share a store and a transformer in a typical small distribution system. Here, we find that the addition of storage has a perceptible impact on transformer sizing. For 100 homes, as we go from no storage to 16.7 kWh of storage, the required transformer size, for a fixed loss probability, goes down by

[2]All homes belong to class 2 (unit size=2.5) presumably.

[3]To convert the unit from Watts to Volt Amperes, we must divide it by the power factor. We set the power factor to 0.8 in our study

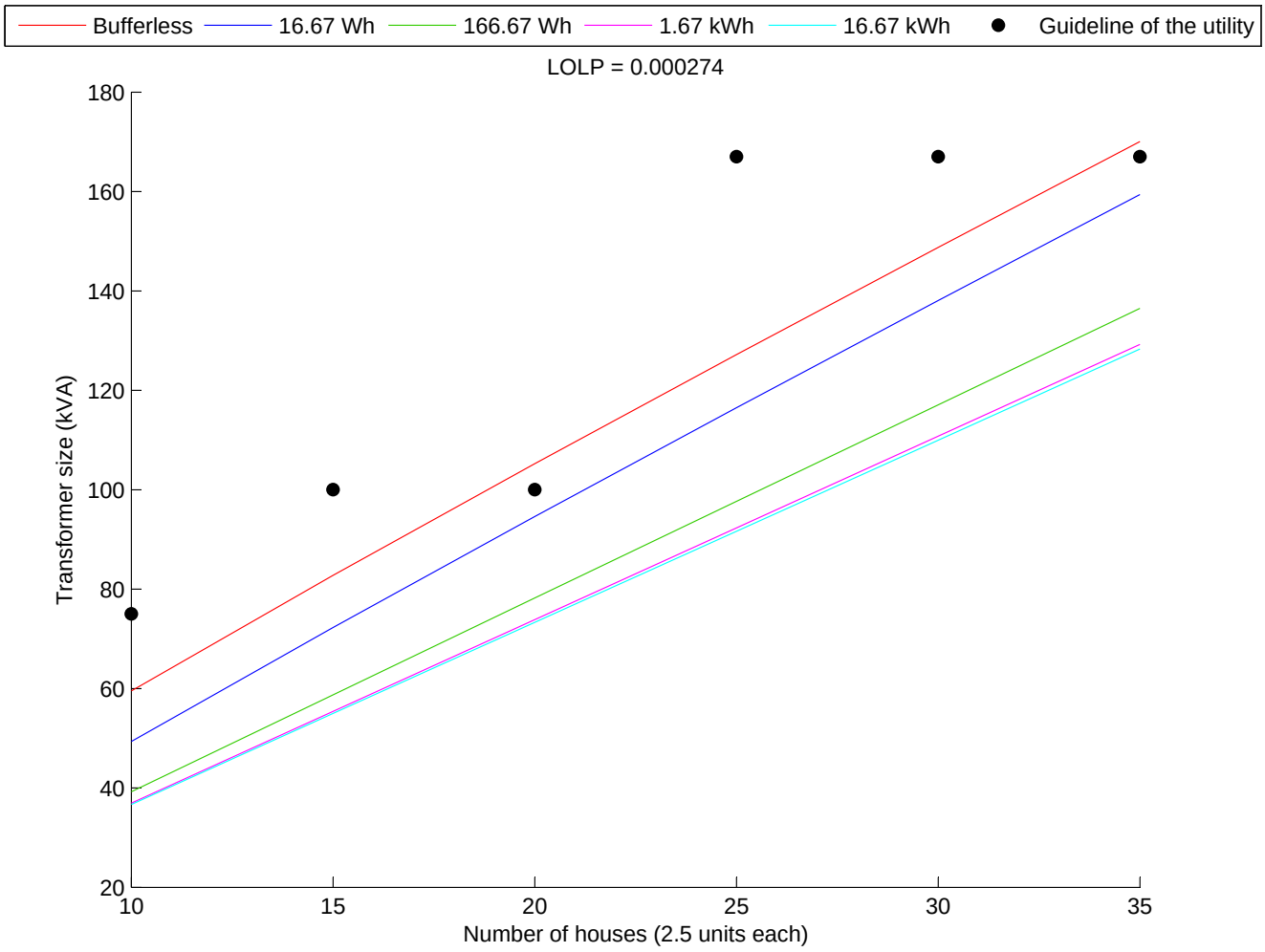


Figure 4.7: Comparison of the teletraffic-based sizing approach and the heuristic sizing approach.

