Demand Response in Smart Grid

by

Kan Zhou
B.Sc., Southeast University, 2008
M.Sc., Southeast University, 2011

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Electrical and Computer Engineering

Demand Response in Smart Grid

by

Kan Zhou
B.Sc., Southeast University, 2008
M.Sc., Southeast University, 2011

Supervisory Committee

Dr. Lin Cai, Supervisor
(Department of Electrical and Computer Engineering)

Dr. Xiaodai Dong, Departmental Member
(Department of Electrical and Computer Engineering)

Dr. Kui Wu, Outside Member
(Department of Computer Science)

Supervisory Committee

Dr. Lin Cai, Supervisor
(Department of Electrical and Computer Engineering)

Dr. Xiaodai Dong, Departmental Member
(Department of Electrical and Computer Engineering)

Dr. Kui Wu, Outside Member
(Department of Computer Science)

ABSTRACT

Conventionally, to support varying power demand, the utility company must prepare to supply more electricity than actually needed, which causes inefficiency and waste. With the increasing penetration of renewable energy which is intermittent and stochastic, how to balance the power generation and demand becomes even more challenging. Demand response, which reschedules part of the elastic load in users' side, is a promising technology to increase power generation efficiency and reduce costs. However, how to coordinate all the distributed heterogeneous elastic loads efficiently is a major challenge and sparks numerous research efforts. In this thesis, we investigate different methods to provide demand response and improve power grid efficiency.

First, we consider how to schedule the charging process of all the Plugged-in Hybrid Electrical Vehicles (PHEVs) so that demand peaks caused by PHEV charging are flattened. Existing solutions are either centralized which may not be scalable, or decentralized based on real-time pricing (RTP) which may not be applicable immediately for many markets. Our proposed PHEV charging approach does not need complicated, centralized control and can be executed online in a distributed manner. In addition, we extend our approach and apply it to the distribution grid to solve the bus congestion and voltage drop problems by controlling the access probability

of PHEVs. One of the advantages of our algorithm is that it does not need accurate predictions on base load and future users' behaviors. Furthermore, it is deployable even when the grid size is large.

Different from PHEVs, whose future arrivals are hard to predict, there is another category of elastic load, such as Heating Ventilation and Air-Conditioning (HVAC) systems, whose future status can be predicted based on the current status and control actions. How to minimize the power generation cost using this kind of elastic load is also an interesting topic to the power companies. Existing work usually used HVAC to do the load following or load shaping based on given control signals or objectives. However, optimal external control signals may not always be available. Without such control signals, how to make a tradeoff between the fluctuation of non-renewable power generation and the limited demand response potential of the elastic load, and to guarantee user comfort level, is still an open problem. To solve this problem, we first model the temperature evolution process of a room and propose an approach to estimate the key parameters of the model. Then, based on the model predictive control, a centralized and a distributed algorithm are proposed to minimize the fluctuation and maximize the user comfort level. In addition, we propose a dynamic water level adjustment algorithm to make the demand response always available in two directions. Extensive simulations based on practical data sets show that the proposed algorithms can effectively reduce the load fluctuation.

Both randomized PHEV charging and HVAC control algorithms discussed above belong to direct or centralized load shaping, which has been heavily investigated. However, it is usually not clear how the users are compensated by providing load shaping services. In the last part of this thesis, we investigate indirect load shaping in a distributed manner. On one hand, we aim to reduce the users' energy cost by investigating how to fully utilize the battery pack and the water tank for the Combined Heat and Power (CHP) systems. We first formulate the queueing models for the CHP systems, and then propose an algorithm based on the Lyapunov optimization technique which does not need any statistical information about the system dynamics. The optimal control actions can be obtained by solving a non-convex optimization problem. We then discuss when it can be converted into a convex optimization problem. On the other hand, based on the users' reaction model, we propose an algorithm, with a time complexity of $O(\log n)$, to determine the RTP for the power company to effectively coordinate all the CHP systems and provide distributed load shaping services.

Contents

List of Tables

List of Figures

List of Abbreviations

CHP	Combined Heat and Power
CHPED	Combined Heat and Power Economic Dispatch
DR	Demand Response
HVAC	Heating Ventilation and Air-Conditioning
MPC	Model Predictive Control
PHEV	Plug-in hybrid electric vehicles
RTP	Real-time Pricing
SOC	State of Charge
UPS	used uninterruptible power systems

ACKNOWLEDGEMENTS

This dissertation has benefited greatly from many people, some of whom I would like to thank here.

To begin with, I would like to express my greatest gratitude to my PhD supervisor Dr. Lin Cai, for her patient guidance and continuous support during my PhD study. Without her help, I cannot reach the current level within such short time. I'm also impressed by her expertise, commitment and enthusiasm for research, as well as the focus on details, which inspire me to pursue perfect during my research.

In addition to my supervisor, I would like to express my sincere appreciation to Prof. Xiao-dai Dong and Prof. Kui Wu as my thesis committee, and Prof. Hao Liang from University of Alberta as my external examiner, for taking time to review my thesis, attending my oral exam and their valuable advices on my research work.

There are also a number of professors, colleagues and friends I need to thank. Specifically, I would like to thank Dr. Jianping Pan for his constructive comments. I also want to acknowledge the help and support from Yuanqian Luo, Zhe Yang, Siyuan Xiang, Jianping He, Min Xing, Lei Zheng, Xuan Wang, Yi Chen, Haoyuan Zhang, Zhe Wei and Yue Li.

Last but not least, I would like to thank my wife Huamei Tian and my parents, for their endless love and support. I'm grateful to have all of you in my life.

Kan Zhou, Victoria, BC, Canada

DEDICATION

To my dearest wife and parents for their endless love

Chapter 1

Introduction

1.1 Background

In the past decades, electricity power generation from fossil fuel, including oil, coal, and natural gas, produces a lot of pollution to the environment all around the world. To reduce these harmful emissions and replace them by clean energy, people are trying to find alternative energy resources which are sustainable and environment friendly. The existing renewable sources include wind, solar and etc. However, the power generation from these renewable sources are usually intermittent and thus cannot be integrated into the current system easily. In addition, the current power system which has severed us for decades is becoming insufficient and inefficient to meet the increasing electricity demand. As a result, voltage sags, blackouts, and overloads are more frequent during the past decades around the world. All of these call for a revolution in the current power grid.

With the help of communication and information technologies, the next-generation electricity power system, called *Smart Grid*, incorporates diversified renewable energy, and is featured with automated and intelligent management to help users and utility companies save cost [94, 62, 39, 40, 38]. Unlike conventional power plants which adjust the power supply according to the change of load, load adjustment is one of the most important new feature in Smart Grid. With smoother load variation, spin reservation can be reduced to save cost and improve efficiency.

Demand response (DR), which allows power generation and load to interact in an automated fashion based on information technology, is the most important method to coordinate demand and flatten load spikes. The main idea of DR is to manage

customers' electricity consumption in response to supply conditions or market prices. This is beneficial to both power companies and users because users can cut their energy bills by delaying elastic load to the time slots when the electricity price is low.

However, the application of demand response also introduces fundamental challenges. Without a good control algorithm, inappropriately controlled devices may lead to new peaks and affect users' comfortableness.

This dissertation is to study different demand response control strategies from various perspectives in Smart Grid. Specifically, we focus on the DR scheduling of Plug-in Hybrid Electric Vehicles (PHEV), Heating Ventilation and Air-Conditioning (HVAC) systems and Combined Heat and Power (CHP) systems from the perspective of both the power company and the end users. The reason we choose them is that they are the most typical elastic loads which are heavily investigated in recent research papers.

1.2 Research Problems

1.2.1 PHEV Charging Scheduling to Flatten Load Peaks

To reduce the dependence on fossil fuel and eliminate harmful emissions to the environment, PHEV has attracted more and more attention. Most vehicle companies have introduced new PHEV models to the market in recent years. In addition to its environment-friendliness, PHEV brings both challenges (due to its high electricity demand) and opportunities (thanks to the elasticity of its demand) to future Smart Grid. Without proper control, the charging of PHEV will create new peaks which are a heavy burden to the power grid.

How to manage the charging of PHEVs so the negative impacts caused by uncontrolled charging can be minimized has become an active research topic [94]. Generally speaking, the existing approaches can be classified into two categories: centralized and decentralized. By using the centralized approach, optimal charging schemes can be obtained through solving optimization problems. It has several disadvantages though.

First, centralized optimization needs several critical information, such as the schedule of arrival and departure of PHEVs, future inelastic electricity demand and power supply information, etc., which is difficult to obtain, particularly if the power supply includes renewable energy. Second, the complexity to solve optimization problems with many variables and constrains can be too high to be scalable. Third,

centralized control is not robust due to the single-point failure problem. Finally, centralized management may affect users' privacy, and it may not be acceptable to some customers.

Decentralized approaches usually rely on real-time pricing (RTP) to coordinate the distributed smart agents. In order to use RTP, the power plants need to broadcast the RTP for the next period of time before the demand scheduling decisions made by the smart agents, which may lead to harmful demand oscillations. For example, if the price is set too low, a large amount of elastic load will be turned on and causes peaks, and vice versa.

The above challenges motivate us to propose a decentralized access algorithm, which can efficiently coordinate all the distributed smart agents to avoid harmful peaks caused by PHEV charging on the high-voltage power grid. However, in this part, we do not consider the impact of PHEV charging on the distribution grid, such as bus congestion and voltage drop. In the next part, we will extend the proposed algorithm to the distribution grid so that both bus congestion and large voltage drop can be avoided even with a large PHEV population.

1.2.2 Randomized PHEV Charging Under Distribution Grid Constraints

For the decentralized PHEV charging algorithm, the previous work mainly focused on the grid constraints at the transport and high-voltage transmission grid [37, 97]. Recent research started to pay attention to the distribution grid. The two most common problems in the distribution grid are bus congestion and voltage drop. As we will discuss in Chapter 3, existing approaches mainly focus on centralized optimization technologies which need accurate predictions and may be difficult to solve within a short time period given a large grid size. In the low voltage grid, some centralized light-weight control algorithms were proposed, but they may not be easily extended to the whole distribution grid with a large population and high PHEV penetration.

According to our literature survey, there still lacks of a distributed scheduling approach for supporting a high PHEV penetration rate and considering the common distribution grid constraints. Therefore, we are motivated to propose a framework to regulate PHEV charging by considering the bus load congestion and voltage drop problems in the distribution grid. Different from the existing algorithms, our algorithm should be decentralized with a low complexity and can be executed in real

time. In addition, it should not rely on any accurate prediction on the load or PHEV arrival time.

In the above two parts, we mainly discuss the application of PHEV in DR. However, there is another category of elastic load, such as HVAC, which is also widely used to provide DR. The main difference between PHEV and HVAC is that the future electricity demand of PHEV is hard to predict as new PHEV will arrive at any time. However, given the number of HVACs and their current states, the future states of HVACs can be predicted based on our control actions and HVAC thermal model. We'll investigate the application of HVAC in the next part.

1.2.3 A Dynamic Water-filling Method for Real-Time HVAC Load Control

Due to its intermittent characteristics, integrating renewable energy into the power grid is challenging. To ensure power grid's stability, the generators need to standby to provide capacity reserve to meet the time-varying demand, which results in a low efficiency.

Demand response, aided by the current information and communication technologies, is anticipated to improve the grid stability and efficiency by interacting with the elastic load at users' side. By changing the elastic load w.r.t. both renewable energy generation and inelastic load variation, demand response can reduce the fluctuation of the non-renewable power generation and thus cut down the power generation cost.

To achieve this goal, existing works can be classified into two categories. In the first category, the authors assumed that how much demand response needed in each time slot is already known. Therefore, the aim of the algorithms is to use demand response to do a load following or load shaping according to a given control signal or control objective [43, 45]. However, in practice, it may be difficult to obtain the optimal control signal, in other words, to know exactly how much demand response is needed for each time slot in the future. As a result, the work in the second category usually assume the availability of some prediction information to help decide how much demand response may be needed. The key problem is that the amount of elastic load that can be adjusted at certain time (we call it "elastic load potential") may be limited. If we use too much elastic load to flatten the non-renewable power generation at the beginning, there may not be enough elastic load to use at a later time. Therefore, a tradeoff must be made between the fluctuation of non-renewable

power generation and elastic load potential.

Existing work, such as [77], usually needs accurate long-term load and renewable energy generation information to obtain the optimal non-renewable energy generation, which is called the water level, for each time slot. So how much elastic load is allowed in each time slot in the future can be obtained by simply calculating the difference between the water level and the predicted non-elastic load. The traditional water filling approach is to make the elastic load in each time slot as close to this difference as possible so that the non-renewable energy generation can reach the optimal value. However, without such accurate long-term estimation, we do not know the optimal water level and thus do not know how much elastic load should be adjusted in each time slot.

To overcome these challenges, we are motivated to propose a novel algorithm which aims to reduce non-renewable energy generation fluctuations while still guarantee user comfort level. To make the problem more practical, we assume only limited amount of elastic load and short-term renewable energy generation prediction are available. The main challenge of the problem is to make a tradeoff between non-renewable energy generation fluctuations and elastic load potential.

Up to now, both randomized PHEV charging and HVAC control algorithms discussed above belong to direct or centralized load shaping. However, it is usually not clear how the users are compensated by providing load shaping services. In the next part of this thesis, we will investigate indirect load shaping based on RTP.

1.2.4 The Scheduling of Combined Heat and Power Systems in Demand Response

Extensive research has been done aiming to reduce the users' electricity bill by taking the advantage of the RTP and the elasticity of certain appliances. However, it has been argued that without an appropriate RTP to coordinate all the elastic loads, these algorithms may lead to new peaks which are undesirable [1]. In order to solve the problem, one approach is to control the elastic load directly by a central controller. For example, in [44, 46] the HVACs can provide load shaping services if the ON/OFF states of each HVAC can be controlled by a control center directly. Others discussed how to determine the RTP to provide indirect load shaping mainly from a game theory perspective. In these papers, the authors usually assumed that the users make decisions according to a certain utility function. However, how to

design appropriate utility functions is still an open problem.

In this thesis, we are motivated to design an indirect load shaping service framework through RTP, which can help both the users and the power companies save cost. Different from the existing game theory approach, the user's reaction model is obtained by minimizing the long-term average cost. In addition, we propose a fast algorithm to determine the optimal real-time price which can effectively coordinate all the CHP system for load shaping services.

1.3 Dissertation Organization

The proposed thesis work is intended to discuss different control algorithms to provide demand response from the perspectives of power companies, customers and micro grid, respectively. In each chapter, we will present the introduction and motivation of the research topic, related works and our proposed methods, including the performance evaluation and future work.

The rest of this thesis is organized as follows. Chapter 2 discusses our research work on how to flatten load peaks in high voltage transmission grids. The design objective is to maximize the power utilization while guarantee that all the PHEVs' batteries are fully charged before their departure.

Chapter 3 considers the PHEV charging problem in a distribution grid, where more practical grid constrains like bus congestion and voltage drop for all the critical buses are restricted to a certain range.

In Chapter 4, we try to reduce the electricity load variation for the conventional power plants by controlling the amount of energy consumed by HVAC systems in each time slot. We first model the temperature evolution process of a room and propose an approach to estimate the key parameters of the model. Second, based on the model predictive control, a centralized and a distributed algorithm are proposed to minimize the fluctuation and maximize user comfort level. In addition, we propose a dynamic water level adjustment algorithm to make the demand response always available in two directions.

In Chapter 5, motivated by the queueing analysis and buffer management solutions in data communication systems, we investigate how to use a battery pack and a water tank to optimize the average cost for the CHP systems by jointly considering the real-time electricity price, renewable energy generation, energy buffer states, etc. On the other hand, based on the users reaction model, we propose an algorithm, with a time

complexity of $O(\log n)$, to determine the RTP for the power company to effectively coordinate all the CHP systems and provide distributed load shaping services.

1.4 Bibliographic Notes

Most of the works reported in this dissertation have appeared in research papers. The works in Chapter 2 have been published in [102]. The works in Chapter 3 have been published in [104]. The works in Chapter 4 have been published in [103], and those in Chapter 5 have been published in [105] and [106].

Chapter 2

PHEV Charging Scheduling to Flatten Load Peaks

2.1 Introduction

The development of PHEV is considered as a promising solution to the worldwide energy and environmental problems [7, 85]. Many automobile manufactures are introducing new models of PHEVs into the market. According to the estimation by the Department of Energy in US, about 1 million PHEVs will be sold by 2015 [36]. The impact of PHEVs on electric power systems, considering its relatively large population and charging load, cannot be ignored. Several studies [97, 17, 66, 75] have shown that, without proper control, the charging of a large number of PHEVs will cause huge peaks on the demand, which is dangerous to the power grid.

How to control users' elastic demand to reduce demand peaks and effectively use renewable energy despite its stochastic characteristics are key objects for smart grids. Existing solutions are either centralized which require accurate future predicted information and have a high computation complexity, or decentralized based on real time pricing (RTP) which may not be deployable immediately.

In this chapter, we introduce a new distributed approach based on a decentralized access technology, which can efficiently coordinate all the distributed smart agents to avoid harmful peaks caused by PHEV charging using the history information only. In addition, our algorithm can provide fast and automatic demand response, taking users' preferences and habits into consideration, and explore the potential of renewable energy despite its stochastic characteristic. Most importantly, it is simple to

deploy without the need to do accurate predictions on future demand and supply.

The main contributions of this chapter are threefold. First, we propose an online decentralized access algorithm for PHEV charging, which can effectively flatten peaks during PHEV charging at night. Meanwhile it can provide demand response intelligently when it is needed during peak hours. Our algorithm is simple and suitable to be executed on embedded systems like smart meters. Second, we determine the values of the control parameters and analyze the performance of the proposed algorithm. Finally, extensive trace-driven simulations using the real data obtained from National Household Travel Survey (NHTS) 2009 [1] and the load of the Electric Reliability Council of Texas (ERCOT) [18] have been conducted to evaluate the performance of the proposed algorithm for PHEV charging and demand response. The results demonstrate the advantages of our proposed solution.

2.2 Related Work

Centralized charging management typically assumed the knowledge of current and/or future demand and supply information [99, 48], the schedule of users [84], or the real-time electricity price [6]. These kinds of information may be difficult to obtain or predict accurately, which motivates the distributed approach.

Recent research [97] showed that, a deterministic on-/mid-/off-peak pricing policy may create new peaks because a large percentage of PHEV owners will choose to charge their vehicles during the off-peak time at a lower price. It was concluded that RTP is necessary with the popularity of PHEV and smart agent who controls the load in each house (include PHEV charging) intelligently [95, 61]. How to schedule the load of PHEV or other elastic load to minimize the overall cost based on the RTP model has been heavily investigated [95, 61, 93, 9]. For example, Vytelingum *et al.* illustrated an agent-based technology to manage micro-storage devices [93]. Wei *et al.* extended this approach by using machine learning in [95]. [61] used Q-learning to predict future electricity price and made a tradeoff between cost and waiting time of users. Chen *et al.* proposed an RTP-based power scheduling scheme to control residential load in [9].

However, if a large portion of Electrical Vehicles (EVs) and smart agents simply shift their load to the low-cost time slots (even with RTP, electricity price can be known beforehand or through prediction [52, 101, 68]), new peaks will appear [66, 67] and lead to undesirable oscillating effects. It is found that, by applying the RTP

mechanism only, demand peaks may not be flattened [66], and a Widrow-Hoff learning mechanism was proposed to gradually adapt the elastic load based on the predicted market prices to the peaks to a certain extent. It still requires a good price prediction, and the time to converge is long. [48] suggested that the smart agents need to report the tentative schedules to the central node back and forth a few times to find a suitable solution. Iordanis *et al.* aimed to minimize the long-term average power grid operation cost using dynamic programming in [34]. In their model, the variations of base load and renewable energy were not considered. [78] tried to reduce the power generation cost by flattening the overall load assuming that each device can be delayed arbitrarily. Briel *et al.* used the accurate future information to shift the elastic loads to specified time periods with different probabilities in [47]. [104] proposed a random access algorithm for PHEV charging focused on the constrains in the distribution grid, while this chapter aims to flatten the charging peaks. In [56], the authors proposed a game theory based algorithm to minimize the peak-to-average ratio of the aggregate load using distributed large batteries. Overall, how to develop a distributed PHEV charging solution without a complicated pricing strategy is still an open issue and motivates this work. The proposed algorithm in this chapter uses historical grid information to coordinate all the PHEVs. It has a low computational complexity and can achieve a performance close to its upper bound.

On the other hand, renewable energy is a promising solution for the shortage of fossil fuel and pollution. A key disadvantage of renewable energy sources is that they are stochastic and not stable. We are also motivated to use the elastic demand of PHEV charging to provide demand response, so the total load can follow the variation of renewable energy supply in real time without accurate predictions.

2.3 System Model and Problem Formulation

There are four main entities in the system: the power company, control center, smart agents and PHEVs.

Power Company: At time t, the power company has a capacity $S(t)$ to generate electricity with relatively low price which may include renewable energy that varies from time to time. If the demand exceeds this capacity, it may be expensive to generate or purchase the additional power (e.g., from gas turbines or import from other power companies). $S(t)$ is above the average electricity load and may be below the peak load when PHEVs are charged without control.

Control Center: Since the higher the load is, the fewer PHEVs have to wait and the less compensation power company should pay, the objective of the control center is to keep the instantaneous load below $S(t)$ while maximize the power utilization at any time t. The corresponding centralized optimization problem is formulated as follows:

$$\max_{L_{PHEV},\omega(k)} \quad u(t) = \frac{L_{base}(t) + L_{PHEV}(t)}{S(t)} \tag{2.1}$$

subject to:

$$L_{base}(t) + L_{PHEV}(t) \leq S(t) \tag{2.2}$$

$$\omega(k) \leq \omega_m(k) \tag{2.3}$$

where $\omega(k)$ is the waited time of PHEV k, $\omega_m(k)$ is the maximum tolerable delay of PHEV k, $L_{base}(t)$ and $L_{PHEV}(t)$ represent the base load without PHEV and the load of charging PHEVs at time t, respectively. To solve this optimization problem we need to know the base load information and the arriving, departure process of PHEVs in the future which are usually unavailable.

For decentralized control without such information, we assume that a proper smart grid communication infrastructure is available between the control center and all smart agents, and the communication delay and packets losses are negligible. The whole communication infrastructure contains three main communication sub-networks: Home Area Networks (HAN), Neighborhood Area Network (NAN), and Wide Area Network (WAN). Each PHEV (and other home devices) is controlled by a smart agent (which is assumed a smart meter) who manages the power supply of the HAN and makes scheduling decisions for elastic-load devices. Several smart meters in a community can form a NAN through either wired or wireless communication. Each data collector manages the bidirectional communication between the control center and a group of homes, collecting smart meter data and transferring information such as the total demand and supply of the last few minutes, control commands and etc.

We consider the system covered by one control center for simplicity. The control center collects two types of information, the current power generation capacity $S(t)$, including non-renewable and renewable energy, and the current demand, including both elastic and inelastic load. In each time slot (in the following, we use a minute as the duration of the time slot), the control center will calculate the ratio of total electricity demand over the total generation capacity of the last slot, γ, and then broadcast this ratio to all the smart agents. Meanwhile, it will monitor the condition

of the whole power grid and adjust some control parameters to assist the scheduling decisions of smart agents if necessary.

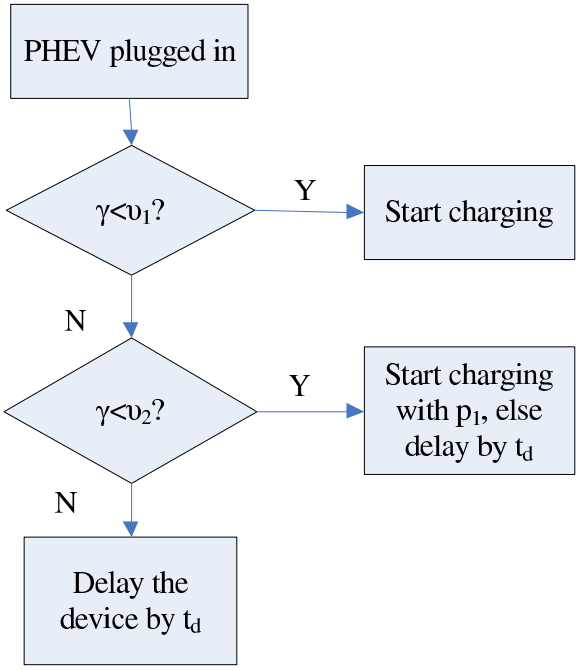
Smart Agent: The smart agent (or, the smart meter) can schedule and manage the power usage in a house. All the elastic load, like washing machine, PHEV, dish washer and thermal loads, can be managed by this smart agent. It can also receive information and instructions from the control center, and use this information to make decisions according to the algorithm described in Section 2.4. All the houses which adopt our algorithms are called volunteers. These volunteers will be compensated by the electricity company depending on the contributions they made. (How to determine the contribution and design the incentive mechanism is an interesting problem for future research.)

PHEV: We assume that a PHEV is plugged-in when it arrives home. Meanwhile, the departure time of the PHEV is set either by the user or by the smart agent according to historical departure time. Then the maximum tolerable delay for PHEV charging is calculated based on the current battery status, charging power and the total parking time. The delay time for PHEV charging should be guaranteed to be less than the maximum tolerable delay time.

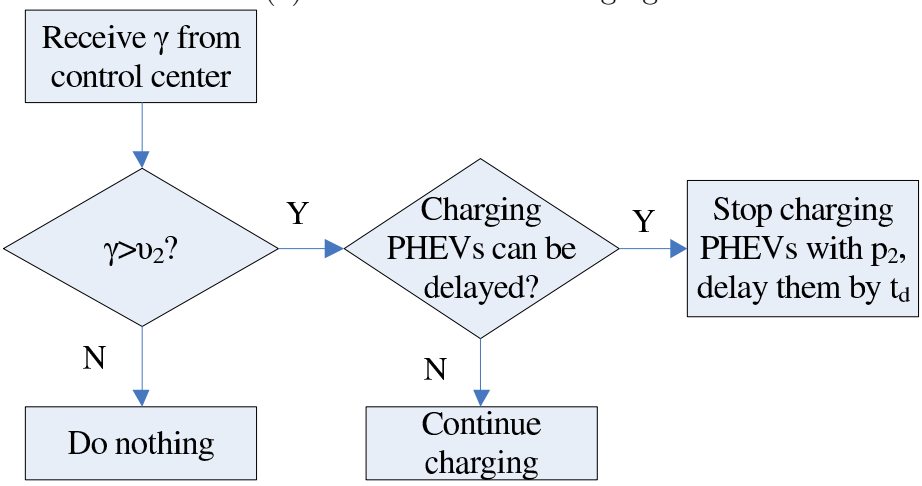
2.4 Proposed Random Access Scheme

The design objective of our algorithm can be summarized as follows. First of all, to ensure low power generation cost, the total power load should be no larger than $S(t)$. Besides, the power utilization $u(t)$ should be maximized. Second, all the PHEVs should be charged within the maximum tolerable delay. Third, users' preferences can be considered. Fourth, providing demand response so the stochastic renewable energy can be efficiently utilized. Finally, the algorithm should not rely on accurate predictions and should be simple enough for real-time control.

Here, we first present a brief overview of the proposed design as shown in Fig. 3.2, and the detailed design and the parameter settings will be discussed in the following subsections. To meet the first and most important design objective, when the current ratio γ is relatively high (larger than threshold one), access by elastic load should be restricted; when γ is even larger than threshold two, demand response should take effect to terminate some charging PHEVs to maintain the load within a safe level. Therefore, when a user plugged the PHEV in, the smart agent will check the recent γ received from the control center. If γ is less than threshold one (denoted by ts1,



(a) Schedule PHEV charging



(b) Demand Response

Figure 2.1: The flow chart for smart agent

with the value of ν_1), it will allow the PHEV start charging immediately; otherwise, it will use a back-off algorithm as follows. When $\nu_1 \leq \gamma < \nu_2$, with probability p_1 the PHEV will start charging immediately. With probability $1 - p_1$, the request will be delayed by t_d.

Once γ is greater than threshold two (denoted by ts2, with the value of ν_2), none of the PHEVs is allowed to start charging unless one reaches the maximum tolerable delay time. The reason is that ts2 represents a level that the demand is very close to $S(t)$ and if the demand keeps increasing, the power generation cost may increase tremendously. In this case, in each slot when $\gamma \geq \nu_2$, a charging PHEV will terminate its charging with a probability p_2 and wait for t_d slots to try again, until γ falls below ν_2 again. By eliminating a part of the PHEV charging load each slot probabilistically, a fast demand response can be achieved.

Notice that the user can always let the smart agent start charging the PHEV immediately, in this case, the PHEV becomes inelastic load and the user will not be compensated by the power company.

Design of the access probability, p_1

Considering the design objectives, p_1 is designed to be a function of the current demand/supply ratio, γ, the value of ts1, ν_1, a parameter δ_1 to reflect the user's preference, and a global parameter κ_1 used by the control center for global adjustment if needed as follows.

$$p_1(\gamma) = \begin{cases} \kappa_1 e^{-\alpha(\gamma - \nu_1)} + \delta_1, & \text{if } \omega < \omega_m; \\ 1, & \text{if } \omega = \omega_m, \end{cases} \tag{2.4}$$

where ω is the current waited time for a PHEV, and ω_m is the maximum tolerable delay time.

In the above design, p_1 decreases exponentially when γ exceeds ν_1, so fewer PHEVs will be allowed to start charging to keep demand below $S(t)$. The global parameter κ_1 can be set by the control center through notification messages, and it is the same for all the users. By increasing κ_1, the control center can increase the probability to admit more PHEV charging load, and vice versa. By default, this value is set to 1 and usually does not need to be changed frequently.

α is the parameter denoting how fast p_1 will become 0 when the current ratio γ is greater than ts1. To ensure $p_1(\nu_1) = 1$ and $p_1(\nu_2) = \varepsilon$ when $\omega < \omega_m$, where ε is a

small positive number, α can be determined as follows:

$$\alpha = \frac{ln\varepsilon}{\nu_1 - \nu_2}. \tag{2.5}$$

On the other hand, if this PHEV charging request is delayed (with probability $1 - p_1$), the delay time t_d can be calculated as follows: if $\omega_m - \omega > t_m$, then t_d is randomly selected from $[0, t_m]$, where t_m is the upper bound for the delay; otherwise, t_d is set as $\omega_m - \omega$.

Design of the charging suspend probability, p_2

p_2 represents the probability to suspend the charging of a PHEV when the total load is larger than threshold two, and the demand response mechanism is triggered. Obviously, p_2 should be small when γ is only slightly larger than ν_2 to avoid suspending too many charging PHEVs which may be unpleasant to users. Also, p_2 should increase rapidly when γ is close to 1. Therefore, we also use an exponential function to design p_2:

$$p_2(\gamma) = \begin{cases} \kappa_2 e^{\phi(\gamma-1)} + \delta_2, & \text{if } \omega < \omega_m; \\ 0, & \text{if } \omega = \omega_m, \end{cases} \tag{2.6}$$

where κ_2 is a global parameter set by the control center to adjust the speed of suspending charging PHEVs, δ_2 is a parameter which represents the preferences of each user, similar to what δ_1 does, and ϕ represents how fast $p_2(\gamma)$ will reach 1. To calculate ϕ, let $p_2(\nu_2) = \varepsilon$ and $p_2(1) = 1$, where ε is a very small positive number, and we have

$$\phi = \frac{ln\varepsilon}{\nu_2 - 1}. \tag{2.7}$$

If a PHEV suspends charging, after waiting t_d, the smart agent will determine whether to let it resume charging or continue keeping it suspended. Similarly, if $\gamma < \nu_1$, the smart agent will start charging the PHEV immediately; if γ is between ν_1 and ν_2, this PHEV will start charging with a probability p_1 or be delayed by t_d with probability $1 - p_1$; else if $\gamma > \nu_2$, the smart agent will suspend the PHEV for another t_d, unless it reaches its maximum tolerable delay. Note that if the demand response is not fast enough, the control center can adjust the global parameter κ_2 to increase the probability to suspend the PHEV charging.

We will describe how to determine ν_1 and ν_2 in the next section.

2.4.1 Other Design Objectives

The main behavior of the smart agent is illustrated in Fig.3.2. Note that, to ensure the maximum tolerable delay of each PHEV, when $\omega = \omega_m$, the PHEV will start or continue its charging whatever γ is, assuming there are enough elastic load to be controlled.

From the above description, the schedules of all the tasks for each house are adjusted automatically and in a distributed manner. A fast response to the power supply change can be achieved. For example, if the wind farm produces more energy, then γ decreases, and more PHEV elastic load can be turned on within a short time. If the base demand (which is the inelastic load) keeps increasing while the generated renewable energy is not sufficient, demand response will take effect when $\gamma > \nu_2$ to decrease the total load by terminating some elastic load to avoid sharp peaks. Therefore, the last three design objectives are also met.

