Techniques for Enabling Operational Efficiency and Privacy Preservation in the Smart Grid

Merwais Shinwari

A Thesis
In the Department of
Concordia Institute for Information Systems Engineering

Presented in Partial Fulfillment of the Requirements
For the Degree of
Master of Applied Science in Information Systems Security
Concordia University
Montreal, Quebec, Canada

May 2012

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Merwais Shinwari

Entitled: Techniques for Enabling Operational Efficiency and Privacy Preservation in the Smart Grid

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science in Information Systems Security

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

Dr. Abdessamad Ben Hamza Chair

Dr. Tarek Zayed Examiner

Dr. Chadi Assi Examiner

Dr. Amr Youssef Supervisor

Dr. Walaa Hamouda Supervisor

Approved by Dr. Mourad Debbabi

Chair of Department or Graduate Program Director

27 June 2012 Dr. Robin Drew

Date Dean of Faculty

Abstract

The processing and communication capabilities of the smart grid provide a solid foundation for enhancing its efficiency and reliability. These capabilities allow utility companies to adjust their offerings in a way that encourages consumers to reduce their peak hour consumption, resulting in a more efficient electrical grid. The smart grid achieves this through the introduction of smart meters; which collect and transmit consumers' detailed power consumption information in an automated way. Despite their benefits, these readings introduce a major privacy threat to residential consumers as they reveal details that could be used to infer information about the activities of the occupants of a home.

In this thesis, we first propose a method for scheduling a community's power consumption such that it becomes almost flat. Our methodology utilizes distributed schedulers that allocate time slots to soft loads probabilistically based on pre-calculated and pre-distributed demand forecast information. This approach requires no communication or coordination between scheduling nodes and the computation performed at each scheduling node is minimal. Obtaining a relatively constant consumption makes it possible to have a relatively constant billing rate and eliminates operational inefficiencies. We also analyze the fairness of our proposed approach, the

effect of the possible errors in the demand forecast, and the participation incentives for consumers.

In the second part of the thesis, we question the need to disclose high frequency readings produced at the home's level. Instead, we propose equipping smart meters with sufficient processing power enabling them to provide their corresponding utility company with a set of well-defined services based on these readings. For demand side management, we propose the collection of high frequency readings at a higher level in the distribution network, such as at local step-down transformers, as this readily provides the accumulated demand of all homes within a branch. In addition, we study the effect of the proposed approach on consumers' privacy, using correlation and relative entropy as measures. We also study the effect of load balancing on consumers' privacy when using the proposed approach. Finally, we assess the detection of appliances using high frequency readings collected for demand side management purposes.

Acknowledgements

First of all, I would like to thank my supervisors Dr. Amr Youssef and Dr. Walaa Hamouda for their vision and continuous support throughout my master's program.

Furthermore, I would like to express special thanks to my parents, Dr. Mohammad Anwar Shinwari and Mrs. Kamila Shinwari for their continuous love and support. I would also like to thank my immediate family members, including Neptune, Daoud, Mahera, Fareed, Zakia, Jamila, Waleed, Zarghoona, Noor Ahmad, and Tasneem for always being there for me.

Last, but not least, I would like to thank my colleagues at Concordia University, specifically, Abdel Alim Kamal, Aleksandar Kircanski, Ashik Faisal, Esam El-Sheh, Hadi Ayoub, Mohammad Raslan, Osama Hayatle, and Roger Zhano for making my time at the University and in Montreal pleasurable and memorable.

Table of Contents

List of Figures

List of Tables

List of Symbols

$\boldsymbol{T}$	Simulation period of one day, i.e. 24 hours.
$\boldsymbol{D(t)}$	Community demand as a function of time
$\boldsymbol{Pr(t)}$	Price rate as a function of time
$\boldsymbol{t_{min}}$	Time of minimum power demand
$\boldsymbol{L}$	Number of sub channels in a communication channel
$\boldsymbol{\lambda_i^{-1}}$	Noise level of sub channel i
$\boldsymbol{P_T}$	Total transmitted power
$\boldsymbol{x_i}$	Power allocated to sub channel i
$\boldsymbol{\mu}$	The water level
$\boldsymbol{H(t)}$	Community hard demand as a function of time
$\boldsymbol{S(t)}$	Community soft demand as a function of time
$\boldsymbol{SD(t)}$	Scheduling Distribution as a function of time
$\boldsymbol{max}$	The maximum value of the hard demand
$\boldsymbol{C(t)}$	Complement of the hard load to max as a function of time
$\boldsymbol{R}$	A threshold for scheduling soft loads
$\boldsymbol{H_k}$	Total hard load of node k
$\boldsymbol{S_k}$	Total soft load of node k

$\boldsymbol{w_k}$	Scheduling threshold for node k
$\Delta_{\boldsymbol{i}}$	Average change in starting time for a scheduled appliance for node i
Δ	Average Δ_i for all nodes
$\boldsymbol{e}$	Error percentage
$\boldsymbol{S}$	Savings introduced by load scheduling if the community does not shifts loads
$\boldsymbol{O}$	Savings introduced by load scheduling if the community shifts loads
$\boldsymbol{-I}$	Cost of inconvenience associated with load scheduling
$\boldsymbol{D}(\boldsymbol{P}\|\|\boldsymbol{Q})$	Kullback Leibler divergence between two probability distributions P and Q

List of Acronyms

CCP	Critical Condition Plan
ECS	Energy Consumption Scheduling
EISA	Energy Independence and Security Act
ESI	Energy Service Interface
HFR	High Frequency Readings
KLD	Kullback Leibler Divergence
NALM	Nonintrusive Appliance Load Monitoring
NIST	National Institute of Standards and Technology
RTP	Real Time Plan
SD	Scheduling Distribution
TOU	Time of Use

Chapter 1

Introduction

In this chapter, we present a brief outline of our research work. We start with an overview section introducing our topic, and then, we present our motivation for selecting this topic. After that, we present our objectives and contributions. Finally we conclude this chapter by presenting the organization of the remainder of this thesis.

1.1 Overview

The architecture of the power grid in use today follows the design proposed by Nikola Tesla over a century ago [1]. At that time, electricity was considered a luxury that was primarily used for lighting. Today, the power grid is a critical infrastructural component which most, if not all, other components rely upon.

Historically, the electric grid was developed to deliver power generated at remote power plants to consumers. At that time, achieving this objective alone was a great success. Issues such as efficiency, greenness, and reliability were not a priority and addressing them was not feasible.

On the other hand, today's technologies allow much more. The smart grid, or the smarter grid, is about improving the efficiency of the existing power grid by employing state-of-the-art technologies [2]. Step by step, the grid will be enhanced to become an interconnected system of systems. This will allow real-time control of resources, better management and consequently, a more efficient system. The smart grid will enable participating parties to operate more efficiently which, in turn, will contribute to reducing both the financial and environmental costs of power generation and consumption [3] [4].

1.2 Motivation

The distribution network of electric grid is undergoing a transformation which aims to improve the efficiency and reliability of the grid by utilizing state of the art communication and computation capabilities. A main part of this is the introduction of smart meters that measure power consumption in detail and transmit these measurements to the utility company. This offers the utility company a clearer vision into the distribution network and forms a platform upon which many energy conservation technologies can be based [5].

The accommodation of energy conservation technologies has long been a critical governmental objective and the move towards a smarter grid is no exception. The upgrade of the distribution network using smart meters is seen as the fundamental step towards increasing energy efficiency and reducing greenhouse gasses. In the United States, for example, the Energy Independence and Security Act of 2007 (EISA) has mandated the move towards a smarter grid and over $8 billion were made available for this cause [6]. Furthermore, the National Institute of Standards and Technology (NIST) has been assigned the responsibility to coordinate the development of an interoperability framework for the move towards the smart grid [6].

The electric grid is considered the largest engineering achievement and is a major infrastructural component upon which almost all other modern infrastructural components rely. The transformation of the grid will have an effect on every single power consumer and therefore, must be conducted in a sound manner. Furthermore, because of the integration of many grid components using digital communications, the security of the future grid, with respect to consumer's privacy, intrusions and cyber-attacks, has become a critical point that must be addressed.

The enhancement of the electric grid offers a wide range of opportunities for research as many challenging open ended problems have not been adequately addressed. This is mainly due to the uniqueness of the problem with respect to scale and interconnectivity. Furthermore, the incorporation of communication and computation technologies in the distribution network opens new horizons with respect to the development of proactive energy management technologies.

1.3 Objectives

In this thesis, we aim to gain a solid understanding of the smart grid and study the main methodologies proposed to achieve its objectives. Specifically, we focus our efforts in two main areas, namely the use of the communication and processing capabilities of smart meters in order to improve the economy of energy production and consumption, and the effect of collecting power consumption information in high frequency readings on consumers' privacy. We also attempt to identify opportunities and propose more efficient and more practical solutions for the two topics stated above. Furthermore, we investigate the effect of these two topics on each other, i.e., the effect of the efficient operation of the grid on consumers' privacy.

1.4 Contributions

Throughout our work, we have made a set of contributions to the topic in consideration. These include the development of a load balancing algorithm for schedulable loads (also called soft loads), the proposal of a privacy enhancing architecture for smart metering and the assessment of the efficient operation of the grid on consumers' privacy. These are introduced below and explained in details in chapters 3 and 4.

Our proposed load balancing algorithm is a centralized water-filling [7] based approach that allows consumers to participate in the load balancing process by scheduling their loads following a statistical distribution. In particular, our proposed algorithm enables consumers to schedule their soft-loads (i.e. loads that do not have operation time constraints) in an efficient way. The algorithm aims to flatten the overall consumption profile of the community resulting in a lower peak-to-average ratio of power demand throughout the day. This reduces the difference in power demand between peak and off-peak hours, facilitating more efficient operation of the grid. Furthermore, the effectiveness of our proposed approach is confirmed through simulations. Different operational aspects of the proposed approach, such as fairness, the effect of possible errors in the demand forecast, and the participation incentives for consumers, are also analyzed.

In the second part of this work, we propose a privacy preserving approach to smart metering. This approach allows utility companies to achieve the objectives of smart metering without requiring the disclosure of consumers' high frequency readings. The approach capitalizes on the communication and processing capabilities of smart meters and is based on transforming meters from data acquisition devices to service providing modules. For demand side management, our approach requires the collection of high frequency readings at a higher level in the distribution network. We also developed a simulation environment to assess the impact of high frequency readings produced for demand side management on the privacy of individual consumers.

Furthermore, using the approaches presented above, we assess the effect of the efficient operation of the grid on consumers' privacy. To do this, we use a model for load balancing with the simulation environment developed for privacy enhancement evaluation.

1.5 Thesis Organization

The remaining of this thesis is organized as follows. The next chapter presents the main preliminaries relevant to the topic at hand. This includes a set of concepts specific to the electric grid that are required to gain a proper understanding of the electric grid, the smart grid, and consequently, this thesis. Chapter 3 presents our proposed load balancing algorithm. This starts by explaining the problem in question, and then presents the proposed solution. After this, we present our simulation results and analysis. Chapter 4 presents our proposed privacy enhancing approach to smart metering. Similarly, the chapter starts by presenting the problem to be addressed. Following that, the proposed solution, the simulation and the analysis of the results are presented. In addition, the effect of using load balancing techniques on privacy is also illustrated throughout this chapter. Finally, in chapter 5 we present our conclusion and possible areas for future work.

Chapter 2

Preliminaries

In this chapter, we present some preliminaries relevant to our work. This includes a set of fundamentals about the existing electric grid, the power consumption profile of a community, and the smart grid.

2.1 The Classical Grid

The electrical grid, called here the classical grid, is a term that refers to the electricity production and distribution infrastructure as a whole [1]. Structurally, the grid is composed of three main components. These are bulk generation plants, transmission networks, and distribution networks. As the name indicates, bulk generation plants produce power in large quantities, typically in the order of hundreds of megawatts. The most common types of these are hydro-electric plants, coal plants and nuclear plants. These are often located in remote areas and the power produced is transmitted to consumption areas through transmission networks. Transmission networks are categorized as high voltage networks, typically in the order of hundreds of kilovolts, and transmit power in bulk over long distances. At substations, the voltage is stepped-down to a distribution network level. Distribution networks are categorized as low voltage networks, typically in the order of tens of kilovolts. At the consumption location, the voltage is stepped-down again to the service voltage required and a meter is installed to measures the amount of power consumed for billing purposes [8].

2.2 The Smart Grid

Although not strictly defined, the smart grid is abroad term used to describe the enhancement of the distribution network of the electrical grid using state of the art technologies [9] [10]. The move towards a smarter grid is expected to enhance the efficiency of energy production and consumption, and the availability and reliability of the service. Furthermore, if consumers participate in energy production, for example using wind turbines or solar panels, the smart grid would enable them to sell any excess generated electricity to the grid.

The power grid can be made smarter in many ways, specifically by improving its efficiency and reliability, supporting distributed generation and storage, and facilitating consumption management for consumers. The move towards a smarter power grid is driven by many factors such as the high fuel prices, increasing energy demands, increasing environmental concerns, and the emergence of electric vehicles.

The move towards a smarter grid requires the development of a platform that enables all participating parties to exchange information and consumption trends and allows them to act upon this information to their greatest interest. In addition to communication networks, and processing facilities, a noticeable component in this enhancement is the introduction of smart meters at consumers' ends. As will be addressed in more details below, this is the major component that enables enhancing the functionality of the grid and the realization of the advantages stated earlier [11].