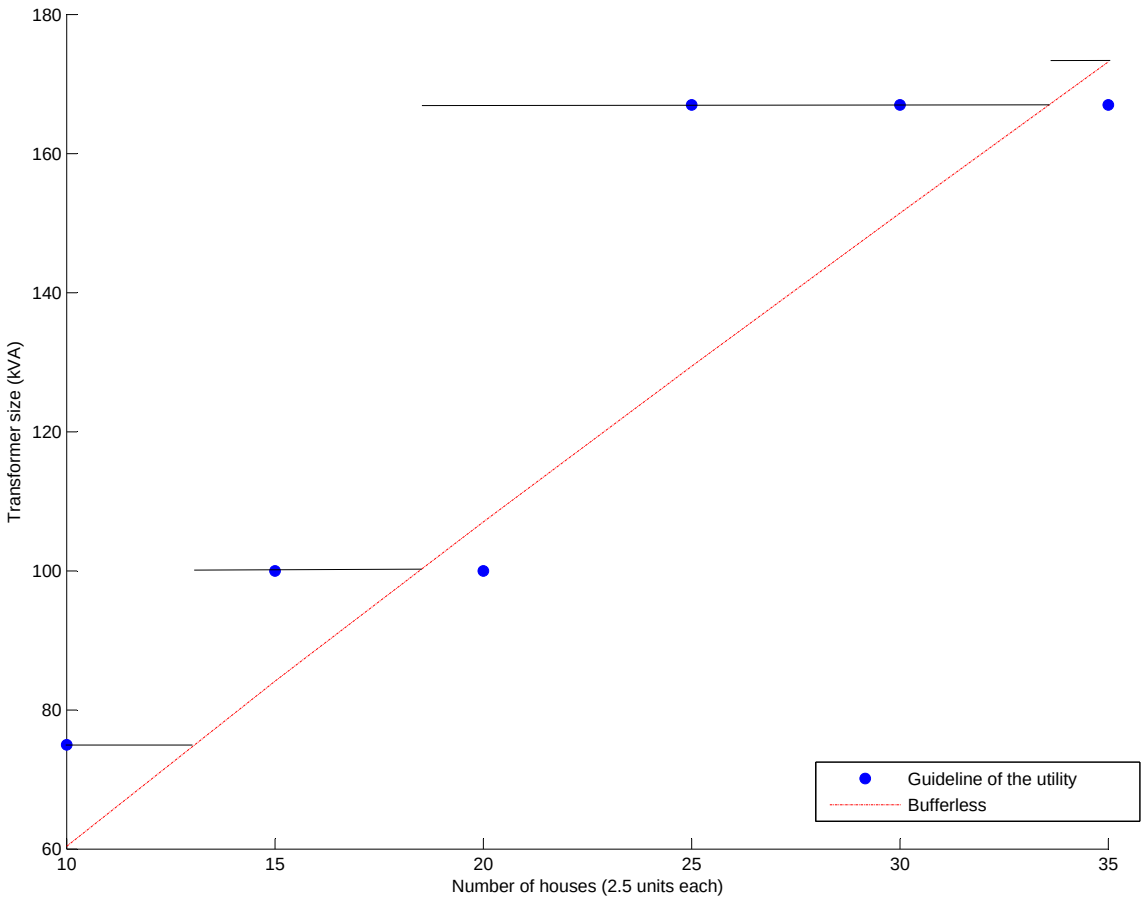


Figure 4.8: Comparison of the teletraffic-based sizing approach and the heuristic sizing approach in the bufferless case when transformer capacity is quantized. The black lines represent quantized values of transformer capacity obtained from the teletraffic-based sizing approach.

about 20%. Interestingly, we find that the transformer size increases fairly sharply as the loss probability measure becomes more stringent. However, by adding storage, we can gain the same level of reliability without increasing the transformer size.