2.4.2 Further Discussion

Usually, the charging period for the PHEVs in the residential area is at night. In the daytime, there may not be enough PHEVs to provide demand response. However, although our algorithm is designed for PHEV charging, it can also be used to manage other elastic loads used during the daytime, such as water heater, wash machine, etc., to provide demand response. The main difference is that the maximum tolerable delay may be different for different appliances, and the amount of the load is also heterogeneous. For example, the maximum tolerable delay for water heater is determined by the current water temperature, environment temperature, and water quantity, etc. In this case, the smart agent can either use artificial intelligence to predict this value or just simply delay the water heater until the water temperature is lower than a predefined threshold. Besides, each appliance can have its own preference values of δ_1 and δ_2 described in (5.11) and (2.6). These parameters can also be adjusted dynamically so different priorities of appliances can be achieved.

For example, the user might prefer the wash machine to be terminated first rather than the water heater. In this case, δ_2 for the water heater might be negative before the wash machine is stopped. Although we can adjust different preferences for different appliances within a house, how to coordinate heterogenous appliances in different places is left for future research.

2.5 Performance Analysis

The proposed algorithm has two main performance metrics: the power utilization at time t and the probability of the total load exceeding $S(t)$. Since the slot duration is short, we assume $S(t)$ has a constant value during one time slot. Similar to [2, 3, 86], we also assume the arrival of PHEV follows a poisson distribution with a maximum arrival rate λ, and the time needed to charge a PHEV is exponentially distributed with parameter μ.

2.5.1 Power Utilization

State n represents that there are n PHEVs charging in this power system. In this part, we use queuing theory to obtain the probability of each state. To simplify the analysis, we assume that the ratio γ is broadcasted in real time and the maximum tolerable delay time for each PHEV is infinity. Therefore L_{PHEV} can be considered as a continuous-time Markov chain. Since the charging power of one PHEV is very small compared to the total power supply, the transition rate from state n to $n+1$ can be approximated as $\lambda p_1(n)$. We have

$$p_1(n) = \begin{cases} 1, & P_c n < \nu_1 S; \\ p_1(\gamma), & \nu_1 S \leq \gamma S = P_c n < \nu_2 S; \\ 0, & \nu_2 S \leq P_c n \leq S, \end{cases} \tag{2.8}$$

where P_c is the average charging power of PHEV, n is in the range of $[0, N)$, N is the maximum number of PHEVs the power system can support, and S is the value of $S(t)$ in the considered time slot.

The transition rate from state $n+1$ to n is $(n+1)\mu$. The balance equation for the steady state situations is:

$$\lambda p_1(n) P_n = (n+1)\mu P_{n+1} \tag{2.9}$$

After algebraic manipulation, we have

$$P_n = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n \prod_{k=0}^{n} p_1(k) P_0. \tag{2.10}$$

With the condition

$$\sum_{n=0}^{N} P(n) = 1, \tag{2.11}$$

we can obtain the probability of each state. Then the expected power utilization can be calculated as follows:

$$E(u) = \frac{1}{S} \sum_{n=0}^{N} P(n) \cdot P_c n. \tag{2.12}$$

2.5.2 The Probability of Total Demand Exceeding $S(t)$

Theorem 1: When the ratio γ is broadcasted in real time, the probability that power demand will exceed S is bounded.

Proof: When the real-time information is available, the arrival/departure of PHEVs can be considered as a continuous Markov chain. Since the charging probability when $\gamma > \nu_2$ equals zero, the only situation that the power load increases is that one of the waiting PHEVs reaches its maximum tolerable delay time. According to our proposed algorithm, once γ is larger than ν_2, all the charging PHEVs will stop with a probability to provide demand response. There are K_1 states below ν_2 and K_2 states between ν_2 and S, and the upper bound of the probability that the demand will exceed S is

$$P_u = \prod_{i=1}^{K_2} (1 - p_2(\gamma_{K_1+i}))^N, \tag{2.13}$$

where γ_{K_1+i} is the ratio when the system is in state $K_1 + i$.

However, the broadcasted ratio γ may not be available in real-time in practice. During a broadcast interval t, there is always a probability that more than the expected number of PHEVs arrive, and thus the power demand may exceed the low-price power generation capacity $S(t)$. This probability is strongly related to ν_1, ν_2 and the PHEV arrival rate. Assume that the maximum arrival rate the system can support is λ, the maximum allowed probability to exceed $S(t)$ during interval t is τ_s, and we can use these criteria to determine ν_1 and ν_2.

Given γ, the number of new PHEVs the system can support without exceeding S is

$$m_1 = \left\lfloor \frac{S \cdot (1 - \gamma)}{P_c} \right\rfloor. \tag{2.14}$$

The probability that the demand will exceed S can be obtained using the following

equation

$$P_s = \sum_{k=m_1+1}^{+\infty} \frac{e^{-\lambda t}(\lambda t)^k}{k!} \cdot \left(\sum_{i=m_1+1}^{k} \binom{k}{i} p(\gamma)^i (1-p(\gamma))^{k-i} \right), \tag{2.15}$$

where $p(\gamma)$ is the charging probability when the ratio equals γ.

Let $P_s < \tau_s$, from (2.15) we can obtain the value of $p(\gamma)$ for each γ. According to our algorithm, ν_1 can be set to the largest γ when $p(\gamma) = 1$. Since the charging probability when $\gamma = \nu_1$ equals 1, to determine ν_2, let $\gamma = \nu_1$, and the probability that the ratio will exceed a certain $\gamma_t > \nu_1$ after slot duration t can be represented as follows:

$$P_e(n > m_2) = 1 - \sum_{k=0}^{m_2} \frac{e^{-\lambda t}(\lambda t)^k}{k!}, \tag{2.16}$$

where

$$m_2 = \left\lfloor \frac{S \cdot (\gamma_t - \gamma)}{P_c} \right\rfloor, \gamma < \nu_2. \tag{2.17}$$

Let $P_e < \tau_e$, where τ_e is a threshold set by control center and $\tau_e > \tau_s$, we can obtain the smallest γ_t which is set as ν_2.

On the other hand, given ν_1 and ν_2 we can calculate the maximum PHEV arrival rate λ_m the system can support with Proposition 1.

Proposition 1: If the probability to exceed ν_2 and S is less than τ_e and τ_s respectively after the slot duration t when $\gamma = \nu_1$, the probability to exceed ν_2 and S is always less than τ_e and τ_s respectively with any other γ not equal to ν_1.

Proof: To prove it, we only have to prove that the power system can support fewest number of arrivals per slot when $\gamma = \nu_1$. We consider the following three situations:

(1) $\gamma < \nu_1$. PHEVs will begin to charge with probability 1. Therefore the number of arrivals in a slot t that the power grid can support is

$$n_{m1} = \left\lfloor \frac{C}{P_c} \right\rfloor = \left\lfloor \frac{S}{P_c}(\nu_2 - \gamma) \right\rfloor > \left\lfloor \frac{S}{P_c}(\nu_2 - \nu_1) \right\rfloor, \tag{2.18}$$

where C is the available power for new PHEVs and P_c is the standard charging power per PHEV.

(2) $\nu_1 \leq \gamma < \nu_2$. PHEVs will begin to charge with probability p_1. Therefore the expected number of PHEVs (denoted as n) the power grid can support is:

$$n \cdot P_c \cdot p_1 \leq C. \tag{2.19}$$

After algebraic manipulation, we have

$$n \leq \frac{S}{P_c}(\nu_2 - \gamma)e^{\alpha(\gamma-\nu_1)}. \tag{2.20}$$

Let $f(\gamma) = \frac{S}{P_c}(\nu_2 - \gamma)e^{\alpha(\gamma-\nu_1)}$. In order to find out the system capacity n_m, we calculate the gradient of $f(\gamma)$:

$$\nabla f(\gamma) = \frac{S}{P_c}(\alpha\nu_2 - \alpha\gamma - 1)e^{\alpha(\gamma-\nu_1)}. \tag{2.21}$$

Substituting (3.12) into (2.21), we have

$$\nabla f(\gamma) = -\frac{S}{P_c}(1 + \frac{\gamma - \nu_2}{\nu_1 - \nu_2}\ln\varepsilon)e^{\alpha(\gamma-\nu_1)}. \tag{2.22}$$

Because ε is a very small positive number, $\nabla f(\gamma) > 0$ when $\nu_1 \leq \gamma < \nu_2$. Thus $f(\gamma)$ is an increasing function in the range $[\nu_1, \nu_2)$. Hence the number of arrivals that the power grid can support is

$$n_{m2} = \min(\lfloor f(\gamma) \rfloor) = \left\lfloor \frac{S}{P_c}(\nu_2 - \nu_1) \right\rfloor. \tag{2.23}$$

(3) $\gamma \geq \nu_2$. In this case, none of the PHEVs will begin to charge, so no matter how many PHEVs arrive, the load will not be affected. Therefore $n_{m3} \to +\infty$.

Among the above three situations, we can find that the power system can support the fewest number of arriving PHEVs when γ equals ν_1. The proof is complete.

In other words, the calculation of λ_m is greatly simplified because we only have to ensure that the probability to exceed ν_2 and S is less than τ_e and τ_s respectively with λ_m when $\gamma = \nu_1$.

In the above analysis, we assume that the base load profile and power generation capacity $S(t)$ do not change. To be more practical, we can further consider the maximum increased load D_m caused by other devices and the maximum reduction in power supply R_m caused by stochastic renewable energy in a slot obtained from historical statistic data, and subtract them from S when calculating ν_1 and ν_2. Therefore, the smaller the interval t is, the smaller D_m and R_m will be, and the system will be more efficient.

In addition, from (2.15) we can find with larger interval time t and fixed P_s, γ, the charging probability $p(\gamma)$ is smaller. Hence ν_1 and ν_2 will also be smaller which

Table 2.1: PHEV types and their key parameters

PHEV Types	Battery Capacity	Max Range	Market Share
Auto	24 kWh	73 miles	49.9%
SUV	37.6 kWh	80 miles	19.4%
Pickup	30 kWh	55 miles	17.8%
Van (and others)	36 kWh	60 miles	12.9%

reduces the power utilization. As a result, a tradeoff must be made between the communication overhead and power utilization.

2.6 Simulation

The objectives of the simulation are twofold: (a) to evaluate the performance of our algorithm on PHEV charging, and (b) to study whether the proposed algorithm can adapt to the change of energy supply by providing automatic demand response.

2.6.1 PHEV Charging

In our simulation, the vehicle data are obtained from National Household Travel Survey (NHTS) in 2009 [1], which gave the travel patterns of light-duty vehicle (LDV) fleet in USA. In highway travel, LDV accounts for 92% of the vehicle miles traveled (VMT) [90], 76% of the energy consumed [59], and 74% of the emitted carbon dioxide [60]. We assume that the PHEV owners' preferences to vehicle types and their driving behaviors will be similar to those of the conventional vehicle owners.

PHEV Type

The charging power and battery capacity are determined by the PHEV type. From the NHTS report, vehicles can be classified into 4 categories, auto, sport utility vehicle (SUV), pick-up trucks and van. To model their charging loads, we use the battery capacity for the 4 EV prototypes in [25], as shown in Table 2.1.

According to the NHTS report, the average daily traveling miles for male and female drivers are 41 miles and 32 miles respectively. The state of the battery for different types of PHEV when they arrive home and begin to charge is different. Assume the numbers of male and female drivers are equal, so the average daily miles are 36.5 miles, corresponding to 50%, 45.6%, 66.3% and 60.8% of the batteries for each

kind of PHEVs respectively. In our simulation, the initial state of the battery follows a truncated normal distribution with the mean described above and the standard deviation equal to 10% of its capacity for simplicity.

The charging power standard is obtained from [23]. Similar to [97], we use a linear battery model in our simulation. According to the charging power standard [23] in the residential area, for a typical household, we set the PHEV charging power to be 2 kW. The time needed to charge a PHEV (T_c) is calculated as follows:

$$T_c = \frac{\text{Battery Capacity} - \text{Battery Remaining Energy}}{\text{Charging Power}}. \tag{2.24}$$

Number of Vehicles

The number of PHEVs in a certain community depends on the population size, the ownership ratio of vehicles, and the PHEV penetration ratio. From the data of Major Travel Indicators of 2009 in USA [1], there are on average 2.50 persons per household while the number of vehicles per household is 1.86, so the vehicle ownership ratio is $1.86/2.5 = 0.744$ per person. Considering the population size of 4000 of the simulated community, the number of vehicles is about 3000. If the PHEV penetration ratio is 0.2, the number of PHEV is 600.

Driving Habits

The start time of charging and the maximum tolerable delay time have a strong relationship with the habits of drivers. Assume that all the PHEVs are connected to the power grid immediately when they arrive home, and the maximum delay is set by the users (or by the smart agent which makes prediction based on history data). According to the analysis of National Household Travel survey [1], the vehicle arriving home time (plug-in time for PHEVs) can be approximated by a truncated normal distribution. In addition, since most of PHEVs begin to charge at night, we assume the arrival and departure time of PHEVs both follow a truncated normal distribution with the mean of 7 pm and 7 am respectively, and the standard deviation of one hour. The charging hours include peaks of electricity base demand, so our simulation can also capture the behaviors of our algorithm in reacting to the peak hours. These assumptions are adopted in others work on the grid integration of PHEVs, such as [75].

Grid Load Profile

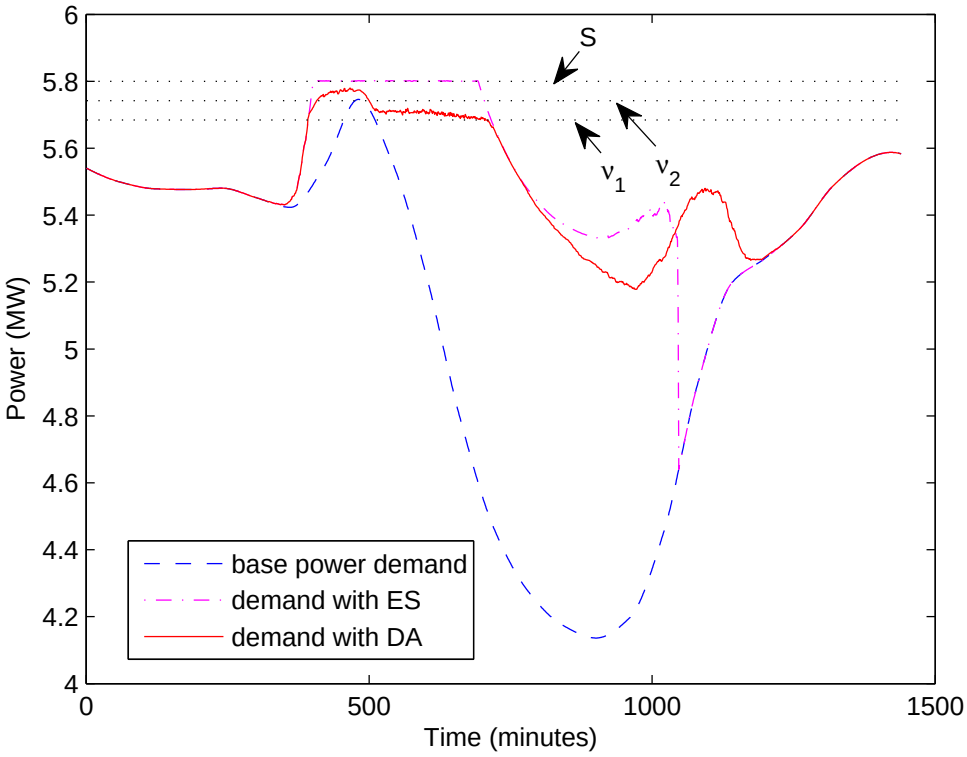
The electricity load profile for the base demand is obtained from the real load measurements in the Reliability Council of Texas (ERCOT) [18]. ERCOT is an isolated and independent electrical system supplier which provides electricity power to 23-million people in Texas. We choose the hourly load profile on March 15th and 16th in 2011 and interpolate the 24 hour data into a curve which consists of 1440 minutes' data of a day.

In our simulation, we scale down the population by considering a community of 4000 people only. The total electricity load from [18] is scaled down correspondingly, and the load from PHEV charging is superimposed into the base load. Each house is equipped with a smart agent which can schedule the elastic load in the house.

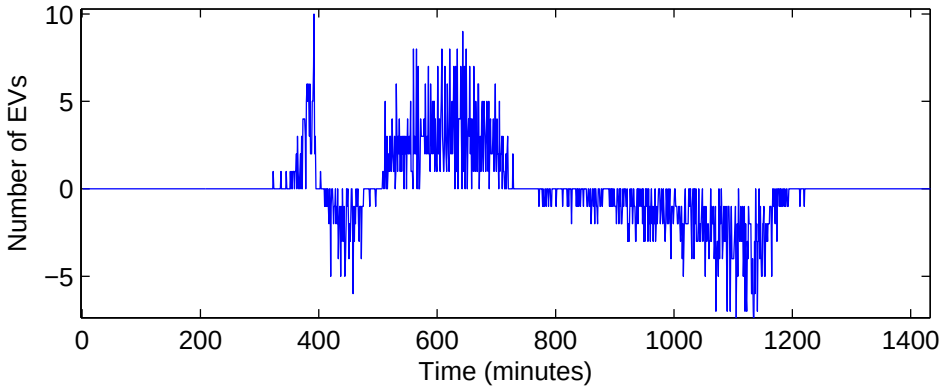
Simulation Results

Since the system model and the design objective of this chapter are difference from the existing work, we compare the performance of the proposed decentralized access (DA) algorithm with that of exhaustive search (ES). We first set the power generation capacity $S(t)$ to be a constant S. The base load, power supply and PHEV arrival process are assumed available to solve the optimization problem (3.2) using exhaustive search. To maximize the power utilization, we also assume that the charging PHEVs can always be stopped at arbitrary time instance which is different from our proposed algorithm that does not allow PHEV charging to stop in the middle unless the ratio is larger than threshold two to protect the battery. The default parameter settings in our algorithm are as follows: $\kappa_1 = \kappa_2 = 1$, δ_1 and δ_2 is uniformly distributed between -0.05 and 0.05, $t_m = 30$ minutes, t_d follows a uniform distribution in the range $[0, t_m]$, $\nu_1 = 0.98$, and $\nu_2 = 0.99$.

First, the number of PHEVs is set to 600. As shown in Fig. 2.2(a), PHEV charging using exhaustive search can utilize all the available power to charge PHEVs. The power utilization is close to 1 at peak time as expected. However, some information used in exhaustive search is not available in practice. From the curve representing the aggregated load using our algorithm, it is effectively flattened during the peak hours. The load is restricted between ν_1 and ν_2. With a quite small performance gap from the exhaustive search result, our algorithm does not need the base load value in the future and PHEV arrival information beforehand and is implemented in a distributed manner which provides both scalability and simplicity. The variation of the number

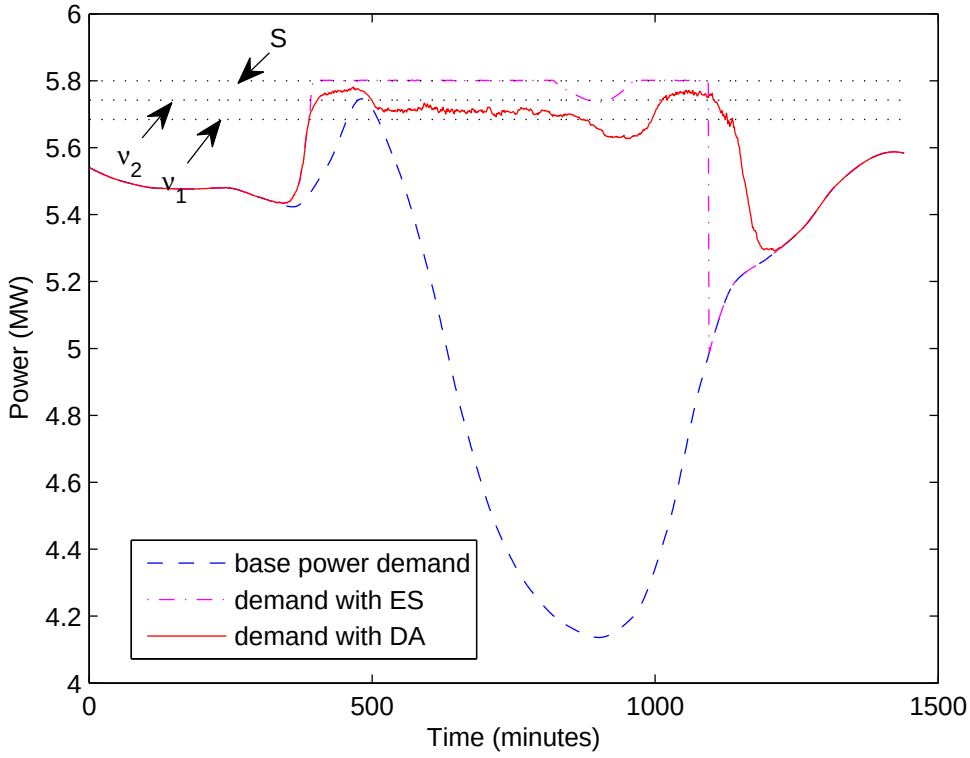


(a) Power demand variation

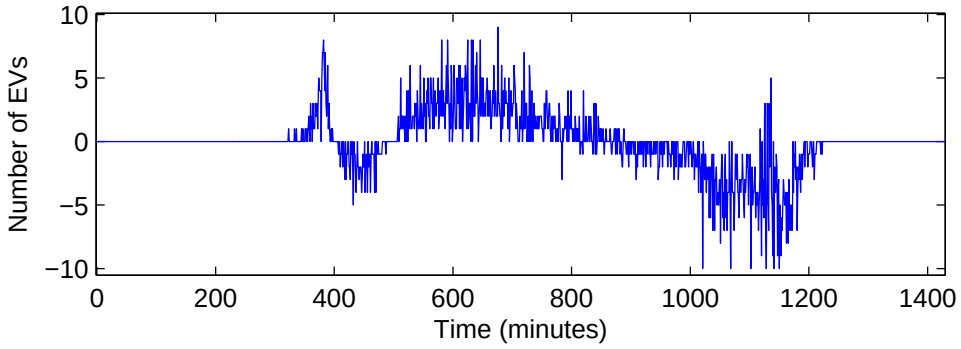


(b) Number of charging PHEVs variation

Figure 2.2: Power demand and number of charging PHEVs, 600 PHEVs



(a) Power demand variation



(b) the Number of charging PHEVs variation

Figure 2.3: Power demand and number of charging PHEVs, 900 PHEVs

of charging PHEVs (the number of those starting to charge minus those stopping to charge) in each minute is shown in Fig.2.2 (b).

From the figures, when the base load is below ν_1, all the arriving PHEVs begin to charge immediately, so the aggregated loads with or without the proposed control overlap. Once the demand exceeds ν_1, a portion of the arrival PHEVs are delayed. During this time, once a PHEV begins to charge, it will continue charging until finish. Thus, the load under control keeps increasing. Once the load meets ν_2, charging PHEVs begin to be suspended which quickly decreases the aggregated load to the base load.

In the peak time period from minute 400 to 690, when there are PHEVs waiting to be charged, the average power utilization for ES and DA are 99.99% and 98.69%, respectively.

Second, we increase the PHEV number to 800, and the results are shown in Figs. 2.3 (a) and (b). From the figures, our algorithm keeps the total demand below ν_2 quite well during peak time although the PHEV number has been increased by 33.3%. During the peak time from minute 400 to 1000, the average power utilization of DA reaches 98.37%, which is only about 1.48% lower than that using exhaustive search. However, at the end of the charging period, the power demand of the proposed algorithm exceeds ν_2 slightly. From Fig. 2.3 (b) we can notice that when the load exceeds ν_2, some PHEVs stopped to provide demand response, but since most PHEVs have finished charging, the number of available PHEVs which can provide demand response is very small, and that is why when the base load increases, the total load is still above ν_2. One possible solution is to allow some charging PHEVs to stop in the middle and provide a chance for other waiting PHEVs to start charging, and giving a higher charging priority to the PHEVs which have shorter remaining tolerable delay time. However, whether stopping charging PHEVs during the charging process very often is harmful to the battery is debatable. A tradeoff must be made between the average number of times a PHEV can be stopped and the available elastic number of PHEVs based on different system states, which is left for future research. Nevertheless, the proposed solution can ensure that the total load is below S for the whole simulation time.

2.6.2 Demand Response by Other Elastic Load

Next, we investigate how our algorithm can make fast and automatic demand response to adapt to the change of renewable energy supply.

Elastic Load Modeling

In this simulation, we use the load profile on March 15th 2011 of ERCOT. Instead of PHEV, we here use some elastic appliances typically operating during daytime. Similar as PHEV, these appliances will not be suspended once they are started unless γ exceeds ν_2. There are 600 elastic appliances, the power of which is chosen from 1 kW, 1.5 kW and 2 kW with equal probability. To generate the ON time periods for these elastic loads, we assume that the probability for each appliance to turn ON follows a Poisson distribution at each half-hour period, with the mean of 2 times during the 13 hours from 6 am to 7 pm. After 7pm, they are turned on with a probability of 0.08 every 30 minutes. Each appliance will run 30 to 90 minutes, and the maximum tolerable delay ranges from 3 to 4 hours. The load of the elastic appliances is superimposed with the base load profile.

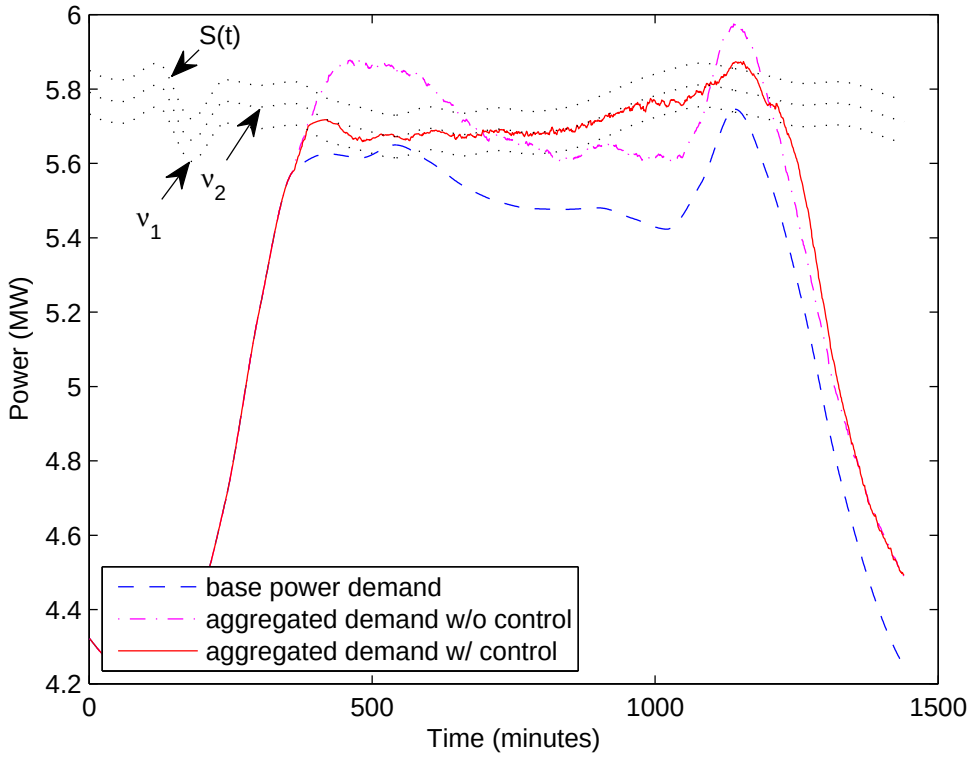
Renewable Energy

The model of the renewable energy in our simulation is obtained from Wind Integration Study [14]. We used the raw data of a typical daily wind generation, scaled it down, and then added it to the original power supply.

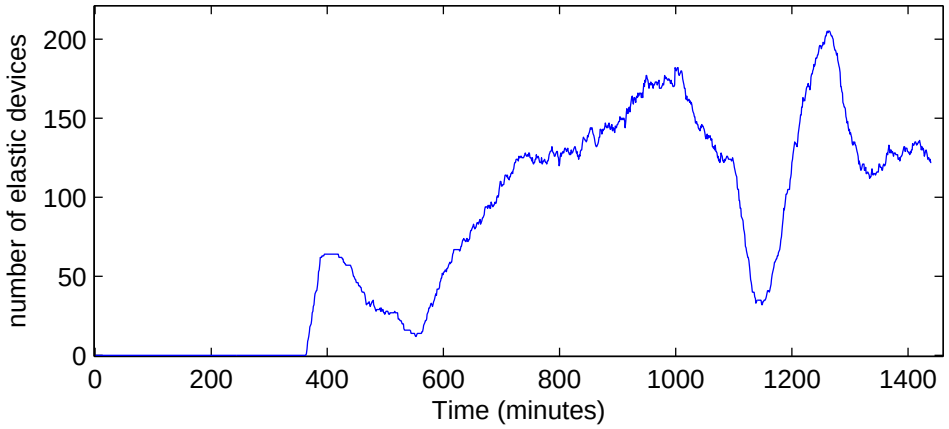
Simulation Results

The objective of this simulation is to test whether our proposed algorithm can follow the changes in renewable energy supply to provide effective demand response. The renewable energy penetration in our simulation is only 3.3%, given the relatively small amount of the elastic load considered. In practice, the base load from real data should also contain the load from elastic appliances, and by adjusting them, a higher renewable energy penetration ratio can be supported.

The simulated power demand with renewable energy is shown in Fig. 2.4 (a). As shown in Fig. 2.4 (a), the controlled load curve can follow the changes in energy supply nicely, while the uncontrolled load exceeds power supply twice. Fig. 2.4 (b) shows the number of elastic appliances being operating in each minute. As expected, when



(a) Power demand and supply variations



(b) Number of turned on appliances

Figure 2.4: Demand response provided by appliances in the daytime

the load exceeds ν_2, most of the appliances are delayed or shut down automatically to provide demand response for the grid. Therefore, two valleys appears in Fig. 2.4 (b), corresponding to the time when demand response is provided to bring down the aggregated load below the supply.

On the other hand, when the renewable energy supply is above the average, our algorithm will turn some delayed appliances ON to efficiently utilize the renewable energy. Unlike other centralized or decentralized algorithms discussed in the literature, our algorithm can provide fast response in a distributed manner and do not need to have an accurate prediction on renewable energy supply.

2.7 Conclusion

In this chapter, we have proposed a decentralized control algorithm for PHEV charging in smart grid to avoid severe new power demand peaks, and it can provide automatic demand response when needed. We have further discussed how to fine tune the algorithm and system parameters, and analyzed the performance bound of the proposed algorithm. By real data trace driven simulations, we have shown that, using the proposed distributed algorithm, without real-time pricing or accurate prediction on power demand and supply, peaks caused by PHEV charging can be controlled to be below the power generation capacity, stochastic renewable energy can be efficiently utilized, and users' preference can be considered. This work has suggested a promising direction on coordinating decentralized smart agents in smart grid.

However, in this chapter, we only consider the influence of PHEV on the high-voltage power grid. Without proper control, the charging of PHEV may also have high impact on the distribution grid, such as bus congestion and voltage drop. In the next chapter, we are going to extend the proposed algorithm to the distribution grid so that both bus congestion and large voltage drop can be avoided even with a large PHEV population.

Chapter 3

Randomized PHEV Charging Under Distribution Grid Constraints

3.1 Introduction

PHEVs are becoming increasingly popular. The energy department of USA estimates that more than one million PHEVs will be sold by the end of 2015 [36]. In addition to its environment friendliness, the adoption of a large number of PHEVs will exert great pressure on the current power grid due to its high power demand [97]. As a result, appropriate actions are needed to eliminate any possible harmful impact, which sparks numerous research efforts.

The previous work mainly focused on the grid constraints at the transport and high-voltage transmission power grid [37, 97]. However, with a high PHEV penetration rate, the existing distribution grids which are built decades ago are more likely to face bus congestion and voltage drop problems. Without proper control, the charging of PHEVs will cause harmful impact on the power distribution grid.

In this chapter, we propose a framework to regulate PHEV charging by considering the bus load congestion and voltage drop problems in the distribution grid. Different from the existing algorithms, our algorithm is decentralized with a low complexity. No complex optimization problem needs to be solved. And it does not rely on any accurate prediction on load or PHEV arrival time and can be executed in real-time. In addition, our approach takes the delay constraints of PHEV charging into consider-

ation. Finally, it can be extended to include other elastic loads to provide automatic demand response to protect the power grid and improve its efficiency.

The main contributions of this chapter are three-fold. First, a decentralized algorithm is introduced, which can efficiently avoid bus congestion and large voltage drop in the distribution grid with charging PHEVs. The smart agents schedule the PHEV charging independently based on the received information of the current grid status from a control center. Second, we analyze the performance of our algorithm and derive the system capacity. Finally, extensive simulation with real data from National House Hold Travel Survey 2009 [1] and the RELOAD database [13] from national energy modeling system are conducted to evaluate the performance of the proposed algorithm on a typical resident area distribution grid [65].