The National Institute of Standards and Technology organizes the smart grid into a framework of seven major components [6]. These are *bulk generation*, *transmission*, *distribution*, *customer*, *operations*, *markets*, and *service providers*. Similar to the classical grid, *bulk generation* covers the production of energy in bulk from various resources and *transmission* addresses transporting the produced energy to the distribution component. Unlike the classical grid, in the smart grid the transmission component also accommodates the possibility of energy generation and storage. The *distribution* component allows the delivery of energy to consumers and also accommodates generation and storage. The *customer* is mainly the energy consumer and is assumed the ability to generate and store electricity. Furthermore, the *customer* is assumed to have a smart meter, an Energy Service Interface (ESI), and smart appliances that consume power taking into consideration information made available by the smart grid. *Operations* interact with all components of the framework and monitor and control most aspects of the smart grid. The *Market* component provides market management and coordinates other components from a commercial perspective in order to ensure competitiveness. Finally, *service provider* conducts billing, addresses emergency issues, installation and maintenance of various grid components [6].

2.3 Community Demand Profile

The power demand of a community varies throughout the day in response to its members' activities. Figure 2-1 shows the power consumption for a community on a typical summer weekday [12]. As shown, the consumption ranges between 13500 MW at about 4AM and 21000 MW at about 6PM with the difference being about 65%. The consumption is lowest in the early hours of the morning, as most people are asleep. Then, it increases throughout the day peaking at about 6PM when people return home from work and add their home loads to the grid [12].

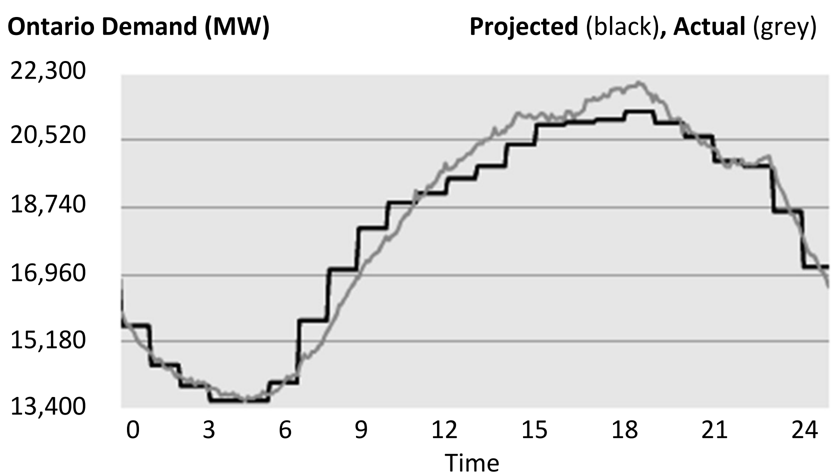

Figure 2-1: Typical daily power demand [12].

Due to the varying demand, utility companies purchase power from producers at varying rates throughout the day. Therefore, power generation will be most expensive during peak hours and less expensive at other times. This happens mainly because of the operational inefficiencies as electricity generators needed to support peak demand hours become idle during off-peak hours. However, to the consumer, the tariff is typically based on an overall average, and therefore, consumers do not have any financial incentive in consuming power with any pattern for a better bulk purchase rate.

Furthermore, the figure shows that it is possible to forecast a community's power consumption with a fair accuracy. In the figure, the black curve shows the forecasted consumption and the grey one shows the actual (measured) consumption. Typical demand forecasting performance indicators that are available in the public domain (e.g., see [13]) show that when projection is made a day ahead, the error is usually within ±10%.

2.4 The Smart Meter

Smart meters are the main components that enable achieving the objectives of the smart grid. Unlike traditional meters that are read manually once a month, smart meters collect and transmit power consumption information in an automated way. This is normally in the form of High Frequency Readings (HFR) that provide real time information on how much power is consumed and at which time [14].

The high frequency readings produced by smart meters enable utility companies to produce bills for their clients following schemes that take into consideration the time of power consumption. Since the bulk purchase rate of power is not constant throughout the day, such schemes enable transferring part of the variance in the rate to the consumer. This, in turn, is expected to encourage consumers to adjust their power consumption for more favorable bills resulting in a more favorable bulk purchase rate, and a more efficient overall power system [14].

Obtaining real time readings from various points in the distribution network provides utility companies with a clearer vision of the network. This facilitates automated fault detection and analysis resulting in a much more reliable and available power grid. Furthermore, in comparison to consumer initiated notifications, the automation of fault detection is expected to reduce the response time to blackouts and brownouts [15].

2.5 Billing Schemes

In the classical grid, power consumption is billed based on the overall cumulative consumption throughout the billing interval, normally a month. Consumers are billed based on a flat rate that is the average production rate. The smart grid enables the use of new plans that encourage consumers to better manage their consumption. The most common of these are the Time of Use plans (TOU), the Critical Condition Plans (CCP), and the Real Time Plans (RTP). TOU plans divide the day into segments with a fixed tariff each. Typically, peak hours are most expensive, off-peak-hours are less expensive, and mid-peak hours are in between. CPPs offer power at a specific rate during normal operation conditions and at a significantly higher rate upon the occurrence of some critical condition, for example, in cold countries during the winter, when the temperature falls below a certain degree. The critical condition is normally communicated to consumers as it occurs and consumers are expected to adjust their consumption accordingly. In RTPs, a fixed tariff is not defined; rather, the rate varies according to the real time supply and demand [14].

Chapter 3

A Load Scheduling Algorithm for the Smart Grid

In this chapter, we present the first of our main contributions to the topic of this thesis. Specifically, we propose a method for scheduling a community's power consumption such that it becomes almost flat. Our methodology utilizes distributed schedulers that allocate time slots to soft loads probabilistically based on pre-calculated and pre-distributed demand forecast information. This approach requires no communication or coordination between scheduling nodes. Furthermore, the computation performed at each scheduling node is minimal. Obtaining a relatively constant consumption makes it possible to have a relatively constant billing rate and eliminates operational inefficiencies. We also analyze the fairness of our proposed approach, the effect of possible errors in the demand forecast, and the participation incentives for consumers.

3.1 The Need for Operational Efficiency

One area of inefficiency arises from the trend of the daily power consumption of a community. As shown in Figure 2-1, the power demand for a community varies throughout the day by about 65%. With the current architecture of the grid, electric power must be consumed the instant it is generated. Because of this, power plants must have the capacity to support peak hour demands. This additional capacity becomes idle during low demand hours which results in operational inefficiency. Furthermore, at peak hours, high demand causes utility companies to purchase power at higher rates. On the other hand, power is typically sold to consumers at a fixed rate. With such a setup, the time at which power is consumed is of no significance to the consumer; and consumers have no incentive in adjusting their power consumption for a better bulk purchase rate.

Utilizing the capabilities of a smarter grid, utility companies are moving towards billing for power on a time of use basis. For example, *Hydro One*, which delivers electricity across the Canadian province of Ontario, is introducing a pricing scheme where each day is divided into on-peak hours, mid-peak hours and off-peak hours. In this scheme, power rates are most expensive during peak hours and less expensive during off-peak hours [12].

3.2 An Overview of Proposed Solutions

Caron and Kesidis [16] proposed a dynamic pricing scheme that encourages consumers to adjust their power use with the objective of getting a flat overall consumption. The authors showed that finding an optimum schedule is NP-hard, and then presented methodologies to study how close one can get to the ideal case. This was done based on the amount of information the consumers are willing to share with their utility company. Furthermore, the authors studied the outcomes based on several scheduling policies, namely, uniform, ALOHA I, ALOHA II, and Time/Slackness. In their work, the authors also compared the performance of these scheduling policies.

Xiong *et al.* [17] proposed an approach based on communication protocols that reduce the power demand in an attempt to produce a uniform power consumption over time. Similar to other approaches, the authors divided power consumption into real-time loads (non-schedulable) and schedulable loads. They also defined a "target" power level and modeled their algorithm to schedule power use such that this target is not exceeded. The algorithm uses a specific structure that consists of three main phases, namely power update phase, power request phase, and power scheduling phase. Furthermore, the authors simulated their approach and showed that it is possible for a consumer to keep their power demand below the defined target.

Gatsis and Giannakis [18] presented a cooperative scheduling approach between the utility company and the consumers. In this approach, loads are classified into loads that must run, loads that must consume a given amount of power (e.g., a re-chargeable battery), and loads that are adjustable in power consumption but the adjustment could cause consumer dissatisfaction (e.g., climate control). This was modeled into a convex optimization problem that was solved using the distributed subgradient method. The authors also presented simulation results that show that it is possible to meet the constraints above in an optimum way.

Chen *et al.* [19] proposed the use of a real-time pricing rate, and formed a Stackelberg game [20] between the utility company and energy management controllers that are to be deployed at each home. The game was setup such that the controllers play the role of the follower and the utility company plays the role of the leader. The authors simulated their proposed methodology and concluded that their approach saves money for the consumers and ensures that rebound peaks do not appear.

Mohsenian-Rad *et al.* [21] presented an autonomous incentive based algorithm for scheduling power consumption. In this scheme, loads are classified into soft (or schedulable) consumption and hard (or non-schedulable) consumption. Soft consumption represents usages that do not have strict time constraints, and hard consumption represents usages that have strict time constraints. The authors also proposed the use of Energy Consumption Scheduling devices (ECS) as a component of smart meters. In this model, an ECS communicates with other ECSs in its neighborhood sharing its scheduling information. Running their proposed distributed algorithm, each ECS computes and broadcasts its optimal schedule. The algorithm repeats until no ECS announces any change of schedule.

The approaches presented above do flatten the overall demand of a community. However, the need to continuously update other nodes with scheduling information and to solve optimization problems poses a great overhead in communications and processing. Furthermore, sharing detailed scheduling information with other consumers presents a major privacy concern [22]. In what follows, we propose a simple heuristic scheduling method that eliminates these requirements. In our proposed method, consumer nodes do not need to share detailed information with each other and do not perform any complex processing. Rather, we utilize statistical information on consumption trends that utility companies could easily make available to their subscribers.

3.3 Proposed Algorithm

In this section, we present our proposed algorithm in details. We first present the effect of scheduling power use, and establish its impact on pricing. We then introduce water-filling, the concept upon which our algorithm is based. After that, we present the logic of our algorithm and the computation of a threshold needed by the algorithm. Finally, we conclude this section by presenting the algorithm in its holistic form.

3.3.1 Scheduling Power Use

If consumers are billed at a constant rate, they will have no incentive to consume power with any pattern. Let $D(t)$ denote the power demand of a community throughout a time period T. For our purpose, we define T to be the duration of one day, i.e., 24 hours. With power consumed at will, $D(t)$ gets a shape as shown in Figure 2-1. Let $Pr(t)$ denote the corresponding price rate as a function of time. Furthermore, let $Pr(t)$ be a function of the instantaneous demand. Therefore, we can define $Pr(t)$ as

$$Pr(t) = F\big(D(t)\big) \tag{3.1}$$

where F is a function that relates price to demand.

With the communication capabilities of the smart grid, distributing real-time pricing information can be easily achieved. If power is priced as a real-time function of demand, it becomes possible for a community to influence the power price rate $Pr(t)$ by adjusting its consumption throughout the day. As shown in [16] and [21], a coordinating community that shifts part of its consumption to off-peak hours can significantly reduce its power bill while consuming the same amount of power.

If power is priced as a function of consumption, and the consumption trend has some minimum value min that occurs at time t_{min}, $Pr(t)$ will also have a minimum value at time t_{min}. Therefore, if consumers are able to shift their loads, possibly by scheduling, it would appear that the most cost effective approach is to target t_{min}. However, as the number of consumers who follow this reasoning increase, the time t_{min} becomes a high demand point and consequently a highly priced point.

Thus, a fair methodology is needed to allocate low-rated hours to consumers without creating new points of high demand. A perfect situation would be one where the demand function becomes constant which also leads to a constant price function. In the following sections, we propose a water-filling based method for scheduling power consumption such that the overall demand becomes relatively constant. We also present simulation examples to show the effectiveness of the proposed methodology.

3.3.2 An Overview of Water-Filling

The classical water-filling algorithm, traditionally used in communications theory [23], solves the problem of maximizing the mutual information between the input and the output of a communication channel that is composed of several sub-channels and that is subject to a global power constraint. In this section, we first present the mathematical model for the channel power allocation problem, its solution (the water-filling result), and explain its implication. We then explain the analogy between the channel power allocation problem and the load scheduling problem illustrating how the water-filling result will be used to build a scheduler for our purpose.

Given a communication channel with multiple sub-channels (*1, 2, 3 ... L*) that are subject to noise levels (λ_1^{-1}, λ_2^{-1}, λ_3^{-1} ... λ_L^{-1}), it is desired to transmit a signal such that the channel capacity (i.e., the information rate) is maximized. The signal to be transmitted is subject to the global power constraint $\sum_{i=1}^{L} x_i \leq P_T$, where x_i is the power allocated to sub-channel i, and P_T is the total power to be transmitted. This problem can be modeled as the following optimization problem

$$\max_{\{x_i\}} \sum_{i=1}^{L} log(1 + x_i \lambda_i) \tag{3.2}$$

subject to

$$\sum_{i=1}^{L} x_i \leq P_T, \tag{3.3}$$

$$x_i \geq 0, \text{and } 1 \leq i \leq L.$$

Using the Lagrange multipliers method, the solution of this problem is given by

$$x_i = (\mu - \lambda_i^{-1})^+ \tag{3.4}$$

where $(x)^+ \overset{\text{def}}{=} max(0, x)$ and μ is a constant chosen to satisfy the power constraint in (3.3).

Equation (3.4) is known as the water-filling result and it implies that for an optimum solution, more transmission power should be allocated to sub-channels with less noise. The constant μ (called the water-level) is selected such that power constraint is satisfied. As shown in Figure 3-1, this capacity achieving solution has the visual interpretation of pouring water over the curve given by the sub-channel noise (i.e. the inverse of the sub-channel gains), and hence the name water-filling or water pouring [24] [23] [25].

Applying the water-filling result for centralized applications where power allocation is performed by a single entity is a straightforward application of (3.4) once the water-level is known [26].