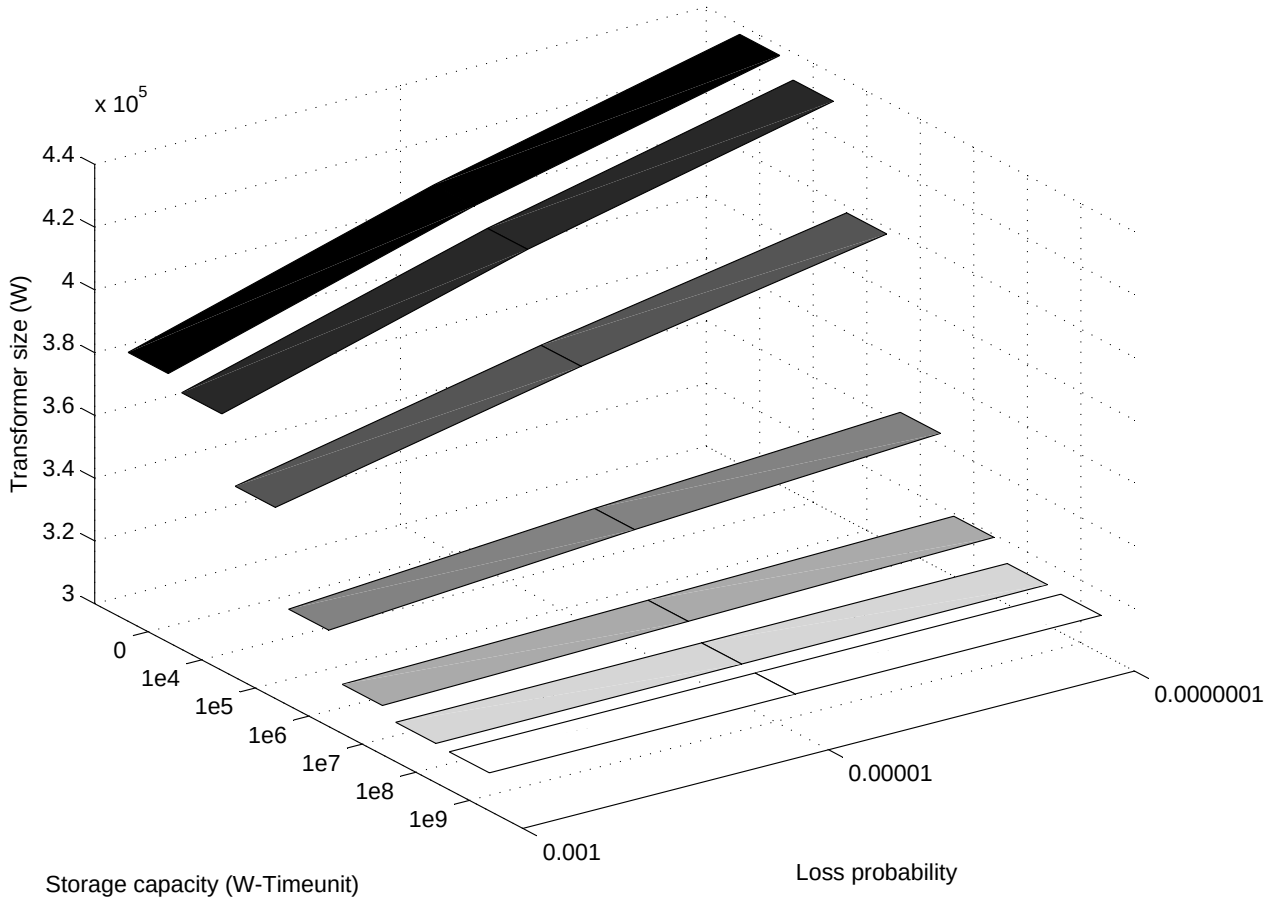


Figure 4.9: The effect of storage size and loss probability on transformer capacity for 100 houses.

We next keep the loss probability fixed and vary both the number of homes n and the storage size (Figures 4.11, and 4.12). For both small and large n the required transformer size increases linearly with the number of houses. This is a straightforward consequence of

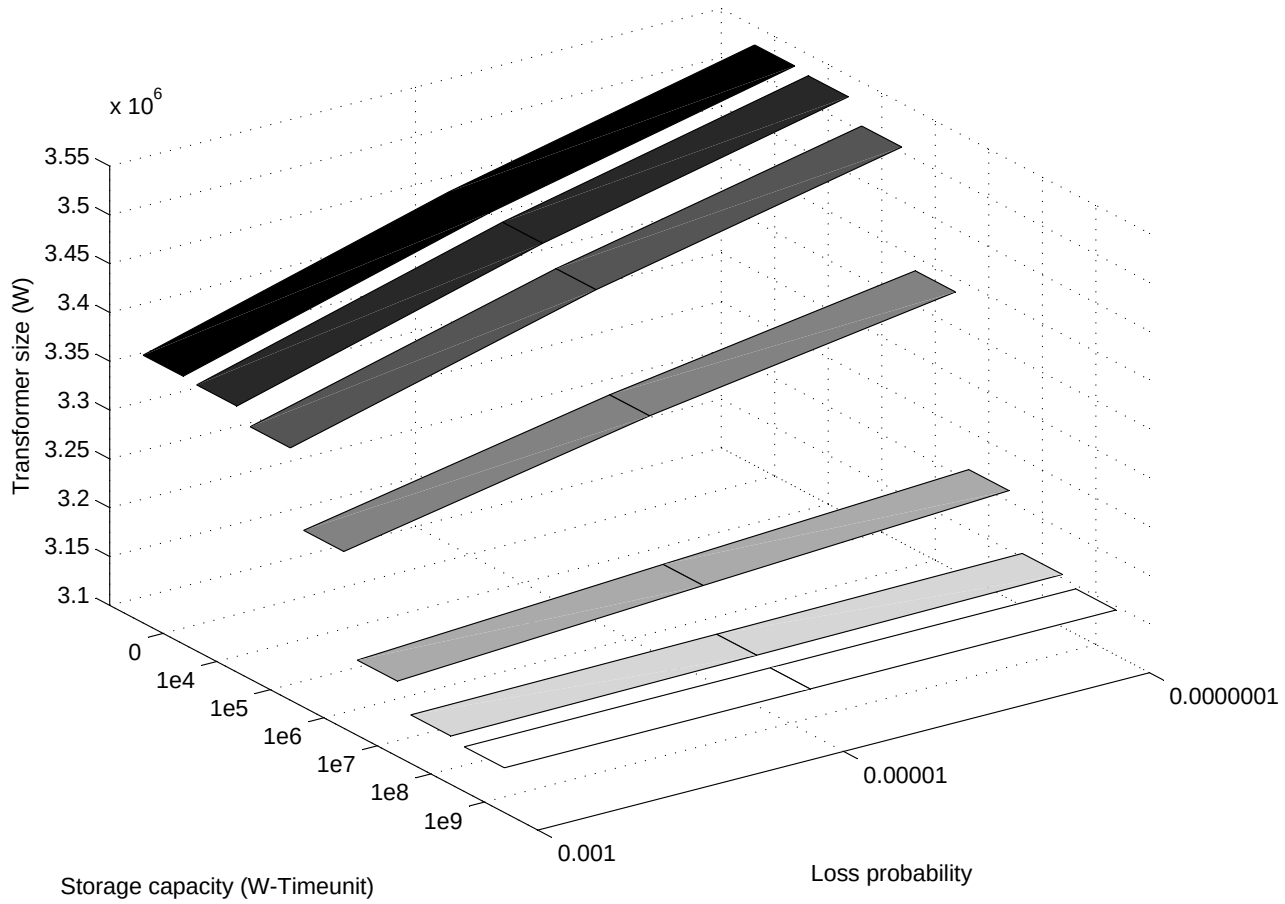


Figure 4.10: The effect of storage size and loss probability on transformer capacity for 1000 houses.

representing a house by its 'effective bandwidth.' We also see that for small values of n, as the storage size increases, the transformer size required decreases significantly, demonstrating that the addition of storage allows us to reduce transformer capacity. However, for large values of n, storage appears to have only moderate effect, a well-known phenomenon in the Internet [5].