3.2 Related Work

J. Taylor *et. al* demonstrated that a high PHEV penetration rate would result in loads exceeding current bus capacity through simulation based on a real data model [88]. In [32], the authors used load flow analysis to show the impact of PHEV on the distribution grid. J. A. *et. al* investigated the impact of PHEV charging on medium voltage grid, considering the bus load congestion, and voltage drop problems [65].

To solve the problems listed above in the distribution grid, O. Sundstrom *et. al* proposed a centralized approach aiming to reach minimum charging cost using an optimization technology. Their model concerns both bus congestion and voltage drop problems in the medium voltage grid [84, 83]. Richardson *et. al* formulated and solved an optimization problem to maximize the energy delivered to all electrical vehicles (EVs) within a certain period of time [71]. Transformer overload and voltage drop of a low-voltage transmission grid are considered, assuming the charging rate of each EV can be adjusted continuously. M. D. Galus *et. al* proposed a hierarchic PHEV scheduling algorithm based on model predictive control and game theory in [19], aiming to avoid transformer overload. In [70], the PHEV charging process can respond to frequency and voltage deviations detected locally. [11] uses stochastic programming to minimize the power losses of the distribution grid. Sortomme et al discussed a method to reduce computation time of minimizing the impacts of PHEV charging on the distribution system's losses in [81]. In [74], each house is assigned an upper bound for power consumption. After reaching this bound, a centralized controller will shut down some devices according to a predefined priority. In [75],

a dynamic upper bound for all the houses based on the rule that the original peak demand without PHEV charging is maintained.

The centralized control used in the previous work may not be easily extended to a large-scale distribution grid with high PHEV penetration, as perfect prediction information may not be available, and the computation time to obtain the control actions by solving complex optimization problems may be long. Also, some users may not want their applications being controlled by others due to privacy reasons. These issues motivate us to design a distributed algorithm which does not need accurate prediction on users' behavior and future load information, and it can be executed online to solve common distribution grid problems.

3.3 System Model

In this section, the topology of the distribution grid, the load profile and PHEV charging patterns are introduced. The PHEV charging profile is modeled based on the data from National House Hold Travel Survey 2009 using a stochastic approach.

3.3.1 Medium Voltage Grid in Our Case Study

Fig. 3.1 shows a typical residential area distribution grid in Portugal [65], corresponding to a semi-urban 15 kV medium voltage grid in a residential area. The triangular shapes in this figure represent the Medium Voltage to Low Voltage (MV/LV) transformers. Each transformer serves 4 neighborhoods including 10 people on average. [1]

Similar to [65, 84, 83], this medium voltage grid is explored using a radial configuration. All the dashed branches are considered to be open. The two round shapes in Fig. 3.1 represent the feeding points. As discussed in [65], this grid may experience two main problems: the buses near the feeding points may reach a high congestion level while the far away ones may encounter the voltage drop problem. [2]

[1] Our algorithm can be scalable to support more people. However the simulation time of the bench-mark algorithm using exhaustive search will be much longer.

[2] Note that the load used in [65] is different from that in this chapter although the same grid topology is used.

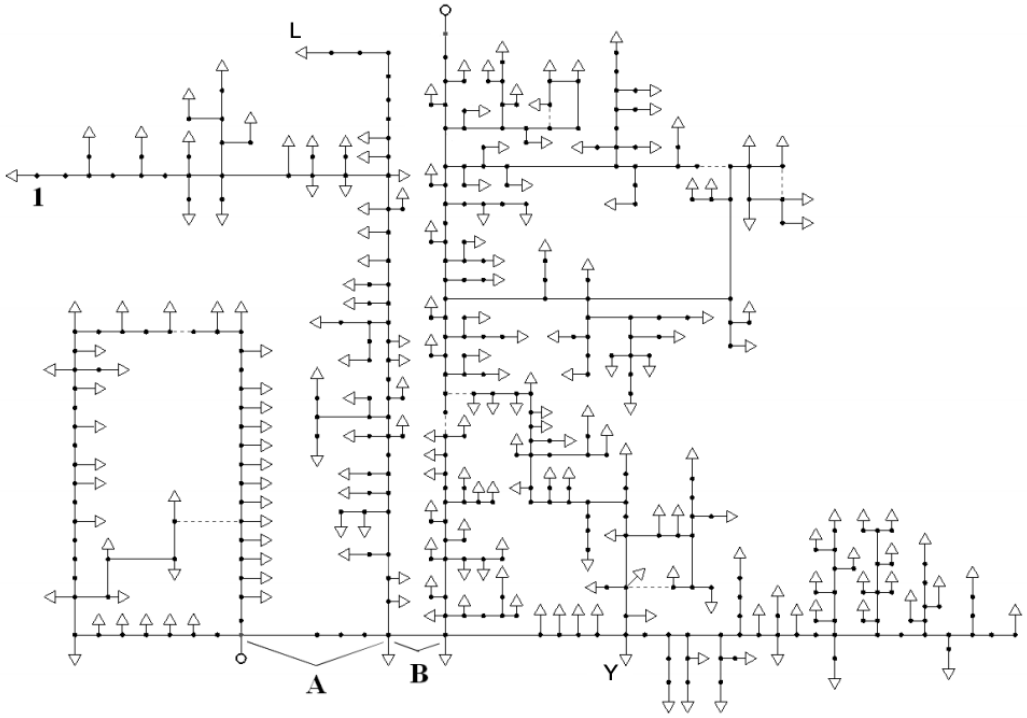


Figure 3.1: Grid Architecture [65]

Table 3.1: PHEV types and their key parameters

PHEV Types	Battery Capacity	Max Range	Market Share
Auto	24 kWh	73 miles	49.9%
SUV	37.6 kWh	80 miles	19.4%
Pickup	30 kWh	55 miles	17.8%
Van (and others)	36 kWh	60 miles	12.9%

3.3.2 Distribution Grid Load

The total load profile for the medium voltage grid used is from [76] which is based on the hourly residential load curves of an average household from the RELOAD database [13] and interpolated using the approach described in [73]. The hourly data is interpolated into the load curve which consists of 1440 minutes of a day.

The load curve consists of 2 different types of residential loads, including both critical loads and elastic loads. Critical loads refer to those that cannot be delayed, such as for cooking and lighting. Elastic loads can be delayed, like for cooling/heating. In this chapter, we assume all the loads except PHEV charging to be critical loads for simplicity. In other words, only PHEV charging is controllable and delayable, although, if needed, other elastic loads can also take part in the demand response process using the proposed framework.

3.3.3 PHEV Charging Modeling

To obtain the PHEV charging profile, we need to know the driving habits, PHEV types and etc. We use the data from National Household Travel Survey (NHTS) 2009 under the assumption that PHEV owners' preferences to vehicle types and their driving behaviors will be similar to the conventional vehicle owners'. From the NHTS report, vehicles can be classified into Auto, Sport Utility Vehicle (SUV), Pick-up trucks and Van.

Their key parameters including the estimated market share are shown in Table 3.1.

In this chapter, the Monte Carlo method is used to simulate the daily driving distance for each PHEV by using the driving data from [91], so the state of charge (SOC) of each vehicle can be determined when they arrive home based on the battery profile and the maximum driving range of that PHEV. Similar to [95], we further assume that the arriving and departure time of PHEVs follows a normal distribution with

the mean of 6pm and 7am respectively and a standard deviation of 1hr, respectively. As our focus is to deal with the impact of PHEV charging during peak time on a residential distribution grid, peak hours are included in the charging period.

In addition to the driving patterns, charging power from [97, 23, 98] is also used to build the PHEV charging model. In this chapter, the charging power of 1.4 kW, 2 kW and 6 kW are chosen with probability 0.45, 0.45 and 0.1, respectively.

Since there are 207 MV/LV transformers in the target grid and each transformer severs 10 people, the population size is 2070. From the data of Major Travel Indicators of 2009 [1], the vehicle ownership ratio is 74.4%. Therefore, the number of vehicles is 1540.

3.4 Problem Formulation

We consider a discrete-time system, *i.e.*, time is divided into slots with a constant duration. We also set the slot duration small enough that the number of PHEVs accessing or leaving the tagged distribution grid per slot is typically no larger than one. The objective of the problem is to maximize the total number of PHEVs that can be charged under the given system capacity by optimizing the charging scheduling vectors

$$X(t) = [X_1(t), X_2(t), ..., X_N(t)], \quad \forall t = 1, 2, ..., T, \tag{3.1}$$

where $X_k(t) \in \{0, 1\}, \quad \forall k = 1, 2, ..., N$, N is the total number of PHEVs and T is the total time slots.

The scheduling needs to consider three constraints. First, each PHEV cannot wait longer than the maximum tolerable delay. Second, the load of each bus cannot exceeds its capacity. Third, the voltage drop of any bus cannot be larger than the maximum allowed voltage drop at any time. This problem can be formulated as follows:

$$\max \quad N \tag{3.2}$$

$$\text{subject to:} \quad \omega(k) \leq \omega_m(k), \tag{3.3}$$

$$f_i(X(t), L^i_{base}(t)) \leq 1, \quad \forall i = 1, 2, ..., M, \tag{3.4}$$

$$f_v(X(t), L^i_{base}(t)) \leq V_{mi}, \quad \forall i = 1, 2, ..., M, \tag{3.5}$$

where M is the number of buses in the distribution grid; $\omega(k)$ is the total waiting

time of PHEV k; $\omega_m(k)$ is the maximum tolerable delay time of PHEV k; V_{mi} is the maximum allowed voltage drop of bus i; L^i_{base} is the base load on bus i; f_i is the mapping function that calculates the loading rate from all the load; f_v is the mapping function that calculates the voltage drop from all the load.

Since it is an integer optimization problem, which is difficult to obtain the optimal solution in polynomial time, and constraints (3.4) and (3.5) are not linear, in this chapter, we aim to obtain a sub-optimal solution in a distributed and real-time manner and compare its performance with the best results using exhaustive search in simulation.

3.5 Proposed Framework

The proposed framework includes three entities: a control center covering one or more medium voltage grids, one smart agent per house, and PHEVs. It is important to note that although our proposed framework is used for PHEV charging, other elastic load such as washing machine, water heater, and air conditioner may also be applicable under this framework to provide demand response.

3.5.1 Control Center

The control center in the grid will monitor two kinds of information: load and voltage drop. All the information can be obtained from sensors distributed in the grid. In this chapter, we assume that an existing smart grid communication infrastructure is available to connect the control center, sensors and smart agents, and the communication delay and packet losses are negligible. We assume that the data from all the sensors are updated in real-time. The influence of grid information update delay on system performance will be discussed in Section 3.8.

From both the historical data and grid topology information, the control center is able to determine which buses are more likely to experience the congestion or voltage drop problems.These are called critical buses, which constitute the congested bus set S_c and voltage drop bus set S_v.

When a PHEV is plugged in, the smart agent which makes scheduling decisions for the PHEV will request a data set from the control center including parameters relevant to the critical buses. The data set contains the loading rate of the most easy-to-congest bus affected by that PHEV (denoted as γ_c), and a voltage ratio γ_v.

At time slot t, the loading rate of bus i is defined as follows (to simplify the notation, we drop t in the following equations):

$$\gamma_c(i) = \frac{\text{current load of bus i}}{\text{maximum allowed load of bus i}}. \tag{3.6}$$

Then γ_c is obtained by choosing the maximum rate of the bus from S_c, which is the most vulnerable to the congestion problem:

$$\gamma_c = \max \gamma_c(i), \quad i \in S_c. \tag{3.7}$$

Similarly γ_v is obtained from all affected buses:

$$\gamma_v(i) = \frac{\text{current voltage drop of bus i}}{\text{maximum tolerable voltage drop of bus i}}. \tag{3.8}$$

$$\gamma_v = \max \gamma_v(i), \quad i \in S_v. \tag{3.9}$$

For example, from [65], for charging load under all the MV/LV transformers, bus A in Fig. 3.1 suffers the severest congestion problem and bus 1 may experience the largest voltage drop. If a PHEV at location L is plugged in, the smart agent at L will request the grid information from the control center which will put $\gamma_c(A)$ as γ_c and $\gamma_v(1)$ as γ_v into a data set and then deliver it to the smart agent.

Meanwhile, the control center will keep monitoring the status of the whole power grid and send instructions to the smart agents to adjust some control parameters which will affect their scheduling decisions if necessary. We will discuss these parameters and actions in Section 3.6.

3.5.2 Smart Agents

The smart agents (or the smart meters) can schedule PHEVs charging. The scheduling decisions are made based on the data sets received from the control center and the algorithms described in Section 3.6.

All the houses adopting our algorithms will receive incentive from the electricity company depending on the contributions they make. (How to determine the contribution and design an incentive mechanism is left for future research.)

It is worth to notice that the user can always let the smart agent charge the PHEV without waiting, in this case the PHEV becomes critical load and the user will not

receive compensation from the power company.

3.5.3 PHEV

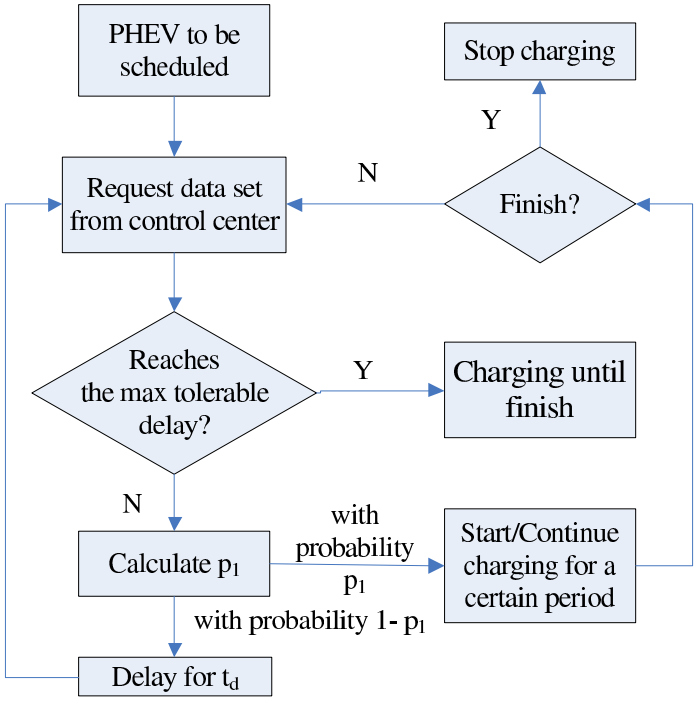
In this chapter, we assume that the PHEVs are plugged in as soon as they arrive home. The departure time can either be set by the user or by the smart agent according to the historical data. Then the smart agent can calculate the maximum tolerable delay time (ω_m) for PHEV charging based on the total parking time (the time between the departure and arrival, denoted by ω_t), state of charge (SOC) and charging power (P_c). Specifically, $\omega_m = \omega_t - (1-SOC)P_B/P_c$, where P_B is the battery capacity. Here we use a linear battery model which is the same as the model used in [97, 74]. In a real system, the PHEV charging behavior can be more complicated. Our algorithm is still applicable so long as the smart agent knows how much time is needed to charge the PHEV and the total parking time of the PHEV.

To satisfy users' requirements, the smart agent should guarantee that the total delay is always less than the maximum tolerable delay time. In addition, to maintain fairness, some charging PHEVs may terminate charging in the middle to yield the charging opportunity to other waiting ones. The smart agents will also assign a higher priority to those PHEVs which have waited for a longer time.

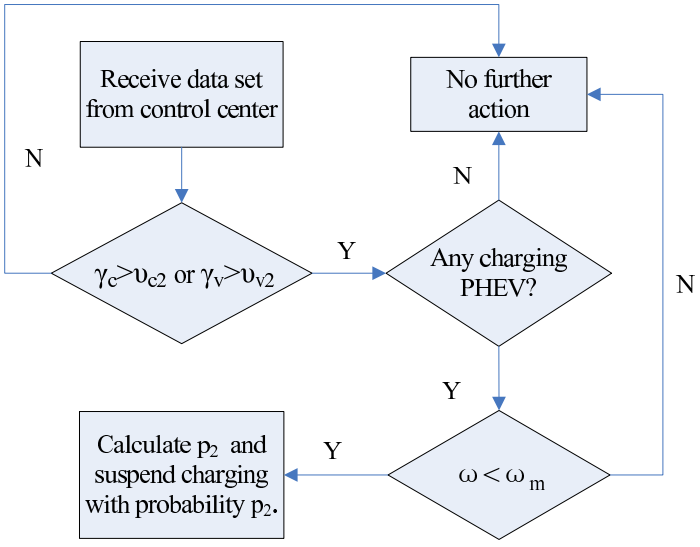
3.6 Random Access Algorithm Design

The design objective of our algorithm can be summarized as follows. First of all, to avoid bus congestion and voltage drop problems in the distribution power grid, (3.4) and (3.5) should be satisfied. Second, fairness should be maintained among all the PHEVs. Third, users' preferences should be taken into consideration. Fourth, (3.3) should be satisfied so that all the PHEVs can be fully charged before their departure. Finally, this algorithm should not rely on future load prediction and is simple enough to be executed online.

The flow chart of our proposed algorithm is presented in Fig. 3.2. Fig. 3.2 (a) describes how the smart agent schedules PHEV charging based on the received information, while Fig. 3.2 (b) shows the process of providing demand response to protect the distribution grid. We will cover the design details in the following subsections. To meet the first design objective, when a PHEV is plugged in, the smart agent will calculate all the access probabilities based on the data set received from the control



(a) Schedule PHEV charging



(b) Demand Response

Figure 3.2: Flow chart for smart agent

center and choose the minimum one to decide the charging probability.

Specifically, from γ_c, and γ_v, the access probability $p_1(\gamma_c)$ and $p_1(\gamma_v)$ can be obtained using the algorithm described below respectively. Then p_1 is set to be the minimum one to decide the charging probability of the PHEV.

$$p_1 = \min\{p_1(\gamma_c), p_1(\gamma_v)\}. \tag{3.10}$$

The access probability is calculated as follows. If both of the received ratios γ_c and γ_v are below the corresponding threshold one (ts1), the PHEV will start charging with probability one; otherwise, if any of these ratios is higher than its corresponding ts1, the charging of the PHEV is restricted based on a back-off algorithm: with probability p_1, this PHEV will start charging immediately; with probability $1-p_1$, it will be delayed by t_d and then try again. If any of the ratios is even higher than the corresponding threshold two (ts2), the charging probability p_1 is set to zero unless the PHEV reaches its maximum tolerable delay time. In addition, in each time slot, the control center will broadcast the ratio when it is larger than the corresponding ts2, to notify all the relevant smart agents until this ratio falls below ts2 again. These smart agents, upon receiving the notification, will suspend the charging PHEVs with a probability p_2 to protect the distribution grid. By stopping charging PHEVs every time slot based on a probability, a fast demand response can be achieved.

To provide an equal chance for each PHEV, every smart agent will acquire the data sets from the control center again after the PHEV has been charged for a period of time and then decide whether to let it continue charging or to suspend it based on the updated probability calculated again from the received ratios. This will provide an opportunity for other waiting PHEVs to start charging. All the waiting PHEVs will obtain a higher priority to charge with a larger waiting time. A suitable charging period is selected to protect the batteries.

However, when a PHEV reaches its maximum tolerable delay time, the smart agent will let it start charging immediately to meet the design objective four.

In the following part, we describe how to design the access probability p_1 and the suspending probability p_2 considering bus congestion and voltage drop in the distribution grid.

3.6.1 Bus Load Congestion

Obviously, the access probability $p_1(\gamma_c)$ should be close to 1 if the ratio γ_c only exceeds ts1 slightly and it must decrease fast when γ_c approaches ts2. Therefore, $p_1(\gamma_c)$ is designed as an exponential function:

$$p_1(\gamma_c) = \begin{cases} \kappa_1 e^{-\alpha_c(\gamma_c - \nu_{c1}) + \beta_c \omega/\omega_m} + \delta_1, & \text{if } \omega < \omega_m, \\ 1, & \text{if } \omega = \omega_m, \end{cases} \tag{3.11}$$

where ω is the current waited time of the tagged PHEV, ω_m represents its maximum tolerable delay, ν_{c1} is the value of threshold one for bus congestion, the parameter δ_1 is used to reflect user's preference, and the global parameter κ_1 is used by the control center for global adjustment if needed.

In the above design, $p_1(\gamma_c)$ decreases exponentially when γ_c increases, which will restrict the number of PHEVs start charging when the bus congestion level is high. Since the smart agent will use the largest $\gamma_c(i)$ to calculate the access probability, all the critical buses can be protected.

The global parameter κ_1 can be set by the control center through notification messages and it is the same for all the smart agents. By decreasing κ_1, the probability to start charging is decreased, so the demand caused by PHEV charging is reduced when γ_c is between ts1 and ts2, and vice versa. By default, κ_1 is set to 1.

The parameter α_c determines how fast $p_1(\gamma_c)$ will decrease when the current loading rate γ_c approaches threshold two for bus congestion (with value ν_{c2}). Considering the design objective, we define $p_1(\nu_{c1}) = 1$ and $p_1(\nu_{c2}) = \varepsilon$ when $\omega < \omega_m$ where ε is a very small positive number. Then, α_c can be expressed as follows:

$$\alpha_c = \frac{\ln\varepsilon}{\nu_{\mathrm{c1}} - \nu_{\mathrm{c2}}}. \tag{3.12}$$

On the other hand, this PHEV charging request may also be delayed with probability $1 - p_1$. If $\omega_m - \omega > t_m$, then the delay time t_d is randomly selected from $[0, t_m]$, where t_m is the upper bound for the delay; otherwise, t_d is set as $\omega_m - \omega$.

From (3.11), PHEVs with the waiting time closer to ω_m have a higher probability to access the grid. This will maintain delay and fairness among all the PHEVs. The maximum tolerable waiting time ω_m can be set by the user or by the smart agent based on the historical data. Here we assume that an incentive mechanism is used so all the users set an appropriate ω_m according to their real needs. The parameter β_c

is used to determine the weight of the waiting time on the charging probability. The value for β_c can be defined as follows: let $p_1(\nu_{c2}) = \rho$, $\kappa_1 = 1$ and $\delta_1 = 0$ in (3.11), and we have: $p_1(\nu_{c2}) = e^{-\alpha_c(\nu_{c2}-\nu_{c1})+\beta_c} = \rho \Rightarrow \beta_c = \ln\rho + \alpha_c(\nu_{c2} - \nu_{c1})$.

To fully utilize PHEV's delay time, ρ should be less than 1; otherwise, every PHEV will get a high probability to start charging before it can be delayed to its maximum tolerable delay time even when γ_c is high. According to our simulation, ρ can be set between 0.3 and 0.6 empirically.

When γ_c is greater than ts2, demand response mechanism will take place. The probability to suspend a charging PHEV is denoted as $p_2(\gamma_c)$. Similarly, $p_2(\gamma_c)$ is also designed as an exponential function:

$$p_2(\gamma_c) = \begin{cases} \kappa_2 e^{\lambda_c(\gamma_c - 1)} + \delta_2, & \text{if } \omega < \omega_m, \\ 0, & \text{if } \omega = \omega_m, \end{cases} \tag{3.13}$$

where κ_2 is another global parameter set by the control center to adjust demand response speed. δ_2 is used to represent users' preferences, similar to δ_1. Both δ_1 and δ_2 can be set by the user or learnt by the smart agent. Of course, they will be included into the incentive mechanism aimed to determine users' contributions.

The parameter λ_c determines the increasing speed of $p_2(\gamma_c)$ when γ_c is above ts2. If we let $p_2(\nu_{c2}) = \varepsilon$ and $p_2(1) = 1$, where ε is a very small positive number, then λ_c can be expressed as:

$$\lambda_c = \frac{\ln\varepsilon}{\nu_{c2} - 1}. \tag{3.14}$$

If a PHEV is delayed or suspended, it will try to access the grid again after t_d slots. Whether its charging request will be approved depends on the ratio at that time.

3.6.2 Voltage Drop

The influence of a charging PHEV on the voltage of a bus is related to the location of that PHEV. For example, in Fig. 3.1, the voltage drop of bus 1 is related to every charging PHEV connected to the buses from the feeder point to bus 1. With the same charging power, the closer the PHEV is to bus 1, the higher the influence it will be. Therefore, the PHEV at location L has a higher effect on the voltage drop of bus 1 than that at location Y. Consequently, we add a weight of location, ϕ, to design the

access probability function:

$$p_1(\gamma_v) = \begin{cases} \frac{\kappa_3}{\phi} e^{-\alpha_v(\gamma_v - \nu_{v1}) + \beta_v \omega/\omega_m} + \delta_3, & \text{if } \omega < \omega_m, \\ 1, & \text{if } \omega = \omega_m, \end{cases} \tag{3.15}$$

where κ_3, α_v, β_v and δ_3 are set similar to κ_1, α_c, β_c and δ_1 respectively. In practice, it is difficult to define an optimal ϕ as it needs perfect grid information. It can be approximated as the ratio of the distance between the PHEV and the feeding point over the distance between the considered critical bus and the feeding point. We will further discuss it in Section 3.8.

Similarly, the suspending probability $p_2(\gamma_v)$ can be designed as follows:

$$p_2(\gamma_v) = \begin{cases} \kappa_4 \phi e^{\lambda_v(\gamma_v - 1)} + \delta_4, & \text{if } \omega < \omega_m, \\ 0, & \text{if } \omega = \omega_m, \end{cases} \tag{3.16}$$

where κ_4, λ_v and δ_4 are set similar to κ_2, λ_c and δ_2 respectively. Therefore, PHEVs closer to the feeding point have a smaller suspending probability for providing demand response.

From the descriptions above, PHEVs are scheduled and charged in a distributed manner, and no prediction is needed. Therefore, the last design objective is met.

Finally, since the smart agent will choose the charging probability calculated from the received data set based on (3.10) to charge the connected PHEV, all the distribution grid components affected by that PHEV are protected.

3.7 Performance Analysis

In this section, the capacity of the proposed algorithm on bus congestion is analyzed using a probabilistic method. The performance analysis on voltage drop is similar and omitted due to the space limitation. To simplify the analysis, we assume the maximum tolerable delay of any PHEV is infinity, and the global control parameters κ_1, κ_2 are set to one. Since Poisson process is an acceptable model if the occurrences are uniformly and independently distributed on an interval of time, the PHEV arriving process is assumed as a Poisson process with an average arrival rate of λ.

In the first part, we will analyze the performance of the proposed algorithm with

real-time grid information, and then we will consider the situation when real-time grid information is not available.

3.7.1 Control Center with Real-time Grid Information

According to the assumption that real-time grid information is available and the slot duration is sufficiently small in Section 3.4, the number of charging PHEVs will increase or decrease at most by one in each slot. Let us consider a certain critical bus A, assume that the maximum loading rate increase caused by one PHEV on bus A is Δ_m. From (3.11), the maximum probability that the loading rate will exceed ts2 caused by one arrival PHEV is

$$p_1(\nu_{c2} - \Delta_m) = e^{-\alpha_c(\nu_{c2}-\Delta_m-\nu_{c1})}. \tag{3.17}$$

Since demand response will start when $\gamma_c > \nu_{c2}$, this probability should be smaller than a threshold p_e, where p_e is defined by the control center. Let $p_1(\nu_{c2} - \Delta_m) \leq p_e$, we can obtain the minimum value of the gap between ν_{c1} and ν_{c2}:

$$\nu_{c2} - \nu_{c1} \geq -\frac{1}{\alpha_c}\ln p_e + \Delta_m. \tag{3.18}$$

Define D_m the maximum loading rate increase caused by the variation of the base load in one slot. Since the charging probability equals 0 when $\gamma_c > \nu_{c2}$, to prevent the loading rate from exceeding one, the gap between ν_{c2} and one should be greater than D_m plus Δ_m. Then we have

$$\nu_{c2} \leq 1 - D_m - \Delta_m. \tag{3.19}$$

Besides, once the loading rate exceeds ts2, the probability to suspend charging PHEVs should be less than a pre-defined threshold p_t so that only a small number of PHEVs are suspended. p_t is set by the control center based on the estimated number of charging PHEVs in the system.

$$p_t \geq p_2(\nu_{c2} + \Delta_m + D_m) = e^{\lambda_c(\nu_{c2}+\Delta_m+D_m-1)}. \tag{3.20}$$

After manipulation, we can obtain another bound for ν_{c2}.

$$\nu_{c2} \leq \frac{1}{\lambda_c} \ln p_t + 1 - \Delta_m - D_m. \tag{3.21}$$

From (3.18), (3.19) and (3.21), we can obtain the upper bounds of ν_{c1} and ν_{c2}.

3.7.2 Control Center Without Real-time Grid Information

However, there is always a time interval between the information update from sensors to the control center in practice. To analyze the performance of the proposed algorithm in a more practical situation, we assume the grid information is updated every t seconds, and the maximum PHEV arrival rate does not exceed λ during the following t seconds.

The expected number of new PHEVs bus A can support without exceeding ts2 is:

$$m_1 = \left\lfloor \frac{\nu_{c2} - \gamma_c}{I_c} \right\rfloor, \tag{3.22}$$

where I_c is the average loading rate increase caused by one PHEV. We further consider the following three situations.

(1) γ_c is below ts1. All the arrival PHEVs will start charging with probability 1. The probability that the number of arriving PHEVs does not exceed m_1 during t is:

$$P(n \leq m_1) = \sum_{k=0}^{m_1} \frac{e^{-\lambda t}(\lambda t)^k}{k!}. \tag{3.23}$$

(2) γ_c is between ν_{c1} and ν_{c2}. The probability for each arriving PHEV to start charging is

$$p_1(i) = e^{-\alpha_c(\gamma_c - \nu_{c1})}. \tag{3.24}$$

The probability that the number of arriving PHEVs is less than or equal to m_1 is

$$P_a = \sum_{k=0}^{m_1} \frac{e^{-\lambda t}(\lambda t)^k}{k!}. \tag{3.25}$$

The probability that the number of PHEVs starting charging is less than m_1, given

the number of arriving PHEV is larger than m_1, is

$$P_b = \sum_{k=m_1+1}^{+\infty} \frac{e^{-\lambda t}(\lambda t)^k}{k!} \cdot \left(\sum_{i=0}^{m_1} \binom{k}{i} {p_1}^i (1-p_1)^{k-i} \right). \tag{3.26}$$

The probability that the total new charging PHEVs' number does not exceed m_1 is

$$P(n \leq m_1) = P_a + P_b. \tag{3.27}$$

(3) γ_c is larger than or equal to ν_{c2}. In this case, all the arriving PHEVs will not start charging, so their arrivals will not affect the bus load.

From (3.23) and (3.27) we can obtain the probability to exceed ν_{c2} for any $\gamma_c < \nu_{c2}$. Similarly, the number of new PHEVs this bus can support without exceeding the bus capacity is

$$m_2 = \left\lfloor \frac{1-\gamma_c}{I_c} \right\rfloor, \tag{3.28}$$

Substituting m_1 by m_2 in (3.23), (3.25), (3.26), (3.27), we obtain the probability that the loading rate does not exceed one.

Given λ and that the probability to exceed ts2 and one should below p_e and p_s respectively, we can obtain the upper bound for ts1 and ts2 by using a reversed process. We omit the details due to the space limitation.

It is worth to mention that ν_{c1} and ν_{c2} can be set dynamically through broadcasting instructions from the control center to all the users according to the changing PHEV arriving rate in different time periods.

On the other hand, given ν_{c1}, ν_{c2}, p_e, p_s, we can obtain the maximum arrival rate λ that the distribution grid can support using the proposed algorithm, which is the capacity of the system.

3.8 Performance Evaluation

The objective of this simulation is to verify the impact of the proposed control algorithm on bus congestion, and bus voltage drop. The simulation is mainly based on real data with an approximated maximum average arriving rate of 30 PHEVs per minute. The maximum tolerable probabilities to exceed threshold two and bus loading rate one/maximum allowed voltage drop are set to 10^{-2} and 10^{-6}, respectively. p_t is set to 0.05. Whenever the ratio of a critical bus exceeds ts2, the specific ratio of the

bus is multicasted to all relevant smart agents every time slot. The situation when the grid information is updated every 60 seconds is also considered in Section 3.8.3.

