On the Use of Teletraffic Theory in Power Distribution Systems

Omid Ardakanian, S. Keshav
David R. Cheriton School of Computer Science
University of Waterloo
Waterloo, ON, Canada
{oardakan, keshav}@uwaterloo.ca

Catherine Rosenberg
Electrical and Computer Engineering Department
University of Waterloo
Waterloo, ON, Canada
cath@uwaterloo.ca

ABSTRACT

Loads on the electrical grid are multiplexed at distribution transformers in the same way that traffic from data sources is multiplexed at a router. This motivates the use of teletraffic theory to size power distribution networks just as it is used to size telecommunication access networks. Specifically, we prove the equivalence between a model of a distribution branch comprised of a transformer and storage that we want to size for a given underflow probability ϵ, and a queuing model that we want to size for a given overflow probability ϵ. Based on this equivalence, we show how existing teletraffic analysis can be applied to size transformers when there is no storage. We compute such sizings using load models obtained from our measurement testbed and load models derived from an electricity demand simulator. 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 by electric utilities, thus validating the use of teletraffic theory.

Categories and Subject Descriptors

G.3 [**Probability and Statistics**]: Queueing theory; I.6.4 [**Simulation and Modeling**]: Model Validation and Analysis

General Terms

Theory, Verification

Keywords

Electrical grid, Distribution transformer sizing, Teletraffic theory

1. INTRODUCTION

The electrical grid is similar to the Internet in many aspects [23]. Both networks have a similar topology; the grid consists of transmission and distribution networks that are analogous to core and access networks in the Internet. The grid serves multiplexed electricity demands of end-customers just as a communication network carries multiplexed traffic from data sources. Similar to Internet traffic, the demand for electricity is variable and exhibits burstiness over multiple time scales. However, unlike the Internet which provides best-effort service, the grid is designed to provide a certain level of reliability at all times. This complicates the problem of sizing lines and transformers in the grid as overestimating their sizes would lead to costly underutilized assets, and, on the other hand, underestimating them would put the reliable operation of the grid at risk. Fortunately, because demand uncertainty is not too high and the physics of transformers permits temporary thermal overloading, grid operators have developed simple sizing guidelines which rely on peak load estimation and forecasting [31, 9].

However, significant changes are underway in the electric power industry including the incorporation of intermittent renewable energy sources, electric vehicles, and storage into the grid. Renewable energy sources will introduce a high level of uncertainty to the supply side [32] and uncontrolled charging of electric vehicles will increase the uncertainty in the demand side [29]. Existing sizing guidelines cannot deal with highly stochastic load and supply let alone storage. Therefore, there is a need to develop new sizing guidelines which simultaneously take into account capacities of lines, transformers, storage, and renewable energy sources.

During the 1990's, a rich body of work in the area of teletraffic theory was developed to model, analyze, and size telecommunication networks such as Asynchronous Transfer Mode (ATM) networks which guarantee certain levels of quality of service [26]. In our work we adapt these techniques to size assets in the electrical grid. We state and prove the Equivalence Theorem that allows us to model a branch of a distribution network as a simple queueing system. The Equivalence Theorem allows us to jointly size transformers, and stochastic sources using teletraffic theory.

In this paper, we demonstrate that a specific area of teletraffic theory, *i.e.*, the large deviations theory developed in the context of ATM networks [17] can be used to size distribution transformers when storage is not available[1]. We make three specific contributions:

- We show that a branch of the distribution network can be modelled as a fluid queuing system which can be analyzed using an equivalent constant service rate fluid queuing model.
- We use large deviations theory to study the behaviour of multiplexed loads in the grid and to size distribution

[1] In [7], based on the Equivalence Theorem proved in this paper, we study more complex sizing problems that include storage.

e-Energy 2012, May 9-11 2012, Madrid, Spain.

transformers.

- We validate the analytical approach to sizing transformers by both comparing it to industry practice and the result of numerical simulation.

The rest of the paper is organized as follows. We present an overview of the electrical grid and specify our assumptions in Section 2. The queueing model for a distribution branch in the grid is derived in Section 3. We study multiplexed loads using teletraffic theory in Section 4. For completeness, we summarize the main points of our previous work [6] in Section 5. Specifically, we discuss our measurement testbed, introduce a classification scheme for the home loads, and explain how we use measurements from our testbed to construct Markovian reference models of electrical loads. We perform the teletraffic-based sizing using these models in Section 6. In Section 7, we present our validation approach. The results of this work are presented in Section 8. We survey related work in Section 9. We discuss our contributions in more detail in Section 10 and conclude in Section 11.

2. BACKGROUND AND ASSUMPTIONS

This section presents a high-level overview of the electrical grid as it relates to our work. The electrical grid consists of three subsystems: generation, transmission, and distribution [27]. Electrical power generators use energy from prime movers 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.

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 at rated voltage and frequency, measured in kilo Volt Amperes or kVA. Although this rating can be exceeded on rare occasions, grid design rules require that no transformer exceeds its rating for more than short time intervals.

Transformers can be expensive. A small pole-top 100kVA single-phase distribution transformer that serves about 20 homes in North America costs around \$3,000 [3]. A typical small utility serving a customer base of 30,000 homes would therefore need to spend \$4,500,000 on poletop distribution transformers alone. High-voltage transformers at substations, which serve thousands of customers, can cost up to \$1,500,000. Therefore, electrical utilities size their transformers to be large enough to meet expected peak loads, but not so large as to be too expensive [16].