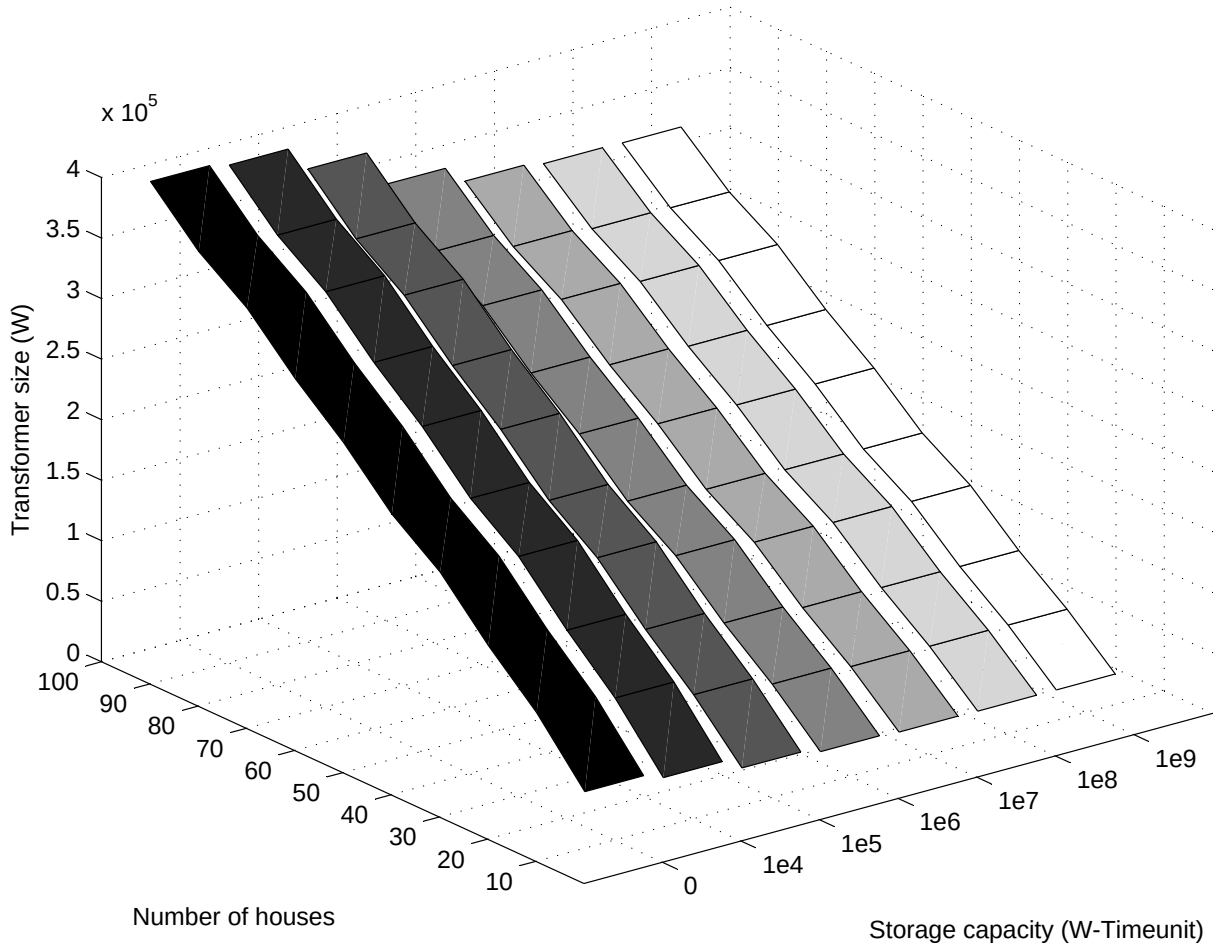


Figure 4.11: The effect of number of homes and storage size on transformer capacity for a fixed loss probability of 10^{-3}(10-100 houses).