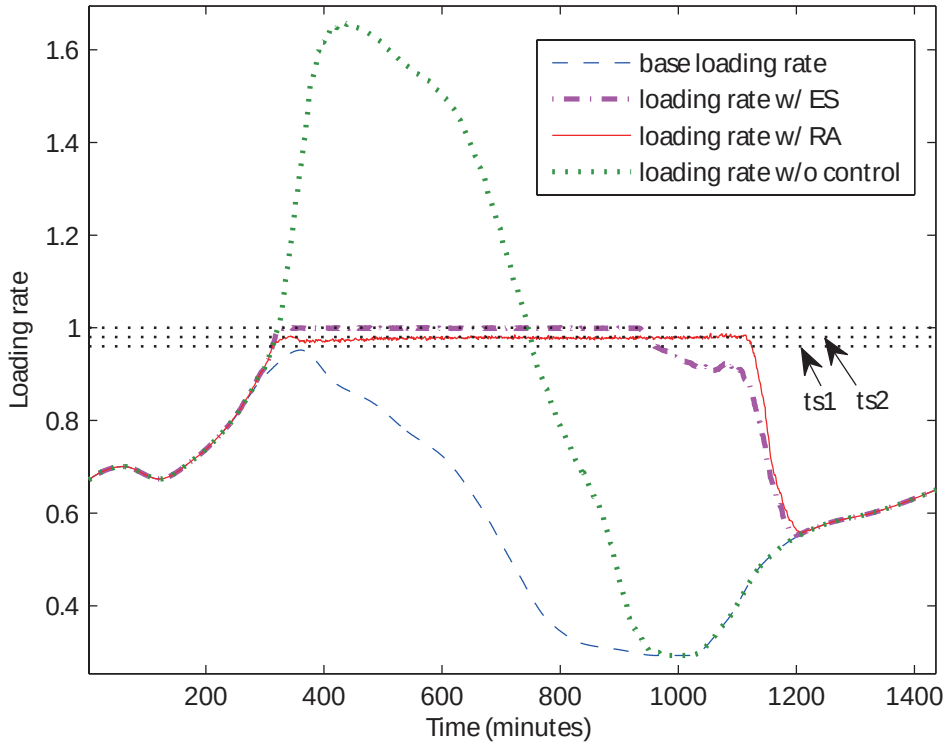
In the following subsections, we only consider one problem at one time under the proposed framework, and assume p_1 equals $p_1(\gamma_c)$ and $p_1(\gamma_v)$ in Section 3.8.1 and Section 3.8.2 respectively. We compare the performance of the proposed random access (RA) algorithm with that obtained from exhaustive search (ES). For exhaustive search, we choose the one with the minimum average PHEV waiting time so it can result in the maximum number of PHEVs being supported without violating the constraints. In the exhaustive search, we assume that PHEVs' charging can be interrupted at any time and any frequency; therefore, ES result can also be considered as the performance upper bound. Similar to [71], we do not consider reactive power or grid losses as they will make the simulation time much longer. On the other hand, in a real system, as the smart agents use the measured data, which reflects the reactive power and grid losses, etc., to calculate the access probabilities, our algorithm can still be applicable.

3.8.1 Bus Load Congestion

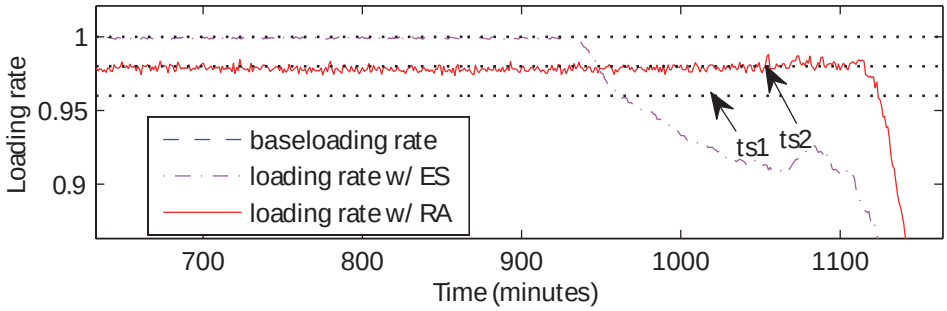
According to the simulation results in [65], bus A suffers the severest congestion problem so its loading rate is delivered to the downstream smart agents when required. In this simulation, ν_{c1} and ν_{c2} are set to 0.96 and 0.98 respectively according to the analysis in Section 3.7.

Fig. 3.3 (a) illustrates the loading rate of bus A with 742 PHEVs. To view the curves more clearly, the most critical time period is zoomed in and shown in Fig. 3.3 (b). This can be considered as the capacity of the proposed algorithm because the loading rate reaches ts2 at the end of the charging period. As shown in the figure, our proposed algorithm can flatten the bus loading rate quite well while the uncontrolled loading rate exceeds one by about 70%.

Fig. 3.4 shows the situation when there are 765 PHEVs which is also the maximum number of PHEVs this distribution grid can support using exhaustive search. Since PHEVs begin to charge immediately whenever they reach their maximum tolerable delay time, the loading rate of RA algorithm exceeds one at the end. Through multiple simulations with different PHEV arriving/departure time, exhaustive search with perfect future information can support about 3% more PHEVs than the proposed algorithm on average in the scenario described in Section 3.3. In other words, the



(a) Loading rate variation



(b) Zoom-in of (a)

Figure 3.3: Loading rate of bus A with 742 PHEVs

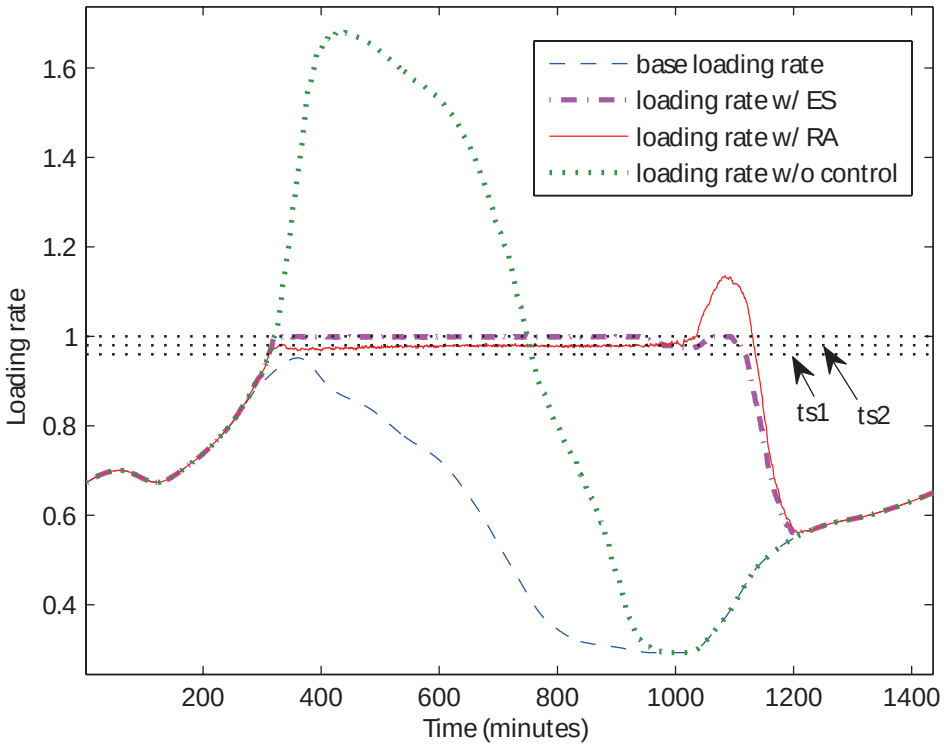


Figure 3.4: Loading rate of bus A with 765 PHEVs

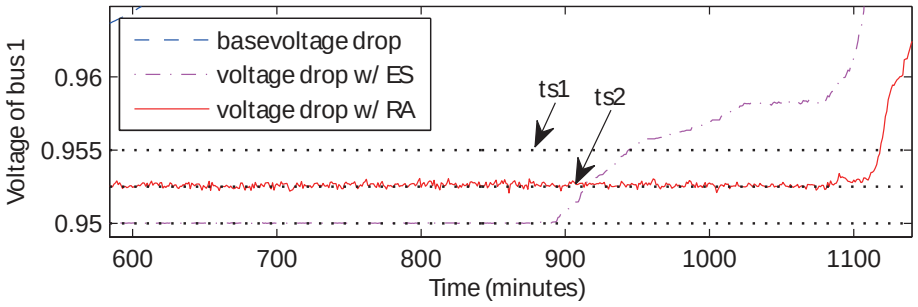


Figure 3.5: Voltage drop of bus 1 with grid topology information

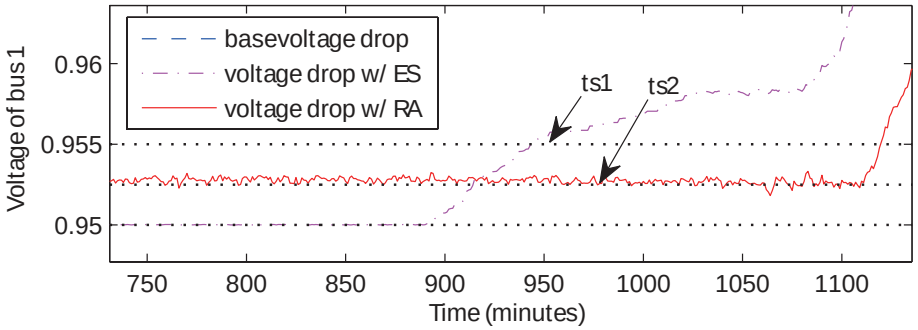


Figure 3.6: Voltage drop of bus 1 without grid topology information

performance of the RA algorithm is very close to the performance upper bound, meanwhile satisfying all the design objectives.

3.8.2 Voltage Drop

In [65], bus 1 has the largest voltage drop, so its voltage ratio is passed down to all the relevant users whenever required. In this simulation, the maximum tolerable voltage drop is 5%, ν_{v1} and ν_{v2} are set to 0.90 and 0.95, respectively.

Figs. 3.5 and 3.6 show the zoomed in simulation results with and without the grid topology information, respectively. The maximum number of PHEVs the distribution system can support are 730 and 727 on average, respectively. Both of the two cases can restrict the voltage drop near ts2. Therefore, we may find that the performance of the random access algorithm is not sensitive to the weight ϕ in (3.15).

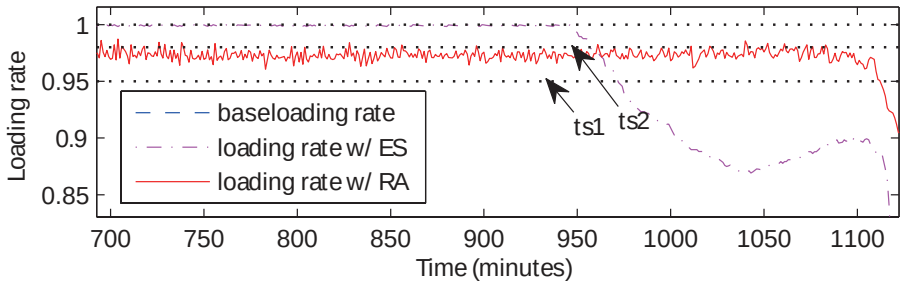


Figure 3.7: Loading rate of bus A with 717 PHEVs: non-real time

3.8.3 Non-real-time Data

In this part, we explore the performance of our algorithm considering that all the data received at the control center is updated every minute. The value for ts1 is decreased to 0.95 based on the algorithm in Section 3.7.

Fig. 3.7 shows the main part of the loading rate of bus A with 717 PHEVs when non-real-time bus loading rate is delivered to smart agents. This can be considered as the capacity of the distribution grid using the RA algorithm with non-real time information. Obviously, the controlled loading rate using the RA algorithm fluctuates more severely. The loading rate even exceeds ts2 several times after minute 700. The reason is that the number of arrival PHEVs varies greatly during the one-minute time interval, and many PHEVs become critical load when reach their maximum tolerable delay time. Nevertheless, the demand response mechanism suspends some charging PHEVs when the loading rate is over ν_{c2} and keeps the loading rate under one all the time. The situation for voltage drop is similar and is omitted due to the space limitation.

3.9 Conclusion

In this chapter, a random access framework has been proposed to coordinate PHEV charging to maximize the number of PHEVs that can be supported considering bus load congestion and voltage drop constraints. Through the simulation on a residential area distribution grid, it has been demonstrated that our algorithm can achieve the performance with a small gap to the best solution. Besides, it can provide demand response efficiently. We also tested our algorithm when real-time grid information is not available, and the result is only about 6% worse than the best one. In addition to regulating PHEV charging, the proposed solution can also be applied

to other elastic devices.

In Chapter 2 and Chapter 3, we use PHEV as the elastic load to provide demand response. One important feature of PHEV is that we do not know the number of future arrivals, nor do we know their future demand or deadline. There is another category of elastic load whose future status can be predicted based on the current status. The utilization of this kind of elastic load to provide demand response can be different. We will cover it in the next chapter.

Chapter 4

A Dynamic Water-filling Method for Real-Time HVAC Load Control

4.1 Introduction

In the current power grid, one area of inefficiency arises from the variation of the power consumption in each day. Generally speaking, in a residential area, the power consumption is typically the lowest early in the morning and the highest when people go to or return from work. In order to make the whole power grid stable, we need to take great effort to make the power generation follow these demand peaks and valleys, which keeps the power generation cost at a high level. If we can flatten the peaks and meanwhile fill the load valleys, the overall power generation efficiency can be improved significantly.

However, we do not know the load in the future, nor do we know how many elastic appliances are available for control and their states during the remaining time of the day. Without such information, it is very challenging to minimize the overall cost from the perspective of the power company. Of course, we can formulate and solve optimization problems using the information from prediction, as what the previous works have done. However, it is still an open issue to find an effective way to predict the future load for a long time with good accuracy. Nevertheless, with the help of energy buffer, such as batteries, the performance of load prediction for a short term is possible and feasible. How to use short-term prediction results of both load and renewable energy generation to guide the control of elastic load scheduling, given limited number of elastic appliances, is the focus of our proposed work. The design

objective will include both user comfortable level constrains and the power generation cost.

The main idea we use is based on Model Predictive Control (MPC). Different from existing works which only use homogeneous load, in this chapter, we will consider heterogeneous loads with different parameters. Besides, we will propose an online learning method for the control center to estimate different parameters for different loads. Furthermore, we extend the centralized algorithm to a distributed one with local controllers. In this way, the computation load of the central controller can be reduced and the system can be more scalable and robust.

The main contributions of this chapter can be summarized as follows. First, we propose a centralized algorithm to control heterogeneous HVACs in a micro-grid. The objective of this algorithm is to reduce non-renewable energy generation fluctuations while still guarantee user comfort level. An approach to estimate heterogeneous HVAC model parameters is also proposed. Second, we extend the centralized algorithm to a distributed one, which has a much lower computational complexity and is more scalable. Third, we further extend the proposed algorithms to support HVAC ON/OFF control modes other than adjusting the HVAC power level. Fourth, since the elastic load potential provided by HVACs is limited compared to the unlimited control time, a dynamic water level adjustment algorithm is proposed to reserve this elastic load potential for future demand response. Finally, extensive simulations using practical data sets obtained from Eirgrid [16] have been conducted to evaluate the performance of the proposed algorithms. The results demonstrate the advantages of the proposed algorithms comparing to existing ones.

4.2 Related Work

Since load fluctuation usually adds cost to power generation and raises requirements on frequency control [80], smoothing or flatting the non-renewable power generation using demand response is one of the most important objectives.

For centralized demand response control, [77] proposed a water-filling approach to flatten the overall load assuming that perfect future load information is available. In [35], Koutsopoulos *et. al* introduced two online demand scheduling policies based on dynamic programming to minimize the long-term average power grid operation cost, without considering the variation of base load and renewable energy. He *et. al* proposed a PHEV scheduling algorithm to minimize the total cost of electrical vehicles

using a sliding window algorithm in [35]. In their algorithms, all the local controllers were coordinated by the same predicted base load model with accurate prediction in the whole time scale. [64] is similar to our work in that it aimed to smooth non-renewable electricity supply by controlling the set point and ON/OFF states of all HVACs. The algorithms proposed above try to avoid or minimize fluctuations, so they all need accurate long-term prediction information to make control decisions. They also assumed the availability of enough elastic load to perform the demand response. On the other hand, there are also some existing works which make control decisions after the power imbalance happens, so no future prediction information is needed. For example, [63, 5, 30] used an energy storage system to provide primary frequency control to the power grid based on the current requirement or historical data. Although the capacity of the primary frequency control system is limited, the secondary frequency control system will help reduce their burden in time. The proposed MPC based algorithms belong to the first group. Taking one step further, our work considers the situation of limited elastic load, and only short-term prediction information is needed.

Thermostatically controlled appliances, such as HVACs have been widely used for demand response in smart grid. In [57, 87], HVAC is used to minimize the economic cost by scheduling its operation time. The application of MPC to HVACs can be found in [100], with an objective to minimize the user discomfort level while keeping the economic cost within a given budget. Karmakar *et. al* introduced an online algorithm which maintained the thermal comfort-bands while keeping the total HVAC load under peak energy consumption constraint [31]. In [43, 45], HVAC is used to provide intra-hour load balancing or load following according to given control signals. Different from the existing work, our control objective is to minimize the fluctuation of the non-renewable power generation without external control signals, which is more challenging.

MPC has been used to solve various control problems in smart grid. [28] proposed an economic MPC algorithm to minimize the total cost of distributed power generation plants. The control actions are adjusting the amount of power generation from each plant. In [20], an aggregator utilizes MPC strategy to track a secondary frequency control signal by controlling heterogenous elastic loads. Different from the previous approaches, the proposed dynamic water level adjustment algorithm will make a tradeoff between the fluctuation of non-renewable power generation and elastic load potential reservation.

4.3 System Model

4.3.1 System Architecture

The investigated system represents a micro-grid with a high renewable energy penetration. It consists of a control center, customers with HVAC installed, and a communication network that connects them together.

The electricity power supply comes from two types of sources: conventional power generators and renewable power generators. Due to its stochastic feature, instantaneous renewable power generation is time-varying, while a good prediction over a short time period is possible [22], especially with the help of large energy buffers (batteries).

The relationship between load and power supply is shown in (4.1), where $S_n(t)$ is the power generation from conventional power plants at time t, S_r is the renewable energy generation, L_b is the non-HVAC load (also called base load), and L_h is the load from all HVACs.

$$S_n(t) = L_b(t) - S_r(t) + L_h(t). \tag{4.1}$$

In (4.1), $L_h(t)$ is the elastic load at time t which can be changed in each time slot; $L_b(t) - S_r(t)$ is the non-elastic load minus renewable energy generation (we call it "remaining load"). Due to the intermittent nature of renewable energy, the renewable power generation always contains a lot of fluctuations which directly affect the remaining load. As a result, the conventional power companies will need high spinning reserve or buy extra frequency regulation service to do the frequency control in the micro-grid which usually has a high cost [80]. Therefore, in this chapter, we are motivated to make $S_n(t)$ change slowly and smoothly by controlling the elastic load in each time slot so the conventional power companies can save cost and improve the system efficiency.

To determine the optimal HVAC load, we also need to know the future base load. Although there is no long-term load prediction algorithm with good accuracy, short-term load prediction algorithms do exist[41]. In this work, we assume that inelastic load can be predicted in a short-term.

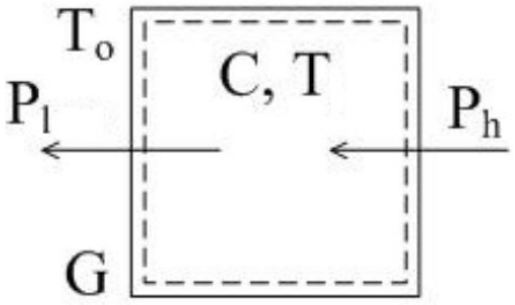


Figure 4.1: HVAC model[69]

4.3.2 HVAC Model

In this chapter, the HVAC model is obtained from [69] and is briefly reviewed as follows.

A simple HVAC model is illustrated in Fig. 4.1. In this model, P_h represents the power consumed by the HVAC in the unit of W; P_l represents the amount of heat transferred outdoors through the house boundaries in the unit of W; T and T_o represent the indoor and outdoor temperature respectively in the unit of K; C represents the effective heat capacity, which is the product of air heat capacity and air quantity in a house, in the unit of J/K; G represents the thermal insulation level of a house and is in the unit of W/K.

In this chapter, time is divided into slots with a fixed duration of t. Define $T_i(k)$ as the indoor temperature of the house with HVAC i at the k-th time slot, and P_h^i, G_i, C_i as the corresponding parameters for this specific house. Then the indoor temperature evolves according to the following equation [69]:

$$T_i(k+1) = \left(T_i(k) - \frac{P_h^i(k)}{G_i} - T_o(k)\right) e^{-\frac{G_i}{C_i}t} + \frac{P_h^i(k)}{G_i} + T_o(k). \tag{4.2}$$

The parameters T_i, P_h^i, G_i, C_i, Q_i in this subsection are all related to the specific house with HVAC i, and in the following, we drop the subscription or superscription of i to simplify the notation.

Since G/C is very small (usually $< 10^{-5}$[69]) and the indoor temperature changes during one slot is typically less than 0.5 degree, we approximate the above nonlinear

model w.r.t. slot duration t by a linear one using Taylor's equations:

$$T(k+1) = T(k) - \Delta T_{off}(k) + QP_h(k), \tag{4.3}$$

where

$$\Delta T_{off}(k) = \frac{Gt}{C}(T(k) - T_o(k)), \tag{4.4}$$

and

$$Q = \frac{t}{C}. \tag{4.5}$$

Q is the conversion coefficient from power to temperature in a time slot for a specific room with the unit of K/J.

Since the tolerable indoor temperature variation is relatively small compared to the difference of $T(k)$ and T_o, ΔT_{off} can be approximated as a constant:

$$\Delta T_{off}(k) = \frac{Gt}{C}(T_r - T_o(k)) \tag{4.6}$$

where T_r is the set point temperature of an HVAC.

In this work, we consider houses and HVACs with different parameters. That is, the parameters C and G for each house may be different, and each HVAC can have a different set point T_r and a different maximum power. In addition, although our HVAC model may not be fully accurate, since the rooms' actual temperatures will be updated in each time slot, error will not be accumulated.

4.4 Centralized Dynamic Water-filling Algorithm

The design objective of the centralized dynamic water-filling algorithm is to reduce the fluctuation of the power demand for conventional power plants by controlling the load of HVACs while guaranteeing HVAC user comfortable requirements.

To give a clear description of the proposed algorithm, we first present the existing MPC framework, then we show the design process of each part.

4.4.1 MPC Framework

Fig. 4.2 shows a typical Model Predictive Control framework. In this figure, the block MPC represents the designed controller, and the block Plant represents the plant model of the specific problem. The variable v represents the measured

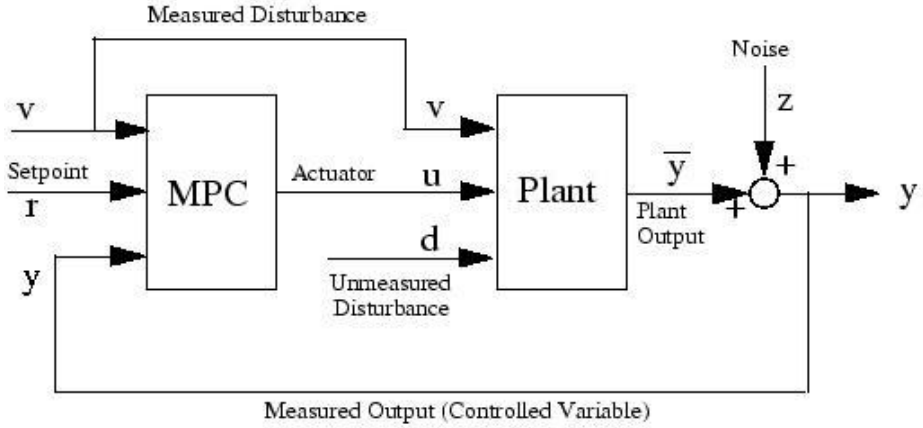


Figure 4.2: Block Diagram of Model Predictive Control Framework [50]

disturbance, r is the set point to the controller, u represents the control actions calculated by the controller, and y is the output of the plant model.

With the help of the plant model, the model predictive controller can calculate a sequence of control actions by solving an optimization problem based on the current plant state, current plant output, control objective function, etc. However, only the control actions in the first time slot are implemented, and others are discarded. The controller then solves the same optimization problem with updated parameters again. This is called the receding horizon principle. More details about Model Predictive Control can be found in [92, 49].

Next we will illustrate how to use MPC to control all the HVACs.

4.4.2 Plant Model Design

(4.3) shows the relationship between the room temperature and the amount of power consumed by HVAC for a single house. Let the indoor temperature of each house and the total load of HVAC be the state of the plant (X), ΔT_{off} for each house be the measured disturbance (V), the input power for each HVAC be the control actions (U), and Y be the output of the plant model. The state space model of the plant is:

$$X(k+1) = AX(k) + B_u U(k) + B_v V(k), \tag{4.7}$$

$$Y(k) = C_x X(k) + D_u U(k) + D_v V(k), \tag{4.8}$$

where

$$X = \begin{bmatrix} T_1 \\ \vdots \\ T_n \\ L_h \end{bmatrix}, \quad U = \begin{bmatrix} P_h^1 \\ \vdots \\ P_h^n \end{bmatrix}, \quad V = \begin{bmatrix} \Delta T_{off}^1 \\ \vdots \\ \Delta T_{off}^n \end{bmatrix}, \tag{4.9}$$

A, B_u, B_v, C_x, D_u, and D_v are coefficients, k is the time slot index, n is the total number of houses with HVACs.

According to (4.3) and the definition of all the variables in this MPC model, we

can determine the coefficients in (4.7) and (4.8) as follows.

$$A = \begin{bmatrix} 1 & & & \\ & 1 & & 0 \\ & & \ddots & \\ 0 & & & 1 \\ & & & & 0 \end{bmatrix}, B_u = \begin{bmatrix} Q_1 & & & \\ & Q_2 & & 0 \\ 0 & & \ddots & \\ & & & Q_n \\ 1 & 1 & \cdots & 1 \end{bmatrix}, \tag{4.10}$$

$$B_v = -A, \quad C = I, \quad D_u = 0, \quad \text{and} \quad D_v = 0. \tag{4.11}$$

In the above equations, the parameters ΔT^i_{off} and Q_i for the house with HVAC i are estimated using the algorithm described in Section 4.4.3.

4.4.3 Heterogenous HVAC Parameters Estimation

To obtain the parameters for heterogeneous HVAC models, there are sensors in each house and a communication network exists between these sensors and the control center.

Let $P_h = 0$ in (4.3), and we can estimate ΔT_{off} when the HVAC is turned off as below:

$$\Delta \hat{T}_{off}(k) = T(k) - T(k+1), \tag{4.12}$$

where $\Delta \hat{T}_{off}$ is the estimated value of ΔT_{off}.

In practice, ΔT_{off} may change during different time slots, so we use the following exponentially weighted moving average (EWMA) algorithm to update ΔT_{off} for an HVAC model.

$$\Delta T_{off}(t_j) = \alpha \cdot \Delta \hat{T}_{off}(t_j) + (1-\alpha) \cdot \Delta T_{off}(t_{j-1}), \tag{4.13}$$

where α is the weight, and t_j is the parameter update time. Note that $\Delta T_{off}(t_j)$ is updated based on the current estimation and its last value, rather than the average of all the former values. The reason is that the status of the room may be different at different time, so the last value may be more accurate. The relationship between k, t and t_j is shown in Fig. 4.3. The time duration between two parameter estimations, $t_j - t_{j-1}$, is determined by the control center. For example, the control center will choose the time slot when the HVAC is off to estimate $\Delta \hat{T}_{off}(t_j)$.

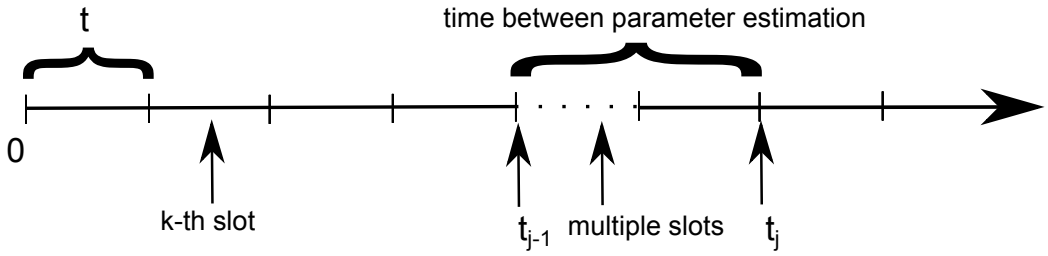


Figure 4.3: Relationship between different time notations

Similarly, from (4.3) we can estimate the value of Q as follows:

$$Q(t_j) = \beta \cdot \hat{Q}(t_j) + (1 - \beta) \cdot Q(t_{j-1}), \tag{4.14}$$

where

$$\hat{Q}(k+1) = \frac{T(k+1) - T(k) + \Delta\hat{T}_{off}(k)}{P_h(k)}, \tag{4.15}$$

and β is the weight.

Through this process, the control center is able to model heterogeneous HVACs by estimating different ΔT_{off} or Q for each HVAC. In the proposed work, we assume that these parameters are already available through estimation.

4.4.4 Controller Design

The objective of our control algorithm is to reduce the fluctuation of the conventional power plants' supply and guarantee user comfort level. Therefore, the objective function of the controller can be formulated as follows:

Problem I (P1)

$$\min_{P_h^i(k)} : \quad \sum_{k \in N} \left\{ (L_h(k) - r_w(k))^2 + \lambda^2 \sum_{i \in S} (T_i(k) - r_i(k))^2 \right\}, \tag{4.16}$$

subject to:

$$\sum_{i \in S} P_h^i(k) = L_h(k), \tag{4.17}$$

$$r_i(k) - \Delta T_l \leq T_i(k) \leq r_i(k) + \Delta T_u, \quad \forall i \in S, \forall k \in N, \tag{4.18}$$

$$0 \leq P_h^i(k) \leq u_i^{\max}, \quad \forall i \in S, \forall k \in N, \tag{4.19}$$

where $r_w(k)$ is the reference value of HVAC load in slot k; $P_h^i(k)$ is the consumed power by HVAC i assigned by the control center in slot k; r_i is the temperature set-point for HVAC i; ΔT_l and ΔT_u represent the maximum allowed temperature decrement and increment from the set point r_i in a house, respectively; $T_i(k)$ is the indoor temperature of the house with HVAC i in slot k; S is the set of all the HVACs; N is the prediction horizon; $u_i^{\max}$ is the maximum power consumption of the i-th HVAC; λ is the weight, and is squared to make the weight always positive.

The first part of the objective function (4.16) represents the deviation of the actual HVAC load from the reference HVAC load (r_w). The second part represents the sum

of temperature deviation from the set-point in each house, which not only represents the influence to users' comfortableness, but also the elastic load potential because the HVAC can only provide one dimensional demand response if the room temperature reaches the upper or lower bound. Of course, we would like the room temperature be close to the set-point so the HVAC can either be turned on or off. This is quite different from the dead band based control policy which does not consider the demand response potential of HVACs [64]. λ is used to make a tradeoff between these two parts. Constraints (4.18) means that the controlled indoor temperature of each house should stay within a certain range of the set point during each time slot. (4.19) ensures that the power of each HVAC is bounded.

By solving this optimization problem, the controller can obtain a sequence of control actions corresponding to each time slot of the prediction horizon. Since the plant model is not accurate and there might be unmeasured disturbance or noise in this system, the actual indoor temperature of each house and the real load may not be the same as predicted after implementing the obtained control actions. Therefore the controller only executes the control actions in the first time slot, then it will update all the parameters and solve the optimization problem again.

4.4.5 Dynamic Water level Adjustment Algorithm

To solve the convex optimization problem **(P1)**, the reference value of HVAC load in slot k, $r_w(k)$, is needed. If $r_w(k)$ is not set appropriately, the controller may not be able to flatten the load effectively. Besides, the energy buffer capacity provided by elastic HVAC load will be consumed when all the indoor temperatures reach their upper or lower bounds. As a result, the HVACs will turn into inelastic load and lose the function of providing demand response.

In addition, since the size of energy buffer is relatively small and limited compared to the remaining load and unlimited control horizon, the value of $r_w(k)$ should not be constant. Instead, $r_w(k)$ should change according to the main trend of the remaining load so that the energy buffer will never be totally full or empty and the HVACs can always perform demand response to reduce the remaining load fluctuation.

To adjust $r_w(k)$ appropriately, we propose a dynamic water level adjustment algorithm stated as follows. In slot k, the sum of $r_w(k)$ and remaining load $L_b(k) - S_r(k)$ is the water level $W_l(k)$, and then we can calculate the reference HVAC load for each

time slot according to

$$r_w(k) = W_l(k) - (L_b(k) - S_r(k)). \tag{4.20}$$

Actually, the water level $W_l(k)$ is the reference value for the total load of the conventional power plants. If we can keep the water level constant, then the total load is constant. However, due to the limited elastic capacity of HVAC, this is impossible. Therefore, we have to change the water level slowly and smoothly to minimize load fluctuations. Assuming a given water level for time slot k (which may not be optimal), by solving the centralized MPC problem **(P1)** we can obtain the predicted system states $X(k+N)$ for time slot $k+N$, where N is the prediction horizon. Then the water level for the next time slot $W_l(k+1)$ can be obtained using Algorithm 1. The main idea is that we adjust the "water" level whenever one of the room temperature may reach the upper or lower bound in the predicted future.