Sizing a transformer is a critical design decision. A utility could potentially save millions of dollars by choosing smaller transformer sizes. On the other hand, underestimating the size of a transformer might lead to transformer overheating which accelerates transformer aging and increases the risk of failure [1]. Thus, transformer sizing must be done such that the grid reliability requirement is met. The grid reliability is typically expressed in terms of the *loss-of-load probability (LOLP)* [27], 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 North America. To comply with this reliability criterion, the transformer should be sized such that the loss of load event occurs less than one-day-in-ten-years. A loss of load *might* happen whenever the aggregate demand exceeds the transformer capacity[2]. We assume, conservatively, that a loss of load *will* happen whenever this condition is met.

Our study focuses on a single distribution branch of the electrical grid associated with a distribution transformer with a nameplate rating of S Volt Amperes capable of supplying C Watts when loaded at its nameplate rating (Figure 1a)[3]. Note that $C = Sf$ where f is the power factor computed for the load supplied by the transformer and is known *a priori*. The transformer is shared by a set of n homes in a residential neighbourhood, indexed by i. Each home places a load of $L_i(t)$ Watts on the system at time t[4]. We call the sum of the home loads at any time as the *aggregate* load at that time. We assume that all homes are located in a small geographical area so that distribution losses can be neglected.

3. A QUEUEING MODEL FOR THE DISTRIBUTION NETWORK

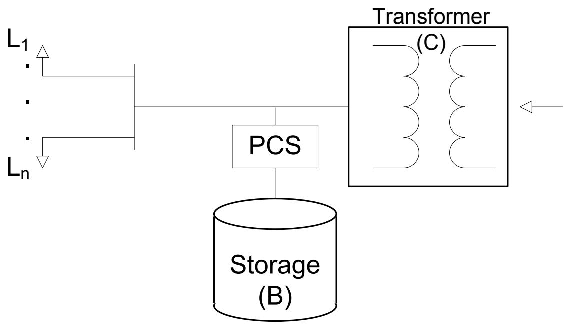
Observe that a branch of the distribution network can be modelled as a fluid queue in which storage can be viewed as a finite buffer with capacity B, the power supplied by the transformer at time t can be viewed as a fluid arrival bringing work to the system with a given peak rate C, and the load of home i can be viewed as a fluid service which consumes infinitesimal units of energy (*i.e.*, electrons) at time t at rate $L_i(t)$ so that the aggregate load of all homes at time t drains the buffer at rate $\sum_{i=1}^{n} L_i(t)$ [5]. We refer to this queueing system as M_1. The critical aspect of this queueing system that we want to quantify is its *underflow* probability, i.e., the probability that a service finds the buffer empty; this corresponds to the loss of load probability in a distribution network equipped with storage. Unfortunately, teletraffic analysis does not deal with this question. However, teletraffic analysis can be used to analyze the buffer *overflow* probability of standard work-preserving constant service rate fluid queue with arbitrary arrivals [22]. Thus, we construct a work-preserving constant service rate fluid queue (Figure 1b), referred to as M_2, such that its buffer size is B, its service rate is C, and the fluid arrival rate of source i at time t is $L_i(t)$ so that fluid arrivals of all sources at time t fill the buffer at rate $\sum_{i=1}^{n} L_i(t)$. Based on the

[2]In case that storage is connected to the distribution feeder (Figure 1a), a loss of load might happen whenever storage is depleted and the aggregate demand exceeds the transformer capacity.

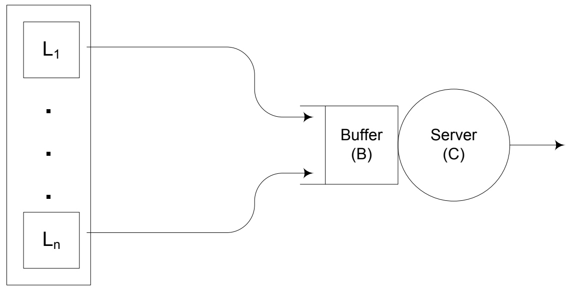
[3]For generality, the figure shows a transformer that is associated with storage of capacity B Watt-hours and a power conversion system, marked 'PCS', that charges storage whenever the aggregate load is smaller than C at rate $C - \sum_i L_i(t)$ and meets demand from storage whenever the aggregate load exceeds C at rate $\sum_i L_i(t) - C$.

[4]$L_i(t)$ represents the total power available to a resistive load. In a three-phase electric power system it is equal to the instantaneous total power consumed by a three phase load.

[5]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 $\rho > 1$.



(a) A branch of the electrical grid supplying power to n homes indexed by i where each home places a load of $L_i(t)$ Watts on the system at time t, the capacity of the battery is B Watt-hours, and the base rating of the transformer is C Watts.



(b) A fluid queue with n sources indexed by i where each source generates traffic at rate $L_i(t)$ bits at time t. The capacity of the buffer is B bits and the channel capacity (*i.e.*, the service rate) is C bits/second.

Figure 1: The storage system and a small network.

above intuition, our plan of attack is to show that we can replace the study of M_1 with the study of M_2, permitting the use of teletraffic analysis.

We denote the amount of work buffered in the M_1 system (i.e., its *workload*) at time t by $W(t)$ and similarly the amount of work buffered in the M_2 system at time t by $\overline{W}(t)$. The *workload trajectory* is defined as a specific realization of the workload process. Our main theoretical result is the following Equivalence Theorem:

Equivalence Theorem *Every workload trajectory in the M_1 queuing system corresponds to an equivalent trajectory in the M_2 queuing system such that $\forall t, W(t) + \overline{W}(t) = B$.*

To prove the theorem, we start with the following lemma.

Lemma 1 *For any fixed t^*, we have $\frac{dW(t)}{dt}|_{t=t^*} = -\frac{d\overline{W}(t)}{dt}|_{t=t^*}$ if $W(t^*)+\overline{W}(t^*) = B$*, and W and $\overline{W}$ are both differentiable at t^*.