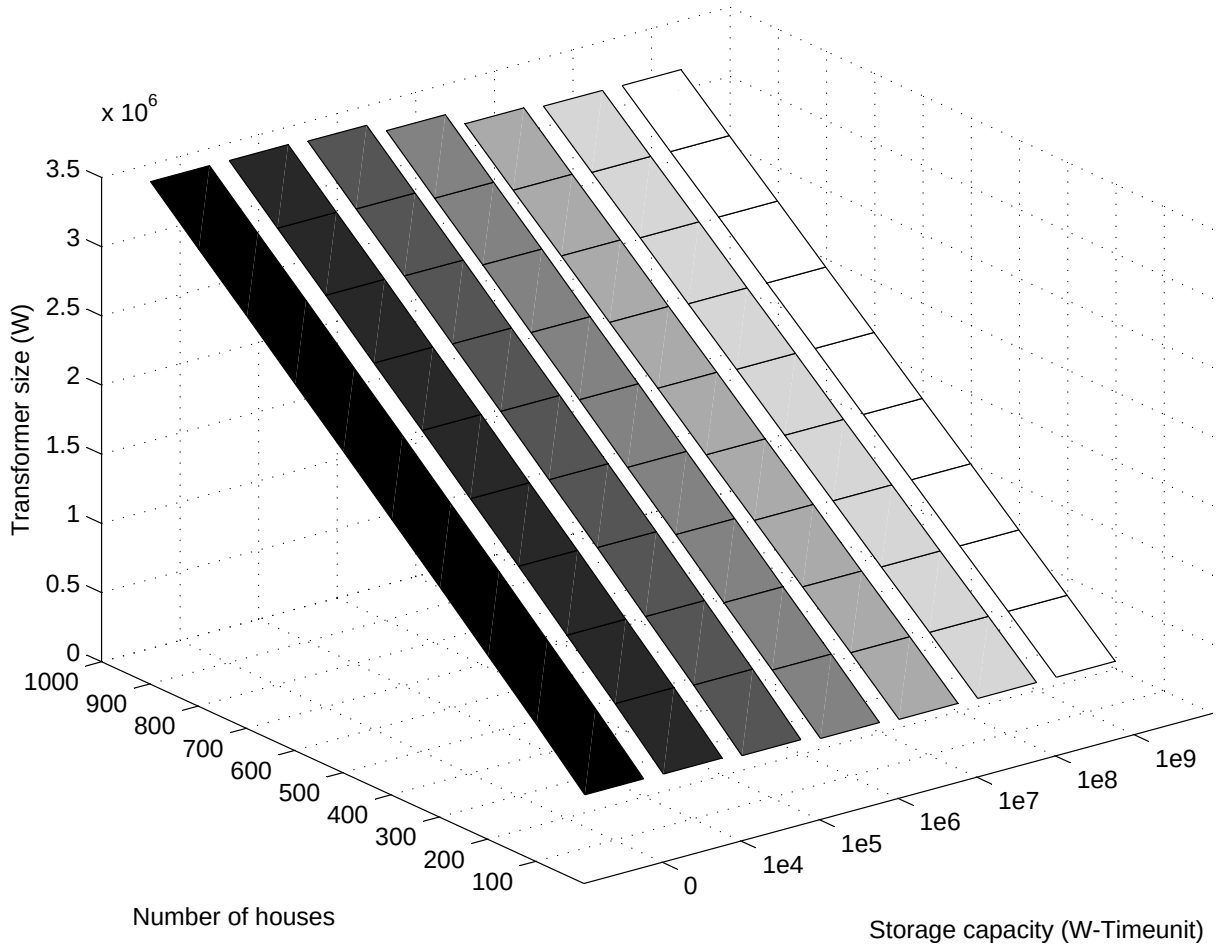


Figure 4.12: The effect of number of homes and storage size on transformer capacity for a fixed loss probability of 10^{-3}(100-1000 houses).

4.2.5 How to Size the Grid?

In the previous sections, we showed high accuracy of the teletraffic-base sizing approach. Now in this section, using our measurements we obtain sizing curves (Figures 4.13,4.14, and 4.15) in presence of storage. Note that LOLP is set to 2.74×10^{-4} that is the industry standard. We believe that these curves can be used by utilities to size either pole-top or substation transformers and storage systems jointly.