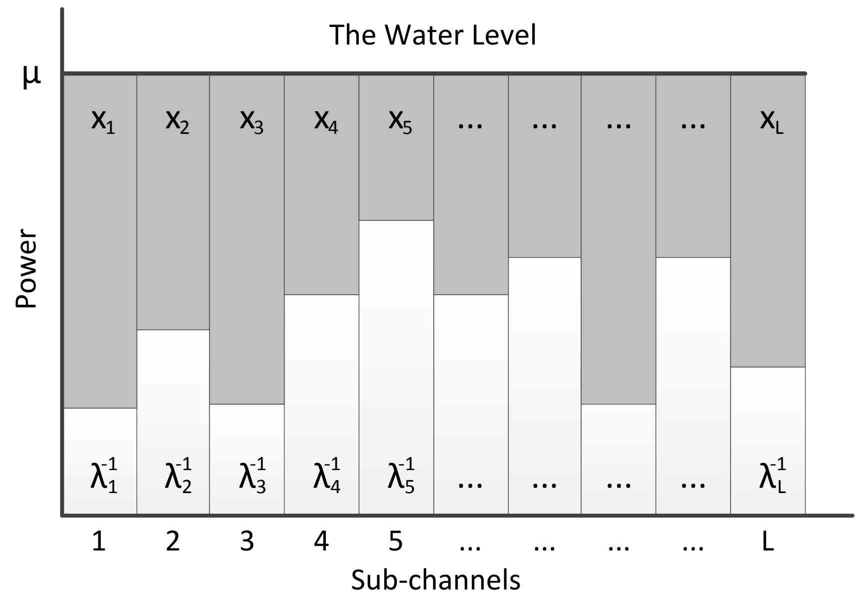

Figure 3-1: Water-filling graphical illustration.

The water-filling approach naturally fits the load balancing problem in consideration. Specifically, the hard loads are analogues to the noise inherent in the channel, the soft loads are analogues to the power that is to be transmitted through the channel, and the objective is to schedule soft loads such that the overall energy consumption becomes flat which yields an efficient system. Thus, similar to the classical capacity maximizing problem above, where more transmission power is allocated to sub-channels with less noise, soft loads can be allocated to the time slots with less hard loads. This reduces the peak-to-average power ratio and, in the event that sufficient soft loads are available, flattens the overall energy consumption profile. As will be explained in section 3.3.3, the main difference between our proposed load balancing approach and the traditional water-filling solution above is that our schedulers allocate soft loads probabilistically following a distribution produced from the water-filling result. This enables our protocol to eliminate the need for continuous communication and synchronization between scheduling nodes. As depicted in Figure 3-2, our proposed algorithm can be summarized as follows:

- *Demand Forecast:* This is the process where the utility company produces a forecast for the community's power demand.
- *Water Level Computation*: This process computes the correct water-level for use with the forecasted demand profile. In our case, as will be explained in the next section, the water-level is set to the lowest point where a constant demand can be achieved.
- *Optimal Allocation Computation*: Once the water-level and the demand forecast are known, the optimal allocation of soft loads can be obtained as a direct application of (3.4).
- *Optimum Allocation Metrics Distribution*: The optimal allocation produced in the previous process is applicable for a central scheduling entity. In our case, this process produces the probabilistic metrics that enable the distributed schedulers to attain the optimum distribution.

- *Scheduling Algorithm Execution*: The scheduling algorithm, presented in detail in section 3.3.5, uses the metrics distributed through the previous process to achieve an overall flat power consumption profile in a distributed way.

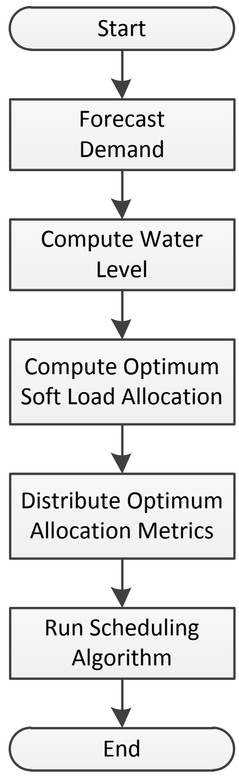

Figure 3-2: Basic steps of proposed algorithm.

3.3.3 Proposed Algorithm

As mentioned previously, a method is needed to allocate consumers' soft consumption to the low priced time slots in a fair way. The allocation strategy has to satisfy the following conditions:

1. Consumers that follow the algorithm pay the same amount for a shifted unit of power regardless of the time slots to which their loads are allocated.

2. It is not possible for a consumer to have a more cost effective allocation.

Both of these conditions will be satisfied if it is possible to achieve and maintain a constant demand. To realize this, we propose allocating soft loads to time slots probabilistically in a way analogous to the water-filling approach described previously. That is, more loads are probabilistically allocated to low demand hours, and less loads are probabilistically allocated to high demand hours. Once an overall constant is achieved, the flat consumption profile is maintained by allocating any additional soft loads uniformly.

Let $H(t)$ denote a function that models the hard power consumption for a community, and let its maximum value be max. Furthermore, let $S(t)$ denote a function that models the soft power consumption of the community. Thus we have

$$D(t) = H(t) + S(t). \tag{3.5}$$

Our objective is to distributively schedule soft loads such that the overall power demand for the community becomes a constant. Therefore, after scheduling, the following equation must be satisfied

$$H(t) + S(t) = D(t) = Constant. \tag{3.6}$$

We then evaluate a Scheduling Distribution (SD) that, when followed, complements the hard loads' consumption such that the total consumption becomes a constant. Similar to water-filling, we relate SD to the difference between the hard loads and the value max. Let $C(t)$ denote a function that complements $H(t)$ to the value max. Thus we have

$$C(t) = max - H(t). \tag{3.7}$$

Having $C(t)$, we define SD as a probability distribution that produces $C(t)$ when followed randomly for a sufficiently large number of times. We can obtain SD by normalizing $C(t)$ such that the area under its curve is 1. Thus we have

$$SD(t) = \frac{C(t)}{\int_0^T C(t)dt}. \tag{3.8}$$

Equation (3.5) shows that the demand $D(t)$ is a sum of both, the hard and the soft power consumption $S(t)$ and $H(t)$, respectively. Therefore, strictly following (3.8), the shape of $D(t)$ will either be a constant, or inclined towards $H(t)$ or $S(t)$. This depends on the ratio of the total hard consumption to the total soft consumption. We summarize the possible outcomes as follows:

1. If the amount of soft consumption is too small compared to the hard consumption, scheduling following SD will have a minimum effect since $H(t)$ will dominate. Therefore, $D(t)$ will have the shape of $H(t)$ rather than a constant. In this case, scheduling can reduce the peak-to-average ratio of the consumption profile, but cannot flatten it.
2. If the amount of soft consumption is too large with respect to hard consumption, scheduling following SD will cause the scheduled loads to dominate, which results in $D(t)$ following the shape of $C(t)$ rather than a constant (i.e. causing an overshot).
3. If the ratio of the soft consumption to the hard consumption equals the ratio of the area under $C(t)$ to the area under $H(t)$, scheduling following the distribution stated will complement $H(t)$ causing $D(t)$ to become flat.

With the increased demand for electric vehicles, which are a form of energy storage and can usually be scheduled, it is expected that the energy consumption of an average home will double [27]. This suggests that having a large amount of soft consumption is a reasonable assumption which eliminates the first outcome above.

Having a large amount of soft consumption, it becomes possible to adjust the second outcome to result in an overall consumption that is constant. This can be done by having each node schedule its soft loads partially following SD and partially following a uniform distribution. More precisely, SD should be followed to a point where it complements the hard consumption to the water-level. After this point, to maintain the overall constant demand, nodes should schedule any remaining soft loads uniformly. For each node, a threshold will define the point of switching from scheduling following SD to scheduling uniformly. This threshold is computed as a factor R of the node's hard consumption. The value of R is to be provided by the utility company and its computation is presented in section 3.3.4.

Therefore, for a given node k with a total hard consumption, h_k, and a total soft consumption, s_k, the threshold is given by w_k where

$$w_k = R \times h_k. \tag{3.9}$$

specifically,

$$w_k = \frac{\int_0^T C(t)dt}{\int_0^T H(t)dt} \times h_k. \tag{3.10}$$

Therefore, each node will schedule an amount, w_k, of its soft loads following SD, and any remaining soft loads, $(s_k - w_k)$, uniformly. Scheduling loads as described above results in an overall flat consumption.

3.3.4 Computation of Threshold

In this section, we describe the method used in computing the ratio needed to calculate the threshold in (3.9). Assuming we have a community with n nodes, let $S_k(t)$ represent the soft power consumption of the k^{th} node as a function of time. Similarly, let $H_k(t)$ represent the hard power consumption for that node. Therefore, the overall soft and the overall hard consumption for all members of the community can be respectively represented by

$$S(t) = \sum_{k=1}^{n} S_k(t). \tag{3.11}$$

and

$$H(t) = \sum_{k=1}^{n} H_k(t). \tag{3.12}$$

For the hard demand curve $H(t)$, its maximum value max is the lowest value that can be used as a water-level that produces a constant demand. This is illustrated in Figure 3-3. Therefore, based on water-filling, we define $C(t)$ as the function that complements $H(t)$ to max, i.e.,

$$C(t) = max - H(t). \tag{3.13}$$

We define the factor R as the ratio of the area under $C(t)$ to the area under $H(t)$. Thus we have

$$R = \frac{\int_0^T C(t)dt}{\int_0^T H(t)dt}. \tag{3.14}$$

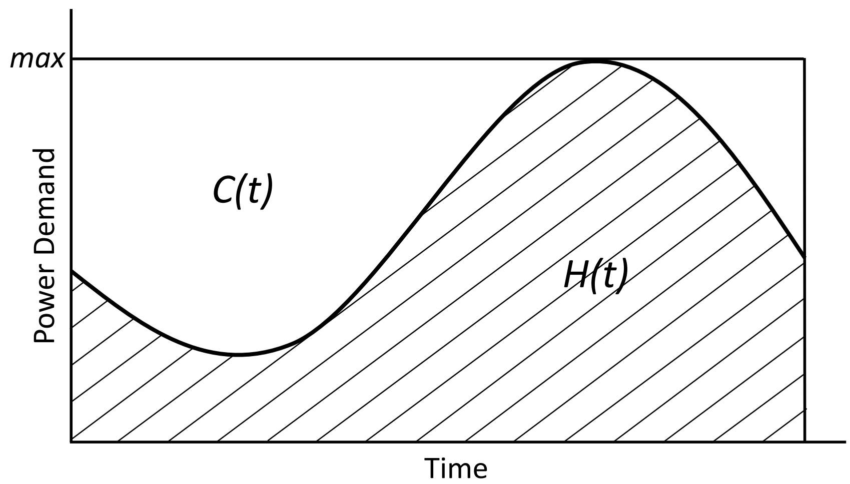

Figure 3-3: Water-level selection illustration.

From (3.12), we see that $H(t)$ was formed by summing each user's hard component. We propose that $C(t)$ be formed in a similar way. Assuming that consumers have a sufficient amount of soft loads, each will dedicate a portion to produce $C(t)$. That is

$$H(t) = \sum_{k=1}^{n} H_k(t). \tag{3.15}$$

$$R \int_0^T H(t)\,dt = R \sum_{k=1}^{n} \int_0^T H_k(t)\,dt. \tag{3.16}$$

$$\frac{\int_0^T C(t)dt}{\int_0^T H(t)dt} \times \int_0^T H(t)\,dt = \int_0^T C(t)\,dt. \tag{3.17}$$

$$\int_0^T C(t)\,dt = \sum_{k=1}^{n} R \int_0^T H_k(t)\,dt. \tag{3.18}$$

Therefore, the amount of soft loads needed to form $C(t)$ can be obtained if each consumer contributes with an amount of soft loads equivalent to a factor R of their hard consumption. Scheduling these soft loads probabilistically following $SD,$ as defined in (3.8) produces $C(t)$, which in turn complements $H(t)$ to the constant max.

3.3.5 Proposed Distributed Algorithm

In this section, we present an algorithm for scheduling a consumer's soft loads throughout the day to achieve the objectives described previously. Our algorithm is based on a probabilistic model that uses forecasted approximations (or projections) of consumption trends rather than real-time computed values. In this section, we describe the environment, the information available to each consumer (scheduling node), and the proposed algorithm. In the following section, we present a simulation of this algorithm for a small residential community.

The main advantage in having participating nodes schedule loads probabilistically following a centrally computed distribution is to eliminate the need for continuous communication and synchronization between nodes. Furthermore, this also reduces the processing required at each node. This approach is possible provided that some basic information related to the overall usage trend can be made available to each scheduling node.

Obtaining perfectly accurate values for $C(t)$ would require gathering and processing the community's demand in real-time. This would cause a large overhead in communications and would have to be made available upfront. Fortunately, a community's overall consumption follows a trend, which, as shown in Figure 2-1, is fairly predictable [12] [13]. Using a projection may introduce a margin of error (see section 3.5.2), but it eliminates the need of gathering and distributing consumer's data in real-time. Furthermore, in practice, loads to be scheduled may not be equal in power demand or in the duration of operation. Therefore, scheduling loads to obtain a perfect complementing function $C(t)$ may not be possible as this would require loads to be infinitely small and have no operation time restrictions whatsoever. Despite these limitations, as shown by our simulations, scheduling following the proposed method significantly reduces the peak-to-average power ratio of the overall demand. This is due to the fact that the power consumption of individual appliances are significantly small in comparison to the overall power demand of a community.

In what follows we assume that the following information can be made available to each scheduling node:

1. A distribution (SD) that, when followed, complements the trend of the hard consumption towards a constant.
2. A ratio R for computing the threshold that indicates how much of the soft load should be scheduled following SD.