Algorithm 1 Water Level Adjustment

Require: $X(k+N)$
$flag \leftarrow 0$
for all $i \in n$ **do**
  if $T_i(k+N) = r_i(k) + \Delta T_u$ **then**
    $flag \leftarrow 1$
    $break$
  else if $T_i(k+N) = r_i(k) - \Delta T_l$ **then**
    $flag \leftarrow -1$
    $break$
  end if
end for
if $flag \neq 0$ **then**
  $c \leftarrow \mu \cdot (\overline{r_w} - P_s)$
  if $|c| > climit$ **then**
    $c \leftarrow climit \cdot flag$
  end if
  $W_l(k+1) \leftarrow W_l(k) - c$
end if

In Algorithm 1, $\overline{r_w}$ represents the average reference value of HVAC load for the following N slots; $P_s = \sum_{i=1}^{n} \frac{\Delta T_{off}^i}{Q_i}$ is the total amount of power needed to counteract all the houses' temperature decrease in each time slot; c is the change to the water level; $climit$ is the maximum allowed water level change in each time slot, which

is determined by the power company; μ is the water level change rate and can be determined empirically.

The relationship of the MPC algorithm and the water level change can be summarized as follows: we use the MPC to predict the system status in the future. If the prediction results show that the total elastic load potential will be depleted in the future, we adjust the water level slightly to avoid the occurrence of this situation. Of course, with longer accurate prediction horizon, we can avoid unnecessary water level adjustment and therefore reduce the fluctuation.

4.5 Distributed Dynamic Water-Filling Algorithm

The centralized dynamic water-filling algorithm described in section 4.4 relies on a centralized controller, which may have high computational complexity when the number of HVACs is large. To be scalable, a distributed architecture is preferable.

4.5.1 Distributed Control Architecture

Different from the centralized algorithm which has only one control center, the distributed algorithm relies on a hierarchical architecture with one central controller and several local controllers. The central controller tries to flatten the total load by adjusting the amount of power used by each local controller. The local controller will assign the amount of power designated by the central controller to each HVAC and maximize the user comfort level. Note that if the population changes overtime, we can simply resize the group [1], and the algorithm still works.

4.5.2 Central Controller Design

Other than flattening the total load, the central controller should guarantee that the amount of power assigned to the local controllers in each time slot will not make any HVAC under that local controller violate the temperature constraints. Since there is no direct control between the central controller and the HVACs, we need a new plant model for the hierarchical MPC problem.

[1] The groups can be resized at two levels. First, the number of HVACs in a group can be increased or decreased. Second, if the number of HVACs in each group changes too much, we can reconfigure all the groups to balance the group size.

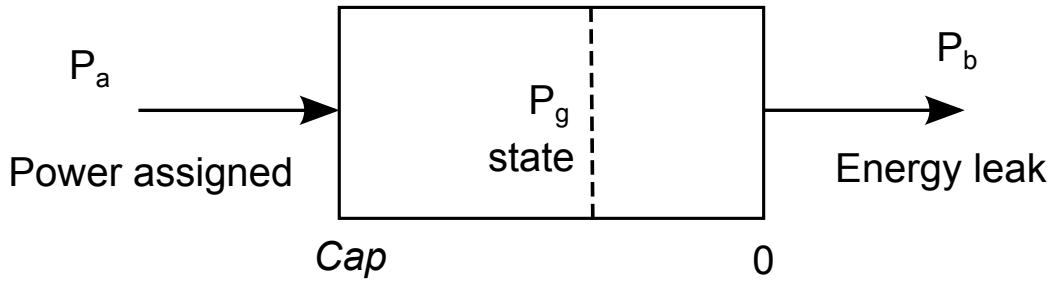


Figure 4.4: Local Group Model

We consider the local controller with all the corresponding HVACs as a group. Each group has different size of the energy buffer provided by its HVACs. The model of a local group is shown in Fig. 4.4. The unit of all the variables in this model is J, and the group index j is omitted to simplify the notation. In each time slot, there is a total amount of energy P_b leak from all the houses, and P_a is the amount of energy assigned by the central controller to this local group. P_g represents the energy buffer state of this group. Then the evolution process of P_g can be shown as follows:

$$P_g(k+1) = P_g(k) + P_a(k) - P_b(k). \tag{4.21}$$

When all the indoor temperatures in the group decrease by ΔT_l from their set-points, the state of the energy buffer is 0; on the other hand, when all the indoor temperatures reach their upper bounds, the state P_g equals Cap. The energy buffer capacity can be obtained from:

$$Cap = (\Delta T_l + \Delta T_u) \sum_{i \in m} C_i, \tag{4.22}$$

where m is the group size, C_i is the effective heat capacity of the house with HVAC i in the unit of J/K, which is described in (4.2).

The central controller must guarantee the states of all the local controllers between 0 and their Cap. In this state space model of the MPC problem, the control actions are the amount of power assigned to all the local controllers. The measured disturbance is the vector containing each P_b for the corresponding group, and the output is the vector including all the group states and the load of all HVACs. The exact state space model formulas are similar to (4.7), (4.8) and are omitted due to space limitation.

The objective function of the central controller is formulated as follows:

Problem II (P2)

$$\min_{P_a^j(k)} : \quad \sum_{k\in N}\left\{(L_h(k)-r_w(k))^2+\lambda^2\sum_{j\in M}(P_g^j(k)-R_j(k))^2\right\}, \tag{4.23}$$

subject to:

$$\sum_{j\in M} P_a^j(k) = L_h(k), \tag{4.24}$$

$$0 \le P_g^j(k) \le Cap_j, \quad \forall j \in M, \forall k \in N, \tag{4.25}$$

$$0 \le P_a^j(k) \le U_j^{\max}, \quad \forall j \in M, \forall k \in N, \tag{4.26}$$

where M is the number of groups; R_j is the reference value of the state for group j, and we set it to half of the energy buffer capacity $Cap_j/2$; $U_j^{\max} = \sum_{i\in m} u_i^{\max}$ is the maximum amount of power allowed to be assigned to local group j in each time slot.

The first part of the objective function (4.23) represents the deviation of the actual HVAC load from the reference HVAC load (r_w). The second part represents the sum of the energy buffer state deviation from the reference state for each local group, which *indirectly* represents the impact on users' comfortableness. λ is used to make a tradeoff between these two parts. Constraint (4.25) requires that the state of each group should be bounded between 0 and Cap. (4.26) ensures that the power assigned to each group is bounded.

Similarly, the central controller will follow the receding horizon principle and update the states of all the groups after each time slot.

4.5.3 Local Controller Design

After the local controller receives the power quota P_a assigned by the central controller for the next time slot, it assigns this amount of power to all the HVACs by maximizing user comfort level. Note that the energy buffer state of the local group is guaranteed to be bounded by the central controller, as a result the temperatures of all the houses will not violate the temperature constraint (4.18).

The local controller determines the amount of power for each house by solving the following optimization problem.

Problem III (P3)

$$\min_{P_h^i(k)} : \quad \sum_{i\in S_l}(T_i(k+1)-r_i(k+1))^2, \tag{4.27}$$

subject to:

$$\sum_{i \in S_l} P_h^i(k) = P_a, \tag{4.28}$$

$$r_i(k+1) - \Delta T_l \leq T_i(k+1) \leq r_i(k+1) + \Delta T_u, \quad \forall i \in S_l, \tag{4.29}$$

$$0 \leq P_h^i(k) \leq u_i^{\max}, \quad \forall i \in S_l, \tag{4.30}$$

$$T_i(k+1) = T_i(k) - \Delta T_{off}^i(k) + Q_i P_h^i(k), \quad \forall i \in S_l, \tag{4.31}$$

where S_l is the set of HVACs under the local controller.

Since the distributed algorithm has a hierarchical architecture, the computation complexity of the central controller can be greatly reduced. However, the central controller cannot control each HVAC directly, neither can it know the exact status of each HVAC. As a result, it may lead to some fairness problems to the HVACs because HVACs under different local controllers may be treated differently. In addition, the control variables for the central controller are reduced (from the number of HVACs to the number of local controllers), so the control precision may not be as good as the centralized algorithm. We will compare the performance of the distributed algorithm with the centralized one in Section 4.7.

4.6 HVAC ON/OFF State Control

In problems **P1**, **P2** and **P3**, we assume that the consumed power for any HVAC in each time slot can be adjusted continuously. However, this may not be true in practice. For instance, some HVACs can only be turned on or off. Therefore, we consider how to change the proposed algorithms to support this kind of control actions.

For the proposed centralized MPC control algorithm, in order to support HVAC ON/OFF control, we can simply replace (4.19) with (4.32) in **P1**.

$$P_h^i(k) \in \{0, u_i^{\max}\}, \quad \forall i \in S, \forall k \in N. \tag{4.32}$$

W.r.t. the distributed MPC algorithm, we can replace (4.30) with (4.33) in **P3**, and let the local controller report the total amount of energy actually used, to the central controller which then updates the original control actions.

$$P_h^i(k) \in \{0, u_i^{\max}\}, \quad \forall i \in S_l, \tag{4.33}$$

This turns the original problems into multiple integer problems (MIP) which usually have much higher computational complexity. Therefore, we propose an heuristic algorithm which can obtain the control actions in polynomial time.

Algorithm 2 Determine HVAC State

Require: P_a

1: $PState[1, 2, \cdots, m] \leftarrow OFF$
2: sort $PState$ according to the difference between room temperature and the set-point from low to high
3: $sum \leftarrow 0$
4: **for all** $i \in m$ **do**
5: **if** $T_i^-(k+1) \leq r_i(k) - \Delta T_l$ **then**
6: $sum \leftarrow sum + P_h^i$
7: $PState(i) \leftarrow ON$
8: **end if**
9: **end for**
10: **for all** $i \in m$ **do**
11: **if** $PState(i) = OFF$ and $sum < P_a$ and $T_i^+(k+1) \leq r_i(k) + \Delta T_u$ **then**
12: **if** $P_g(k) \leq R(k)$ **then**
13: $sum \leftarrow sum + P_h^i$
14: $PState(i) \leftarrow ON$
15: **else if** $sum + P_h^i < P_a$ **then**
16: $sum \leftarrow sum + P_h^i$
17: $PState(i) \leftarrow ON$
18: **end if**
19: **end if**
20: **end for**

In Algorithm 2, m is the number of HVACs in the group; $PState$ is an vector which stores the state of all the HVACs; $T_i^-(k+1)$ represents the temperature of house i in the next time slot if the state of the HVAC is OFF; $T_i^+(k+1)$ represents the temperature of house i in the next time slot if the state of the HVAC is ON; $R(k)$ is the reference value of the state in the current time slot.

Line 2 guarantees that the houses with lower temperature have a higher priority. Lines 4 to 9 of Algorithm 2 set the state of HVAC whose house temperature is very close to the lower bound to be ON. Line 2 guarantees that the houses with lower temperature have a higher priority. If the current group state $P_g(k)$ is smaller than the reference state $R(k)$, we allow one more HVAC to be turned on (from lines 12 to 14). In this case, the total amount of power used may be greater than P_a. Otherwise we let sum always be smaller than P_a (from lines 15 to 17).

After obtaining the ON/OFF states of each HVAC, the local controller will report the actually used energy and the actual group state to the central controller. The central controller will update the model states and move to the next time slot.

4.7 Performance Evaluation

In this section, we evaluate the performance of the proposed algorithms and compare them with the algorithm in [64] (we call it "SPDW" algorithm, which is the combination of the first letters of all the authors' names, for easy reference) and the uncontrolled one. To achieve a flatten overall non-renewable power generation, the SPDW algorithm tries to minimize the difference of the load between the current and the last time slot by adjusting the set-points of HVACs. The parameters of our algorithms are tuned using the existing approaches for MPC [92], which aims to minimize (4.34). After tuning, the values of the weights in our simulation are set to 3.3 and 0.28 for centralized MPC algorithm (CMPC) and distributed MPC algorithm (DMPC), respectively. The control interval is set to 2 minutes. The communication latency between the controllers and HVACs is negligible and can be ignored. For the data gathering time delay, we assume that the users will send their states to the controller at the beginning of each time slot. Without considering the communication latency and the transmission error, all the data should be obtained within seconds. With respect to the computation time of the proposed algorithms, the proposed CMPC algorithm does need a long time. However, the proposed DMPC is much faster. For example, To apply the DMPC algorithm to a community with 1000 HVACs, we can divide it into 20 groups with 50 HVACs in each group. The computation of either the central controller and the local controller for one time slot is below 0.4s. Since the room temperature will not change significantly within several seconds, the influence of the data gathering time and the computation time of the proposed algorithm are both tolerable. The wind energy generation data and users' load are obtained from Eirgrid [16]. While these data have a resolution of 15 minutes, we use shape-preserving piecewise cubic interpolation [51] to interpolate them into a resolution of 2 minutes. Besides, we scaled them down to fit a micro-grid with a population size of 1000 and let the wind energy penetration take about 50% of the total load.

For the HVAC and house model, we assume that each house only has one room and one HVAC for simplicity, and the decrement of room temperature in each time slot follows a normal distribution, with an average of 0.2 Celsius, and a standard

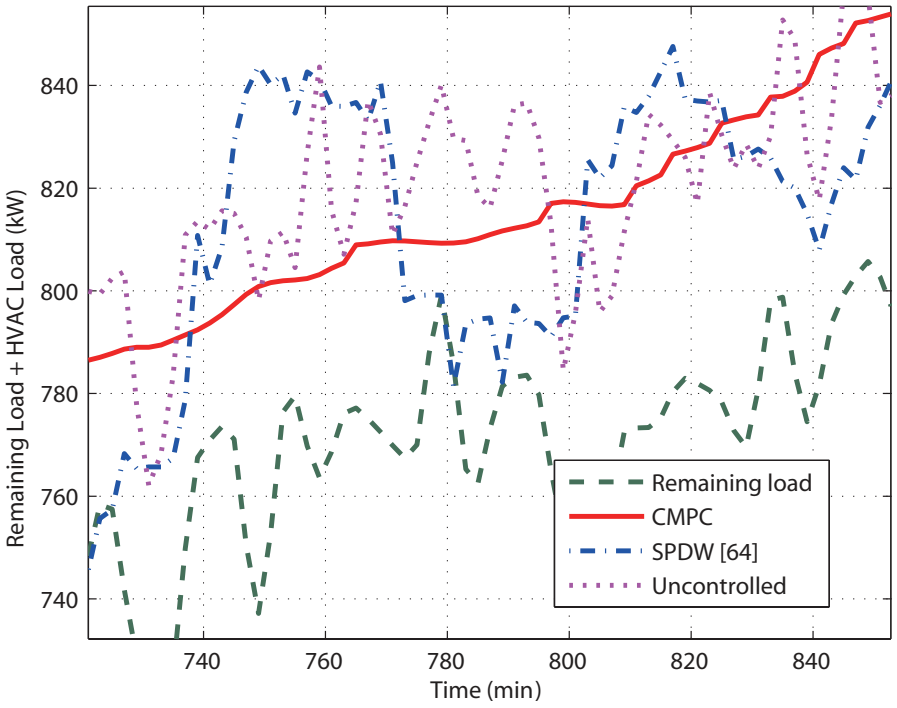
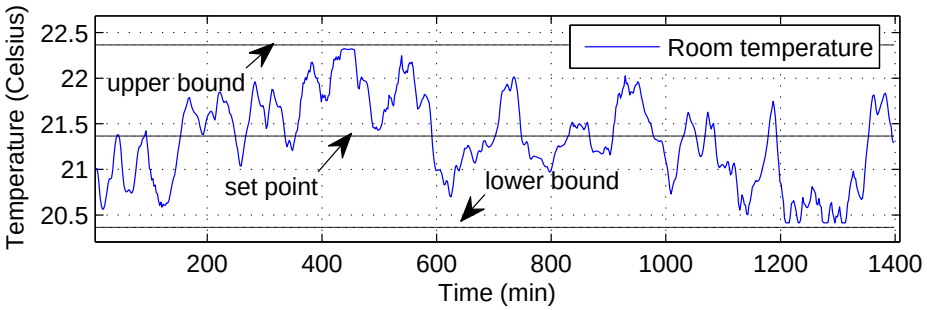


Figure 4.5: Load for conventional power plants

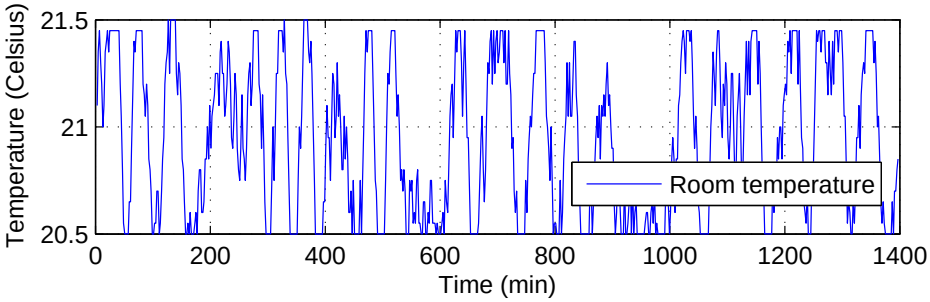
deviation of 0.03 Celsius. The parameter Q also follows a normal distribution, with an average value of 1.65×10^{-4} and a standard deviation of 0.25×10^{-4}. The maximum power of each HVAC is uniformly distributed between $4kW$ and $6kW$. Note that the HVAC power used in our simulation may be different from that in practice, but it will not affect the effectiveness of the proposed algorithms. The parameters ΔT_l and ΔT_u are both set to one Celsius. The indoor set-point is uniformly distributed from 20 Celsius to 22 Celsius. Initially, all the indoor temperatures are scatted around their set-points by at most 0.5 degree.

In this simulation, the prediction horizon in the proposed CMPC is set to 30 minutes. Due to the computation complexity of the controller, the number of controlled HVAC in this simulation is only 40. We will show the simulation result with more HVACs using the DMPC later, which is much faster. The reason is that DMPC has a hierarchical architecture, so either the central controller or the local controller has much fewer variables to optimize in each time slot.

Fig. 4.5 shows a typical time period of the power provided by conventional power



(a) Room temperature of CMPC



(b) Room temperature of SPDW [64]

Figure 4.6: Room temperature

plants. As illustrated in the figure, both SPDW and CMPC can make the load smoother and flatter. Notice that the SPDW algorithm can flatten the load for a while (from minute 750 to 770), then the load suddenly decreases and is kept flat for another period of time (from minute 770 to 800) before another increase. The reason is that the SPDW algorithm will keep the load as flat as possible by adjusting the ON/OFF state of all the HVACs until the elastic capacity of all the HVACs is no longer enough to provide further demand response. Then it will make a dramatic load change to push the HVAC away from the temperature bound so they can continue to provide demand response. For example, around minute 770, the majority of HVACs have reached the temperature upper bound so they are all turned off which make the load decrease tremendously. The opposite situation happens around minute 800. The proposed CMPC algorithm predicts the status of all the rooms in the future and changes the water level beforehand so that the overall load change is much smoother.

To measure the performance of different algorithms numerically, we define a criterion "average fluctuation (AF)" to represent the amount of load fluctuation as follows

(similar criterion can be found in [45]):

$$AF = \frac{1}{N-1} \sum_{i=1}^{N-1} \frac{|data(i+1) - data(i)|}{data(i)}, \tag{4.34}$$

where N is the total number of time slots; $data(i)$ is the data in the i-th slot; the denominator is used to normalize the data difference.

The average fluctuation for CMPC, SPDW and uncontrolled cases are 0.0028, 0.0123 and 0.0166, respectively. We find the proposed algorithm has a much smaller AF value which mainly because of the following three reasons. First, if the HVAC buffer capacity may not be enough in the future, the proposed algorithm will get prepared by changing the water level gradually, so the average fluctuation can be reduced. Second, the HVAC models in the proposed algorithm are assumed to use any amount of energy during the time slot while the SPDW algorithm only controls the ON/OFF states of the HVACs. We will discuss the performance of the proposed distributed MPC algorithm with ON/OFF support (DMPCOF) later. Third, the HVACs in the proposed algorithm are controlled directly by the central controller while the SPDW algorithm only control the set-point of all HVACs which indirectly affect the ON/OFF states of HVACs.

Fig. 4.6 (a) shows the temperature variation of a typical room using the proposed CMPC algorithm. The set-point is not an integer because we set random set-point for each house. As can be seen, CMPC can effectively guarantee user comfort level by restricting room temperature's variation within a range of the set-point. Fig. 4.6 (b) represents the set-point of all the controlled HVACs by SPDW. Note that the individual house temperature is allowed to deviate from the set-point for 0.5 Celsius, so the temperature variation range of a house is the same as that in CMPC. Comparing these two figures, we will find the temperature variation of the proposed algorithm is much smoother which leads to a higher user comfort level.

For the DMPC and DMPCOF algorithms, we set the number of local controllers to be 4, and each local controller manages 8 to 12 HVACs, with a total number of 40. The prediction horizon is also 30 minutes.

Fig. 4.7 illustrates the zoomed-in load of SPDW, CMPC, DMPC and DMPCOF algorithms. The load of CMPC is slightly below the others due to different water levels. Since the control actions and system model are a bit different, the water level of CPMC and DMPC may not be the same all the time. The curve of DMPCOF

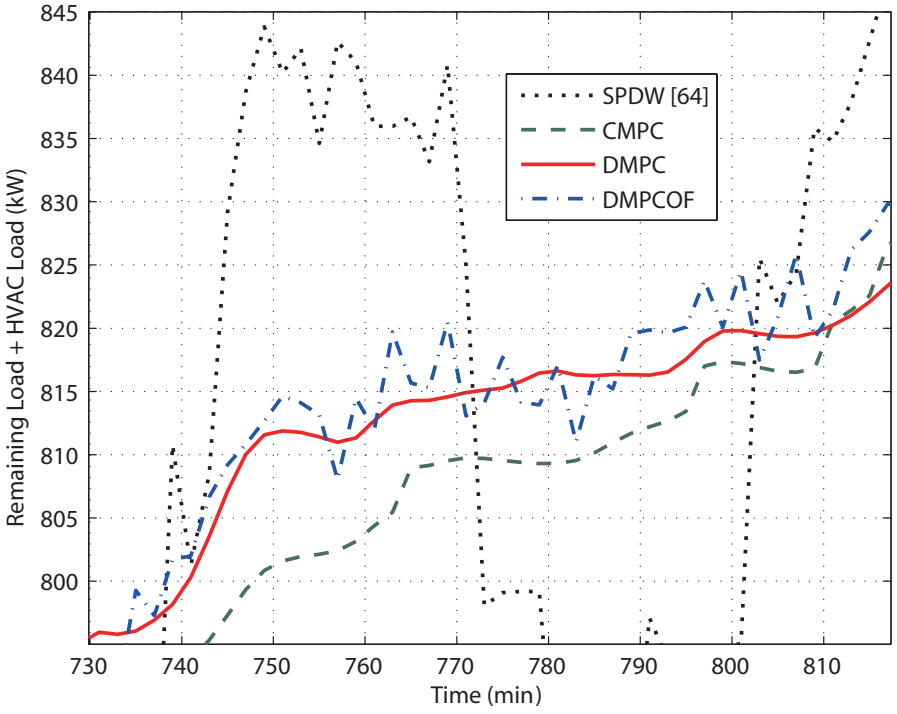


Figure 4.7: Load comparison of proposed three algorithms

Table 4.1: Average fluctuation

Prediction Horizon	DMPC with 40 HVACS	DMPCOF with 40 HVACS	DMPCOF with 80 HVACS
10min	0.0032	0.0058	0.0052
20min	0.0030	0.0052	0.0048
30min	0.0029	0.0050	0.0046

contains more fluctuations because the load change is discrete rather than continuous. However, the fluctuation is still much smaller than that of SPDW.

The AF of DMPC and DMPCOF are 0.0029 and 0.0050 respectively, and slightly larger than that of CMPC but still much smaller than that of SPDW and the uncontrolled cases. The reason why the AF of CMPC is smaller may be that DMPC has fewer control variables for the central controller.

In addition, we define the user comfortable influence factor (UIF) as the root mean square of all the room temperature deviation from the set-point.

We find the UIF of DMPC and DMPCOF are about 8.23% and 72.4% larger than that of CMPC. The reason is that CMPC minimizes UIF directly while DCMP and DCMPOF minimize it indirectly by controlling the energy buffer state for each local controller. Besides, the HVACs of DCMPOF can only be turned ON or OFF which makes its UIF even larger.

With a different prediction horizon, the total load for the conventional power plants is anticipated to be different. Fig. 4.8 shows the situation when the prediction horizon is 10 minutes and 30 minutes, respectively. The total number of HVACs are both 40.

From Fig. 4.8, the curve DMPC-30 (corresponding to 30-minute prediction horizon) is smoother than that of DMPC-10 (corresponding to 10-minute prediction horizon) because it contains fewer ups and downs. The reason is that with a longer prediction horizon, the controller has more information about the future load change and thus can get prepared earlier.

Table 4.1 shows the average fluctuation for DMPC, DMPCOF with 40 HVACs and DMPCOF with 80 HVACs (4 groups, each group has 16 to 24 HVACs). Obviously, the average fluctuation of DMPCOF is larger than that of DMPC with the same prediction horizon, and with a longer prediction horizon the average fluctuation is smaller. We also notice that with more HVACs, the average fluctuation of DMPCOF is smaller under the same prediction horizon too. The reason is that with more

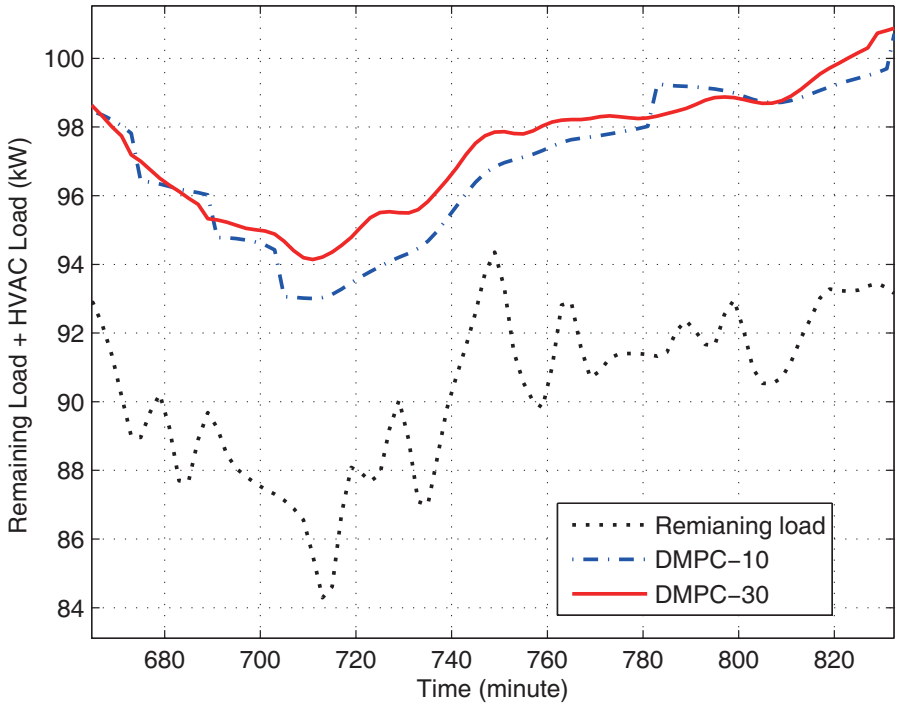


Figure 4.8: Load with different prediction horizon and HVAC number

HVACs to be controlled in a group, the amount of actual HVAC load can have a better chance to be closer to the reference HVAC load value.

4.8 Conclusion

In this chapter, we have proposed two algorithms to control HVACs for demand response based on MPC. The centralized approach directly controls all the HVACs while the distributed approach uses a hierarchical architecture. Both of them can effectively reduce load fluctuations while keeping all the room temperature within a range of the set-point. Moreover, the proposed distributed algorithm has been extended under a more practical assumption that each HAVC can only support ON and OFF.

Chapter 5

Optimal Combined Heat and Power System Scheduling

5.1 Introduction

CHP systems can generate both electricity and thermal energy simultaneously from a single fuel source, and can achieve a much higher energy efficiency than generating electricity and heat separately.

In this chapter, we will consider the CHP system scheduling problem from the perspective of both the users and the power company, assuming that the electricity price changes in real time. On one hand, the users can reduce their energy bill by using the battery pack and the water tank of the CHP system as the energy buffer. We first formulate the queueing models for the CHP systems, and then propose an algorithm based on the Lyapunov optimization technique which does not need any statistical information about the system dynamics. Different from existing works which only consider the electricity flow, the system we considered includes both the electricity flow and the thermal flow. The main challenge is that these two flows are coupled by the CHP system, and can inter-play with each other.

On the other hand, given the scheduling policy of the CHP systems at the users' side. How to set an appropriate RTP so that all the CHP systems can be coordinated to provide load shaping service is a big challenge. In order to determine the desired RTP, we need to solve a nested optimization problem which is very challenging. However, based on the feature of the specific problem, we propose a binary search algorithm which can find the optimal RTP in $O(\log n)$ time.

The contributions of this chapter are three-fold. First, we propose a comprehensive model from the perspective of a commercial customer, which incorporates both the electricity and thermal energy queues. We investigate the relationship of these two queues to minimize the average cost. Second, we propose an algorithm to approximately achieve the optimal average cost, considering the limited capacities of the battery pack and the water tank. The algorithm does not require any statistical information of the system dynamics such as electricity and hot water demands, etc. To obtain the optimal scheduling decision, we discuss when we can use the specific features of the problem to turn a non-convex optimization problem into a convex one which can be solved in real time. Third, we discuss how to set the appropriate RTP to coordinate all the CHP systems indirectly to provide load shaping services. The time complexity of the proposed searching algorithm is $O(\log n)$.

5.2 Related Work

Thanks to the ubiquitous communications technologies, it is possible to optimize the provisioning and delivery of various energy sources to achieve a higher efficiency [87, 21]. To provide both electricity and heat economically, the design and operation strategies of CHP systems have been well investigated. [26] discussed operating strategies, such as heat and electricity load following, for three micro-CHP technologies. [53] evaluated four typical operation modes in a hotel based on measured electric and heating loads. [15] analyzed the utilization of micro-CHP systems in conjunction with domestic household appliances. [42] analyzed the cost for different fuel-cell systems. These works tried to find the most cost-effective strategies from a system view, and do not consider the detailed control policies. The Combined Heat and Power Economic Dispatch (CHPED) problem, first raised in [72], aimed to find the optimal operation point of CHP with minimum energy cost such that both electricity and heat demands were met. A two-level strategy to separate the objective function and constraints was adopted in [72]. Besides the traditional mathematical approaches, evolutionary computation techniques were used to improve the performance [82]. However, in the CHPED problems, optimization was performed to minimize the cost in each time slot. No energy buffer was used to minimize the long-term cost. In addition, it did not consider the stochastic nature of energy demand.

Various optimization technologies have been used to optimize the cost of a smart grid system with energy buffers. T. Chang *et al.* used the dynamic programming and

decomposition approach to minimize the total cost from the perspective of each user and the whole micro-grid, respectively [8]. In their work, the distributions of all the stochastic variables such as the load and real-time price information were assumed to be available from historical data. [33] proposed a threshold based energy storage control policy that minimizes the long-term average grid operational cost based on dynamic programming. Model predictive control (MPC) has also been applied to obtain the optimal control policies. For example, T. G. Hovgaard *et al.* proposed an economic MPC algorithm to minimize the total cost of distributed power generation plants by using large cold rooms as the energy buffer. Different from these approaches, our proposed algorithm tries to minimize the long-term average cost without the need to estimate the statistical system dynamics from historical data.

There are also several works which use the Lyapunov optimization technique to construct low complexity energy storage management policies. M. J. Neely *et al.* minimized the time average cost from the perspective of one user, and guaranteed the worst-case delay for each elastic load in [55]. In [89], the authors used uninterruptible power systems (UPS) in the data center to reduce the electricity bill in a real-time price environment. Their model did not consider renewable energy sources. Guo *et al.* investigated how to use a household battery to minimize the average electricity cost, considering both inelastic and elastic load in [24]. Instead of guaranteeing the worst-case delay, [29] guaranteed that the percentage of the delayed elastic load was less than a threshold. These works discussed above only considered one energy buffer, however, the system model discussed in our work includes two energy buffers, the battery pack and the water tank, which are correlated by the CHP system. With two dependent queues, the system model is more complicated and we need to solve a non-convex optimization problem to obtain the optimal control policy. In addition, we illustrate the relationship between the capacity of these two energy buffers and the minimum required capacity to achieve the optimal performance. This chapter focuses on the problems closely related to the unique features of the CHP systems. Some well-studied applications of Lyapunov optimization in smart grid, such as elastic load queue, worst-case delay, etc, are not discussed here due to the space limitation.

How to determine an appropriate real-time price to coordinate all the "selfish" users is also a challenging task. [96] proposed a pricing scheme to stimulate a large group of electrical vehicle users to provide frequency regulation based on game theory. Y. Gao extended their work by considering different users' preferences under the presence of information asymmetry using contract theory [21]. One problem in the

above work is that the users' preferences may keep changing over time, and are highly related to the state-of-charge of the electrical vehicles' battery. Therefore, the performance gain of the above user preference learning algorithm may be limited. In order to obtain a better user reaction model to price signals, C. Chen *et al.* proposed an iterative methods based on a leader and follower level game theory that need frequent information exchange, which may lead to a high communication overhead [10]. In this chapter, different from the game theory approach, we propose a fast algorithm to determine the optimal real-time price which can effectively coordinate all the CHP system for load shaping services.