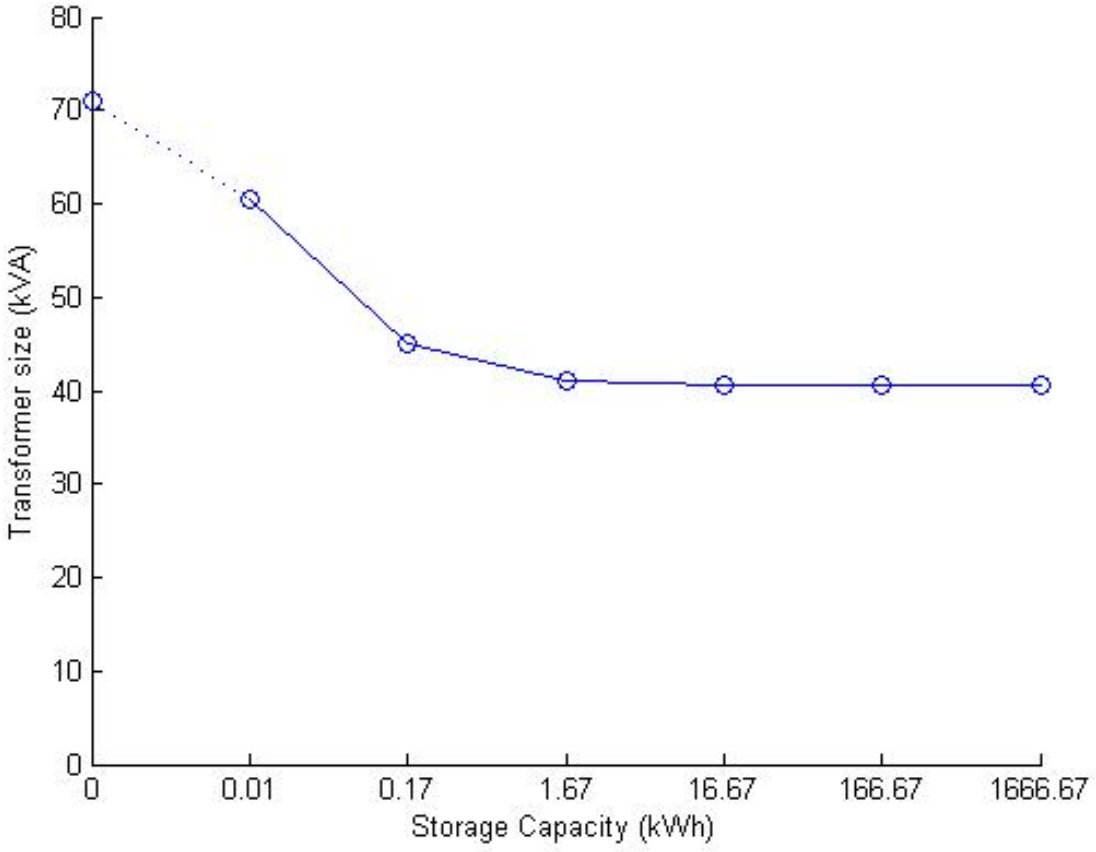


Figure 4.13: Equivalent pairs of (B, C) values computed for 10 houses for the industry standard LOLP of 2.74×10^{-4}.

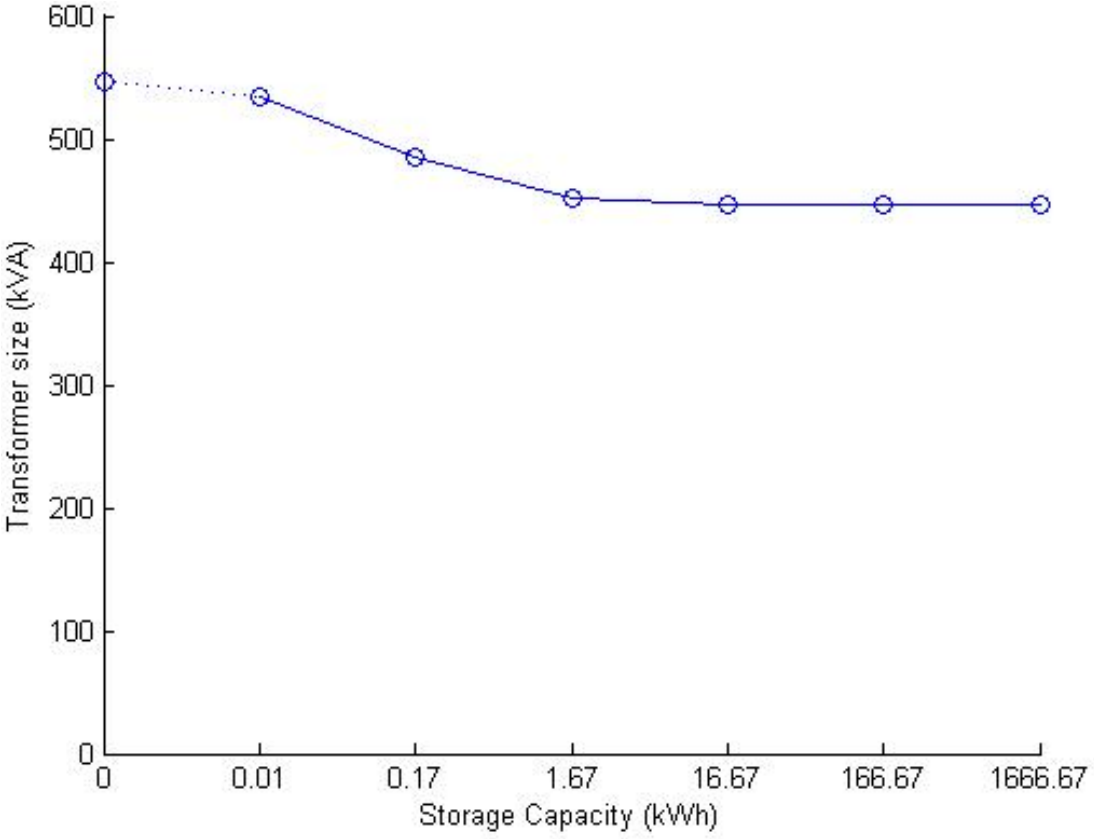


Figure 4.14: Equivalent pairs of (B, C) values computed for 100 houses for the industry standard LOLP of 2.74×10^{-4}.