The algorithm below reflects the methodology discussed in section 3.3.3. At each scheduling node, the initialization phase acquires the centrally computed scheduling information. At this point, an accumulator is defined to keep track of how much load had already been scheduled. If the total amount of scheduled loads is less than the threshold, the next load will be scheduled following SD and the accumulator will be incremented by the weight of this load. When the threshold is exceeded, the next item to be scheduled will be scheduled uniformly across all time slots.

Scheduling Algorithm

```
Initialization
    Acquire SD
    Acquire Threshold
    Reset Accumulator
For each Soft Load, weighted W
    If (Accumulator < Threshold)
        Schedule Soft Load following SD
        Accumulator = Accumulator + W
    Else
        Schedule Soft Load uniformly
        Accumulator = Accumulator + W
    End If
End For
```

Algorithm 3-1: Scheduling soft loads to flatten the overall power consumption profile.

3.4 Algorithm Assessment

We simulate the algorithm presented previously to assess its effectiveness. In this section, we present the simulation environment developed, followed by the simulation results.

3.4.1 Simulation Environment

To demonstrate the effectiveness of our proposed algorithm, we simulate the basic power consumption of a residential community. Initially, we simulate the normal use pattern where each home consumes power without scheduling any loads. These initial rounds resemble historic data used in demand forecasting by the utility company. Next, we use the simulated data to compute SD as described in section 3.3.3. Finally, we simulate the community's power consumption when following the proposed algorithm and present our results.

To do this, we first define a selection of appliances available in most homes and assign them their typical power ratings. Table 3-1 shows some of the basic attributes of these appliances. We approximate the operation time for each appliance in the second column. The third column shows whether this load can be shifted or not based on its nature of use, i.e., if it is a soft load. The last column shows if the time slots have to be allocated continuously. That is, if the nature of the appliance dictates that all the required slots must be scheduled back to back. This is only considered for schedulable appliances.

In our simulation, we divide the scheduling interval into 30-minute time slots and the appliances are turned on for the duration of a multiple of these slots. Each appliance is turned on probabilistically at a time slot either following a uniform or a normal distribution. The details of the distributions used for each appliance are shown in Table 3-2.

Appliances that are modeled by a uniform distribution are constrained with a time frame. That is, they can only operate during that time frame. Furthermore, appliances that are modeled by a normal distribution are constrained with a mean value. Therefore, most of the time, these appliances will be turned on at the time that corresponds to the mean. If an appliance is to be simulated for two operating periods, we list it twice. For example, we assume that kitchen appliances are used twice a day and, consequently, they have two entries with two different mean values.

Table 3-1: Simulated household appliances.

Appliance	*Operation Time*	*Soft Load*	*Continuous Use*
Clothes Dryer	1 hour	Yes	Yes
Electric Vehicle	5 hour	Yes	No
Clothes Washer	0.5 hour	Yes	Yes
Climate Control	5 hour	No	N/A
Water Heater	6 hour	No	N/A
Range (1st run)	1 hour	No	N/A
Range (2nd run)	1 hour	No	N/A
Electronics	5 hour	No	N/A
Lighting	6 hour	No	N/A
Fridge	24 hour	No	N/A
Kitchen App. (1st run)	1 hour	No	N/A
Kitchen App. (2nd run)	1 hour	No	N/A

Following Table 3-1 and Table 3-2, simulating the power consumption for a community of a thousand homes (nodes) results in the overall power demand shown in Figure 3-4. In this figure, the upper curve shows the total average consumption for the simulated community, i.e., the summation of the soft consumption and the hard consumption. This is the average of 50 rounds with a mean of 524, a standard deviation of 226.8, and a peak-to-average power ratio of 1.47. The figure also shows the hard consumption component of the total use.

The hard consumption component is used to compute SD as described in (3.7) and (3.8). SD, shown in Figure 3-5, is computed once based on an average case and is made available to all nodes. As the figure shows, a node scheduling loads following this distribution has a higher probability of scheduling at 5AM in comparison to 6PM. Moreover scheduling loads is not limited to off-peak hours; rather, it extends throughout the day with varying probabilities.

Table 3-2: Appliance operation time selection methodology

Appliance	*Starts Following*
Clothes Dryer	Normal dist. mean at 5PM
Electric Vehicle	Normal dist. mean at 6PM
Clothes Washer	Normal dist. mean at 5PM
Climate Control	Uniform dist. over 24 hours
Water Heater	Uniform dist. from 8AM to 12AM
Range (1st run)	Normal dist. mean at 1PM
Range (2nd run)	Normal dist. mean at 6PM
Electronics	Uniform dist. from 3PM to 1AM
Lighting	Uniform dist. from 8AM to 1AM
Fridge	Uniform dist. over 24 hours
Kitchen Appliances (1st run)	Normal dist. mean at 1PM
Kitchen Appliances (2nd run)	Normal dist. mean at 6PM

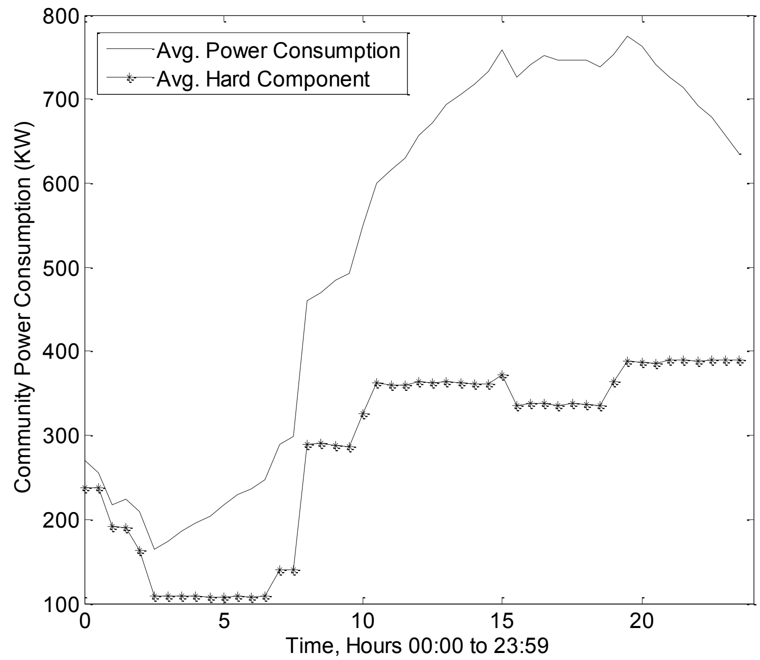

Figure 3-4: Simulated community power consumption.

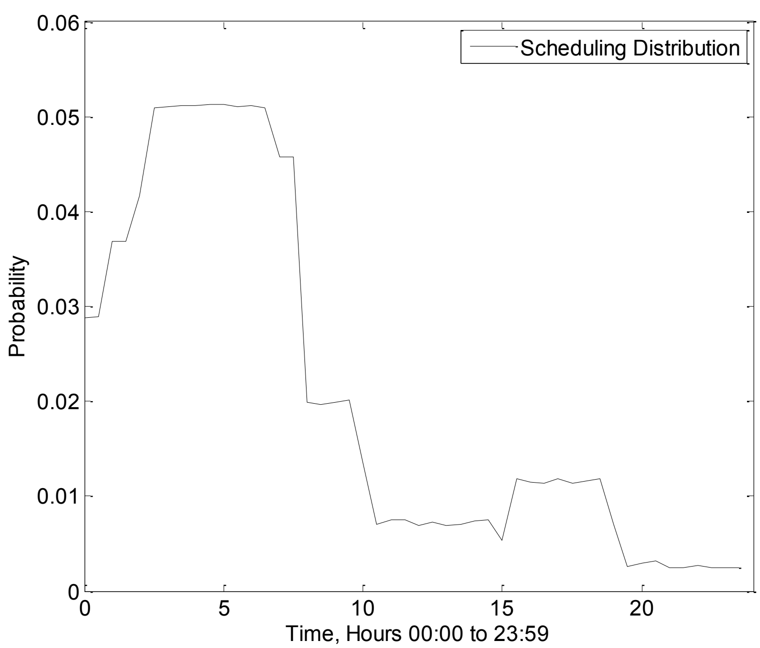

Figure 3-5: Simulated scheduling distribution.

3.4.2 Simulation Results

The results corresponding to the case when following our scheduling algorithm are shown in Figure 3-6. The figure shows the overall hard loads of the community being low in the early hours of the morning and increasing throughout the day. This is the outcome of a single round with a mean of 523, a standard deviation of 15.8, and a peak-to-average ratio of 1.069. The figure also shows the overall scheduled consumption and the total use. As depicted in the figure, the distributed schedulers clearly counter the hard consumption by scheduling most loads towards the early morning hours. Furthermore, the schedulers prevent a peak-hour from appearing at low demand hours and maintain an overall relatively flat consumption. Our approach significantly reduces the peak-to-average ratio (to 1.069 from 1.47) and the standard deviation (to 15.8 from 226) with a relatively small processing and communication overhead. Furthermore, it ensures that consumers pay relatively the same amount (per unit of power) regardless of the time slot their loads were allocated.

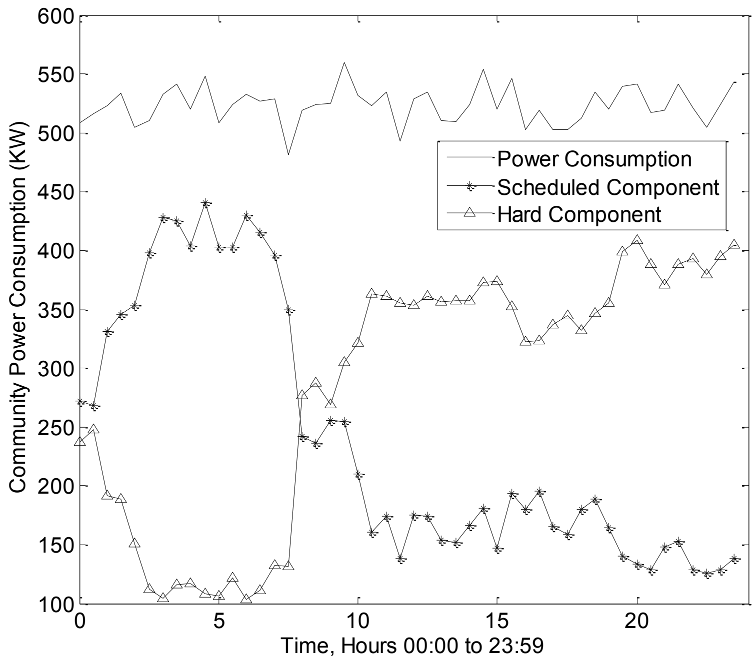

Figure 3-6: Proposed algorithm simulation (single round).

Figure 3-7 shows the average power consumption profile obtained by repeating the simulation for 50 runs. As shown in the figure, the average power consumption is much closer to a constant with a mean of 524, a minimal peak-to-average ratio of 1.021, and a standard deviation of 6.1. Figure 3-8 shows the breakdown of how the community's loads were scheduled in the average case of 50 rounds. As shown, a portion of the soft loads are scheduled following SD and any remaining loads are scheduled uniformly. Furthermore, the figure shows that scheduling nodes can be enhanced to accommodate further operation time constraints. For example, in the portion scheduled uniformly, a scheduler could locally swap loads to accommodate additional time constraints.

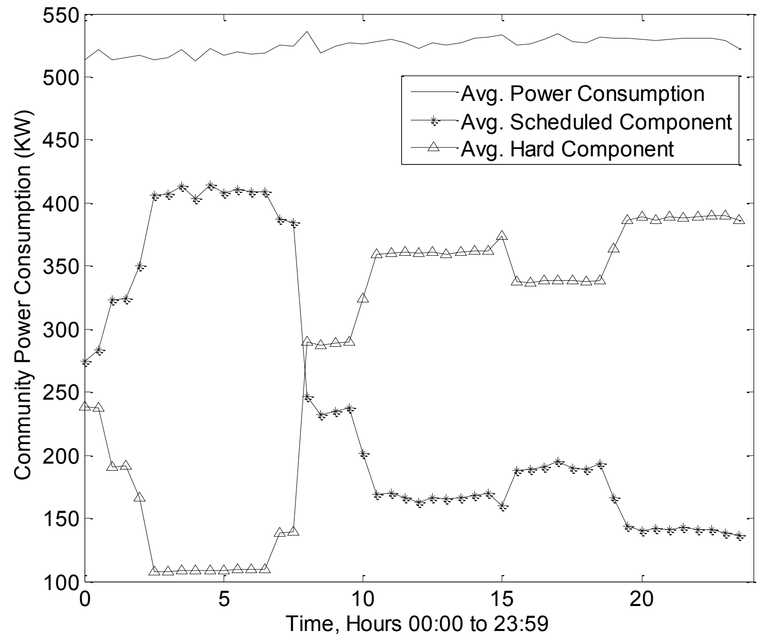

Figure 3-7: Proposed algorithm simulation (average case).

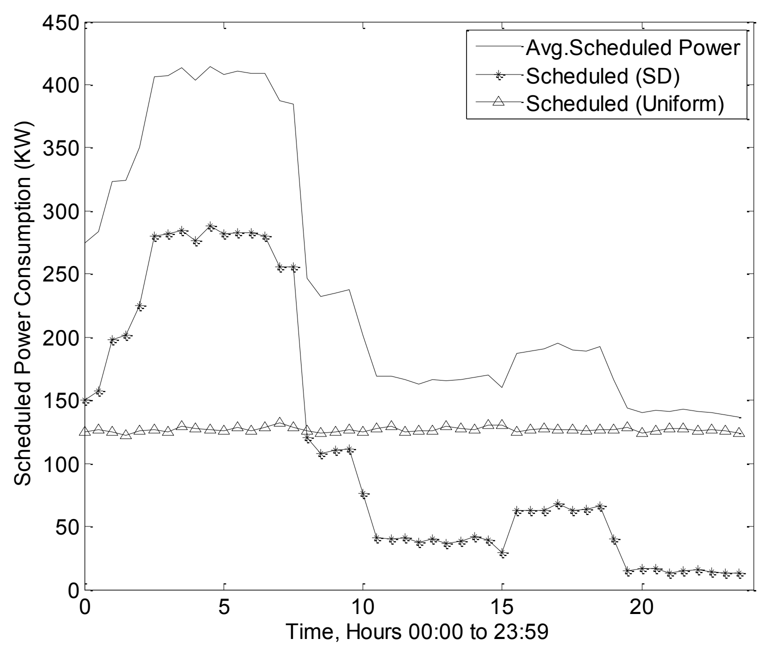

Figure 3-8: Scheduled loads average distribution illustration.