5.3 System Model

We consider how to minimize the average energy cost using the CHP device and energy buffers. The mathematical models of this system are discussed in this section.

5.3.1 System Architecture

Fig. 5.1 gives an overview of the CHP system, such as the one used in a hotel. $L_e(t)$ and $L_w(t)$ represent the electricity and hot water demands from users in each time slot, respectively, which are stochastic. $L_e(t)$ can be met by the electricity discharged from the battery $D(t)$ or bought from the power grid $G_l(t)$. $L_w(t)$ is met by the hot water stored in the water tank.

In each time slot, the CHP device can generate electricity, in the amount of $\eta_{ce}P_c(t)$, to charge the battery and hot water, in the amount of $\eta_{cg}P_c(t)$, to fill the water tank, where $P_c(t)$ is the amount of the natural gas consumed by the CHP, η_{ce} is the conversion efficiency from natural gas to the amount of the electricity charged to the battery, and η_{cg} is the conversion efficiency from natural gas to the amount of hot water. Meanwhile, if the battery is full or the grid electricity price is high, the electricity generated from the CHP, in the amount of $\eta_{co}P_c(t)$, can be sold back to the grid with the conversion efficiency η_{co}. The parameter $r(t)$, ranging from 0 to 1 is used to make a tradeoff between the amount of electricity used to charge the battery and that sold to the grid.

Note that we did not let the power generated from the CHP supply the user's electricity demand $L_e(t)$ directly in the above model to simplify the analysis. The reason is that we assume the electricity price bought from and sold to the power grid

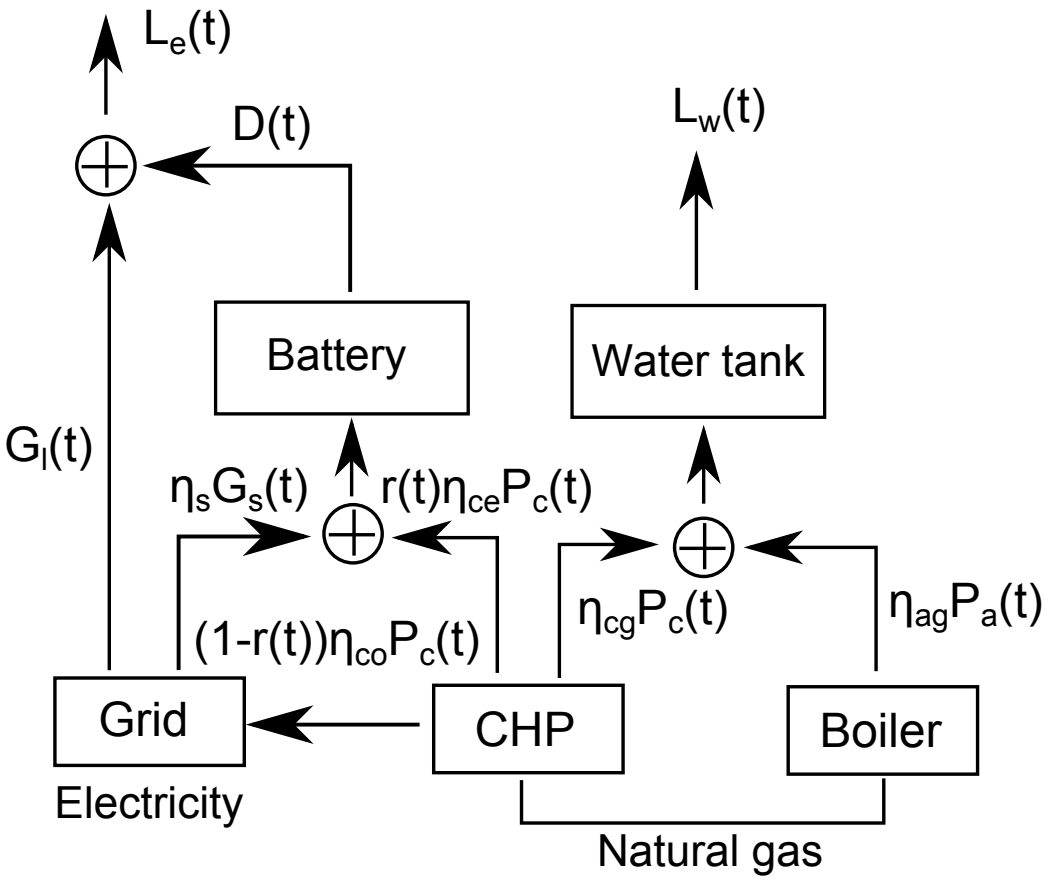


Figure 5.1: The CHP system uses natural gas

are the same, so whether the electricity is used to supply the user's demand directly or sold back to the grid does not affect the total energy cost (if we do not sell the electricity, less electricity is bought from the grid).

Since the electricity price in the real-time electricity market changes according to the supply and demand, in this chapter we assume the real-time electricity price $C_e(t)$ for the next time slot is known ahead of time. $C_e(t)$ is bounded in the range $[C_{e,\min}, C_{e,\max}]$. On the other hand, the price of the natural gas does not change frequently and the percentage of the change is usually not large, so we assume it is constant in each time slot. However, the proposed algorithm is still applicable if we also consider the real-time gas price because the control decisions of the proposed algorithm are made upon the current system states in each time slot, including the natural gas price.

To minimize the average energy cost in the long term, in each time slot the controller determines the amount of electricity $G_l(t)$ and $G_s(t)$ bought from the grid to supply the electricity demand and charge the battery, the amount of the natural gas $P_c(t)$ consumed by the CHP and the amount of the natural gas $P_a(t)$ consumed by the boiler. It also needs to determine the value of $r(t)$ which specifies the dispatch of the generated electricity from the CHP to the battery. The parameter η_s in Fig. 5.1 represents the battery charging efficiency, and η_{ag} represents the conversion efficiency from natural gas to the amount of hot water using the boiler.

The intuition is that the controller discharges the battery and makes the CHP generate more electricity to meet the high electricity demand or sell to the grid to earn profit when the electricity price is high. On the contrary, the controller charges the battery using the electricity from the grid when the electricity price is low. This problem is challenging because we do not know the distribution of the electricity and hot water demand, nor do we know the distribution of the real-time electricity price. Of course, we can use dynamic programming to estimate the distribution of these stochastic variables, but it is usually computationally complex and may have the "curse of dimensionality" problem [4].

5.3.2 Electricity Queueing Model

In practice, although the lifetime of the battery may be influenced by the charging and discharging process, etc, we do not take them into account. Besides, we assume that the state of charge (SOC) of the battery, viewed as the energy queue of the

battery, is linear to simplify our analysis. However, the proposed algorithm will not be largely affected if we incorporate more complicated battery models because the proposed algorithm only needs to know the current battery status to make control decisions.

The SOC level of the battery $B(t)$ evolves according to the following equation:

$$B(t+1) = B(t) - D(t) + \eta_s G_s(t) + r(t)\eta_{ce} P_c(t). \tag{5.1}$$

Obviously, in any slot t, the battery needs to have the following capacity and charge/discharge constraints.

$$0 \leq B(t) \leq B_{\max}, \tag{5.2}$$

$$0 \leq D(t) \leq D_{\max}, \tag{5.3}$$

$$0 \leq \eta_s G_s(t) + r(t)\eta_{ce} P_c(t) \leq C_{char}, \tag{5.4}$$

where $B_{\max}$ is the capacity of the battery, $D_{\max}$ is the maximum discharge rate of the battery, and C_{char} is the maximum charge rate of the battery.

The amount of electricity drawn from the grid in one time slot is also bounded by $P_{e,\max}$:

$$0 \leq G_l(t) + G_s(t) \leq P_{e,\max}, \tag{5.5}$$

$$0 \leq G_l(t) \leq G_{l,\max}, 0 \leq G_s(t) \leq G_{s,\max}, \tag{5.6}$$

where $G_{l,\max}$ and $G_{s,\max}$ are the upper bound of $G_l(t)$ and $G_s(t)$, respectively. Since the grid can meet the commercial power demand most of the time, we assume $P_{e,\max} \geq L_{e,\max}$ where $L_{e,\max}$ is the upper bound of $L_e(t)$.

5.3.3 Water Queueing Model

The water tank discussed here is assumed an ideal one, so we do not consider heat leakage. A more practical water tank model can easily be applied as we can consider the amount of the heat needed to reheat the water tank as the additional heat demand in the form of hot water from the users.

The amount of hot water stored in the water tank, which is the queue length of

the water tank, evolves according to the following equation:

$$W(t+1) = W(t) - L_w(t) + \eta_{cg} P_c(t) + \eta_{ag} P_a(t), \tag{5.7}$$

where $W(t)$ is the water level in the water tank in slot t.

Since the amount of the water stored in the water tank should always be bounded by the size of the water tank, we have: $0 \leq W(t) \leq W_{\max}$, where $W_{\max}$ is the capacity of the water tank. In addition, since we assume the hot water demand in each time slot will not exceed $L_{w,\max}$, to make sure that users' demand can always be met even in the worst-case situation, i.e., the hot water demand is always $L_{w,\max}$, we assume the following constraint holds:

$$L_{w,\max} \leq \eta_{ag} P_{a,\max}, \tag{5.8}$$

where $P_{a,\max}$ is the maximum amount of the natural gas used by the boiler in each time slot, and $L_{w,\max}$ is the upper bound of the hot water demand in each time slot.

5.3.4 Control Objective

In each time slot, the total energy cost for the CHP system is the sum of the electricity and natural gas cost minus the amount of the electricity sold to the grid:

$$f(t) = C_e(t)\{G_l(t) + G_s(t) - (1 - r(t))\eta_{co} P_c\} + C_g\{P_c + P_a\}, \tag{5.9}$$

where C_g is the natural gas price.

The control objective is to find a control policy determining the amount of the electricity and natural gas dispatched in each time slot, so as to minimize the long-term average energy cost.

$$f_{avg} = \lim_{T \to \infty} \frac{1}{T} \sum_{i=0}^{T-1} \mathbb{E}\{f(i)\}. \tag{5.10}$$

5.4 The CHP System Scheduling Algorithm

In this section, we assume the electricity and hot water demands in each time slot $L_e(t)$, $L_w(t)$ are independent. The proposed algorithm in this section will solve the following problems. First, given the current states of the CHP system, including the

electricity and hot water demand, battery and water tank storage level, electricity price in the current time slot, etc, how to obtain the optimal control decisions with a low computational complexity and can adapt to the stochastic system dynamics while still provide a good performance? Second, what is the minimum capacity of the battery pack and water tank we should have to achieve a given performance requirement? Third, since the CHP can generate both electricity and heat, the battery pack and water tank queues specified in (5.1) and (5.7) are dependent. What is the relationship between the capacity of the battery pack and that of the water tank?

According to the system architecture and control objective described in Section 5.3, the problem can be formulated as the following stochastic optimization problem.

Problem One (P-I)

$$\min_{D(t),r(t),G_l(t),G_s(t),P_c(t),P_a(t)} P_1 = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{f(t)\}, \tag{5.11}$$

subject to

$$B(t+1) = B(t) - D(t) + \eta_s G_s(t) + r(t)\eta_{ce} P_c(t), \tag{5.12}$$

$$W(t+1) = W(t) - L_w(t) + \eta_{cg} P_c(t) + \eta_{ag} P_a(t), \tag{5.13}$$

$$0 \le B(t) \le B_{\max}, \tag{5.14}$$

$$0 \le W(t) \le W_{\max}, \tag{5.15}$$

$$L_e(t) = G_l(t) + D(t), \tag{5.16}$$

$$0 \le \eta_s G_s(t) + r(t)\eta_{ce} P_c(t) \le C_{char}, \tag{5.17}$$

$$0 \le r(t) \le 1, P_c(t), G_l(t), G_s(t) \ge 0, \tag{5.18}$$

$$0 \le D(t) \le D_{\max}. \tag{5.19}$$

The above problem cannot fit into the stochastic optimization framework directly mainly because of the battery and water tank capacity constraints (5.14) and (5.15). Specifically, stochastic optimization can only guarantee the average energy generation equals the average consumption in the long term, but cannot provide a hard bound on the difference between the generation and consumption in any time slot. To solve this problem, we take the expectation on the two sides of (5.12) and (5.13), which leads to the following relaxed problem:

Problem Two (P-II)

$$\min_{D(t),r(t),G_l(t),G_s(t),P_c(t),P_a(t)} P_1 = \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\{f(t)\}, \tag{5.20}$$

subject to

$$\overline{D(t)} = \eta_s \overline{G_s(t)} + \eta_{ce}\overline{r(t)P_c(t)}, \tag{5.21}$$

$$\overline{L_w(t)} = \eta_{cg}\overline{P_c(t)} + \eta_{ag}\overline{P_a(t)}, \tag{5.22}$$

and (5.16), (5.17), (5.18), (5.19).

P-II fits the stochastic optimization framework, so we can solve it using existing algorithms [54, 10]. Obviously, only when the solutions to P-II can meet the constraints (5.14) and (5.15) for $\forall t \in T$, they are also feasible to P-I. To reach this objective, we define two constants θ and ε. The intuition is that by adjusting these two constants appropriately, we can make the solutions to P-II also be feasible to P-I.

To start, we define two queues $E(t)$ and $X(t)$:

$$E(t) = B(t) - \theta, \tag{5.23}$$

$$X(t) = W(t) - \varepsilon. \tag{5.24}$$

The constants θ and ε are two queue offsets, which are used to guarantee that the two queues $B(t)$ and $W(t)$ are bounded.

From (5.12) and (5.13), we can obtain the queueing dynamics:

$$E(t+1) = E(t) - D(t) + \eta_s G_s(t) + r(t)\eta_{ce}P_c(t), \tag{5.25}$$

$$X(t+1) = X(t) - L_w(t) + \eta_{cg}P_c(t) + \eta_{ag}P_a(t). \tag{5.26}$$

We then define the Lyapunov function $Q(t) = \frac{1}{2}E(t)^2 + \frac{1}{2}X(t)^2$. The conditional one-slot Lyapunov drift is:

$$\Delta(t) = \mathbb{E}\{Q(t+1) - Q(t)|E(t), X(t)\}. \tag{5.27}$$

Here, the battery queue and the water tank queue are of equal weight. Our algorithm can be extended if we assign different weights to them.

According to (5.25) and (5.26), by squaring both sides, we have:

$$\begin{aligned}\Delta(t) \leq & 0.5\max[(\eta_s G_{s,\max}+\eta_{ce}P_{c,\max})^2, D_{\max}^2] \\ & - E(t)[D(t)-\eta_s G_s(t)-r(t)\eta_{ce}P_c(t) \\ & + 0.5\max[(\eta_{cg}P_{c,\max}+\eta_{ag}P_{a,\max})^2, L_{w,\max}^2] \\ & - X(t)[L_w(t)-\eta_{cg}P_c(t)-\eta_{ag}P_a(t)] \\ = & B - E(t)[D(t)-\eta_s G_s(t)-r(t)\eta_{ce}P_c(t)] \\ & - X(t)[L_w(t)-\eta_{cg}P_c(t)-\eta_{ag}P_a(t)],\end{aligned}$$

where $P_{c,\max}$ is the maximum amount of the natural gas that can be used by the CHP in each time slot, and B is a constant and defined as

$$\begin{aligned}B = & 0.5\max[(\eta_s G_{s,\max}+\eta_{ce}P_{c,\max})^2, D_{\max}^2] \\ & + 0.5\max[(\eta_{cg}P_{c,\max}+\eta_{ag}P_{a,\max})^2, L_{w,\max}^2].\end{aligned} \tag{5.28}$$

According to the stochastic optimization framework, in order to make the two queues $E(t)$ and $X(t)$ mean rate stable, we must minimize the drift $\Delta(t)$. In addition, our control objective is to minimize the average cost. So we use a constant V to represent the tradeoff between these two objectives. Then the drift plus penalty function can be written as follows.

$$\begin{aligned}& \Delta(t) + V\mathbb{E}\{f(t)\} \\ & \leq B - E(t)\mathbb{E}\{D(t)-\eta_s G_s(t)-r(t)\eta_{ce}P_c(t)|E(t)\} \\ & \quad - X(t)\mathbb{E}\{L_w(t)-\eta_{cg}P_c(t)-\eta_{ag}P_a(t)|X(t)\} \\ & \quad + V\mathbb{E}\{C_e(t)\{G_l(t)+G_s(t)-(1-r(t))\eta_{co}P_c\} \\ & \quad + C_g\{P_c+P_a\}\}.\end{aligned} \tag{5.29}$$

We then substitute $G_l(t)$ in (5.29) according to (5.16), and after some manipula-

tion we can obtain:

$$\begin{aligned}
&\Delta(t) + V\mathbb{E}\{f(t)\} \\
&\leq B + V\mathbb{E}\{C_e(t)L_e(t)|E(t)\} - \mathbb{E}\{L_w(t)X(t)|X(t)\} \\
&\quad - \mathbb{E}\{D(t)[E(t) + VC_e(t)]|E(t)\} \\
&\quad + \mathbb{E}\{G_s(t)[\eta_s E(t) + VC_e(t)]|E(t)\} \\
&\quad + \mathbb{E}\{P_c(t)[r(t)\eta_{ce}E(t) + \eta_{cg}X(t) \\
&\quad - (1 - r(t))\eta_{co}VC_e(t) + C_gV]|E(t), X(t)\} \\
&\quad + \mathbb{E}\{P_a(t)[\eta_{ag}X(t) + VC_g]|X(t)\}.
\end{aligned} \tag{5.30}$$

Based on the "min-drift" principle of the Lyapunov optimization approach, the main idea of the proposed algorithm is to minimize the right-hand side (RHS) of (5.30) over all the feasible control policies in each time slot. In other words, at the beginning of each time slot, we observe the system states $B(t)$, $W(t)$, $L_e(t)$, $L_w(t)$, $C_e(t)$, determine the value of $B + V\mathbb{E}\{C_e(t)L_e(t)|E(t)\} - \mathbb{E}\{L_w(t)X(t)|X(t)\}$, and then solve the following problem:

Problem Three (P-III)

$$\min_{D(t),r(t),G_l(t),G_s(t),P_c(t),P_a(t)} \quad G_s(t)H_s(t) + P_c(t)H_c(r(t)) + P_a(t)H_a(t) - D(t)H_d(t), \tag{5.31}$$

subject to

$$L_e(t) = G_l(t) + D(t), \tag{5.32}$$

$$0 \leq \eta_s G_s(t) + r(t)\eta_{ce}P_c(t) \leq C_{char}, \tag{5.33}$$

$$0 \leq r(t) \leq 1, P_c(t), G_l(t), G_s(t) \geq 0. \tag{5.34}$$

$$0 \leq D(t) \leq D_{\max}, \tag{5.35}$$

where

$$\begin{aligned}
H_s(t) =& \eta_s E(t) + VC_e(t), \quad H_a(t) = \eta_{ag}X(t) + VC_g, \\
H_d(t) =& E(t) + VC_e(t), \quad H_c(r(t)) = H_r(t)r(t) + H_b(t), \\
H_r(t) =& \eta_{ce}E(t) + \eta_{co}VC_e(t), \\
H_b(t) =& \eta_{cg}X(t) - \eta_{co}VC_e(t) + C_gV.
\end{aligned} \tag{5.36}$$

Note that (5.31) contains the product of $P_c(t)$ and functions of $r(t)$, so P-III is a non-convex optimization problem because its Hessian matrix is not always positive

definite. When looking into the structure of P-III, we can find that $D(t)$ and $P_a(t)$ can be easily obtained according to the value of $H_d(t)$ and $H_a(t)$. If $H_d(t) \geq 0$, then $D(t) = \min\{D_{\max}, L_e(t)\}$; otherwise $D(t) = 0$. If $H_a(t) \leq 0$, then $P_a(t) = P_{a,\max}$; otherwise $P_a(t) = 0$. Therefore, we only have to solve the following subproblem:

$$\min_{D(t),r(t),G_l(t),G_s(t),P_c(t),P_a(t)} : \quad G_s(t)H_s(t) + P_c(t)[H_r(t)r(t) + H_b(t)], \tag{5.37}$$

subject to: (5.33) and (5.34).

Suppose (5.33) is not active and $0 < r(t) < 1$. Since $H_c(r(t))$ is a linear function of $r(t)$, then we can always increase or decrease $r(t)$ to make (5.37) smaller. Therefore, either (5.33) is active or $r(t)$ equals 0 or 1.

Suppose (5.33) is active. We can replace $G_s(t)$ using (5.33) in (5.37) and get:

$$\begin{aligned} \min_{D(t),r(t),G_l(t),P_c(t),P_a(t)} : \quad & (\eta_{co} - \frac{\eta_{ce}}{\eta_s})VC_e(t)P_c(t)r(t) \\ & + \frac{C_{char}}{\eta_s}[\eta_s E(t) + VC_e(t)] + P_c(t)H_b(t). \end{aligned} \tag{5.38}$$

Obviously, since $C_e(t), P_c(t) \geq 0$, if $\eta_{co} \geq \frac{\eta_{ce}}{\eta_s}$, then $r(t) = 0$; otherwise $r(t)$ should be as large as possible. If $\eta_{ce}P_{c,\max} \leq C_{char}$, then $r(t)$ can be 1 and we can use this fact to convert P-III to a linear optimization problem by substituting $r(t) = 0$ and $r(t) = 1$ into P-III, respectively, and choose the minimum value. Otherwise, $r(t)$ should be in the range of $[C_{char}/\eta_{ce}P_{c,\max}, 1]$, and since this range is not too large, we can use a search algorithm to obtain the optimal solution.

Next we need to prove that the solutions to P-III are also feasible to P-I. In other words, the solutions to P-III can meet constraints (5.14) and (5.15) for $\forall t \in T$.

Theorem 1: Suppose θ and ε are defined in (5.39) and (5.40), respectively,

$$\theta = \frac{VC_{e,\max}}{\eta_s} + \min\{D_{\max}, L_{e,\max}\}, \tag{5.39}$$

$$\varepsilon = \frac{VC_g}{\eta_{ag}} + L_{w,\max}. \tag{5.40}$$

Then through minimizing P-III, we can have the following results:

$$0 \leq B(t) \leq \theta + C_{char}, \quad \forall t \in T, \tag{5.41}$$

$$
\begin{aligned}
0 \leq W(t) \leq \max\{&\varepsilon + \frac{\eta_{ce}\theta - VC_g}{\eta_{cg}} + \eta_{cg}P_{c,\max}, \\
&\varepsilon + \frac{\eta_{co}VC_{e,\max} - VC_g}{\eta_{cg}} + \eta_{cg}P_{c,\max}, \\
&\varepsilon + \eta_{cg}P_{c,\max} + \eta_{ag}P_{a,\max}\}, \quad \forall t \in T,
\end{aligned} \tag{5.42}
$$

given that the above relationships are satisfied at $t = 0$.

Proof. First, we use induction to prove the upper bound of $B(t)$ and $W(t)$. Since it holds when $t = 0$, we assume it also holds at time slot t.

1) Suppose $B(t) \leq \theta$. Since $0 \leq \eta_s G_s(t) + r(t)\eta_{ce}P_c(t) \leq C_{char}$, we can have $B(t+1) \leq \theta + C_{char}$.

2) Suppose $B(t) > \theta$. According to (5.23), $E(t) > 0$. So we have $H_s(t) > 0, H_d(t) > 0$. To minimize P-III, it must be $G_s(t) = 0$ and $D(t) = \min\{D_{\max}, L_e(t)\}$. Besides, we can find that $H_c(r(t))$ is an increasing function with $r(t)$, so $P_c(t)H_c(r(t))$ reaches its minimum value when $r(t) = 0$. From all the above we can see whenever $B(t) > \theta$, the battery will discharge and do not charge. Therefore $B(t+1) \leq B(t) \leq \theta + C_{char}$.

3) Suppose $W(t) \leq \varepsilon$, it is obvious that $W(t+1) \leq \varepsilon + \eta_{cg}P_{c,\max} + \eta_{ag}P_{a,\max}$.

4) Suppose $W(t) > \varepsilon$, then $X(t) > 0$ according to (5.24), and $H_a(t) > 0$. To minimize P-III, $P_a(t)$ must be 0. Next we consider the following two cases. For the first case, if $\eta_{ce}E(t) + \eta_{co}VC_e(t) < 0$, $H_c(r(t))$ reaches its minimum value $-\eta_{ce}\theta + \eta_{cg}[W(t) - \varepsilon] + VC_g$ when $r(t) = 1$ and $B(t) = 0$. Therefore, when $W(t) > \varepsilon + \frac{\eta_{ce}\theta - VC_g}{\eta_{cg}}$, $H_c(r(t)) > 0$ and $P_c(t)$ is set to 0 to minimize P-III. In this case, $W(t+1) \leq W(t)$. On the other hand, $W(t+1) \leq W(t) + \eta_{cg}P_{c,\max} \leq \varepsilon + \frac{\eta_{ce}\theta - VC_g}{\eta_{cg}} + \eta_{cg}P_{c,\max}$.

   For the second case, if $\eta_{ce}E(t) + \eta_{co}VC_e(t) \geq 0$, $H_c(r(t))$ reaches its minimum value $-\eta_{co}VC_{e,\max} + \eta_{cg}[W(t) - \varepsilon] + VC_g$ when $r(t) = 0$. Therefore, when $W(t) > \varepsilon + \frac{\eta_{co}VC_{e,\max} - VC_g}{\eta_{cg}}$, $H_c(r(t)) > 0$ and $P_c(t)$ is set to 0 to minimize P-III. In this case, $W(t+1) \leq W(t)$. On the other hand, $W(t+1) \leq W(t) + \eta_{cg}P_{c,\max} \leq \varepsilon + \frac{\eta_{co}VC_{e,\max} - VC_g}{\eta_{cg}} + \eta_{cg}P_{c,\max}$.

From the above analysis, we can find that both $B(t)$ and $W(t)$ are upper bounded. So we can determine the capacity of the battery $B_{\max}$ and the capacity of the water tank $W_{\max}$ according to these upper bounds. Note that both $B_{\max}$ and $W_{\max}$ are

functions of the parameter V, so we can make a tradeoff between the energy buffer capacity and the average cost. On the other hand, with a given battery pack or water tank capacity, we can obtain the corresponding parameter V.

Second, we also use induction to prove the lower bound of the battery and the water tank.

5) Suppose $B(t) \geq \min\{D_{\max}, L_{e,\max}\}$, obviously $B(t+1) \geq 0$.

6) Suppose $0 \leq B(t) < \min\{D_{\max}, L_{e,\max}\}$, substitute (5.39), (5.23) and (5.24) in (5.36) we can have $H_d(t) < 0$,$H_s(t) < 0$. In order to minimize P-III, we must have $D(t) = 0$ and $G_s(t) \geq 0$. Since $P_c(t) \geq 0$, $B(t+1) \geq B(t) \geq 0$.

7) Suppose $W(t) \geq L_{e,\max}$, obviously $W(t+1) \geq 0$.

8) Suppose $W(t) < L_{e,\max}$, substitute (5.39), (5.23) and (5.24) in (5.36) we can have $H_a(t) < 0$. In order to minimize P-III, we must have $P_a(t) = P_{a,\max}$. According to (5.8), we have $L_w(t+1) \geq L_w(t) \geq 0$.

Since both the SOC of the battery pack and the water level in the water tank are bounded, the solution to P-III is also feasible to P-I. □

Theorem 2: If $C_e(t)$, $L_e(t)$, $L_w(t)$ are independent over slots, then the expected cost using the proposed algorithm over time is within bound B/V of the optimal cost. In other words

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{f^{'}(t)\} \leq P_1^* + B/V, \tag{5.43}$$

where $f^{'}(t)$ represents the energy cost in one time slot using the proposed algorithm, and P_1^* is the optimal solution to P-I.

To achieve P_1^*, we need to know the distributions of the stochastic variables $C_e(t)$, $L_e(t)$, and $L_w(t)$, which are difficult to obtain. Therefore, the proposed algorithm can provide a low-complexity approach to achieve a performance deviated no more than $O(1/V)$ from the optimal one.

Proof. Assume $(D^*(t), r^*(t), G_l^*(t), G_s^*(t), P_c^*(t), P_a^*(t))$ is the optimal policy to achieve P_1^*. Since the proposed algorithm is obtained by minimizing the RHS of (5.29), the value of the RHS of (5.29) should be no larger than that using the optimal policy.

Then we have:

$$\begin{aligned}
&\Delta(t) + V\mathbb{E}\{f^{'}(t)\} \\
\leq\ & B + V\mathbb{E}\{C_e(t)L_e(t)|E(t)\} - \mathbb{E}\{L_w(t)X(t)|X(t)\} \\
&- \mathbb{E}\{D^*(t)[E(t) + VC_e(t)]|E(t)\} \\
&+ \mathbb{E}\{G_s^*(t)[\eta_s E(t) + VC_e(t)]|E(t)\} \\
&+ \mathbb{E}\{P_c^*(t)[r^*(t)\eta_{ce}E(t) + \eta_{cg}X(t) \\
&- (1 - r^*(t))\eta_{co}VC_e(t) + C_gV]|E(t), X(t)\} \\
&+ \mathbb{E}\{P_a^*(t)[\eta_{ag}X(t) + VC_g]|X(t)\} \\
\leq\ & B + VP_1^*.
\end{aligned} \tag{5.44}$$

Taking the expectation on both sides, and summing over $t \in \{0, 1, 2, \cdots, T-1\}$, then we obtain

$$\mathbb{E}\{Q(T) - Q(0)\} + \sum_{t=0}^{T-1} V\mathbb{E}\{f^{'}(t)\} \leq TB + TVP_1^*. \tag{5.45}$$

Dividing both sides by TV, letting $T \to \infty$, and using the fact that both $Q(T)$ and $Q(0)$ are finite, we have:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{f^{'}(t)\} \leq P_1^* + \frac{B}{V}. \tag{5.46}$$

□

5.5 CHP Using Renewable Energy

In the previous section, we discussed the CHP system using the natural gas as the fuel. However, environmental concerns and the rising cost of fossil fuels make people to consider using renewable energy. Fortunately, some new CHP technologies can make use of certain renewable energy sources, such as Biomass and geothermal energy to generate both electricity and heat. In addition, as discussed in the CHPED problems, we can also adjust the ratio of generated electricity and heat in each time slot to optimize the energy cost. For example, we can use more geothermal energy to generate electricity and less to boil water if the electricity price is high, and vice versa. In this section, we discuss how to adjust the CHP generation ratio to minimize the average

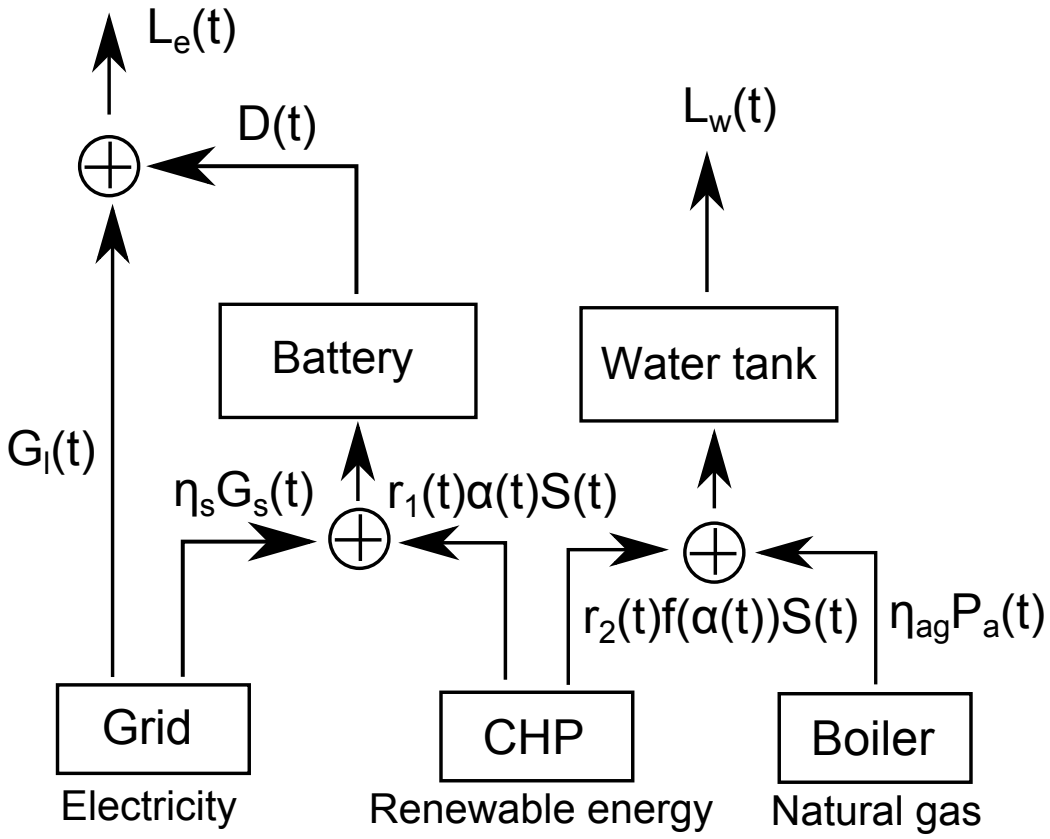


Figure 5.2: CHP using renewable energy

energy cost without the knowledge of the distribution of renewable energy generation, electricity and heat demands by using the Lyapunov optimization technique.