Proof: First consider the rate of change of the workload of M_1 at time t. The buffer is filled/drained at the rate $C - \sum_i L_i(t)$, until it becomes full/empty. We write this as:

$$dW(t)/dt = \begin{cases} 0 & \text{if } W(t) = B \text{ and } C > \sum_i L_i(t), \\ 0 & \text{if } W(t) = 0 \text{ and } C < \sum_i L_i(t), \\ C - \sum_i L_i(t) & \text{otherwise} \end{cases} \tag{1}$$

Similarly, consider the rate of change of the workload of M_2 at time t. The buffer is filled/drained at the rate $\sum_i L_i(t) - C$, until it becomes full/empty. We write this as:

$$d\overline{W}(t)/dt = \begin{cases} 0 & \text{if } \overline{W}(t) = B \text{ and } C < \sum_i L_i(t), \\ 0 & \text{if } \overline{W}(t) = 0 \text{ and } C > \sum_i L_i(t), \\ \sum_i L_i(t) - C & \text{otherwise} \end{cases} \tag{2}$$

Comparing equations (1) and (2), it can be concluded that $\frac{dW(t)}{dt}|_{t=t^*} = -\frac{d\overline{W}(t)}{dt}|_{t=t^*}$ given that $W(t^*) + \overline{W}(t^*) = B$, and W and $\overline{W}$ are both differentiable at t^*. We note that W and $\overline{W}$ may be merely right- or left-differentiable at countably many points; however, the lemma holds for the right- or left-derivatives of W and $\overline{W}$ at these points. □

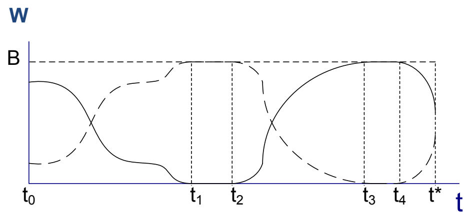


Figure 2: Workload of M_1 and M_2 systems. The dashed line and the solid line represent $\overline{W}(t)$ and $W(t)$ respectively.

Let the initial workload state in M_1 be $W(t_0)$. Then, in M_2, we set the corresponding initial state to $\overline{W}(t_0) = B - W(t_0)$. In the following lemma we show that the evolution of the workload in M_1 and M_2 is such that we always have $W(t) + \overline{W}(t) = B$.

Lemma 2 *If $W(t_0) + \overline{W}(t_0) = B$, then we have $W(t^*) + \overline{W}(t^*) = B$ for all $t^* > t_0$.*

Proof: Let $t_{2i-1},\ i = 1, 2, \cdots$, be the i^{th} time in the interval $[t_0, t^*]$ that storage becomes either full or empty and stays persistently in this state until t_{2i} (Figure 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 $\bar{t}_{2i}$.

Since $W(t)$ and $\overline{W}(t)$ admit a derivative at all but countably many points on $[t_0, t^*]$ (this is because W and $\overline{W}$ might not be differentiable at t_i and $\bar{t}_i$ for $i = 1, 2, \cdots$ respectively), a generalized version of the Fundamental Theorem of Calculus allows us to write:

$$W(t^*) - W(t_0) = \int_{t_0}^{t^*} dW(t)$$

$$\overline{W}(t^*) - \overline{W}(t_0) = \int_{t_0}^{t^*} d\overline{W}(t)$$

Using Lemma 1, we obtain $dW(t)/dt = -d\overline{W}(t)/dt$ at any fixed point t that the derivative can be defined since it is assumed that $W(t_0) + \overline{W}(t_0) = B$. Thus, it can be readily seen that $W(t^*)-W(t_0) = -(\overline{W}(t^*)-\overline{W}(t_0))$. Since we have $W(t_0)+\overline{W}(t_0) = B$, we can conclude that $W(t^*)+\overline{W}(t^*) = B$ and the proof is complete. □

Proof of the Equivalence Theorem

It follows from Lemma 2 that there is a one-to-one mapping from trajectories of the M_1 queuing system to trajectories of the M_2 queuing system and that $\forall t, W(t) + \overline{W}(t) = B$. □

In light of the Equivalence Theorem, we postulate that the stationary buffer underflow probability in M_1 can be approximated by the stationary overflow probability in the equivalent fluid queue, M_2. The latter probability has been thoroughly investigated in teletraffic theory.

In the following section, due to lack of space, we only focus on the bufferless case, *i.e.*, we assume that the distribution branch has no storage. Specifically, we investigate ways to compute the overflow probability which corresponds to the transformer overloading probability.

4. TELETRAFFIC ANALYSIS

This section briefly states standard results from the theory of large deviations that enables us to study the asymptotic behaviour of a tail probability of the sum of independent random variables. Specifically, we want to compute approximations for the overflow probability (*i.e.*, the probability that the aggregate load is larger than the transformer capacity, C) under the assumption that the arrivals are Markovian. We validate our use of teletraffic theory in Section 7.

We make the technical assumption 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). Typically, our requirement is that the overflow probability in the original system is less than a desired small value ϵ, which corresponds to the LOLP target of 2.7×10^{-4}.