3.5 Algorithm Analysis

In this section, we present our analysis of the proposed approach from various perspectives. Specifically, we analyze the fairness of the algorithm, the effect of possible errors in demand forecasting, and the participation incentives for consumers from a game theoretic perspective.

3.5.1 Fairness

The fairness of the proposed approach is an important factor that would determine the willingness of consumers to follow the scheduling algorithm. To assess the fairness, we use the difference between the starting time for soft load devices without any scheduling and after following the scheduling algorithm as a measure of inconvenience amongst all users. A scheduling algorithm would be considered unfair if some users would have a significantly larger difference compared to others. On the other hand, if all users have relatively the same average change in time, then the algorithm would be considered fair.

More precisely, let Δ_i denote the average absolute value of change in the starting time of the scheduled soft loads belonging to the i^{th} user (reference to the case without scheduling). In other words, Δ_i is a measure of the inconvenience caused to the i^{th} user for opting to follow the scheduling algorithm. Let Δ denote the average of Δ_is for all users. Figure 3-9 below shows the average deviation of Δ_i from Δ obtained by simulating our scheduling algorithm with a thousand users for a thousand runs.

As shown in the figure, on average, the discrepancy among users is very low, i.e., the soft loads of most users are shifted by almost the same amount of time which reflects the fairness of the proposed approach.

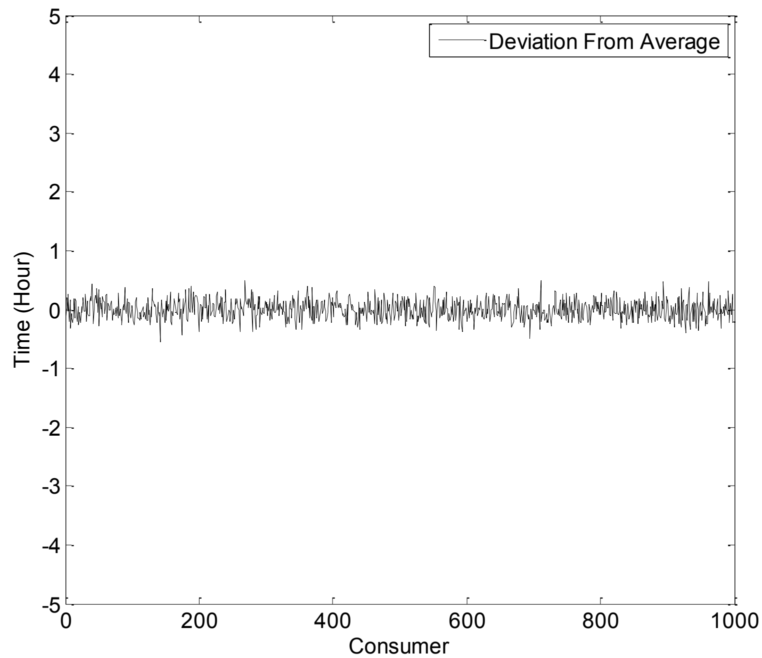

Figure 3-9: Fairness of proposed algorithm.

3.5.2 Demand Forecast

The proposed algorithm relies on historical power consumption data to forecast the consumption for the upcoming day. This information is used to compute SD that guides scheduling nodes in distributing their loads throughout the day. Therefore, the effectiveness of the scheduling algorithm depends on the accuracy of the forecasted demand.

As previously stated and shown in Figure 2-1, utility companies have the ability to forecast the power demand of a community with good accuracy. Furthermore, typical demand forecasting performance indicators that are available in the public domain (e.g., see [13]) show that when the projection is made a day ahead, the error is usually within ±10%.

To evaluate the robustness of our proposed approach against errors in the demand forecast, we introduce a uniformly distributed error between e and $-e$ to the forecasted data; where e is determined as a percentage of the instantaneous consumption at a given time slot. This can be seen as an extreme case as it introduces random errors at all time slots. Figure 3-10 shows the effect of error on the peak-to-average ratio of the community's demand after scheduling.

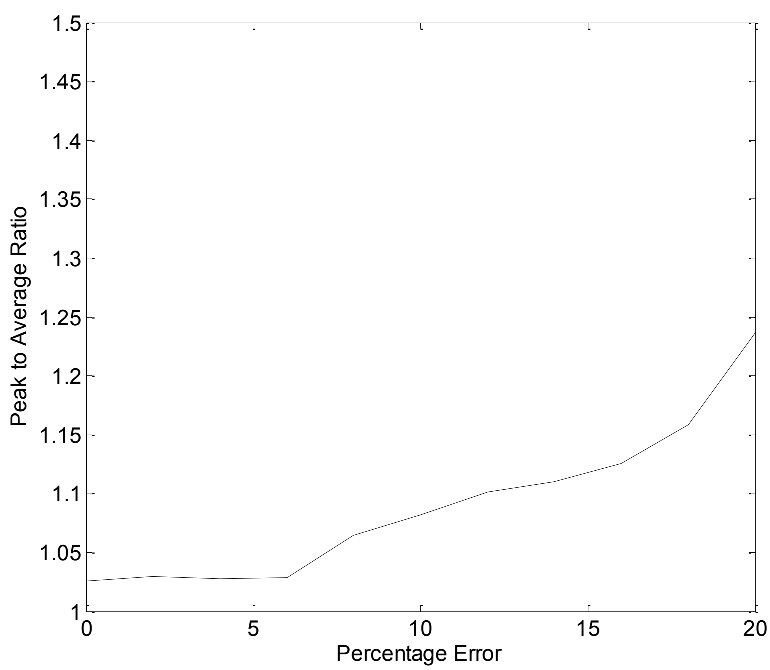

Figure 3-10: Effect of error in demand forecast.

As the figure shows, the peak-to-average ratio of the demand profile after scheduling remains below 1.1 when the error is bounded to ±10% which reflects the effectiveness and robustness of using forecasted data in scheduling soft loads. Furthermore, the figure shows that the peak-to-average ratio of the scheduling community is not affected by errors up to about 5%. This can be attributed to the inherent difference between runs used to produce SD and runs used to simulate the community's consumption, i.e., due to the probabilistic nature of our proposed approach and simulations.

3.5.3 Participation Incentive

To further examine the consumer's participation in the proposed scheme, we analyze the incentives of participation from a game theoretic perspective. We take two main factors into account, namely the action of the community and the inconvenience introduced by scheduling.

A consumer's utility from scheduling consumption is directly related to the community's trend of power usage. If the community does not shift its consumption, a consumer would have a greater incentive in shifting loads towards lower demand hours. On the other hand, if sufficient consumers shift their consumption, a consumer would have less incentive in shifting his/her loads.

Allocating utilities for savings in power costs and possible inconvenience of the scheduling process is a great challenge. In what follows, we focus on the outcome for an individual in the proposed setup. The consumer's incentive would be in reducing his/her electric bill, taking the convenience of this reduction as a main factor. That is, a consumer could decide to schedule or not to schedule soft loads in an environment where the community schedules or does not schedule. In our analysis, we use the following notation. If the community does not shift loads, the savings introduced by the consumer's action are represented by S, and savings missed are represented by $-S$. Furthermore, if the community does shift its loads, the introduced savings are represented by O, and savings missed by $-O$. We represent the inconvenience of scheduling by $-I$. Thus, the possible combinations, summarized in Table 3-3, are:

1. The consumer follows the scheduling algorithm in an environment that follows the algorithm. In this case, the consumer's soft loads draw power at their allocated time. This results in the near optimum overall consumption, which offers the community the best possible rate. We represent the incentive of this outcome by $O - I$.
2. The consumer does not follow the scheduling algorithm in an environment that follows the algorithm. In this case, the consumer misses the slot allocated by the algorithm, and consumes power at a different time resulting in increased demand at that time. Because of the loss, we represent this outcome by $-O$.
3. The consumer follows the scheduling algorithm in an environment that does not follow the algorithm. In this case, the consumer's soft loads are mostly moved toward low demand hours. This allows the consumer to introduce savings in power cost. We represent the incentive of this outcome by $S - I$.
4. The consumer does not follow the scheduling algorithm in an environment that does not follow the algorithm. In this case, neither the consumer, nor the

community benefit from the opportunity. Due to the loss, we represent the outcome by $-S$.

Table 3-3: Scheduling outcome for a consumer

	Consumer Schedules	*Consumer Does Not Schedule*
Community Schedules	O-I	-O
Community Does Not Schedule	S-I	-S

With $S > O$, analyzing the previous table shows that if $O > I/2$, scheduling loads is a dominant strategy. That is, the left column would always be greater than the right column. Therefore, given the constraint above, the choice "to schedule" would be a strongly dominant strategy and consequently, a Nash equilibrium [20] [28].

3.6 Conclusion

The communication and computation capabilities of the smart grid make it possible to have a dynamic pricing scheme where the rate for power is a direct function of the instantaneous demand. Throughout this chapter, we have shown that it becomes possible for a community to adjust its power consumption controlling the price rate and eliminating operational inefficiencies in addition to benefiting from lower power costs.

Within this context, we proposed a methodology that allows consumers to shift part of their soft loads to off-peak hours in a probabilistic way that results in a relatively constant overall power consumption profile. Our simulation results confirm the efficiency and fairness of the proposed scheduling algorithm.

Chapter 4

Privacy Preservation in the Smart Grid

High frequency power consumption readings produced by smart meters introduce a major privacy threat to residential consumers as they reveal details that could be used to infer information about the activities of the occupants of a home. In this chapter, we question the need to disclose high frequency readings produced at the home's level. Instead, we propose equipping smart meters with sufficient processing power enabling them to provide the utility company with a set of well-defined services based on these readings. For demand side management, we propose the collection of high frequency readings at a higher level in the distribution network, such as at the local step-down transformers, as this readily provides the accumulated demand of all homes within the branch. Furthermore, we study the effect of the proposed approach on consumers' privacy, using correlation and relative entropy as measures. We also study the effect of load balancing on consumers' privacy when using the proposed approach. Finally, we assess the detection of appliances using high frequency readings collected for demand side management purposes.

4.1 The Privacy Impact of Smart Metering

By their nature, electrical appliances consume power in specific patterns which produce detectable signatures. For example, a standard incandescent lamp constantly consumes a fixed amount of power during its operation, whereas a refrigerator consumes most of its power during its cooling cycles, when the compressor is running, and significantly less power during its idle cycles. Such patterns can be used to produce signature libraries which can be used for appliance detection and identification [29] [30].

Given a library of power consumption signatures of appliances, and the detailed power consumption of a home, this home's consumption can be decomposed and individual appliances can be detected using Nonintrusive Appliance Load Monitoring technologies (NALM) [31]. Figure 4-1 shows an example of appliance detection using power signatures. As shown, many appliances can be identified through their distinct power consumption patterns. That is, with NALM technologies, high frequency readings produced by smart meters offer a window into the activities of homes' occupants. This includes the identification of appliances and any other information possibly inferable from the appliances used. Furthermore, by observing the real time power consumption of a given home, an intruder can identify when the occupants are awake/asleep or whether the home is occupied or not [32].

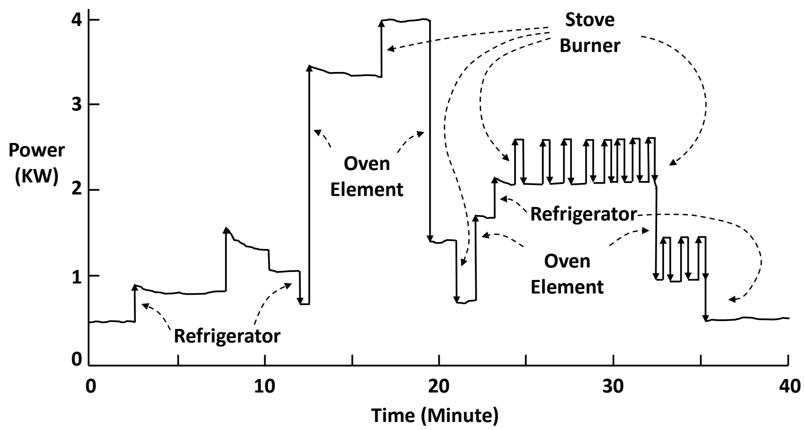

Figure 4-1: Appliance signature detection [31].

Although high frequency readings produced by smart meters enable improving the efficiency of the electric grid, they introduce privacy threats that were not present in the classical grid. In this chapter, we present an alternative approach to smart metering with the objective of maintaining its advantageous functionality while preserving consumers' privacy.

4.2 An Overview of Proposed Solutions

The privacy impact of collecting power consumption readings in high frequency is well known and widely studied. Reports such as [33], [34], and [35] indicate that the privacy concerns of smart metering must be taken into consideration and addressed at the design stage rather than as a later addition. Furthermore, many researchers have proposed various approaches to address this problem. In this section, we present a selection of the main contributions in this area highlighting the core ideas proposed.