The system architecture is shown in Fig. 5.2. The electricity generation ratio from renewable energy is $\alpha(t)$, which ranges from α_{min} to $\alpha_{\max}$, and the hot water generation ratio is a function of $\alpha(t)$, i.e., $f(\alpha(t))$, where f is a decreasing function. $r_1(t)$ and $r_2(t)$, ranging from 0 to 1, are used to avoid battery and water tank overflow. Since the renewable energy in a household is typically very limited, we assume $S_{\max}\alpha_{\max} \leq C_{char}$, where $S_{\max}$ is the maximum renewable energy available in one time slot. We will discuss the situation without this constraint later. We use a similar battery and water tank model as those in Section 5.3, and the optimization problem can be formulated as follows:

Problem Four (P-IV)

$$\min_{D(t),r(t),G_l(t),G_s(t),\alpha(t),P_a(t)}: \quad P_4 = \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\{C_e(t)[G_s(t)+G_l(t)]+C_gP_a(t)\} \tag{5.47}$$

subject to

$$B(t+1) = B(t) + \eta_s G_s(t) + r_1(t)S(t)\alpha(t) - D(t), \tag{5.48}$$

$$W(t+1) = W(t) + r_2(t)S(t)f(\alpha(t)) + \eta_{ag}P_a(t) - L_w(t), \tag{5.49}$$

$$0 \leq B(t) \leq B_{\max}, \tag{5.50}$$

$$0 \leq W(t) \leq W_{\max}, \tag{5.51}$$

$$L_e(t) = G_l(t) + D(t), \tag{5.52}$$

$$0 \leq \eta_s G_s(t) + r_1(t)S(t)\alpha(t) \leq C_{char}, \tag{5.53}$$

$$S_{\max}\alpha_{\max} \leq C_{char}, \tag{5.54}$$

$$0 \leq G_l(t) + G_s(t) \leq P_{e,\max}, \tag{5.55}$$

$$0 \leq r_1(t), r_2(t) \leq 1, G_l(t), G_s(t) \geq 0, \tag{5.56}$$

$$0 \leq D(t) \leq D_{\max}, \tag{5.57}$$

$$\alpha_{min} \leq \alpha \leq \alpha_{\max}. \tag{5.58}$$

After we relax P-IV, the drift plus penalty function after manipulation can be

written as follows:

$$\begin{aligned}
&\Delta(t) + V\mathbb{E}\{C_e(t)[G_s(t) + G_l(t)] + C_g P_a(t)\} \\
&\leq B' + V\mathbb{E}\{C_e(t)L_e(t)|E(t)\} - \mathbb{E}\{L_w(t)X(t)|X(t)\} \\
&\quad - \mathbb{E}\{D(t)[E(t) + VC_e(t)]|E(t)\} \\
&\quad + \mathbb{E}\{G_s(t)[\eta_s E(t) + VC_e(t)]|E(t)\} \\
&\quad + \mathbb{E}\{\alpha(t)[E(t)r_1(t)S(t)]|E(t)\} \\
&\quad + \mathbb{E}\{f(\alpha(t))[X(t)r_2(t)S(t)|X(t)]\} \\
&\quad + \mathbb{E}\{P_a(t)[\eta_{ag}X(t) + VC_g]|X(t)\},
\end{aligned} \tag{5.59}$$

where B' is a constant and defined as

$$\begin{aligned}
B' =& \frac{1}{2}\max[(\eta_s G_{s,\max} + \alpha_{\max} S_{\max})^2, D_{\max}^2] \\
&+ \frac{1}{2}\max[(\eta_{cg} P_{c,\max} + f(\alpha_{min}) S_{\max})^2, L_{w,\max}^2].
\end{aligned} \tag{5.60}$$

Our algorithm is to minimize the RHS of (5.59), i.e., to solve P-V.

Problem V (P-V)

$$\begin{aligned}
\min_{D(t),r(t),G_l(t),G_s(t),\alpha(t),P_a(t)} \quad & G_s(t)H_s(t) + \alpha(t)H_h(r_1(t)) + f(\alpha(t))H_f(r_2(t)) \\
& + P_a(t)H_a(t) - D(t)H_d(t),
\end{aligned} \tag{5.61}$$

subject to (5.52) (5.53) (5.54) (5.55) (5.56) (5.57) (5.58),
where $H_h(r_1(t)) = E(t)r_1(t)S(t)$ and $H_f(r_2(t)) = X(t)r_2(t)S(t)$.

Notice that P-V is also a non-convex optimization problem. However, $r_1(t)$ and $r_2(t)$ must be 0 or 1. To prove it, let's consider the following situations:

1) Suppose $E(t) \geq 0$, then both $G_s(t)$ and $r_1(t)$ should be 0 to minimize P-V and avoid battery overflow.

2) Suppose $E(t) < 0$ but $H_s(t) \geq 0$. We must have $G_s(t) = 0$ and $r_1(t) = 1$ to minimize P-V, and constraint (5.53) will not be violated due to (5.54).

3) Suppose $H_s(t) < 0$, we have $E(t) < 0$. Then constraint (5.53) is active or $r_1(t) = 1$, otherwise we can increase $r_1(t)$ to further minimize P-V. Substituting

$G_s(t)$ with $(C_{char} - r_1(t)S(t)\alpha(t))/\eta_s$, and expanding $H_s(t)$ in (5.61), we have:

$$\begin{aligned} \min: \quad & -VC_e(t)/\eta_s S(t)\alpha(t)r_1(t) \\ & +f(\alpha(t))H_f(r_2(t))P_a(t)H_a(t) - D(t)H_d(t), \end{aligned} \tag{5.62}$$

Since (5.54) guarantees that (5.53) will not be violated when $r_1(t) = 1$, we have $r_1(t) = 1$ to minimize (5.62).

4) Suppose $X(t) \geq 0$. Since $S(t) \geq 0$ and $0 \leq r_2(t) \leq 1$, we have $H_f(t) \geq 0$. Because $f(\alpha(t)) \geq 0$, $r_2(t)$ must be 0 to minimize P-V.

5) Suppose $X(t) < 0$. Since $S(t) \geq 0$ and $0 \leq r_1(t) \leq 1$, we have $H_f(t) < 0$. Because $f(\alpha(t)) \geq 0$, $r_2(t)$ must be 1 to minimize P-V.

By determining the values of $r_1(t)$ and $r_2(t)$ based on $E(t)$ and $X(t)$, we can convert P-V to a linear optimization problem with five variables which can be easily solved in real time. Note that constraint (5.54) is critical to make this problem convex. Without this constraint, we have to search $r_1(t)$ in the range of $[C_{char}/S(t)/\alpha_{\max}, 1]$ to obtain the optimal solution.

Theorem 3: Suppose θ and ε are defined in (5.39) and (5.40), respectively. Then through minimizing P-V, we can have the following results:

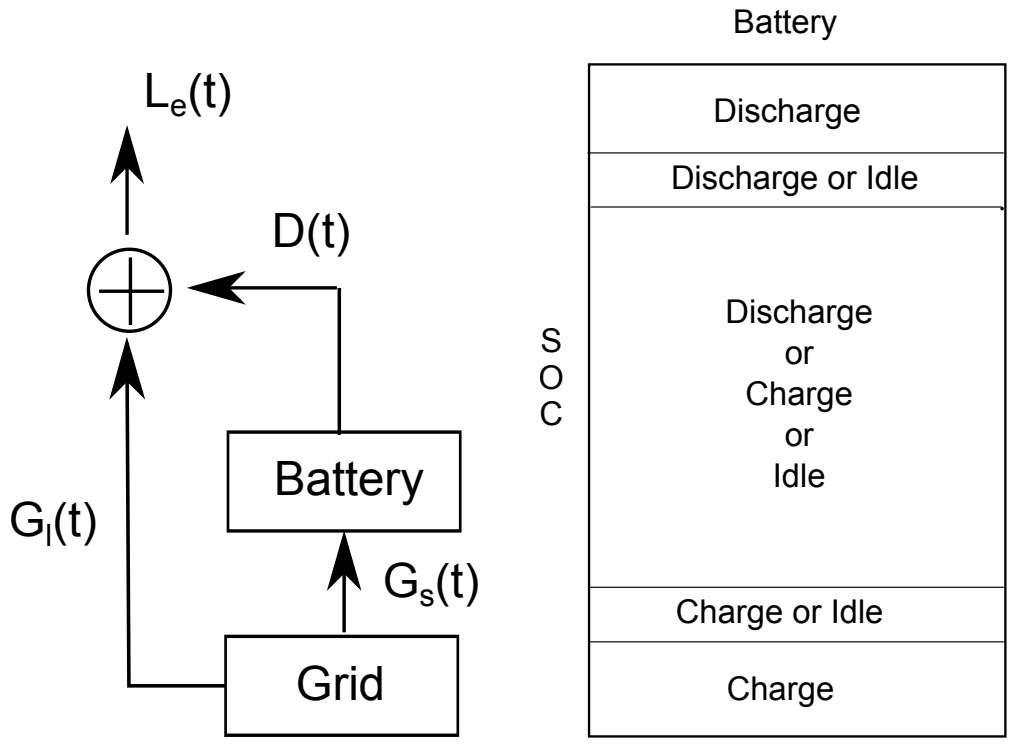
$$0 \leq B(t) \leq \theta + C_{char}, \quad \forall t \in T, \tag{5.63}$$

$$0 \leq W(t) \leq \varepsilon + \eta_{ag} P_{a,\max} + S_{\max} f(\alpha_{min}), \quad \forall t \in T. \tag{5.64}$$

given that the above relationships are satisfied at $t = 0$, and

$$\frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}\{f^{''}(t)\} \leq P_4^* + B'/V, \tag{5.65}$$

where $f^{''}(t)$ represents the energy cost in one time slot using the proposed algorithm to minimize P-V, and P_4^* is the optimal solution to P-IV. The proof of Theorem 3 is similar to that of Theorems 1 and 2 and thus is omitted due to space limitation.



(a) System model (b) Battery model

Figure 5.3: Simplified Battery System

5.6 Determine the Real-time Price

From Section 5.4, we notice that the CHP systems make charging or discharging decisions based on the current battery SOC and water tank level, the power and hot water demand in the current time slot, and the current real-time electricity price. Therefore, by setting appropriate real-time price, the power company can control the charging and discharging process of all the CHP systems and make them provide load shaping service. In this section, we aim to design an algorithm to determine this real-time price to coordinate all the CHP systems. The following assumptions are made. First, the control center in the power company can obtain the status of all the CHP systems, including the current battery SOC and water tank level, the power and hot water demand in the next time slot through two-way communications. Second, there is a sufficient number of CHP systems to provide the load shaping service.

To get some insights into the CHP system reaction model w.r.t. different electricity price and system status, we first consider a simplified system as shown in Fig. 5.3(a).

By using a similar approach as that in Section 5.4, the charging or discharging process of the simplified system is determined by solving the following optimization problem.

$$\min_{D(t),G_s(t)} G_s(t)[\eta_s E(t) + VC_e(t)] - D(t)[E(t) + VC_e(t)], \tag{5.66}$$

subject to:

$$0 \leq B(t) \leq \theta + C_{char}, \tag{5.67}$$

where

$$\theta = \frac{VC_{e,\max}}{\eta_s} + \min\{D_{\max}, L_{e,\max}\}, \tag{5.68}$$

$$E(t) = B(t) - \theta, \quad 0 < \eta_s \leq 1. \tag{5.69}$$

To obtain the behavior of this simplified system, we consider the following cases.

1) $B(t) \geq \theta$. To minimize (5.66), we have $G_s(t) = 0$, and $D(t) = D_{\max}$. In other words, the battery will always discharge.

2) $B(t) \leq \min\{D_{\max}, L_{e,\max}\}$. To minimize (5.66), we have $G_s(t) = C_{char}$, and $D(t) = 0$. In other words, the battery will always charge.

3) $\theta - VC_e(t)/\eta_s \leq B(t) \leq \theta - VC_e(t)$. To minimize (5.66), we have $G_s(t) = 0$, and $D(t) = 0$. In other words, the battery will remain idle. i.e., neither charge nor discharge.

4) $\theta - VC_e(t) \leq B(t) < \theta$. To minimize (5.66), we have $G_s(t) = 0$. In other words, the battery will discharge or remain idle.

5) $\min\{D_{\max}, L_{e,\max}\} < B(t) \leq \theta - VC_{e,\min}(t)/\eta_s$. To minimize (5.66), we have $D(t) = 0$. In other words, the battery will charge or remain idle.

Notice that $G_s(t)$ and $D(t)$ cannot be positive at the same time. We can prove it by contradiction. If $G_s(t)$ and $D(t)$ are both greater than 0, from (5.66) we have $\eta_s E(t) + VC_e(t) < 0$ and $E(t) + VC_e(t) > 0$. Therefore, it must be $\eta_s E(t) < E(t) < 0$. Since $0 < \eta_s \leq 1$, we know it is impossible.

Fig. 5.3(b) illustrates the physical meaning of the above cases. The whole battery can be divided into five regions. In the upper two regions, the battery will not charge, while in the lower two regions, the battery will not discharge. If the SOC of the battery stays in the middle region, we can control the charging or discharging

process of the simplified system by setting an appropriate electricity price between $C_{e,\min}$ and $C_{e,\max}$.

We use superscript i to represent the i-th user. In each time slot, the amount of electricity bought from the grid for user i is $G_l^i(t)+G_s^i(t) = L_e^i(t)-D^i(t)+G_s^i(t)$. Then, the total electricity load in that time slot is $\sum L_e^i(t) - \sum D^i(t) + \sum G_s^i(t)$. Assume that the target load of the load shaping service is $L_t(t)$, at the beginning of each time slot, the control center obtains the total electricity demand of all users $\sum L_e^i(t)$ through two-way communications. The amount of load shaping service should be $L_t(t) - \sum L_e^i(t)$. Our objective is to set an appropriate real-time price $C_e(t)$ so that the amount of demand response from the simplified battery system in time slot t $(\sum G_s^i(t) - \sum D^i(t))$ should be as close to the amount of load shaping service needed $(L_t(t) - \sum L_e^i(t))$ as possible. Therefore, we can formulate the following optimization problem:

$$\min_{D^i(t),G_s^i(t)} |\sum L_e^i(t) - \sum D^i(t) + \sum G_s^i(t) - L_t(t)|, \tag{5.70}$$

where $L_e^i(t)$, $L_t(t)$ are already known, and $D^i(t)$ and $G_s^i(t)$ are determined by solving the optimization problem (5.66).

This is a nested optimization problem. Since the objective function of the CHP system is non-convex, we cannot use the existing bilevel programming to solve it. However, we can utilize some unique features of this problem to obtain the solution with a time complexity of $O(\log(n))$, where n is the number of real-time prices $C_e(t)$ can be chosen from.

Notice that $G_s(t)$ is a non-increasing function of $C_e(t)$, and $D(t)$ is a non-decreasing function of $C_e(t)$, thus $-\sum D^i(t) + \sum G_s^i(t)$ in (5.70) is a non-increasing function of $C_e(t)$. To prove it, we first look at (5.66). If $C_e(t)$ increases, since $V > 0$, both $\eta_s E(t) + VC_e(t)$ and $E(t) + VC_e(t)$ will increase. Therefore, to minimize the objective function, $G_s(t)$ will remain the same when $\eta_s E(t) + VC_e(t) \leq 0$ or $G_s(t) = C_{char}$ when $\eta_s E(t) + VC_e(t) > 0$, $D(t)$ will remain the same when $E(t) + VC_e(t) \leq 0$, and $D(t) = 0$ when $E(t) + VC_e(t) > 0$.

Since $\sum L_e^i(t) - L_t(t)$ in (5.70) is already known at the beginning of each time slot, $\sum L_e^i(t) - \sum D^i(t) + \sum G_s^i(t) - L_t(t)$ in (5.70) is a non-increasing function of $C_e(t)$ in each time slot. Therefore, we can find the appropriate real-time price by using a binary search. Let $g(C_e(t)) = \sum L_e^i(t) - \sum D^i(t) + \sum G_s^i(t) - L_t(t)$ when the current electricity price is $C_e(t)$, the searching algorithm is described in Algorithm 3.

The parameter τ represents the minimum price resolution in our search. Lines

Algorithm 3 Searching the electricity price

Require: $C_{e,\min}$, $C_{e,\max}$

$C_{e,high} = C_{e,\max}$
$C_{e,low} = C_{e,\min}$
$C_{e,mid} = (C_{e,high} + C_{e,low})/2$
while $C_{e,mid} - C_{e,min} > \tau$ OR $C_{e,max} - C_{e,mid} > \tau$ **do**
  if $g(C_{e,min}) = g(C_{e,max})$ **then**
    return $C_{e,low}$ and $C_{e,high}$
  end if
  if $g(C_{e,mid}) < 0$ **then**
    $C_{e,high} = C_{e,mid}$
  else
    $C_{e,low} = C_{e,mid}$
  end if
  $C_{e,mid} = (C_{e,high} + C_{e,low})/2$
end while
return $\min\{g(C_{e,max}), g(C_{e,mid}), g(C_{e,min})\}$ and the corresponding C_e

5 to 7 mean that we return the price range if $g(C_e(t))$ does not change within that range.

There are three points we need to notice. First, since the amount of charging or discharging of each simplified system is discrete, we cannot guarantee that the total load will be exactly the same as the target load. Second, the electrical price $C_e(t)$ obtained from Algorithm 3 may not be the only solution, because the value of $\sum L_e^i(t) - \sum D^i(t) + \sum G_s^i(t) - L_t(t)$ in (5.70) may remain the same even with different electricity price. For example, assuming there is only one simplified system, and we need it to discharge. Any electricity price which can make $E(t) + VC_e(t) > 0$ will reach this goal. Therefore, Algorithm 3 will also return the price range if found. This may lead to another problem: the utility company may always choose the highest price to maximize its profit. However, how to design an appropriate mechanism to effectively restrict the behavior of the utility company is an interesting problem left for future research. Third, there are some tricks to simplify the calculation of $g(C_e(t))$. From the previous discussion of the behavior of the simplified system, the battery will remain idle if $\theta - VC_e(t)/\eta_s \leq B(t) \leq \theta - VC_e(t)$. In other words, given $B(t)$, the battery will remain idle if $C_e(t)$ changes between a certain range $[C_{e,l}(t), C_{e,h}(t)]$. If $C_e(t) < C_{e,l}(t)$, the battery will charge by C_{char}. On the other hand, if $C_e(t) > C_{e,h}(t)$, the battery will discharge by $D_{\max}$. Therefore, with the value of $C_{e,l}(t)$ and $C_{e,h}(t)$ calculated beforehand, we can obtain the value of $-D^i(t) + G_s^i(t)$ in $O(1)$ time.

With respect to the CHP system, we have a more complicated model with a dependent hot water flow queue. However, the fundamental idea to find the real-time price to coordinate all the CHP systems is similar. The higher the price is, the more likely the CHP system will sell the electricity to the grid and use the electricity stored in the battery. Let $V(C_e(t)) = \sum[G_s^i(t) + G_l^i(t) - (1 - r^i(t))\eta_{co}P_c^i(t)] - L_t(t)$ represent the difference between the total electricity demand and the target load. Then $V(C_e(t))$ is a non-increasing function of the real-time price $C_e(t)$. The proof is straightforward and is omitted due to the space limitation. Therefore, we can use a similar binary search algorithm to find the appropriate real-time price. The algorithm is exactly the same as Algorithm 3 by replacing $g(C_e(t))$ with $V(C_e(t))$.

Notice that $L_e(t), L_w(t), G_l(t)$ are not needed to solve P-III. Therefore, instead of solving the optimization problem for each CHP system in every time slot, we can build a look up table with only three input parameters $C_e(t), E(t), X(t)$ and five outputs $G_s(t), P_c(t), P_a(t), r(t), D(t)$ for quick search, assuming homogeneous CHP systems.

5.7 Performance Evaluation

We evaluate the performance of the proposed algorithms using real data. We consider a small hotel with 20 rooms equipped with a battery pack and a water tank. Their capacity are calculated based on the results in Theorem 1. In addition, we assume $\eta_{co}\eta_s >= \eta_{ce}$ and $S_{\max}\alpha_{\max} \leq C_{char}$, so both P-III and P-IV can be solved by convex optimization. If these conditions are not met, we can simply search within a small region to obtain the optimal result.

5.7.1 Simulation Setup

The real-time electricity price data we used in this simulation are obtained from [58]. The electricity sell-back price is assumed to be the same as the purchase price. The natural gas price is assumed to be a constant, $\$5.5/MMBtu$, which is obtained from [27]. Meanwhile we used the wind-power generation data from [105] which has a time resolution of 15 minutes as the renewable energy source for the CHP system. The wind power is scaled down so that the maximum wind power generation equals $12kWh$ per hour. The time slot duration is set to be 15 minutes too. We assume that the original water temperature is 20 Celsius, so it need $111.11Btu$ to heat one liter of water to 70 Celsius. The charging efficiency η_s is set to 0.95, and the boiler efficiency η_{ag} is set to 0.8. The default efficiency of the CHP is assumed to be 75% with 30% to generate electricity and 45% to generate heat. Since $1kWh = 3.41kBtu$, we can have $\eta_{co} = 0.088kWh/kBtu$ and $\eta_{ce} = 0.0836kWh/kBtu$. For the CHP with a variable generation ratio, we assume the total CHP efficiency is still 75%, and the efficiency to generate electricity ranges from 20% to 40%. Since the average power consumption is about $0.8kWh$ per hour per user [12], we assume the electricity demand in one hour is uniformly distributed between 0 and $L_{e,\max} = 32kWh$, while the hot water demand is also assumed to have a uniform distribution between 0 and $L_{w,\max} = 200L/h$. We fix the other parameters as follows (per hour): $D_{\max} = 30kWh$, $C_{char} = 20kWh$, $G_{l,\max} = G_{s,\max} = 32kWh$, $P_{c,\max} = 0.05MMBtu$, and $P_{a,\max} = 0.01MMBtu$.

5.7.2 Benchmark Algorithm

We compare the performance of our algorithms with the situation without energy buffer, which is similar to the CHPED problem [72]. However, since the electricity and gas price model in the CHPED problem is different from ours, we did some changes to

make them comparable. In each time slot, the controller chooses the control actions by solving the following optimization problem.

Benchmark Algorithm I (B-I)

$$\min_{G_l(t),P_c(t),P_a(t)} \quad C_e(t)G_l(t) + C_g[P_c(t) + P_a(t)], \tag{5.71}$$

subject to

$$G_l(t) + \eta_{ce}P_c(t) \geq L_e(t), \tag{5.72}$$

$$\eta_{cg}P_c(t) + \eta_{ag}P_a(t) \geq L_w(t), \tag{5.73}$$

$$G_l(t), P_c(t), P_a(t) \geq 0. \tag{5.74}$$

The objective of the optimization problem is to minimize the total cost in each time slot. Constraints (5.72) and (5.73) mean the electricity and heat generation in each time slot should be no less than the electricity and heat demand.

For the CHP using renewable energy, we have another corresponding benchmark algorithm.

Benchmark Algorithm II (B-II)

$$\min_{G_l(t),P_a(t),\alpha(t)} \quad C_e(t)G_l(t) + C_gP_a(t), \tag{5.75}$$

subject to

$$G_l(t) + \alpha(t)S(t) \geq L_e(t), \tag{5.76}$$

$$f(\alpha(t))S(t) + \eta_{ag}P_a(t) \geq L_w(t), \tag{5.77}$$

$$\alpha_{min} \leq \alpha \leq \alpha_{\max}, \tag{5.78}$$

$$G_l(t), P_a(t) \geq 0. \tag{5.79}$$

The objective and constraints of B-II are similar as those of B-I, with the difference that we have different control actions and include renewable energy into the problem formulation.

5.7.3 Cost Saving using CHP

Fig. 5.4 shows the average cost in one time slot for different parameter V. Due to the inherent exponential convergence property [79], the average cost decreases

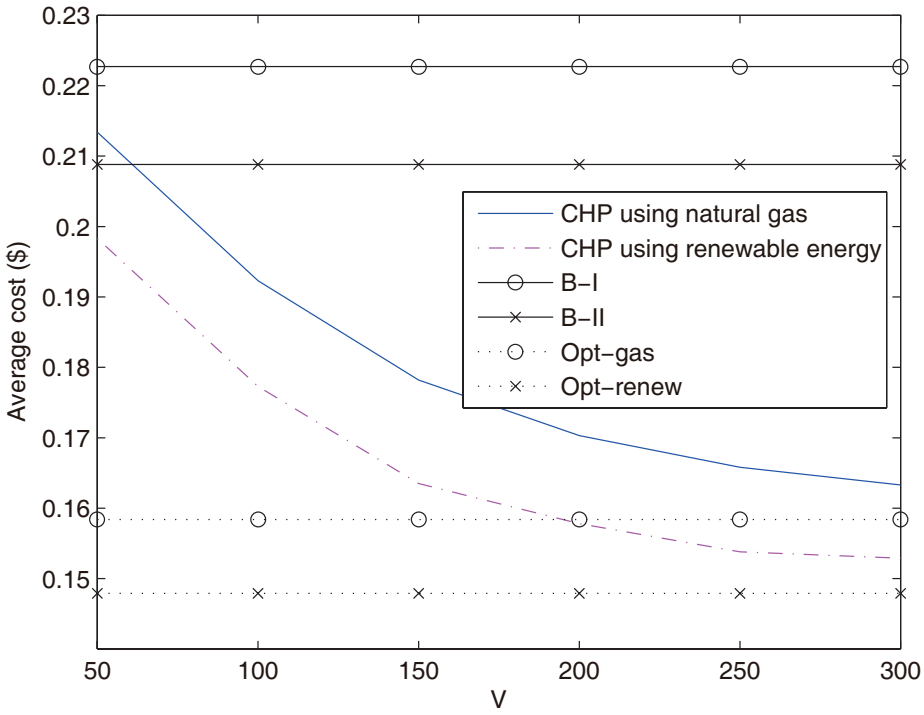


Figure 5.4: Average cost with different V.

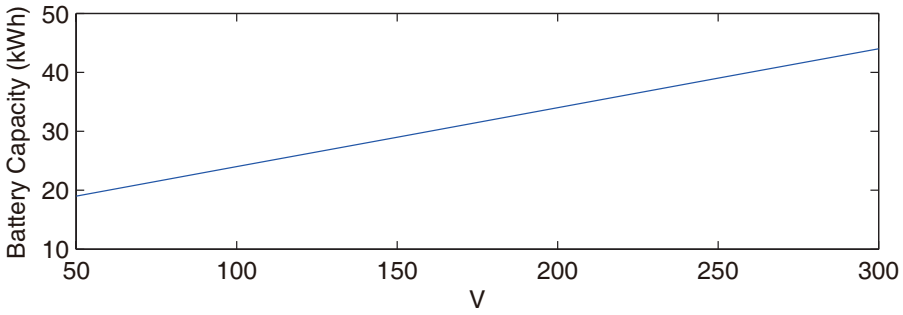


Figure 5.5: The relationship between battery capacity and V

exponentially. The average cost of Benchmark Algorithm I and II in one time slot are \$0.2227 and \$0.2088, respectively. We can see the CHP system using natural gas can save up to 26.54% while the CHP system using renewable energy can save up to 28.63%. In other words, with the help of energy buffers, the saving of using either CHP system can reach approximately $\$2,000$ annually. The saving mainly dues to the following reasons. First, with an energy buffer, the controller can use the electricity stored in the battery and make the CHP generate more electricity or even sell to the grid to make a profit when the electricity price is high and purchase electricity to charge the battery when the electricity price is low. Second, if the renewable energy source can generate more electricity and heat than required, we can store them in the energy buffer for future use.

In Fig. 5.4, the optimal average cost in one time slot for the CHP using natural gas (Opt-gas) and renewable energy (Opt-renew) are about \$0.1584 and \$0.1479, respectively. To obtain these results, we assume that the battery pack and water tank capacities are infinite and we can charge/discharge as much energy as possible to/from the battery when the electricity price is low/high. However, in practice, due to constraints (5.3) and (5.4), these values may not be achievable and are just provided here to show the possible bounds.

Fig. 5.5 shows the relationship between the required battery capacity pack and the parameter V. As was discussed in Theorem 1, the required battery capacity increases linearly with the increase of V. The capacity of the water tank has a similar relationship and is omitted due to the space limitation. With $V = 200$, the battery pack and water tank capacity are $34kWh$ and $419L$, respectively. These values are quite reasonable for a small hotel with about 20 rooms.

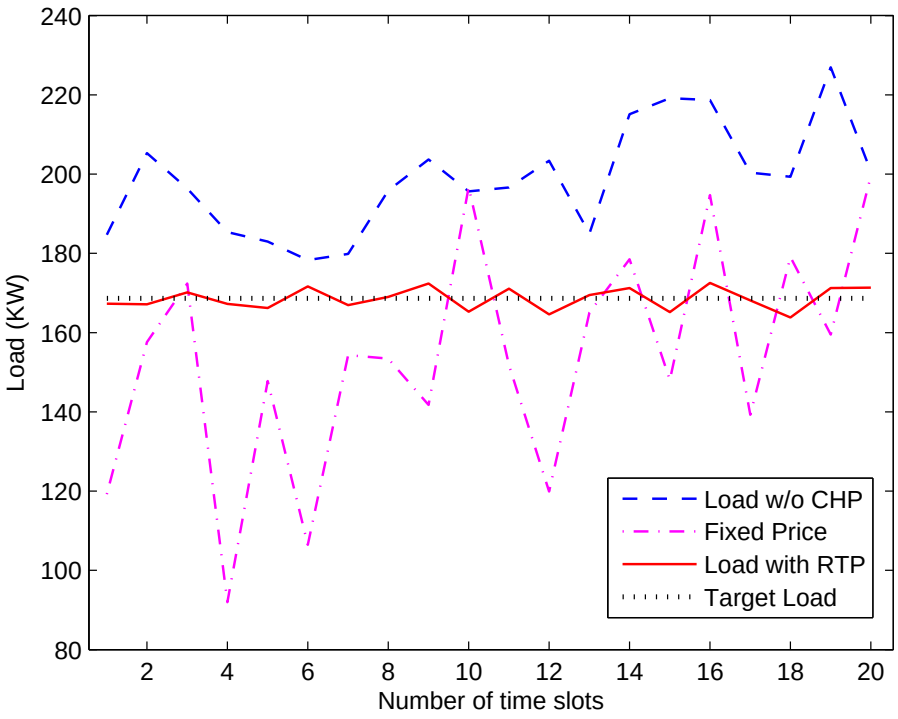


Figure 5.6: Load with different price

5.7.4 Load Shaping by Setting RTP

With battery packs and water tanks, CHP systems can be considered as energy buffers to provide load shaping services. In this simulation, we consider a scenario with 50 homogeneous CHP systems aiming to provide a five-hour load shaping service, consisting 20 time slots with a 15-minute slot duration. The electricity and hot water demands in each time slot for each CHP system are uniformly distributed between 0 and its corresponding maximum value. We set the target load as a constant value which is $30KW$ below the average electricity demand. The real-time electricity price used to coordinate all the CHP systems is obtained using the algorithm described in Section 5.6. The load using CHP systems but with a constant electricity price, which is set to the middle of the maximum and minimum electricity price, is also simulated.

Fig. 5.6 shows the performance of the load shaping service. Since the CHP systems will generate electricity using natural gas, the average load using a fixed price is lower than the average load without CHP systems. Although CHP systems can help users

Table 5.1: The relationship of d_t and sd_t with PLR

PLR	0%	2%	4%	6%	8%	10%
average d_t	3.6725	4.0220	5.0241	6.0691	6.3459	7.1039
max d_t	6.7312	8.5417	15.7296	17.4501	24.3477	37.6074
average sd_t	4.7918	5.2102	6.1082	6.7285	7.3751	8.4507

save money, the fluctuation of the total load with CHP systems and a fixed electricity price is higher than that without using CHP systems. This will add extra cost to the power company as more frequency regulation or spinning reservation services may be needed. By setting the real-time price according to the proposed algorithm, we can reduce the fluctuation of the total load with CHP systems and make the total load close to the target load. Therefore, the combination of CHP systems and the RTP searching algorithm can help both users and the utility company reduce their cost.

5.7.5 Influence of Communication Packet Loss

Each CHP system needs the RTP information at the beginning of every time slot to determine its operation. However, the communication links between the power company and the users may not be reliable, and can lead to packet losses. In this section, we are going to explore the relationship between the performance of the proposed algorithm and the packet loss rate (PLR). We assume that if the user does not receive the RTP information for any time slot, it will use the RTP of the last time slot to determine the operation of the CHP system. In order to measure the influence of the communication packet losses on the performance of our algorithm, we define the average deviation and standard deviation of the total load to the target load as d_t and sd_t respectively.

$$d_t = \frac{1}{T}\sum_{t=0}^{T} |L_a(t) - L_t(t)|, \tag{5.80}$$

$$sd_t = \sqrt{\frac{1}{T}\sum_{t=0}^{T} (L_a(t) - L_t(t))^2}, \tag{5.81}$$

where T is the total number of time slots, and $L_a(t)$ and $L_t(t)$ represent the total electricity load and target load at time slot t, respectively.