We can write our requirement as:

$$\log P(Y \geq C) \leq -\beta = \log \epsilon \tag{3}$$

Following Kelly [22], we use Chernoff's bound to obtain an upper bound on (and an approximation for) the overflow probability:

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

where $s \geq 0$ is a free parameter and $\log E[e^{sY}]$ is the logarithm of the moment-generating function of Y. The *effective bandwidth* [22] of a source with the stationary fluid generation rate Y is defined as:

$$\alpha(s) = \frac{1}{s} \log E[e^{sY}] \tag{4}$$

An improved approximation for the loss probability can be derived using the approach of El Walid *et al* [17]:

$$P(Y \geq C) \sim \frac{e^{s^*(\alpha(s^*)-C)}}{s^*(2\pi\sigma^2(s^*))^{\frac{1}{2}}} \quad as \quad C \to \infty \tag{5}$$

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

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

Hence, given the aggregate load Y, C can be computed so that the overflow probability is less than ϵ.

We are also interested in understanding how a mix of loads can impact the sizing of the transformer. Assume that loads belong to N classes where all loads in a class are i.i.d. and loads from different classes are mutually independent. Then, if $\alpha_i(s)$ is the effective bandwidth of a home in class i and n_i is the number of homes in class i, the aggregate effective bandwidth is

$$\alpha(s) = \sum_{i=1}^{N} n_i \alpha_i(s) \tag{6}$$

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), will be:

$$Capacity \;\; region = \{C | \inf_s \{s(\sum_{i=1}^{N} n_i\alpha_i(s) - C)\} \leq -\beta\} \tag{7}$$

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. In the following sections, we size distribution transformers by using this formula and validate our sizing approach using our measurements.

5. MEASUREMENT AND CLASSIFICATION OF HOME LOADS [6]

Obtaining an accurate model for electricity demand of a house is essential in estimating the overflow probability and sizing transformers consequently. In this section we present our testbed and explain how we obtain load measurements. These measurements are used in Sections 6 and 7 to construct load models and perform numerical simulations respectively. We also introduce a home classification scheme developed by an electric utility. This classification allows us to construct different models for different types of houses.

5.1 Obtaining Real Demand Workloads

Our first step is to obtain real measurements of electrical load. In this paper, we focus on residential loads rather than commercial or industrial loads.

Detailed models for residential loads have been presented in the power engineering, environmental studies, and civil engineering literature [30, 8, 10, 28]. However, these models suffer from two problems. First, the data sets on which these models are based are not publicly available. Second, to the best of our knowledge, existing models group all homes into a single class. Our measurements show significant differences in demand behaviour at different homes. Therefore, it would be better to model each class of home differently, which is the approach that we follow in our analysis. This is also the approach followed by electric utilities.

To obtain our own load data set, we built a testbed to measure aggregate loads at 20 homes. We deployed measurement nodes at 19 houses and one home-based small business covering a range of living area sizes, number of occupants, appliances, and energy consumption patterns. For the purpose

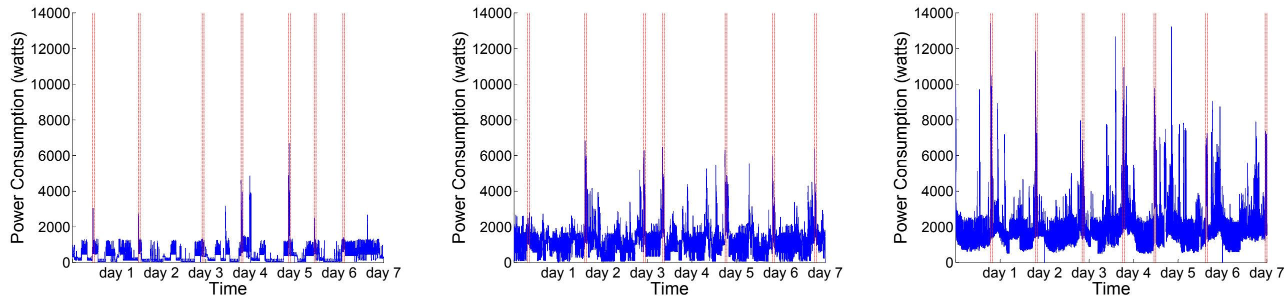


Figure 3: Load measurements from houses in three classes for one week with busy hours marked by vertical lines.

of our small pilot study, we used a convenience sample rather than a stratified random sample. Our methodology generalizes to samples chosen using standard population sampling techniques [13].

Each measurement node consists of a Current Cost Envi device [2] and a netbook. The Envi device measures the power consumption of a house every six seconds and stores it locally in flash memory[6]. 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 are anonymized before being stored in a secure directory on the file server.

Typical loads from three of four types of houses for one week are shown in Figure 3, with the busy hours (defined in Section 6.1) marked with vertical lines.

5.2 Classifying Home Loads

Home electricity loads are highly variable and depend on factors such as the number of occupants, the time of day, the season, mean household income, and the types of appliances commonly in use in the geographical area. Given this variability, choosing a classification for home loads 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 obtained such a parametrization, specifically used for transformer sizing, from a major utility in our area (Table 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 shows the four classes so obtained.

6. LOAD MODELLING AND TELETRAFFIC-BASED SIZING

This section describes our approach to sizing transformers using teletraffic theory. Our overall approach is to construct Markovian reference models of electrical loads (one per class) from the real measurements of the loads. We then compute the effective bandwidth of each Markovian model so obtained and use it in the teletraffic-based sizing of transformers.

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

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

Class	Unit value	Number of houses
1	1.2	8
2	2.5	7
3	3.5	3
4	4.5	2

Table 2: Number of homes in our experiment within each class.

An overview of our approach is shown in Figure 4. We explain the details of this approach in the remainder of this section.

6.1 Assumptions for Teletraffic-based Sizing

Using teletraffic theory to size transformers has several advantages over an empirical approach based on numerical simulations. Applying the theory allows us to readily compute the effect of varying the number of homes, or the proportion of the homes in each class, without having to re-compute or re-measure the aggregate load and run onerous numerical simulations.