Kalogridis *et al.* [36] propose masking the power signature of appliances using a rechargeable battery. In particular, the authors propose the use of an energy routing device that controls power flow from the grid to the home and from/to a rechargeable battery following a water-filling based algorithm. This either charges or discharges the battery in a way that masks some details of the home's power demand. Additional work presented in [37] attempts to quantify the privacy offered by the battery solution, and concludes that privacy preservation increases as the battery size gets larger. Although this approach does mask part of the consumption profile, introducing a large rechargeable battery and a power routing device presents a hindrance to consumers. Furthermore, this solution does not offer as much privacy as consumers had before the introduction of smart meters. For example, information deducible from a home's general consumption pattern such as "when did the occupants wake up" or "is the house occupied" can still be attained even with the deployment of this solution.

Efthymiou and Kalogridis [22] argue that although high frequency readings may be needed for operational purposes, there is no need to attribute them to specific consumers. Consequently, the authors propose the use of two sets of readings: one in high frequency, and the other in low frequency. The high frequency readings are to be collected anonymously with the help of an escrow service and are provided to the utility company. Since the consumers' identities are not associated with these records, the consumption and usage characteristics cannot be traced to a specific consumer. The lower frequency readings are to be bounded to their respective consumers and used for billing purposes. Since these do not capture detailed power consumption information, they are not a threat to consumers' privacy. Although this method may seem effective, the use of an escrow service simply transfers the trust problem from the utility company to the escrow service provider, and therefore, does not provide a fundamental solution to the original problem.

In [38], Tomosada and Sinohara propose that smart meters transmit synthetically produced data that shares the same statistical properties of the real readings instead of transmitting the readings themselves. The authors argue that since this virtual demand shares the same statistical properties with the real demand, it can be used for demand side management when averaged over multiple users. In their work, the authors propose a methodology for producing virtual demand from the real demand and conclude that this approach preserves the consumer's privacy. Although this method produces correct statistics, other characteristics critical to demand side management could be lost, for example, the peak value and the time at which this peak value occurs.

4.3 Proposed Approach to Smart Metering

In this section, we present our proposed approach in addressing the privacy concerns introduced by smart meters. We start by presenting our vision for the metering infrastructure, by proposing an alternative architecture. We then introduce and describe the use and functionality of home meters and zone meters in some details.

4.3.1 Metering Architecture

Smart meters are typically used as distributed data acquisition devices. That is, the meters only produce and transmit high frequency readings to the utility company. The utility company, in turn, centrally processes this data producing bills for its subscribers based on the time electricity was consumed. Figure 4-2 illustrates this view of smart meters functionality.

With this approach, high frequency readings are present at the meter, in transmission and in storage at the utility company's processing facility. Having this data at all these points maximizes the potential attack surface for an attacker. This way, the attacker needs to identify some vulnerability in *any* of these points to be able to access the detailed consumption records. Furthermore, if an attacker is able to identify and exploit some vulnerability at the central processing facility, the impact would be devastating as hundreds of thousands of records could be compromised in a single breach. Verifiably securing large distribution networks, communication networks and processing facilities is practically infeasible. Furthermore, an attempt to secure such interconnected systems would be a tedious task that is almost impossible to implement flawlessly.

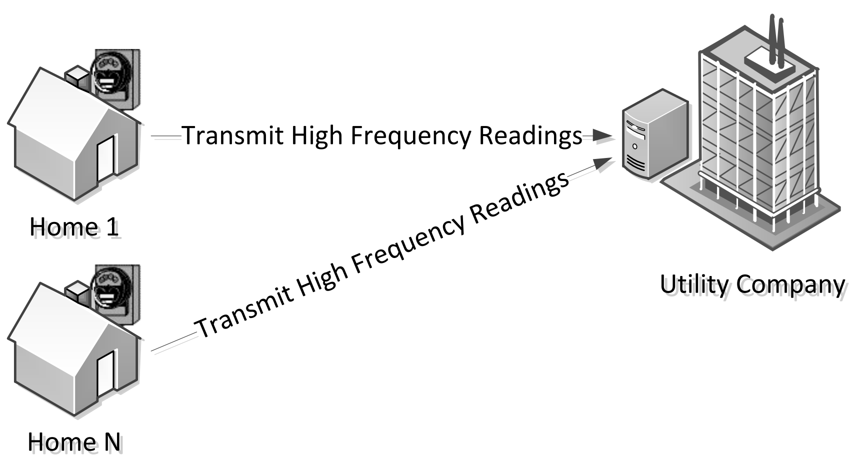

Figure 4-2: Home meters as data acquisition devices.

For the purpose of our work, we categorize consumer-oriented data collected by the utility company into two basic types, namely *subject data* and *community data*. We define *subject data* as data collected from and identifiable to a single consumer. We assume that actions taken based on this type of data will only affect its associated consumer. On the other hand, we define *community data* as data identifiable to a group rather than a single consumer. Furthermore, we assume that actions taken on a large scale, i.e., on the community as a whole, are based on this type of data.

We propose the deployment of smart meters in a way that segregates these two types of data offering each an appropriate level of protection. Therefore, from a privacy perspective, *subject data* would have a higher level of protection in comparison to *community data*. To do so, we propose the use of two sets of meters positioned at different locations in the distribution network, namely *home meters* and *zone meters*.

4.3.2 Home Meters

We propose that *home meters* function as service providing modules rather than data acquisition devices. That is, assuming that each meter is equipped with sufficient processing power, the meters are to offer the utility company a set of well-defined services computed over the consumer's high frequency readings. In addition, *home meters* are not to disclose the collected high frequency readings to any party. Furthermore, the services offered by the meter must be developed on a need to know basis.

With this approach, the meters become the entities that perform all required processing on their respective consumers' data and only the outcomes of processing, i.e., the final results, are made available to the utility company. This allows *home meters* to provide the desired functionality while eliminating the need to disclose users' high frequency readings, consequently, preserving the consumers' privacy. Furthermore, this introduces a point of control on the type of information the utility company gains access to.

The services provided by a *home meter* would depend on the protocols/functions it implements. For example, for billing purposes, meters would implement a billing protocol that starts by receiving an authenticated request from the utility company to produce the consumer's bill for a given period. The meter, in turn, uses the high frequency readings from its internal storage to compute the amount owed in dollars based on a pre-agreed upon pricing scheme. The final result of the process (the amount in dollars) would be encrypted, digitally signed by the meter and transmitted to the utility company. This would provide the utility company with the desired information ensuring that it was produced by the meter.

Besides billing, other useful functionalities can be easily implemented. For example, meters could report their operation status or fault codes by periodically transmitting a status message that can be protected using cryptographic techniques. Figure 4-3 illustrates the use of *home meters* as service providing modules.

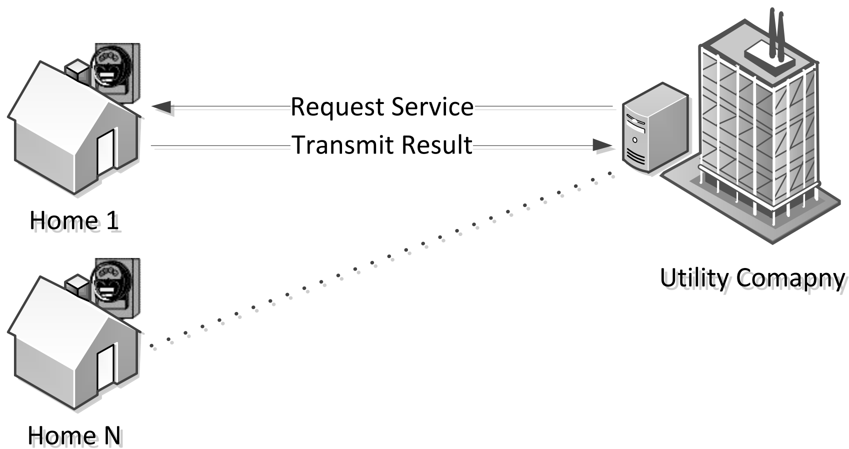

Figure 4-3: Home meters as service providing modules.

It should be noted that, in our proposed approach, *home meters* must be trusted as they are the only components that hold and directly process their respective consumer's high frequency power consumption readings. Since a meter is a relatively small device that implements a limited set of protocols, it is feasible to build verifiably secure meters. This can be achieved through the use of trusted computing platforms [39] [40], and the use of formally verifiable designs.

With this approach, since each consumer's data is only present at a single point, the attack surface is significantly reduced. That is, if an attacker is successful in penetrating a given meter, only the records belonging to the associated customer would be compromised. This is a significant security advantage in comparison to the use of meters as distributed collection nodes and centrally processing the data collected.

4.3.3 Zone Meters

As stated above, *home meters* do not disclose unprocessed power consumption readings; rather they provide specific information computed from this data. Even though it is possible to produce accurate *time of use* based consumption bills, information on consumption trends and the time of power consumption would be masked. This introduces a hindrance to the operations of utility companies as such information is aggregated for demand side management. To address this, we propose the use of an additional small set of meters placed at a higher level in the distribution network; typically at local step-down transformers.

Instead of securely producing accumulative readings through an escrow service or using cryptographic approaches, we take advantage of the already existing topology of the distribution network. By observing that the demand at a step down station is the accumulation of the demand of all homes supported by this station, measuring the power consumption at this level is equivalent to aggregating the power demand of individual homes in this zone. Therefore, by collecting measurements at this location using *zone meters*, we readily obtain the accumulative demand of all homes within a given branch. Figure 4-4 illustrates the use of two sets of meters at the home and the zone level.

Although high frequency readings produced by *zone meters* are not directly attributable to a single consumer, they produce readings of the composite demand of all homes supported by their branch which is a function of the consumption of each home. In the following section we analyze the privacy impact of the proposed approach and the visibility of a home's demand through readings produced by *zone meters*. We also specifically consider the impact of load balancing on consumers' privacy when using the proposed approach.

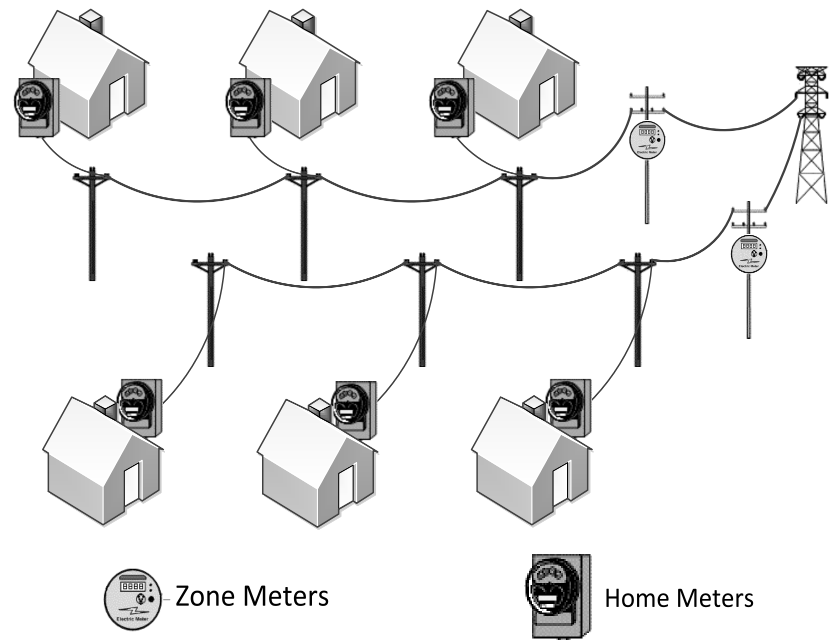

Figure 4-4: Smart metering at two levels.

4.4 Privacy Assessment

In our proposed approach, high frequency readings produced by *home meters* are not disclosed to the utility company or any other party. This eliminates the direct privacy impact introduced by these readings. Although *zone meters* introduce indirect threats, the superposition of signatures from different homes introduces an obfuscation effect on individual signatures which lessens the disclosed information. In this section, we assess the obfuscation gained by overlapping home signatures. The number of homes supported by a single *zone meter* is a main factor in the level of signature overlap; therefore we consider this as a primary factor in our simulation. Furthermore, since time of use billing schemes are expected to flatten the overall demand of a community which affects the overlapping of signatures, we take this into consideration as well.

4.4.1 Simulation Environment

We produced our simulation environment using a set of appliances with distinct power signatures similar to [31] and using measurements from [41]. Each home is allocated a set of appliances and the operation time of each appliance is selected randomly. The simulation is conducted over a period of 24 hours. Furthermore, to assess the impact of load balancing, two sets of results are produced for each simulation scenario. The first represents the case where power is consumed at will. This type of consumption results in the appearance of peak demand hours as is the case with the classical power grid [12]. The second represents the case where consumers shift part of their consumption to off-peak hours using load balancing techniques such as those described in [16], [19], and [21]. This results in a relatively flat consumption profile for the community as a whole. Figure 4-5 shows the simulated power demand for a sample home with a sampling interval of one minute. As depicted in the figure, the consumption patterns of many appliances can be easily identified.

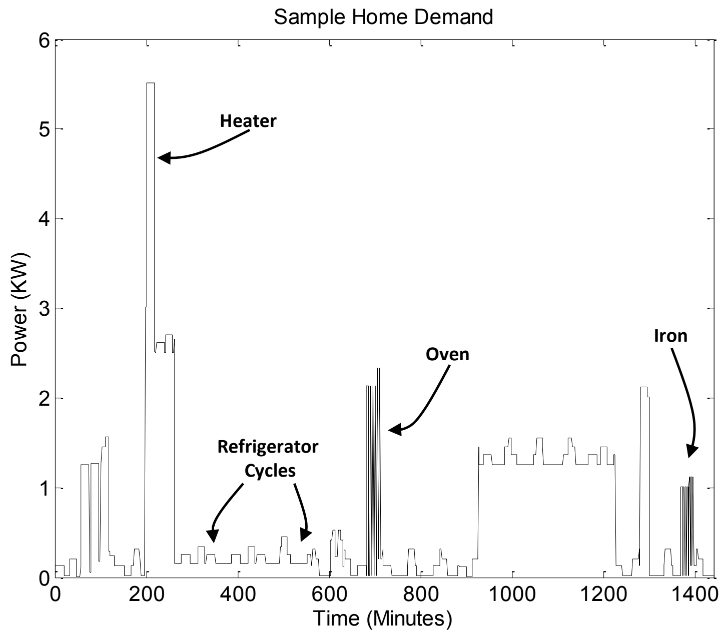

Figure 4-5: Sample home demand simulation.