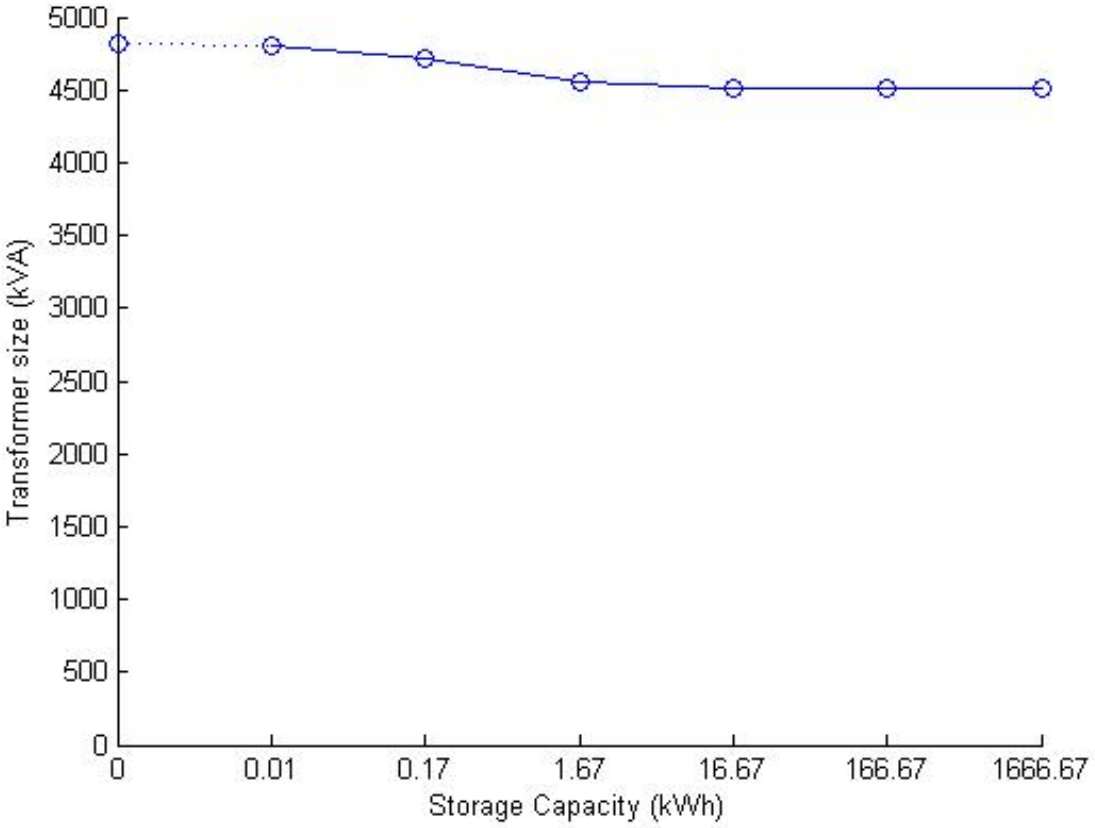


Figure 4.15: Equivalent pairs of (B, C) values computed for 1000 houses for the industry standard LOLP of 2.74×10^{-4}.

Chapter 5

Conclusions and Future Work

5.1 Limitations and Future Work

Our work has made several simplifying assumptions. First, we have already noted that teletraffic design rules are meant to be used in the asymptotic regimes for the number of houses, storage capacity, and transformer capacity. Although these asymptotes are arguably achieved for storage capacity (10^7 Watt-TimeUnits) and transformer size (10^4 VA) in our measurement study, they are certainly not achieved for the number of houses. Therefore, we caution the use of these rules for small distribution networks: they are far more applicable deeper in the distribution tree. Unfortunately, lacking data from a sufficiently large number of houses, we were forced to apply our techniques to the small-n regime.

Second, network buffers can be filled and drained with no perceptible degradation. In contrast, the lifetime of a battery may depend on the depth of discharge, especially for Lithium-ion cells [21]. This is due to the buildup of electrolytic deposits in the battery, and degradation of the anode [27]. In addition, the lifetime also may depend on the temperature of use and the exact charging voltage. These effects are difficult to quantify, making the use of teletraffic analysis only one piece of a complex puzzle. Incorporation of battery dynamics into the system model without overly complicating the analysis is an exciting area for future work.

Third, although our work was motivated by the need to re-examine design rules for time-varying generation from renewable energy sources, this paper does not deal with this issue. We believe, however, that $G/G/1/B$ teletraffic analysis can be used to study time-varying Markovian generation systems.

Despite these limitations, we believe that the use of teletraffic analysis to model and size electrical grids represents an exciting area of multi-disciplinary work. We hope to use our approach in the future to answer questions such as:

- We currently study the problem of sizing a single branch of distribution systems with one transformer and one storage system. If we have a hierarchy of storage devices deployed in different locations, from homes to distribution and transmission systems, how should we size all these devices?
- If home owners also own electric vehicles so that there is storage at each home, is shared storage in the distribution system necessary or cost effective?
- Transformers could be overloaded for a short period of time; that is why there are short-term ratings for transformers [1]. How should we change the model to take into account these short-term ratings?
- If a home generates electricity according to a stochastic process that models wind or photovoltaic generation, how does this affect the sizing and operations of the in-home store and shared store?
- The buffer overflow approximations which constitute the basis of our work are only valid for large buffers and for many sources. How can we solve the sizing problem when we are not in the asymptotic regime (for instance, sizing a single battery for a single house)?
- What could be said if homes are not statistically independent?

5.2 Concluding Note

We revisit the rules for sizing elements of the electrical grid motivated by the replacement of ageing infrastructure and the anticipated increase in storage deployment. Instead of modeling loads by their peak values and using linear optimization, which is the standard approach in power systems, our work presents a new approach to define design rules for distribution systems. The basis of our work is the Equivalence Theorem, which states that a battery in the electrical grid can be modeled as a buffer in a network. This permits us to apply teletraffic analysis to size the electrical grid. We validate our approach by using our

[1] The first and the second ratings are usually 1.33 and 1.66 times the base rating respectively.

own measurement data as well as synthetic data. Our results show that our approach is in good agreement both with numerical simulations and with industry practice. Moreover, we present parsimonious models for home loads that can be used in different problem domains.