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

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 electrical demand during the *busy hour* (defined

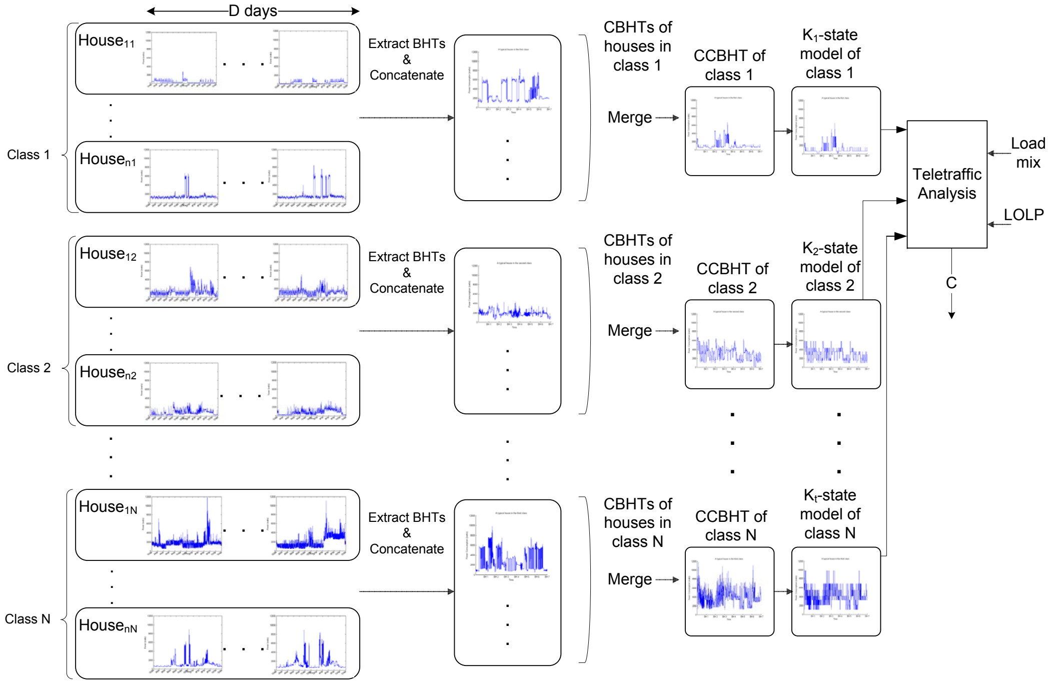


Figure 4: An overview of our teletraffic-based sizing approach.

below) at each home is a conservative upper bound on its demand.

4. The cumulative busy hour trace (CBHT) of a home (defined below) represents the typical busy hour demand of a home.
5. 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, where the value of k is not necessarily the same for all classes. This implicitly assumes that busy hour behaviour is stationary and ergodic.
8. Asymptotic limits can be used even for the fairly small number of homes and CCBHTs in our study.

We 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 using the entire day and therefore provides a sufficient cushion against measurement bias and lack of complete measurement data[7]. Note that it is also a common practice among electric utilities to perform statistical analysis of the load demand using the half-hourly or hourly electrical demand with the highest energy consumption [20].

Our methodology for the teletraffic-based sizing is as follows. First, based on Assumption 3, 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. Usually, the busy hour happens during the peak hours, i.e., 7am-11am and 5pm-9pm during the winter[8]. 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 home for a specific number of days to obtain the cumulative busy hour trace (CBHT) for that period[9]. This represents the typical peak demand of the home according to Assumption 4. Third, based on Assumption 5, we concatenate CBHTs of homes in the same class to get the concatenated CBHT (CCBHT) of that class, which represents its busy hour behavior. We use concatenation to make sure that a CCBHT reproduces the loads of all homes within that class. Figure 5 shows the typical CCBHT of three of four classes in our measurement study. Fourth, as described in Section 6.2 we use Assumption 7 to extract a Markov model for each class. These Markov models are building blocks of the teletraffic sizing algorithm described in Section 6.3.

We jointly validate Assumptions 4-8 by comparing the loss duration predicted by teletraffic analysis to the one computed by numerical simulation in Section 8. Note that the last three assumptions are technical assumptions needed for teletraffic analysis.

6.2 A Markovian Model for the Load Demand

Home loads arise from the superposition of loads from different electrical appliances [30]. Both the literature and our observations suggest that each appliance can be modelled as a multi-level ON-OFF source. An appliance i consumes

[7]Note that using busy hour traces instead of entire 24-hour traces results in smaller pairwise correlations between CCBHTs, in line with Assumption 6.

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

[9]Note that the length of this period is not necessarily equal to the length of load traces used in numerical simulation.

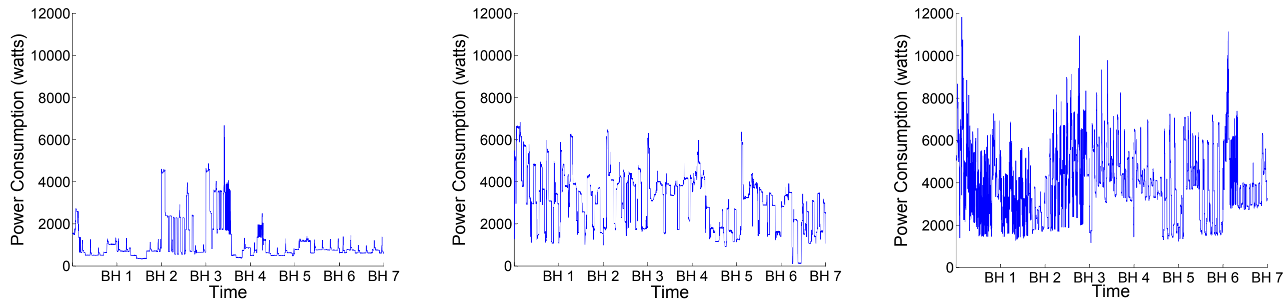


Figure 5: CCBHTs for three classes for one week.