To simulate the readings produced by *zone meters*, a variable number of homes are simulated and accumulated. Figure 4-6 illustrates the aggregate demand of a community of 50 homes. As depicted in the figure, the overlapping of signatures of different homes distorts the appliance signatures. The figure also reflects the effect of load balancing on the overall consumption of the community. As shown, the use of load balancing (bottom) results in a flatter overall demand with a lower peak to average ratio in comparison to its absence (top). This results in a more uniform level of overlapping between appliance signatures.

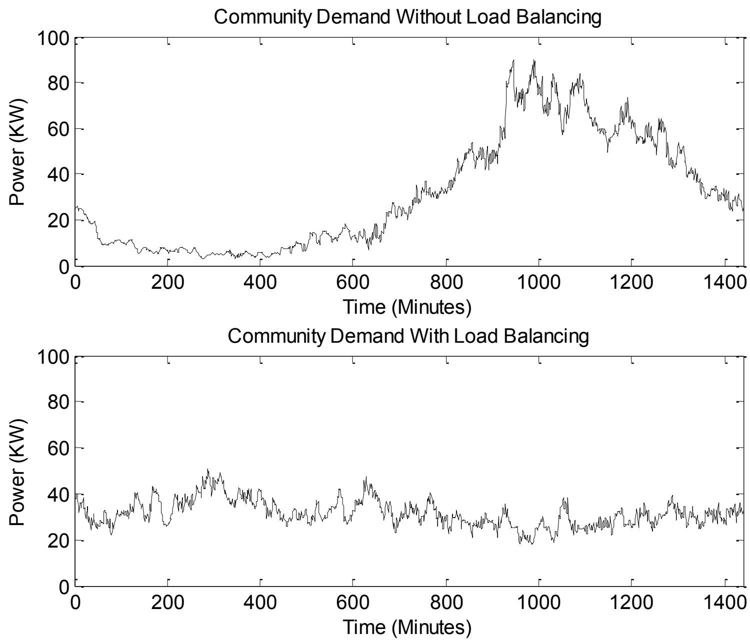

Figure 4-6: Community demand simulation.

4.4.2 Home Demand Obfuscation

Figure 4-7 shows the decrease in correlation between a home's demand and the overall community consumption as the number of homes in the community is increased. This figure represents the average case of 100 iterations while increasing the number of homes up to 50. The decreasing trend in the figure can be attributed to the distortion caused to the demand signature from other homes. As the number of homes increase, so does the level of distortion. Furthermore, the figure shows that the use of load balancing techniques further reduces the correlation value resulting in better privacy. This is because load balancing results in even overlapping and eliminates the low overlap between signatures during low demand hours.

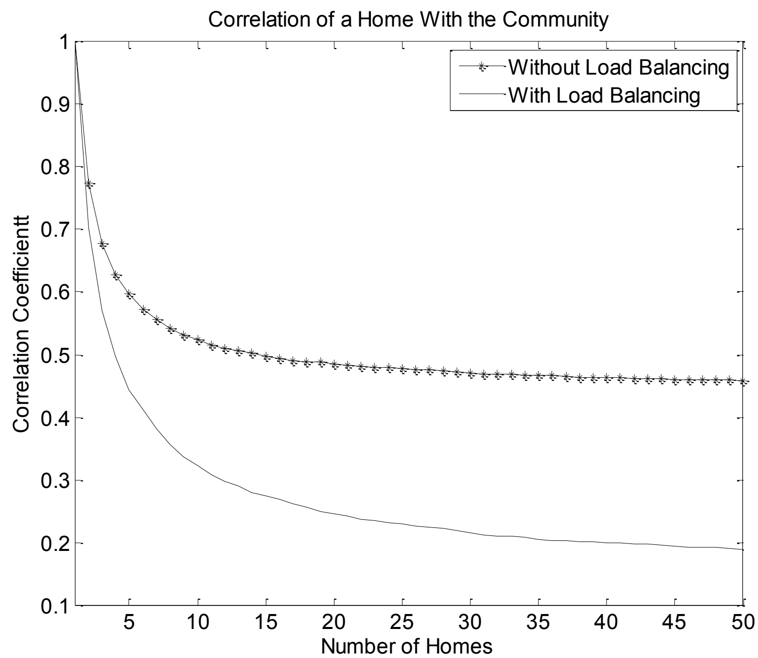

Figure 4-7: Home and community signature correlation.

Figure 4-8 shows the correlation value in the average case of 100 iterations for each home when simulating a community of 50 homes. The results indicate that the signatures of all simulated homes were distorted to a similar level.

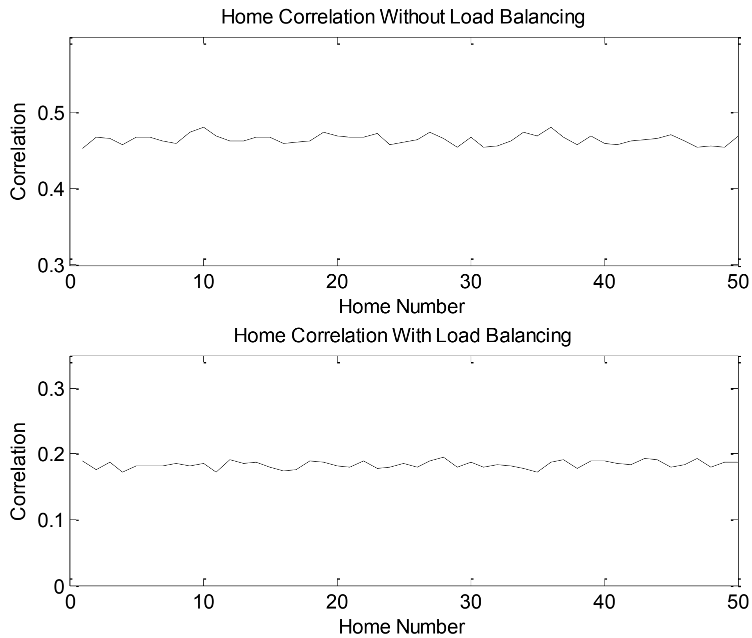

Figure 4-8: Community correlation average cases.

As an assessment of the difference between the probabilistic distributions in the data sets, we compute the Kullback Leibler Divergence (KLD) [23] (also known as the relative entropy) while increasing the number of homes in the community. This is a well-known information theoretic measure that can be used to quantify the relationship between two signals. Given two signals with probability distributions P and Q, the Kullback Leibler divergence can be defined as

$$D(P||Q) = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} \, dx. \tag{4.19}$$

As shown in Figure 4-9, the value of the Kullback Leibler divergence is zero for a single home, indicating identical distributions, and grows rapidly to saturate when accumulating about 10 homes. This figure represents the average case of 100 iterations while increasing the number of homes up to 50. This indicates that accumulating a relatively small number of homes would have a good effect on masking individual homes' consumption profiles. Furthermore, as depicted in the figure, the use of load balancing helps achieve better privacy protection.

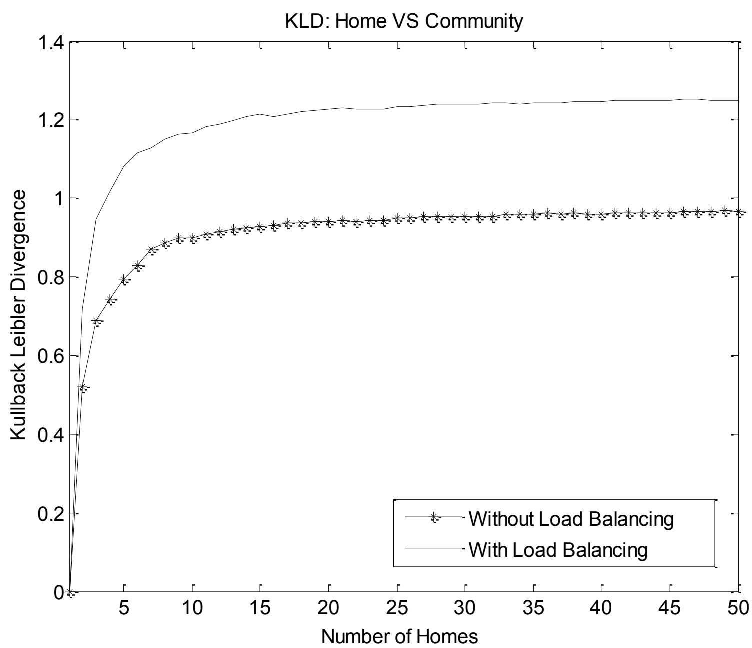

Figure 4-9: KLD of home and community signatures.

4.4.3 Appliance Detection

Assuming the availability of a library of appliance signatures [41], we assess the detection of the operation of appliances using cross correlation from readings produced by *zone meters* while increasing the number of homes in the community. Unlike previous simulations, the appliance to be detected is only run once by one member of the community. This ensures that a false detection is not caused by a duplicate signature of the appliance in question.

This case was simulated allowing the target appliance (i.e., the one to be detected) to be turned on randomly, following a uniform distribution throughout the 24 hour simulation interval with a step size of one minute; i.e., a total of 1440 possible time slots. We define a correct detection as one where the precise time slot was identified.

Figure 4-10 shows the percentage of correct detections for a sample appliance as the number of homes in the community is increased. The figure represents the average case of 100 iterations while increasing the number of homes up to 50. As the figure shows, the use of load balancing techniques causes a more rapid deterioration in detection accuracy, i.e., it achieves better privacy protection. Figure 4-11 shows the error in detection of the same simulation, in time slots, as a function of the number of homes in the community. As shown, the detection becomes more distant from the real start time as the number of homes is increased. Furthermore, the use of load balancing causes an increase in the detection error, which implies better preservation of privacy.

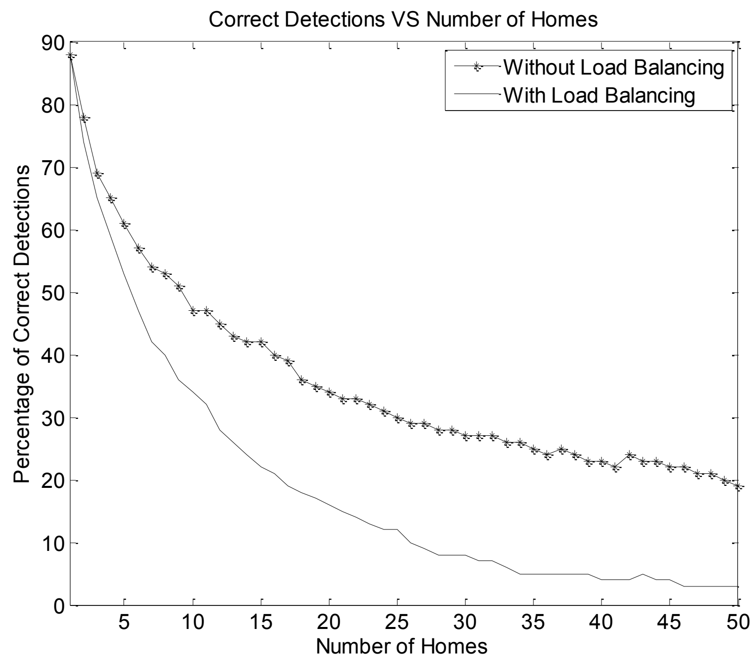

Figure 4-10: Sample appliance detection.

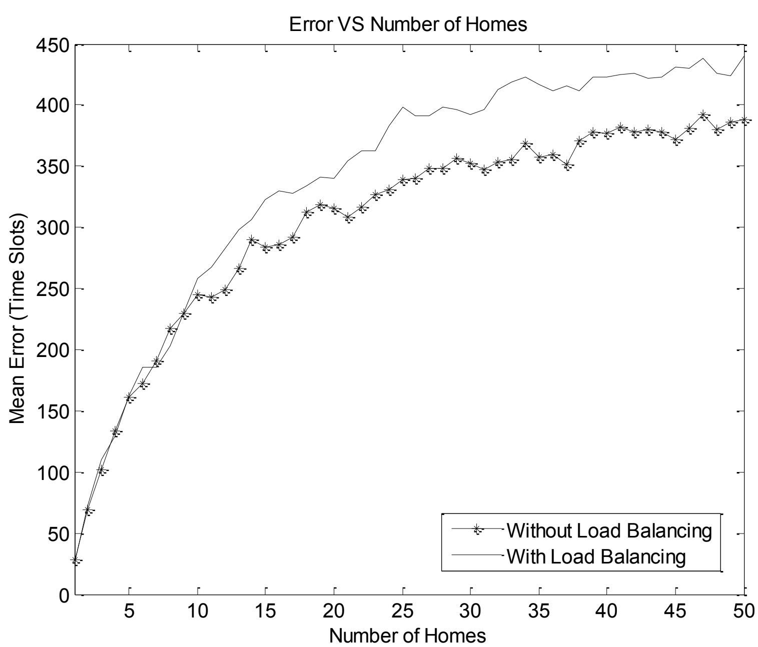

Figure 4-11: Mean error in appliance detection.

4.5 Approach Assessment

In this section, we present our analysis of the proposed approach to smart metering. The key advantages considered are the elimination of disclosing high frequency readings while achieving the objectives of smart metering, the reduction of the attack surface and the effect of the overall efficient operation on consumers' privacy.