References

[1] Current cost. `http://www.currentcost.com/`, January 2011. 9

[2] General electric transformer pricing. `http://www.geindustrial.com/catalog/buylog/20_BL.pdf`, January 2011. 2

[3] D. Anick, D. Mitra, and M. M. Sondhi. Stochastic theory of a data handling system with multiple sources. *Bell System Technical Journal*, 61:1871–1894, 1982. 31

[4] K. Aoki, K. Nara, T. Satoh, M. Kitagawa, and K. Yamanaka. New approximate optimization method for distribution system planning. *IEEE Trans. Power Systems*, 5(1):126–132, 1990. 4

[5] G. Appenzeller, I. Keslassy, and N. McKeown. Sizing router buffers. In *Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 281–292. ACM, 2004. 48

[6] M. Armstrong, M.C. Swinton, H. Ribberink, I. Beausoleil-Morrison, and J. Millette. Synthetically derived profiles for representing occupant-driven electric loads in Canadian Housing. *J. of Building Performance Simulation*, 2(1):15–30, 2010. 8

[7] A. Capasso, W. Grattieri, R. Lamedica, and A. Prudenzi. A bottom-up approach to residential load modeling. *IEEE Trans. Power Systems*, 9(2):957–964, 2002. 8

[8] W.G. Cochran. *Sampling techniques*. John Wiley & Sons Inc, 1977. 9

[9] MK Deshmukh and SS Deshmukh. Modeling of hybrid renewable energy systems. *Renewable and Sustainable Energy Reviews*, 12(1):235–249, 2008. 4

[10] KC Divya and J. Østergaard. Battery energy storage technology for power systems–An overview. *Electric Power Systems Research*, 79(4):511–520, 2009. 4

[11] W. El-Khattam, YG Hegazy, and MMA Salama. An integrated distributed generation optimization model for distribution system planning. *IEEE Trans. Power Systems*, 20(2):1158–1165, 2005. 1, 4

[12] Anwar Elwalid, Daniel Heyman, T. V. Lakshman, Debasis Mitra, and Alan Weiss. Fundamental bounds and approximations for atm multiplexers with applications to video teleconferencing. *IEEE JSAC*, 13(6):1004–1016, 1995. 30, 31

[13] Anwar Elwalid and Debasis Mitra. Effective bandwidth of general markovian traffic sources and admission control of high speed networks. *IEEE/ACM ToN*, 1(3):329 –343, June 1993. 29

[14] Frank Kelly. *Notes on Effective Bandwidth*, pages 141–168. Oxford University Press, 1996. 26, 28, 30

[15] W. Kempton and J. Tomic. Vehicle-to-grid power fundamentals: Calculating capacity and net revenue. *J. Power Sources*, 144(1):268–279, 2005. 4

[16] WR Lachs and D. Sutanto. Uncertainty in electricity supply controlled by energy storage. In *Proc. Energy Management and Power Delivery*, volume 1, pages 302–307. IEEE, 1995. 4

[17] B.S. Lee and D.E. Gushee. Massive Electricity storage. *AIChE White Paper, AIChE Government Relations Committee*, 2008. 1, 4

[18] Ravi Mazumdar. *Performance Modeling, Loss Networks, and Statistical Multiplexing*. Synthesis Lectures on Communication Networks. Morgan & Claypool Publishers, 2009. 22

[19] Alexandra Meier. *Electric Power Systems: A Conceptual Introduction*. Wiley-IEEE Press, 2006. 2, 3, 5

[20] J.V. Paatero and P.D. Lund. A model for generating household electricity load profiles. *International J. of Energy Research*, 30(5):273–290, 2006. 8

[21] S.B. Peterson, J. Apt, and JF Whitacre. Lithium-ion battery cell degradation resulting from realistic vehicle and vehicle-to-grid utilization. *J. Power Sources*, 2009. 54

[22] K. Ponnambalam, YE Saad, M. Mahootchi, and AW Heemink. Comparison of methods for battery capacity design in renewable energy systems for constant demand and uncertain supply. In *7th Intl. Conf. on the European Energy Market*, pages 1–5. IEEE, 2010. 4

[23] I. Richardson, M. Thomson, D. Infield, and C. Clifford. Domestic electricity use: A high-resolution energy demand model. *Energy and Buildings*, 2010. 7, 8, 38, 42

[24] J. W. Roberts. Performance evaluation and design of multiservice networks. COST 224 Final Report, Commission of the European Communities, October 1992. 8

[25] A. Roy, S.B. Kedare, and S. Bandyopadhyay. Application of design space methodology for optimum sizing of wind-battery systems. *Applied Energy*, 86(12):2690–2703, 2009. 4

[26] P. Sen, B. Maglaris, N.-E. Rikli, and D. Anastassiou. Models for packet switching of variable-bit-rate video sources. *Selected Areas in Communications, IEEE Journal on*, 7(5):865 –869, June 1989. 13

[27] J. Vetter, P. Novak, MR Wagner, C. Veit, K.C. Moller, JO Besenhard, M. Winter, M. Wohlfahrt-Mehrens, C. Vogler, and A. Hammouche. Ageing mechanisms in lithium-ion batteries. *J. Power Sources*, 147(1-2):269–281, 2005. 54

[28] Ward Whitt. *Stochastic-Process Limits - An Introduction to Stochastic-Process Limits and Their Application to Queues*. Springer Series in Operations Research and Financial Engineering. Springer, 2002. 26, 31