P_i^k watts when in the kth ON state, and 0 when it is OFF. Therefore, it is reasonable to model the power consumption of a home (i.e., a random superposition of different appliances) as a k-state continuous-time Markov process. However, this still leaves the assignment of power levels to the Markov states and choosing the value of k open.

To address these issues, 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 substituting 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. Then we use the goodness-of-fit metric introduced in [6] to find the minimum number of states necessary for representing the home load of a class in a period.

6.3 The Teletraffic-based Sizing Algorithm

Given the set of Markov models, one for each class, using teletraffic theory to compute sizing (*i.e.*, the value of C for a given ϵ) 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 a given value of k 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 state i to 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, the moment generating function of the stationary power consumption of a Markovian source is then

$$M(s) = \sum_i \pi_i e^{sr_i} \tag{8}$$

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 (see Equation 4). El Walid *et al* prove that the effective bandwidth of a Markovian source is the maximal real eigenvalue of the matrix $R_d - \frac{1}{z}Q$, where $R_d = diag(R)$ [18].

Finally, in the fourth step, using Equation (6) we compute the aggregate effective bandwidth of a given load mix[10] and we use the approximation in (5), to find the minimum value of C such that for this load mix, the loss probability is less than a specific value.

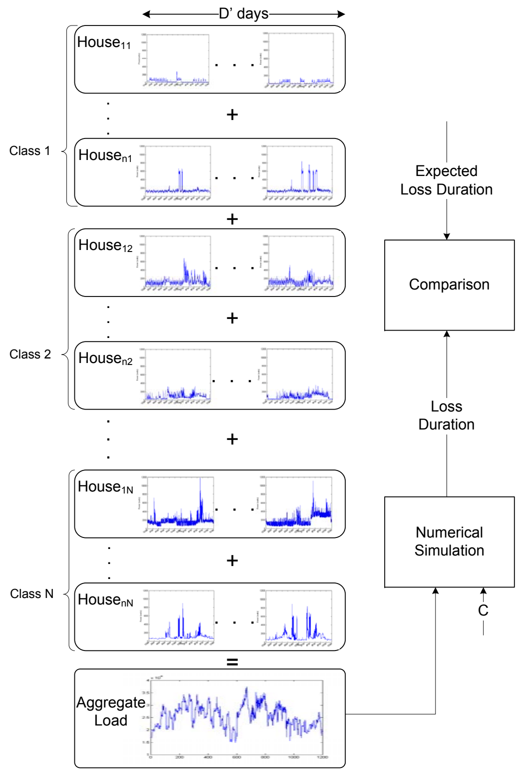


Figure 6: An overview of our validation approach.

To sum up, once we have constructed the Markov models (one per class), we can use these models in teletraffic-based sizing of distribution transformers for any residential neighbourhood in our geographical area. The only thing we need to know about this neighbourhood is its load mix.

7. VALIDATION

We have already described our assumptions about the nature of the electricity demand (including stationarity and pairwise independence) needed to use teletraffic theory. Are the results of teletraffic analysis really applicable to the electrical grid? This section describes our approach to answering

[10] A load mix specifies the number of houses in each class.

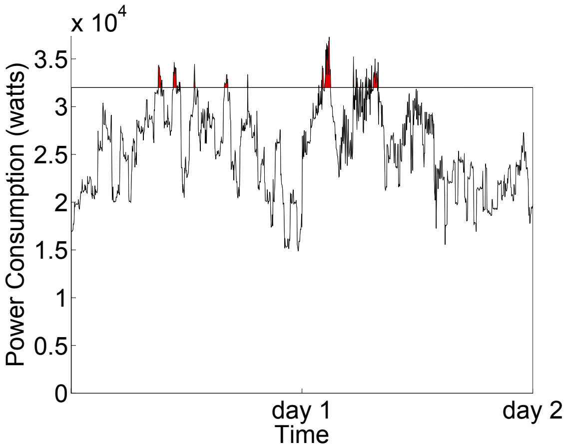


Figure 7: An example of the aggregate workload over a period of two days. The shaded areas above the horizontal line represent the times when demand is larger than the transformer size, that is 32.4 kilo-Watts.

this critical question. An overview of our approach is shown in Figure 6.

Our general approach is to use our load measurements as well as synthetic load traces generated by the simulator developed at the University of Loughborough [30] (described in Section 8) in a numerical simulation to determine the duration of load disruption corresponding to a transformer sizing obtained using teletraffic theory. We then validate our use of teletraffic theory by checking that this duration is indeed lower than the LOLP target.

To find the duration of load disruption we use the following methodology. First, we sum the load traces over a period of time (preferably different from the period used for creating the load models) to find the aggregate power consumption. A typical aggregate load is shown in Figure 7. Second, given the aggregate load, a straightforward numerical simulation suffices to determine the aggregate duration of power outage corresponding to a particular transformer sizing. This simulation compares the aggregated load with the transformer size and records the transformer overloading durations, allowing us to empirically estimate the loss of load probability. The result of this numerical simulation is compared with the LOLP target used in teletraffic-based sizing in Section 8.

8. RESULTS

We present our results in two parts. First, we validate our use of teletraffic theory by comparing the aggregate duration of power outage, for a particular sizing, obtained using numerical simulations and teletraffic analysis. Second, we compare the sizing obtained from our model with what is obtained by using the guideline of a major electricity utility in our geographical area.