4.5.1 Elimination of Detailed Readings Disclosure

The use of smart meters as data acquisition devices requires high frequency readings to be processed by an entity other than the meter, possibly, a central data processing center. On the other hand, with smart meters performing as service providing modules, there no longer remains a need for meters to disclose high frequency readings. Instead, meters locally process the high frequency readings collected and only reveal the final result of pre-defined operations to the utility company.

Although this architecture requires meters to be trusted, this is more practical and cost effective than requiring the entire system to be trusted. This is because meters are relatively small components with limited functionality, and therefore can be built in a verifiably secure manner.

4.5.2 Achievement of Smart Metering Objectives

In the literature, the collection of high frequency readings of power consumption is seen as the fundamental enabler for the advantages of the smart grid, and many works are based on this. However, in our work, we question if this is necessary. In fact, as presented, we believe that given the capabilities of smart meters, it is possible to achieve the smart metering objectives without the need to disclose consumers' power consumption details. Therefore, we propose that smart meters collect but not disclose this data. Instead, the meters are to provide the utility company with a set of well-defined services computed over the consumer's private data.

4.5.3 Reduction of Attack Surface

In the traditional approach, smart meters collect consumers' power readings in high frequency and transmit them to the utility company for central processing. This way, these readings are present at the meter, in transmission, and possibly in storage at the utility company. Therefore, all these systems must be secured to ensure that users' data is not compromised. Furthermore, from an attacker's perspective, having this data in all possible locations maximizes the attack surface. That is, the attacker needs to identify and exploit some vulnerability at any of these points to gain access to the private data. On the other hand, with our approach, a given consumer's data is only present in their meter and is not disclosed or transmitted to any party. This significantly reduces the attack surface as it eliminates the presence of this data in transmission and in storage at the utility company. Furthermore, in the event that a specific meter is compromised, only the respective consumer's data would be affected. On the other hand, if all consumers' data is stored in a central database and this database is compromised, the outcome would be devastating as all records will be affected.

4.5.4 Increase of Privacy with Efficient Operation

The quest for efficient operation of the electric grid introduced smart meters and the privacy concerns that accompany them. Therefore, it would appear that the more efficient operation of the grid requires compromises on the privacy front. However, in our work, we show that when following our proposed approach for smart metering, the more efficient operation of the grid results in increased consumer privacy. In addition to not disclosing consumers' direct high frequency readings, readings produced by *zone meters* are less relatable to a consumer's data when the community consumes power in a way that reduces the peak-to-average ratio of the overall demand.

4.6 Conclusion

In this chapter, we showed that it is possible to achieve the objectives of smart metering without compromising the privacy of residential consumers. In our approach, *home meters* function as service providing modules rather than data acquisition devices; this allows them to provide the desired functionality, such as *time of use* billing, while eliminating the need to disclose users' high frequency readings to the utility company. Collecting high frequency readings is done at a higher level such as at the local step-down transformer which allows utility companies to achieve their operational objectives. While our approach requires *home meters* to be trusted, ensuring this requirement is more feasible than attempting to secure the meters, the transmission network, and all the processing facilities. Furthermore, we also showed that our approach achieves better privacy protection when consumers opt to use load balancing techniques. Our results imply that, with the approach proposed, the more efficient operation of the grid can result in better privacy protection for individual customers.

Chapter 5

Conclusions and Future Works

Throughout our work, we show that by utilizing the capabilities of smart meters and information producible by utility companies, it is practically possible for members of a community to adjust their power consumption in a way that eliminates operational inefficiencies and allows them to benefit from lower power costs. Specifically, we have proposed a scheduling algorithm that enables a community to flatten their demand profile. Our algorithm eliminates the need for continuous communication between nodes in quest for an optimum solution while minimizing the processing required by the meter. When followed, the algorithm results in a relatively constant demand in a way that is fair to all members of the community. This allows the elimination of the operational inefficiencies related to power generation and consumption, resulting in a greener and more cost effective grid.

Furthermore, we have shown that achieving the objectives of smart metering do not necessarily require compromises in consumer's privacy. In this context, we propose that smart meters be architected as service providing modules, rather than data acquisition devices. With this approach, utility companies can obtain the information they need, through the well-defined functions implemented by the meters. For demand side management, we propose that high frequency readings be collected at a higher level in the distribution network. As shown by our simulations, accumulating a relatively small number of homes, this approach maintains the consumer's privacy and provides the utility company with the desired information. In addition, this approach offers a layer of security against attacks, as it reduces the attack surface. Our simulations have also indicated that the more efficient operation of the grid does not necessarily require introducing compromises to privacy. In fact, the more efficient operation resulted in higher privacy when using the approach proposed.

It should be noted, though, that throughout our analysis of our proposed scheduling algorithm, we have considered appliances to be either completely shiftable (soft loads), or completely non-shiftable (hard loads). In reality, some appliances can only accommodate small tolerance in their operation time, and therefore, cannot be considered completely shiftable. One possible area of research on this front would be to model the constraints associated with such appliances and study their effect on the peak-to-average ratio of the scheduling process.

Another possible area for improvement would be to compare the proposed scheduling algorithm with other similar approaches with the objective of quantifying the effectiveness, and the communication and processing overheads. Similar approaches include the use of a greedy algorithm instead of a water-filling based algorithm, continuously updating a central point of control with real-time data instead of using forecasts. Another point to investigate would be to study the possibility of having each node flatten its own demand. It would appear that the approaches stated above could be practical, however, it would be important to study the feasibility based on the soft appliances of a home. That is, to study if an average home would have a sufficient amount of soft loads that can be scheduled such that each home's power demand becomes a constant.

We would also like to point that our analysis of the proposed privacy enhancing approach to smart metering is based on simulated appliance signatures. Our results would be more significant if conducted over real signatures using power consumption information as produced by smart meters. This is because the signature details and the metering frequency are key factors for the correct detection of appliances. Another area for improvement is to identify, and quantify the effectiveness of other measures for appliance detection. We mainly used cross-correlation for this, but we believe that the robustness of the proposed approach would be further proven if other, more sophisticated NALM techniques were used.

Although we have established the foundations for enhancing the privacy of smart meters, we have not identified or developed secure protocols for this purpose. Consequently, this work can be further enhanced by identifying the precise purpose of utility companies in collecting high frequency readings, and then developing secure protocols based on the approach presented herein, that achieve this purpose. We believe that this would be a tangible contribution with an immediate impact on the industry.

Bibliography

[1] P. Schewe, The Grid: A Journey Through the Heart of Our Electrified World, National Academy Press, 2007.

[2] US Department of Energy, "ENERGY.GOV," 2009. [Online]. Available: http://energy.gov/. [Accessed April 2012].

[3] E. Santacana, G. Rackliffe, L. Tang and X. Feng, "Getting Smart," *IEEE Power and Energy Magazine,* pp. 41-48, 2010.

[4] S. Collier, "Ten Steps to a Smarter Grid," *IEEE Rural Electric Power Conference (REPC '09),* pp. B2 - B2-7, 2009.

[5] D.-Q. Sun, J.-w. Zhengb, T. Zhangc, Z.-J. Z. a, H.-T. Liud, F. Zhaod and Z.-J. Qiue, "The Utilization and Development Strategies of Smart Grid and New Energy," *Power and Energy Engineering Conference (APPEEC),* pp. 1 - 4, 2010.

[6] Office of the National Coordinator for Smart Grid Interoperability, "NIST Framework and Roadmap for Smart Grid Interoperability Standards," National Institute for Standards and Technology (NIST), 2010.

[7] D. Palomar and J. Fonollosa, "Practical algorithms for a family of waterfilling

solutions," *IEEE Transactions on Signal Processing,* vol. 53, no. 2, pp. 686 - 695, 2005.

[8] Books LLC, Electrical Grid: Electric Grid Interconnections, Electric Power Blackouts, Electric Power Transmission Systems, Smart Grid, 2011.

[9] S. M. Kaplan and F. Sissine, Smart Grid: Modernizing Electric Power Transmission and Distribution, Thecapitol.Net, 2009.

[10] F. P. Sioshansi, Smart Grid: Integrating Renewable, Distributed & Efficient Energy, Academic Press, 2011.

[11] C. W. Gellings, The Smart Grid: Enabling Energy Efficiency and Demand Response, CRC Press, 2009.

[12] "Ontario Demand and Market Prices," The Independent Electricity System Operator (IESO), [Online]. Available: http://www.ieso.ca/. [Accessed July 2010].

[13] "IESO Demand Forecasting Performance Indicators," Demand Forecast Deviations Working Group, the Independent Electricity System Operator (IESO), [Online]. Available: http://www.ieso.ca/. [Accessed April 2012].

[14] J. Girvan, "Consumer Council of Canada," March 2009. [Online]. Available: http://www.consumerscouncil.com.

[15] R. Krishnan, "Meters of tomorrow [In My View]," *IEEE Power and Energy Magazine,* pp. 94-96, March 2008.

[16] S. Caron and G. Kesidis, "Incentive-Based Energy Consumption Scheduling Algorithms for the Smart Grid," *First IEEE International Conference on Smart Grid*

Communications (SmartGridComm), pp. 391 - 396, 2010.

[17] G. Xiong, C. Chen, S. Kishore and A. Yener, "Smart (In-Home) Power Scheduling for Demand Response on the Smart Grid," *IEEE PES Innovative Smart Grid Technologies (ISGT),* pp. 1 - 7, 2011.

[18] N. Gatsis and G. Giannakis, "Cooperative Multi-Residence Demand Response Scheduling," *45th Annual Conference on Information Sciences and Systems (CISS),* pp. 1 - 6, 2011.

[19] C. Chen, S. Kishore and L. Snyder, "An Innovative RTP-Based Residential Power Scheduling Scheme for Smart Grids," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* pp. 5956 - 5959, 2011.

[20] M. J. Osborne and A. Rubinstein, A Course in Game Theory, The MIT Press, 1994.

[21] A. H. Mohsenian-Rad, V. Wong, J. Jatskevich and R. Schober, "Optimal and Autonomous Incentive-Based Energy Consumption Scheduling Algorithm for Smart Grid," *Innovative Smart Grid Technologies (ISGT),* pp. 1-6, 2010.

[22] C. Efthymiou and G. Kalogridis, "Smart Grid Privacy via Anonymization of Smart Metering Data," *First IEEE International Conference on Smart Grid Communications (SmartGridComm),* pp. 238 - 243, 2010.

[23] T. M. Cover and J. A. Thomas, Elements of Information Theory, Wiley-Interscience, 2006.

[24] C. Shannon, "Communication In The Presence Of Noise," *Proceedings of the IEEE (PROC),* vol. 86, no. 2, pp. 447-457, 1998.

[25] J. Proakis and M. Salehi, Digital Communications, McGraw-Hill, 2007.

[26] D. Palomar and J. Fonollosa, "Practical Algorithms for a Family of Waterfilling Solutions," *IEEE Transactions on Signal Processing,* vol. 53, no. 2, pp. 686 - 695 , 2005.

[27] A. Ipakchi and F. Albuyeh, "Grid of the Future," *IEEE Power and Energy Magazine,* vol. 7, no. 2, pp. 52 - 62, 2009.

[28] P. D. Straffin, Game Theory and Strategy, The Mathematical Association of America, 2006.

[29] J. Liang, S. Ng, G. Kendall and J. Cheng, "Load Signature Study—Part I: Basic Concept, Structure, and Methodology," *IEEE Transactions on Power Delivery,* vol. 25, no. 2, pp. 551-560, 2010.

[30] J. Liang, S. Ng, G. Kendall and J. Cheng, "Load Signature Study—Part II: Disaggregation Framework, Simulation, and Applications," *IEEE Transactions on Power Delivery,* vol. 24, no. 2, pp. 561 - 569, 2010.

[31] G. Hart, "Nonintrusive Appliance Load Monitoring," *Proceedings of the IEEE,* vol. 20, no. 12, pp. 1870 - 1891 , 1992.

[32] E. Quinn, "Privacy and the New Energy Infrastructure," Social Science Research Network (SSRN), 2009.

[33] A. Cavoukian, "Privacy by Design: Achieving the Gold Standard in Data Protection for the Smart Grid," Information and Privacy Commissioner, Ontario, Canada, 2010.

[34] A. Cavoukian, "Operationalizing Privacy by Design: the Ontario Smart Grid Case

Study," Information and Privacy Commissioner, Ontario, Canada, 2011.

[35] A. Cavoukian, "Smart Privacy for the Smart Grid: Embedding Privacy into the Design of Electricity Conservation," Information and Privacy Commissioner, Ontario, Canada, 2009.

[36] G. Kalogridis, C. Efthymiou, S. Denic, T. Lewis and R. Cepeda, "Privacy for Smart Meters: Towards Undetectable Appliance Load Signatures," *First IEEE International Conference on Smart Grid Communications (SmartGridComm),* pp. 232 - 237, 2010.

[37] G. Kalogridis, Z. Fan and S. Basutkar, "Affordable Privacy for Home Smart Meters," *Ninth IEEE International Symposium on Parallel and Distributed Processing with Applications Workshops (ISPAW),* pp. 77 - 84, 2011.

[38] M. Tomosada and Y. Sinohara, "Virtual Energy Demand Data: Estimating Energy Load and Protecting Consumers' Privacy," *IEEE PES Innovative Smart Grid Technologies (ISGT),* pp. 1 - 8, 2011.

[39] C. Mitchell, Trusted Computing, Inspec, 2005.

[40] S. Pearson, "Trusted Computing Platforms, the Next Security Solution," Trusted Systems Lab, HP Laboratories, 2002.

[41] I. Richardson, M. Thomson and D. G. Infield, "A High-Resolution Domestic Building Occupancy Model for Energy Demand Simulations," *Energy and Buildings,* vol. 40, no. 8, pp. 1560-1566, 2008.