Table. 5.1 shows the average of 50 runs' deviation, standard deviation and maximum deviation of the total load to the target load. As we can see, all the three cri-

terions increase with the increase of the PLR. Therefore, the power company should keep the PLR as low as possible.

5.8 Conclusions

In this chapter, motivated by the queueing analysis and buffer management solutions in data communication systems, we have proposed an approach to minimize the average energy cost for two different types of CHP systems. Our system model includes renewable energy, real-time price, stochastic energy demand and energy buffers with finite capacity. Since the battery queue and the water tank queue are dependent, we need to solve a non-convex optimization problem to obtain the optimal control actions. By using the Lyapunov optimization techniques, our schemes can achieve a near-optimal performance, which will deviate no more than $O(1/V)$ from the optimal solution.

On the other hand, we discussed how to set an appropriate real-time price to coordinate all the CHP systems to provide load shaping service. Different from the existing works which mainly determine the RTP from the perspective of game theories, the proposed algorithm uses binary search to find the optimal RTP with a time complexity of $O(\log n)$. Extensive simulation shows that the use of CHP systems can reduce the cost of both users and the utility company with the proposed scheduling and pricing algorithms.

Chapter 6

Contributions and Future Work

6.1 Contributions

In this dissertation, we have discussed and analyzed four demand response application scenarios from the perspective of both users and the power company. The following outlines the contributions we have achieved.

- In Chapter 2, we have proposed an online decentralized access algorithm for PHEV charging, which can effectively flatten load peaks during PHEV charging process. The main idea is that when the current electricity load is higher than a certain threshold, we will restrict the access of elastic load, such as the charging of PHEV; when the electricity load is even higher, some elastic load is turned off automatically with a probability so that the total load can decrease to a safe region. The simulation results show that the stochastic renewable energy can be efficiently utilized even without prediction.

- In Chapter 3, we have introduced a decentralized random access algorithm, which can efficiently avoid bus congestion and large voltage drop in the distribution grid with charging PHEVs. The smart agent in each house makes PHEV charging scheduling decisions independently based on the received information of the current distribution grid status from a control center. The proposed algorithm does not require any prediction information about users' future actions. We have also analyzed the performance of our algorithm and derive the system capacity.

- In Chapter 4, we have proposed a centralized algorithm to control heteroge-

neous HVACs in a micro-grid. The objective of this algorithm is to reduce non-renewable energy generation fluctuations while still guarantee user comfort level. We then extend the centralized algorithm to a distributed one, which has a much lower computational complexity and is more scalable. We have also extended the proposed algorithms to support HVAC ON/OFF control modes other than adjusting the HVAC power level. Finally, since the elastic load potential provided by HVACs is limited compared to the unlimited control time, a dynamic water level adjustment algorithm is proposed to reserve this elastic load potential for future demand response.

- In Chapter 5, we first proposed a comprehensive economic model using CHP systems from the perspective of a commercial customer, which incorporates both the electricity and thermal energy queues. We investigate the relationship of these two queues to minimize the average cost. Second, we proposed an algorithm to approximately achieve the optimal average cost, considering the limited capacities of the battery pack and the water tank. The algorithm does not require any statistical information of the system dynamics such as electricity and hot water demands, etc. To obtain the optimal scheduling decision, we discussed when we can use the specific features of the problem to turn a non-convex optimization problem into a convex one which can be solved in real time. Finally, we discussed how to set the appropriate RTP to coordinate all the CHP systems indirectly to provide load shaping services for the power company. The time complexity of the proposed searching algorithm is $O(\log n)$.

6.2 Future Work

Some of the further research issues are listed as follows.

- In Chapter 2, we did not consider the communication delay and packet loss. On one hand, if the PHEVs do not receive the broadcasted information for that time slot, they have to keep their current charging status. On the other hand, if the received information is always delayed by one or more time slots, the behavior of the PHEVs will also be delayed. As a result, the total electricity load may not be the same as expected. To guarantee the effectiveness of the proposed algorithm, we need to investigate the relationship between the percentage of packet loss and communication delay and the performance of the decentralized

access algorithm. To achieve this goal, we need to include the estimated packet loss rate and percentage of communication delay into our analytical model and assume these PHEVs operate differently as we expected to guarantee the worst case performance. This is an important issue needs further investigation before the proposed algorithm can be used in practice.

- In Chapter 3, the locations of customers will lead to different levels of congestions to the network and thus affect the access probabilities of their PHEVs. This may lead to a fairness problem and requires further study together with other pricing and economic policies. In addition, the distribution grid in this chapter has a radial topology, for a meshed distribution grid, things are much more complicated. However, the main idea of the proposed algorithm is still useful to control the PHEV access if the critical bus is congested. These are important further research issues. In practice, several other problems in the distribution grid should be considered, such as medium voltage to low voltage (MV/LV) transformer overload. Finally, how to make a tradeoff between the complexity of the control algorithm and the distribution grid efficiency is left for future research.

- In Chapter 4, there are several open issues left behind. First, how to extend the current MPC algorithms to control other types of elastic load requires further investigation. The key difference is that the initial states of HVACs are already known, while for other types of elastic load, such as PHEV, the arrival time and departure time in the future may not be available. Second, the water level change rate μ in Algorithm 1 is determined empirically and is a constant in this chapter. If we can adjust μ w.r.t. history statistic information, such as peak time etc., we may achieve an even better performance. Third, the computational complexity will increase with more control variables. Therefore, how to make a tradeoff between the number of groups, the size of each group, load fluctuation, and the influence to user comfort level is also an important problem left to future research.

- In Chapter 5, although the proposed searching algorithm can help to find the optimal RTP, there may be multiple RTPs which can achieve the same effect. Therefore, it is possible that the utility company will always choose the maximum RTP to maximize its profit. How to effectively regulate the actions of

the utility company is a problem left for future research. Second, since the target load will affect the RTP for all the CHP systems, and thus affect the profit of the utility company, how to select an appropriate target load is also an interesting problem needs further investigation. Finally, in this chapter, we assume that there are enough CHP systems to provide load shaping services. However, in the real world, we cannot make a finite number of CHP systems provide infinite load shaping services. Once all the batteries of CHP systems are low or high, they can no longer provide demand response. Therefore, we need to monitor the status of all the CHP systems and change the target load accordingly. This is also an important issue left for future research.

Appendix A

Notations

For easy reference, the symbols used in this thesis are summarized in Table A.1.

Table A.1: Notations

Symbol	Description
$B(t)$	the SOC level of the battery
$B_{\max}$	the capacity of the battery
c	the change to the water level
C	the effective heat capacity in a house
Cap	the energy buffer capacity of the local controller
$C_e(t)$	electricity price at time t
C_{char}	the maximum charge rate of the battery
$climit$	the maximum allowed water level
	change in each time slot
$D(t)$	the amount of electricity discharged from the battery
C_g	the natural gas price
$D_{\max}$	the maximum discharge rate of the battery
G	the thermal insulation level of a house
$G_l(t)$	the amount of electricity bought to supply the user demand
$G_s(t)$	the amount of electricity bought to charge the battery
I_{ki}	the loading rate brought by PHEV k on bus i
I^i_{base}	the loading rate brought by base load on bus i
k	the index of time slot
L_b	the non-HVAC load

Table A.1: (continued)

Symbol	Description
L_h	the total load of HVAC
$L_{base}(t)$	the base load without PHEV
$L_{PHEV}(t)$	the load of charging PHEVs at time t
$L_e(t)$	the electricity demands from users
$L_t(t)$	the target load of load shaping service
$L_w(t)$	the hot water demands from users
$L_{w,\max}$	the upper bound of the hot water demand
m	number of HVACs under a local controller
M	the number of local controller or
	the number of buses in the distribution grid
N	the prediction horizon in terms of slot
N_u	total number of users
N_s	total number of slots
P_a	the amount of power assigned from
	the central controller to the local group
$P_a(t)$	the amount of the natural gas consumed by the boiler
P_b	the amount of power of a local group
	leak to the outside environment
$P_c(t)$	the amount of natural gas consumed by the CHP
$P_{a,\max}$	the maximum amount of the natural gas used by the boiler
P_g	the state of the local group
P_h	the power consumed by the HVAC
$PState$	an vector used to store the state of all the HVACs
p_1	charging probability obtained based on γ
p_{1c}	charging probability obtained based on γ_i
p_{1v}	charging probability obtained based on γ_v
p_2	charging suspend probability obtained based on γ
p_{2c}	charging suspend probability obtained based on γ_i
p_{2v}	charging suspend probability obtained based on γ_v
$P_c(t)$	the amount of the natural gas consumed by the CHP
Q	the conversion coefficient from power to temperature
	for a specific room with the unit of k/J

Table A.1: (continued)

Symbol	Description
r_i	the ideal temperature set by the user for HVAC i
R_i	the set-point of a local group
r_w	the reference HVAC load
$r(t)$	a tradeoff between the amount of electricity used to charge the battery and that sold to the grid
S	the set of all the HVACs
S_l	the set of HVACs under the local controller
S_n	the power generation from non-renewable power plants
S_r	the renewable energy generation
$S(t)$	power generation capacity
t_d	the PHEV delay time after the failure of each random access attempt
t_m	the maximum value of t_d,
t	time slot duration
Δt	one-slot Lyapunov drift
T	the number of time slots
$T_i^-(k+1)$	the temperature of house i in the next time slot if the state of the HVAC is OFF
$T_i^+(k+1)$	the temperature of house i in the next time slot if the state of the HVAC is ON
ΔT_l	the maximum allowed temperature decrement from the set point in a house
ΔT_u	the maximum allowed temperature increment from the set point in a house
ΔT_{off}	the decreased temperature for an HVAC during one time slot
T_o	the outdoor temperature
T_i	the indoor temperature of house i
$u_i^{\max}$	maximum amount of power allowed to be assigned to an HVAC in each time slot
U	the manipulated input variable of the MPC model
$U_i^{\max}$	maximum amount of power allowed to be assigned to local controller in each time slot

Table A.1: (continued)

Symbol	Description
$U(t)$	power utilization at time t
V_{ki}	the voltage drop brought by PHEV k on bus i
V_{mi}	the maximum allowed voltage drop of bus i
V_{base}^{i}	the voltage drop brought by base load on bus i
V	the measured input disturbance of the MPC model
W_l	water level
$W(t)$	the water level in the water tank in slot t
$X(k)$	the state of the MPC model at time slot k
$Y(k)$	the output of the MPC model at time slot k
λ	the weight of uers' comfort level
σ	standard deviation of the prediction error
η	the ratio of σ and the maximum remain load
	change during the prediction horizon
ν_1	the value of threshold one for power utilization
ν_{1i}	the value of threshold one for bus congestion
ν_{1v}	the value of threshold one for voltage drop
ν_2	the value of threshold two for power utilization
ν_{2i}	the value of threshold two for bus congestion
ν_{2v}	the value of threshold two for voltage drop
δ_1,δ_2,δ_3,δ_4	parameters represent users' preference
κ_1,κ_2,κ_3,κ_4	global control parameters to adjust charging or suspend probability
α	a parameter denotes the changing rate of p_1
$\alpha(t)$	the electricity generation ratio from renewable energy
ϕ	a parameter denotes the changing rate of p_2
ω	the weight between the battery queue and
	the water tank queue
$\omega(k)$	waited time of PHEV k
$\omega_m(k)$	the maximum tolerable delay of PHEV k
γ	the current power usage ratio
γ_i	the current bus load ratio
γ_v	the current bus voltage drop ratio
$\gamma_i(k)$	the loading rate of bus k

Table A.1: (continued)

Symbol	Description
$\gamma_v(k)$	the voltage drop ratio of bus k
η_{ce}	the conversion efficiency from natural gas to electricity charged to the battery
η_{cg}	the conversion efficiency from natural gas to hot water
η_{co}	the electricity conversion efficiency sold back to the grid
η_{ag}	the conversion efficiency from natural gas to hot water
η_s	the battery charging efficiency

Bibliography

[1] Federal Highway Administration. Summary of Travel Trends: 2009 National Household Travel Survey. http://nhts.ornl.gov/2009/pub/stt.pdf, 2009.

[2] M. Alizadeh, A. Scaglione, and R.J. Thomas. Direct load management of electric vehicles. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*,, pages 5964–5967, May 2011.

[3] I.S. Bayram, G. Michailidis, and M. Devetsikiotis. Electric power resource provisioning for large scale public ev charging facilities. In *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 133–138, Oct 2013.

[4] D. P. Bertsekas. *Dynamic Programming and Optimal Control, 2nd ed.* Athena Scientific, 2000.

[5] T. Borsche, A. Ulbig, M. Koller, and G. Andersson. Power and energy capacity requirements of storages providing frequency control reserves. In *IEEE Power and Energy Society General Meeting*, pages 1–5, 2013.

[6] Y. Cao, S. Tang, C. Li, P. Zhang, Y. Tan, Z. Zhang, and J. Li. An optimized ev charging model considering tou price and soc curve. *IEEE Transactions on Smart Grid*, 3(1):388–393, March 2012.

[7] J. T. Chambers. Fleet electrification roadmap: Revolutionizing transportation and achieving energy security. `http://www.rmi.org/Content/Files/Fleet\%20Electrification\%20Roadmap.pdf`, 2010.

[8] T. Chang, M. Alizadeh, and A. Scaglione. Real-time power balancing via decentralized coordinated home energy scheduling. *IEEE Transactions on Smart Grid*, 4(3):1490–1504, Sept 2013.

[9] C. Chen, S. Kishore, and L.V. Snyder. An innovative rtp-based residential power scheduling scheme for smart grids. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5956–5959, May 2011.

[10] C. Chen, S. Kishore, and L.V. Snyder. An innovative rtp-based residential power scheduling scheme for smart grids. In *IEEE ICASSP*, pages 5956–5959, May 2011.

[11] K. Clement-Nyns, E. Haesen, and J. Driesen. The impact of charging plug-in hybrid electric vehicles on a residential distribution grid. *IEEE Transactions on Power Systems*, 25(1):371–380, Feb 2010.

[12] California Energy Commission. Residential Appliance Saturation Study. http://www.energy.ca.gov/appliances/rass/.

[13] Energy Systems Consulting. RELOAD Database Documentation and Evaluation and Use in NEMS. http://www.onlocationinc.com/LoadShapesReload2001.pdf, 2001.

[14] EnerNex Corporation. Wind Integration Study. http://www.uwig.org/XcelMNDOCStudyReport.pdf, 2004.

[15] M. Dentice, M. Sasso, S. Sibilio, and L. Vanoli. Micro-combined heat and power in residential and light commercial applications. *Applied Thermal Engineering*, 23(10):1247–1259, 2003.

[16] EirGrid. Wind energy generation & System demand data sets. http://www.eirgrid.com/operations/systemperformancedata/.

[17] P. Fairley. Speed bumps ahead for electric-vehicle charging. *IEEE Spectrum*, 47(1):13–14, Jan 2010.

[18] FERC. Annual Electric Balancing Authority and Planning Area Report. http://www.ferc.gov/docs-filing/forms.asp, 2009.

[19] M. D. Galus, R. A. Waraich, and G. Andersson. Predictive, distributed, hierarchical charging control of PHEVs in the distribution system of a large urban area incorporating a multi agent transportation simulation. In *17th Power Systems Computation Conference*, 2011.

[20] Matthias D. Galus, S. Koch, and G. Andersson. Provision of load frequency control by PHEVs, controllable loads, and a cogeneration unit. *IEEE Transactions on Industrial Electronics*, 58(10):4568–4582, 2011.

[21] Y. Gao, Y. Chen, C. Wang, and K.J.R. Liu. A contract-based approach for ancillary services in v2g networks: Optimality and learning. In *IEEE INFOCOM 2013*, pages 1151–1159, April 2013.

[22] G. Giebel. The state-of-the-art in short-term prediction of wind power: A literature overview. http://www.risoe.dk/rispubl/vea/veapdf/ANEMOS_giebel.pdf.

[23] R. Graham et al. Comparing the benefits and impacts of hybrid electric vehicle options. *Electric Power Research Institute (EPRI), Palo Alto, CA, Report*, 1000349, 2001.

[24] Y. Guo, M. Pan, and Y. Fang. Optimal power management of residential customers in the smart grid. *IEEE Transactions on Parallel and Distributed Systems*, 23(9):1593–1606, Sept 2012.

[25] P. Han, J. Wang, Y. Han, and Y. Li. Resident plug-in electric vehicle charging modeling and scheduling mechanism in the smart grid. *Mathematical Problems in Engineering*, 2014, 2014.

[26] AD Hawkes and MA Leach. Cost-effective operating strategy for residential micro-combined heat and power. *Energy*, 32(5):711–723, 2007.

[27] B. Hedman. CHP Policy Analysis and 2011-2030 Market Assessment. Technical report, ICF International, Inc., June, 2012.

[28] T.G. Hovgaard, K. Edlund, and J. Bagterp Jorgensen. The potential of economic mpc for power management. In *49th IEEE Conference on Decision and Control (CDC),*, pages 7533–7538, 2010.

[29] Y. Huang, S. Mao, and R.M. Nelms. Adaptive electricity scheduling in microgrids. *IEEE Transactions on Smart Grid*, 5(1):270–281, Jan 2014.

[30] C. Jin, N. Lu, S. Lu, Y. Makarov, and R.A. Dougal. Coordinated control algorithm for hybrid energy storage systems. In *2011 IEEE Power and Energy Society General Meeting*, pages 1–7, 2011.

[31] G. Karmakar, A. Kabra, and K. Ramamritham. Coordinated scheduling of thermostatically controlled real-time systems under peak power constraint. In *2013 IEEE 19th Real-Time and Embedded Technology and Applications Symposium (RTAS)*, pages 33–42, 2013.

[32] A. Karnama. Analysis of integration of plug-in hybrid electric vehicles in the distribution grid. http://www.diva-portal.org/smash/get/diva2:610066/FULLTEXT01.pdf, 2009.

[33] I. Koutsopoulos, V. Hatzi, and L. Tassiulas. Optimal energy storage control policies for the smart power grid. In *IEEE SmartGridComm 2011*, pages 475–480, Oct 2011.

[34] I. Koutsopoulos and L. Tassiulas. Optimal control policies for power demand scheduling in the smart grid. *IEEE Journal on Selected Areas in Communications*, 30(6):1049–1060, July 2012.

[35] I. Koutsopoulos and L. Tassiulas. Optimal control policies for power demand scheduling in the smart grid. *IEEE Journal on Selected Areas in Communications*, 30(6):1049–1060, 2012.

[36] K.Sikes. Plug-in hybrid electric vehicle market introduction study: Final report. http://info.ornl.gov/sites/publications/files/Pub14078.pdf, 2010.

[37] S. Letendre and R. A.Watts. Effects of plug-in hybrid electric vehicles on the vermont electric transmission system. In *Proc. Transp. Res. Board Annu. Meet.*, Jan 2009.

[38] H. Liang, B. Choi, A. Abdrabou, W. Zhuang, and X. Shen. Decentralized economic dispatch in microgrids via heterogeneous wireless networks. *IEEE Journal on Selected Areas in Communications*, 30(6):1061–1074, July 2012.

[39] H. Liang, B. Choi, W. Zhuang, and X. Shen. Optimizing the energy delivery via v2g systems based on stochastic inventory theory. *IEEE Transactions on Smart Grid*, 4(4):2230–2243, Dec 2013.

[40] H. Liang, B. Choi, W. Zhuang, and X. Shen. Stability enhancement of decentralized inverter control through wireless communications in microgrids. *IEEE Transactions on Smart Grid*, 4(1):321–331, March 2013.

[41] K. Liu, S. Subbarayan, R.R. Shoults, M.T. Manry, C. Kwan, F.L. Lewis, and J. Naccarino. Comparison of very short-term load forecasting techniques. *IEEE Transactions on Power Systems*, 11(2):877–882, 1996.

[42] A. Lokurlu, T. Grube, B. Höhlein, and D. Stolten. Fuel cells for mobile and stationary applications cost analysis for combined heat and power stations on the basis of fuel cells. *International journal of hydrogen energy*, 28(7):703–711, 2003.

[43] N. Lu. An evaluation of the hvac load potential for providing load balancing service. *IEEE Transactions on Smart Grid*, 3(3):1263–1270, 2012.

[44] N. Lu. An evaluation of the hvac load potential for providing load balancing service. *IEEE Transactions on Smart Grid*, 3(3):1263–1270, 2012.

[45] N. Lu and Y. Zhang. Design considerations of a centralized load controller using thermostatically controlled appliances for continuous regulation reserves. *IEEE Transactions on Smart Grid*, 4(2):914–921, 2013.

[46] N. Lu and Y. Zhang. Design considerations of a centralized load controller using thermostatically controlled appliances for continuous regulation reserves. *IEEE Transactions on Smart Grid*, 4(2):914–921, 2013.

[47] P. Scott M. Briel and S. Thiebaux. Randomized Load Control: A Simple Distributed Approach for Scheduling Smart Appliances. http://users.cecs.anu.edu.au/ thiebaux/papers/ijcai13-rlc.pdf.

[48] Z. Ma, D. Callaway, and I. Hiskens. Decentralized charging control for large populations of plug-in electric vehicles. In *2010 49th IEEE Conference on Decision and Control (CDC),*, pages 206–212, Dec 2010.

[49] J. M. Maciejowski. *Predictive Control With Constraints*. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[50] MathWorks. Block Diagram of Model Predictive Control Framework. http://www.mathworks.com/help/mpc/ug/how-plants-are-used-in-model-predictive-control.html.

[51] MathWorks. Interpolation. http://www.mathworks.com/moler/interp.pdf.

[52] A.-H. Mohsenian-Rad and A. Leon-Garcia. Optimal residential load control with price prediction in real-time electricity pricing environments. *IEEE Transactions on Smart Grid*, 1(2):120–133, Sept 2010.

[53] K. Nanaeda, F. Mueller, J. Brouwer, and S. Samuelsen. Dynamic modeling and evaluation of solid oxide fuel cell–combined heat and power system operating strategies. *Journal of Power Sources*, 195(10):3176–3185, 2010.

[54] M. Neely. *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan & Claypool, 2010.

[55] M.J. Neely, A.S. Tehrani, and A.G. Dimakis. Efficient algorithms for renewable energy allocation to delay tolerant consumers. In *IEEE SmartGridComm 2010*, pages 549–554, Oct 2010.

[56] H. K. Nguyen, J.B. Song, and Z. Han. Demand side management to reduce peak-to-average ratio using game theory in smart grid. In *2012 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 91–96, March 2012.

[57] H. T. Nguyen, D. Nguyen, and L. B. Le. Home energy management with generic thermal dynamics and user temperature preference. In *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 552–557, 2013.

[58] The Electric Reliability Council of Texas. Real-time-price data. http://www.ercot.com/.

[59] Bureau of Transportation Statistics U.S. Department of Transportation. National Transportation Statistics. Table 4-6. http://www.bts.gov/publications/national_transportation_statistics/html/table_04_06.html.

[60] Bureau of Transportation Statistics U.S. Department of Transportation. Transportation Statistics Annual Report 2008, Table 1-1-1. http://www.bts.gov/publications/transportation_statistics_annual_report/2008/html/chapter
2008.

[61] D. O'Neill, M. Levorato, A. Goldsmith, and U. Mitra. Residential demand response using reinforcement learning. In *2010 First IEEE International Con-*

ference on Smart Grid Communications (SmartGridComm), pages 409–414, Oct 2010.

[62] C. Ou, H. Liang, and W. Zhuang. Investigating wireless charging and mobility of electric vehicles on electricity market. *IEEE Transactions on Industrial Electronics*, PP(99):1–1, 2014.

[63] A. Oudalov, D. Chartouni, and C. Ohler. Optimizing a battery energy storage system for primary frequency control. *IEEE Transactions on Power Systems*, 22(3):1259–1266, 2007.

[64] S. Parkinson, D. Wang, C. Crawford, and N. Djilali. Comfort-constrained distributed heat pump management. *Energy Procedia*, 12:849 – 855, 2011.

[65] J.A. Peas Lopes, F.J. Soares, and P.M.R. Almeida. Identifying management procedures to deal with connection of electric vehicles in the grid. In *2009 IEEE Bucharest PowerTech*, pages 1–8, June 2009.

[66] Sarvapali D. Ramchurn, P. Vytelingum, A. Rogers, and N. Jennings. Agent-based control for decentralised demand side management in the smart grid. In *The 10th International Conference on Autonomous Agents and Multiagent Systems*, pages 5–12, 2011.

[67] Sarvapali D Ramchurn, P. Vytelingum, A. Rogers, and N. R. Jennings. Putting the'smarts' into the smart grid: a grand challenge for artificial intelligence. *Communications of the ACM*, 55(4):86–97, 2012.

[68] M. Ranjbar, S. Soleymani, N. Sadati, and A. M. Ranjbar. Electricity price forecasting using artificial neural network. In *International Conference on Power Electronics, Drives and Energy Systems, 2006. PEDES '06.*, pages 1–5, Dec 2006.

[69] A. Rautiainen, S. Repo, and P. Jarventausta. Using frequency dependent electric space heating loads to manage frequency disturbances in power systems. In *IEEE Bucharest PowerTech*, pages 1–6, 2009.

[70] R. J. Rei, F.J. Soares, P.M.R. Almeida, and J.A. Peas Lopes. Grid interactive charging control for plug-in electric vehicles. In *2010 13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 386–391, Sept 2010.

[71] P. Richardson, D. Flynn, and A. Keane. Optimal charging of electric vehicles in low-voltage distribution systems. In *2012 IEEE Power and Energy Society General Meeting*, pages 1–1, July 2012.

[72] F.J. Rooijers and R. A M Van Amerongen. Static economic dispatch for co-generation systems. *IEEE Transactions on Power Systems*, 9(3):1392–1398, Aug 1994.

[73] S. Shao, M. Pipattanasomporn, and S. Rahman. Challenges of PHEV penetration to the residential distribution network. In *Power Energy Society General Meeting, 2009. PES '09. IEEE*, pages 1–8, July 2009.

[74] S. Shao, M. Pipattanasomporn, and S. Rahman. Demand response as a load shaping tool in an intelligent grid with electric vehicles. *IEEE Transactions on Smart Grid*, 2(4):624–631, Dec 2011.

[75] S. Shao, M. Pipattanasomporn, and S. Rahman. Grid integration of electric vehicles and demand response with customer choice. *IEEE Transactions on Smart Grid*, 3(1):543–550, March 2012.

[76] S. Shao, T. Zhang, M. Pipattanasomporn, and S. Rahman. Impact of tou rates on distribution load shapes in a smart grid with PHEV penetration. In *2010 IEEE PES Transmission and Distribution Conference and Exposition*, pages 1–6, April 2010.

[77] M. Shinwari, A. Youssef, and W. Hamouda. A water-filling based scheduling algorithm for the smart grid. *IEEE Transactions on Smart Grid*, 3(2):710–719, 2012.

[78] M. Shinwari, A. Youssef, and W. Hamouda. A water-filling based scheduling algorithm for the smart grid. *IEEE Transactions on Smart Grid*, 3(2):710–719, June 2012.

[79] J. Slotine and W. Li. *Applied Nonlinear Control*. Prentice Hall, 1991.

[80] O. SOLBERG. A new wind turbine control method to smooth power generation. Master's thesis, Department of Energy and Environment, Chalmers University of Technology, 2012.

[81] E. Sortomme, M.M. Hindi, S.D.J. MacPherson, and S.S. Venkata. Coordinated charging of plug-in hybrid electric vehicles to minimize distribution system losses. *IEEE Transactions on Smart Grid*, 2(1):198–205, March 2011.

[82] P. Subbaraj, R. Rengaraj, and S. Salivahanan. Enhancement of combined heat and power economic dispatch using self adaptive real-coded genetic algorithm. *Applied Energy*, 86(6):915–921, 2009.

[83] O. Sundstrom and C. Binding. Planning electric-drive vehicle charging under constrained grid conditions. In *2010 International Conference on Power System Technology (POWERCON)*, pages 1–6, Oct 2010.

[84] O. Sundstrom and C. Binding. Flexible charging optimization for electric vehicles considering distribution grid constraints. *IEEE Transactions on Smart Grid*, 3(1):26–37, March 2012.

[85] N. Tanaka. Technology roadmap: Electric and plug-in hybrid electric vehicles. http://www.iea.org/publications/freepublications/publication/name-3851-en.html, 2011.

[86] Wanrong Tang, Suzhi Bi, and Ying Jun Zhang. Online speeding optimal charging algorithm for electric vehicles without future information. In *Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference on*, pages 175–180, Oct 2013.

[87] M. Tasdighi, H. Ghasemi, and A. Rahimi-Kian. Residential microgrid scheduling based on smart meters data and temperature dependent thermal load modeling. *IEEE Transactions on Smart Grid*, 5(1):349–357, Jan 2014.

[88] J. Taylor, A. Maitra, M. Alexander, D. Brooks, and M. Duvall. Evaluation of the impact of plug-in electric vehicle loading on distribution system operations. In *IEEE 2009 Power Energy Society General Meeting, PES '09.*, pages 1–6, July 2009.

[89] R. Urgaonkar, B. Urgaonkar, M. J. Neely, and A. Sivasubramaniam. Optimal power cost management using stored energy in data centers. In *SIGMETRICS*, pages 221–232. ACM, 2011.

[90] Federal Highway Administration U.S. Department of Transportation. Highway Statistics 2006. Table VM-1. http://www.fhwa.dot.gov/policy/ohim/hs06/htm/vm1.htm, 2006.

[91] Federal Highway Administration U.S. Department of Transportation. National household travel database. http://nhts.ornl.gov/download.shtml, 2009.

[92] Ton J.J. van den Boom and Ton C.P.M. Back-x. Lecture Notes for the MPC DISC Course. http://citeseerx.ist.psu.edu/viewdoc/download?rep=rep1&type=pdf&doi=10.1.1.216.3393.

[93] P. Vytelingum, Thomas D. Voice, Sarvapali D. Ramchurn, A. Rogers, and Nicholas R. Jennings. Agent-based micro-storage management for the smart grid. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1 - Volume 1*, AAMAS '10, pages 39–46, Richland, SC, 2010. International Foundation for Autonomous Agents and Multiagent Systems.

[94] M. Wang, H. Liang, R. Zhang, R. Deng, and X. Shen. Mobility-aware coordinated charging for electric vehicles in vanet-enhanced smart grid. *IEEE Journal on Selected Areas in Communications*, 32(7):1344–1360, July 2014.

[95] C. Wei, H. Hu, Q. Chen, and G. Yang. Learning agents for storage devices management in the smart grid. In *2010 International Conference on Computational Intelligence and Software Engineering (CiSE)*, pages 1–4, Dec 2010.

[96] C. Wu, H. Mohsenian-Rad, and J. Huang. Vehicle-to-aggregator interaction game. *IEEE Transactions on Smart Grid*, 3(1):434–442, March 2012.

[97] D. Wu, D. C. Aliprantis, and L. Ying. Load scheduling and dispatch for aggregators of plug-in electric vehicles. *IEEE Transactions on Smart Grid*, 3(1):368–376, March 2012.

[98] D. Wu, D.C. Aliprantis, and K. Gkritza. Electric energy and power consumption by light-duty plug-in electric vehicles. *IEEE Transactions on Power Systems*, 26(2):738–746, May 2011.

[99] J. Xu and V.W.S. Wong. An approximate dynamic programming approach for coordinated charging control at vehicle-to-grid aggregator. In *2011 IEEE*

International Conference on Smart Grid Communications (SmartGridComm), pages 279–284, Oct 2011.

[100] Z. Yu, L. McLaughlin, L. Jia, M.C. Murphy-Hoye, A. Pratt, and L. Tong. Modeling and stochastic control for home energy management. In *2012 IEEE Power and Energy Society General Meeting*, pages 1–9, 2012.

[101] J. Zhao, Z. Dong, Z. Xu, and K. Wong. A statistical approach for interval forecasting of the electricity price. *IEEE Transactions on Power Systems*, 23(2):267–276, May 2008.

[102] K. Zhou and L. Cai. A decentralized access control algorithm for PHEV charging in smart grid. *Energy Systems*, pages 1–20, 2013.

[103] K. Zhou and L. Cai. A Dynamic Water-filling Method for Real-Time HVAC Load Control Based on MPC. Technical report, E&CE at University of Victoria, 2014.

[104] K. Zhou and L. Cai. Randomized PHEV charging under distribution grid constraints. *IEEE Transactions on Smart Grid*, 5(2):879–887, March 2014.

[105] K. Zhou, J. Pan, and L. Cai. Optimal combined heat and power system scheduling in smart grid. In *IEEE INFOCOM 2014*, pages 2831–2839, April 2014.

[106] K. Zhou, J. Pan, and L. Cai. Indirect Load Shaping for CHP Systems through Real-Time Price Signals. *IEEE Transactions on Smart Grid*, April 2015.