8.1 Comparing Results from Numerical Simulation and Teletraffic Theory

We use both teletraffic theory and numerical simulations to compare the expected aggregate duration of power outage for the set of 20 homes in our measurement study. Our teletraffic-based sizing results are from the concatenation of the busy hour traces extracted from a week of measurement.

LOLP	Transformer size (kVA)	Teletraffic theory (Seconds)	Numerical simulation (Seconds)
10^{-3}	103.25	1209.6	0
2.74×10^{-4}	107.27	331.4	0
10^{-5}	116.03	12.1	0
10^{-7}	125.75	0.1	0

Table 3: Loss duration for 14 days of measurements conducted in 20 houses. Teletraffic analysis is based on busy hour traces.

For particular values of LOLP and the load mix of Table 2, we compute the value of C, as shown in Table 3. We convert the transformer size obtained from teletraffic analysis from Watt to Volt-Ampere (VA) by dividing it by the power factor; following convention we set the power factor to 0.9 for this residential neighbourhood [27].

Given the value of C, we use numerical simulation to compute the actual duration of load disruption. The length of load traces used in the numerical simulation is 14 days. This is a fair evaluation because these days do not overlap the days used to extract the model parameters. Table 3 compares the loss of load duration predicted using the two techniques for different values of LOLP. We see that the predictions from theory are an upper bound on the simulation results.

To compensate for the limited duration of our trace and to additionally validate our approach when the number of homes is relatively large, we synthetically generate the electricity demand for 100 homes for 100 days using a 1-minute-grain simulator developed at the University of Loughborough [30]. This electricity demand generator has been shown to closely approximate real domestic demands. To generate this synthetic trace, we choose all homes to have four occupants and a randomly selected mix of appliances. All other values were those set by default including the occupancy pattern. The subsequent modelling and analysis of this data set is 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 modelling. For the industry standard LOLP of 2.74×10^{-4}, the transformer size should be 636.35 kVA according to the teletraffic analysis. The overflow duration determined by the numerical simulation is 0, whereas the teletraffic analysis predicts 39.46 minutes of overflow. We again see that predictions from teletraffic analysis are consistent with (and overestimate) ground truth.

8.2 Comparing Our Sizing with Industry Practice

The transformer sizing rules used by a major utility in our geographical area are shown in Table 4. We now compare the sizing obtained by using our analysis and these rules.

The total unit value of the 20 homes in our study is 46.6. Thus, for the industry standard LOLP of 2.74×10^{-4} the transformer size is 100 kVA. From Table 3, we predict that for the same LOLP with no storage, the transformer size should be 107.27 kVA. This is in excellent agreement with the heuristics used by the utility. This indicates that a careful load modelling based on measurements matches heuris-

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: Transformer sizing rules used by a major utility.

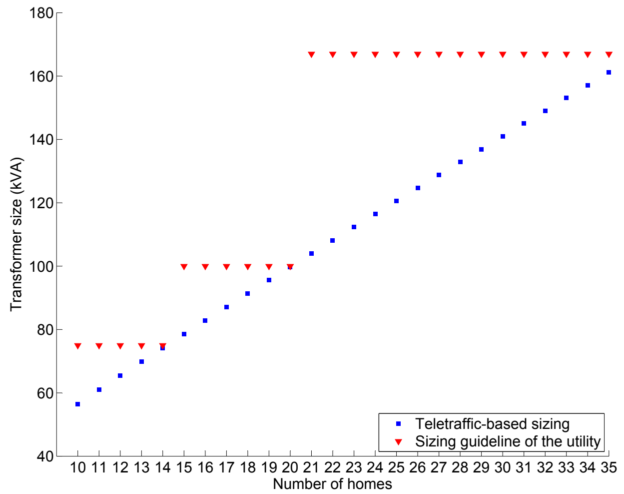


Figure 8: Comparison of the teletraffic-based sizing with the sizing guideline of a utility for a homogeneous load population.

tics developed over decades of field experience, validating our analysis. Note that had the sum of unit values been even slightly larger (greater than 50), the industry heuristic would have advocated a size of 167 kVA, which would have been about 56% greater than what is strictly necessary to meet the LOLP, according to our analysis.

We also compare our proposed sizing with the sizing suggested by the utility for a homogeneous load population. For this purpose, we consider different neighbourhoods comprised of 10 to 35 homes, all from Class 2, *i.e.*, the unit value of each home is 2.5. We use the Markov model constructed for this class in Section 6.2 to compute the aggregate effective bandwidth and the corresponding transformer sizing for different number of homes using the industry standard LOLP. Figure 8 depicts that the sizing obtained by using teletraffic theory closely matches with the utility's sizing rules (Table 4).

9. RELATED WORK

Teletraffic theory has been used extensively in call admission control in ATM networks [24, 18, 17, 12], and in sizing buffers in Internet routers [5, 19]. We used that same approach in this paper and in [7] to size transformers and storage in the grid. However, our work is different from their work in that the requirements of the grid are different from QoS requirements of the Internet. For example, the grid does not care about delay and jitter.

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 when loads are modelled using only their peak values, ignoring temporal variations [4, 16]. Another set of approaches are based on using the estimated load profile of a transformer to model the temperature rise and estimate the transformer loss of life. The load profile that is used in this process is the estimated load of the two peak days in summer and winter [31], the estimated load profile based on long-time measurements of transformer loading such that the probability that the load demand exceeds this load profile is less than a given value [21], or the estimated load using fuzzy load models [11].

Storage can be used both to smooth out variations in demand, as well as variations in supply, especially in the context of variable-rate generation by wind turbines and photovoltaic cells: see Divya and Ostergaard [15] and Deshmukh et al [14] 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. The line of work closest to ours is by Le Boudec and Tomozei [25]. They use min-plus system theory to size the battery, and schedule its operation such that it can be guaranteed that the inflexible load is always satisfied. However, this is different from our work as their model does not include transformers and renewable sources.

10. DISCUSSION AND FUTURE WORK

Our work has made a number of 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, and transformer capacity. Although the conditions under which using asymptotic results are valid are arguably achieved for the 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, although our work was motivated by the need to re-examine transformer sizing guidelines when storage is available and the need to develop a storage sizing guideline to firm up intermittent renewable generation, this paper does not deal with this issue. We believe, however, that other teletraffic models can be used to study grid-connected renewable generation and distributed energy storage 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. Specifically, in this paper we do not deal with the problem of storage sizing in the electrical grid. However, the Equivalence Theorem allows us to compute the storage underflow probability [7] which allows us to jointly size transformers and storage in the distribution network. We also hope to use our approach in the future to answer questions such as:

- 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?
- How much fast-response storage must be installed at distributed generation sites to make intermittent energy sources dispatchable?

11. CONCLUSION

We study multiplexed loads in the electrical grid and revisit the rules for sizing a transformer in a distribution network. Instead of modelling loads by their peak values our work presents a new approach to define design rules for distribution systems. The basis of our work is the Equivalence Theorem, which permits us to apply teletraffic analysis to size transformers, storage, and renewable energy sources in the grid. We validate our approach by using our own measurement data as well as synthetic data. Our results show that our approach is in good agreement both with numerical simulations and industry practice.

12. REFERENCES

[1] IEEE Guide for Loading Mineral-Oil-Immersed Transformers. *IEEE Std C57.91-1995*, 1996.

[2] Current cost. http://www.currentcost.com, January 2011.

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

[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.

[5] G. Appenzeller, I. Keslassy, and N. McKeown. Sizing router buffers. In *ACM SIGCOMM 2004*, pages 281–292, 2004.

[6] O. Ardakanian, S. Keshav, and C. Rosenberg. Markovian models for home electricity consumption. In *Proc. ACM SIGCOMM Green Networking Workshop*, 2011.

[7] O. Ardakanian, S. Keshav, and C. Rosenberg. Sizing the electrical grid. Technical Report CS-2011-18, University of Waterloo, July 2011.

[8] M. Armstrong, M. 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.

[9] R. Broadwater, A. Sargent, A. Yarali, H. Shaalan, and J. Nazarko. Estimating substation peaks from load research data. *Power Delivery, IEEE Transactions on*, 12(1):451 –456, jan 1997.

[10] 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.

[11] R.-F. Chang, R.-C. Leou, and C.-N. Lu. Distribution transformer load modeling using load research data. *Power Delivery, IEEE Transactions on*, 17(2):655 –661, apr 2002.

[12] G. Choudhury, D. Lucantoni, and W. Whitt. Squeezing the most out of atm. *Communications, IEEE Transactions on*, 44(2):203 –217, feb 1996.

[13] W. Cochran. *Sampling techniques*. John Wiley & Sons Inc, 1977.

[14] M. Deshmukh and S. Deshmukh. Modeling of hybrid renewable energy systems. *Renewable and Sustainable Energy Reviews*, 12(1):235–249, 2008.

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

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

[17] A. Elwalid, D. Heyman, T. V. Lakshman, D. Mitra, and A. Weiss. Fundamental bounds and approximations for ATM multiplexers with applications to video teleconferencing. *IEEE JSAC*, 13(6):1004–1016, 1995.

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

[19] M. Enachescu, Y. Ganjali, A. Goel, N. McKeown, and T. Roughgarden. Routers with very small buffers. In *IEEE INFOCOM 2006*, pages 1 –11, april 2006.

[20] R. Herman and J. Kritzinger. The statistical description of grouped domestic electrical load currents. *Electric Power Systems Research*, 27(1):43 – 48, 1993.

[21] J. Jardini, H. Schmidt, C. Tahan, C. De Oliveira, and S. Ahn. Distribution transformer loss of life evaluation: a novel approach based on daily load profiles. *Power Delivery, IEEE Transactions on*, 15(1):361 –366, jan 2000.

[22] F. Kelly. *Notes on Effective Bandwidth*, pages 141–168. Oxford University Press, 1996.

[23] S. Keshav and C. Rosenberg. How internet concepts and technologies can help green and smarten the electrical grid. *SIGCOMM Comput. Commun. Rev.*, 41:109–114, January 2011.

[24] G. Kesidis, J. Walrand, and C.-S. Chang. Effective bandwidths for multiclass markov fluids and other atm sources. *IEEE/ACM ToN*, 1(4):424 –428, aug 1993.

[25] J.-Y. Le Boudec and D.-C. Tomozei. A demand-response calculus with perfect batteries. In *Workshop on Network Calculus (WoNeCa 2012)*, 2012.

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

[27] A. Meier. *Electric Power Systems: A Conceptual Introduction*. Wiley-IEEE Press, 2006.

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

[29] K. Qian, C. Zhou, M. Allan, and Y. Yuan. Modeling of load demand due to ev battery charging in distribution systems. *Power Systems, IEEE Transactions on*, 26(2):802 –810, may 2011.

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

[31] K. Schneider and R. Hoad. Initial transformer sizing for single-phase residential load. *Power Delivery, IEEE Transactions on*, 7(4):2074 –2081, oct 1992.

[32] L. Xie, P. Carvalho, L. Ferreira, J. Liu, B. Krogh, N. Popli, and M. Ilic? and. Wind integration in power systems: Operational challenges and possible solutions. *Proceedings of the IEEE*, 99(1):214 –232, jan. 2011.