Understanding Deregulated Retail Electricity Markets in the Future: A Perspective from Machine Learning and Optimization

by

Tao Chen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical and Computer Engineering)
in the University of Michigan-Dearborn
2018

Doctoral Committee:

Assistant Professor Wencong Su, Chair
Associate Professor Yi-Su Chen
Associate Professor Sridhar Lakshmanan
Assistant Professor Samir Rawashdeh

ACKNOWLEDGEMENTS

Firstly I would like to thank to Prof. Wencong Su, for his guidance and support through all the three years of my PhD study. It was a great pleasure to learn academics and much more than academics from him. Thanks for numerous discussion and idea exchange happening on his cosy office. His sharp advices and remarkable help concluding all the publications contained in this dissertation. I really appreciate it!

For willing to be on my doctoral committee and for helpful comments on my research, I thank Yi-Su Chen, Sridhar Lakshmanan, and Samir Rawashdeh. I also thank Luis Ortiz and Jian Hu for their kindly acceptance of the invitation. Furthermore, I am greatly thankful to the excellent cooperation with all my coauthors, many colleagues in Umich and GEIRINA, and Pertti Järventausta for providing dataset and hosting for visiting back Finland. Of course, there is a huge research community of many people and friends who supported me deeply. I don't have the space or time to thank everyone I met in my doctoral journey, but they've all made me become myself.

This thesis is dedicated to my family. To my grandparents, Rongxiu Ye and Xianfu Chen. To my parents, Jinghua Jiang and Yimin Chen, for their unconditional love and always support. To my wife, Yajuan Cao, without her forever love and support, this dissertation would not have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

On top of Smart Grid technologies and new market mechanism design, the further deregulation of retail electricity market at distribution level will play a important role in promoting energy system transformation in a socioeconomic way. In today's retail electricity market, customers have very limited ”energy choice,” or freedom to choose different types of energy services. Although the installation of distributed energy resources (DERs) has become prevalent in many regions, most customers and prosumers who have local energy generation and possible surplus can still only choose to trade with utility companies. They either purchase energy from or sell energy surplus back to the utilities directly while suffering from some price gap. The key to providing more energy trading freedom and open innovation in the retail electricity market is to develop new consumer-centric business models and possibly a localized energy trading platform. This dissertation is exactly pursuing these ideas and proposing a holistic localized electricity retail market to push the next-generation retail electricity market infrastructure to be a level playing field, where all customers have an equal opportunity to actively participate directly. This dissertation also studied and discussed opportunities of many emerging technologies, such as reinforcement learning and deep reinforcement learning, for intelligent energy system operation. Some improvement suggestion of the modeling framework and methodology are included as well.

CHAPTER I

Introduction

1.1 Background

[1] Although electricity market deregulation has been underway since the United Kingdom opened a power pool in April 1990 [2], competitive forces in the U.S. electricity market have been largely silent since the early-2000s California electricity crisis. Then, since the 2010s, many power sector reforms and new market mechanism designs have been under intense discussion again due to the emerging smart grid technologies plus some innovative information technology (IT) business models and an Internet-inspired commercial paradigm [3]. However, most research on the electricity market still focuses on the wholesale market, particularly the bidding process and financial transmission rights [4] [5]. The development of the retail electricity market seldom borrows much experience from such bulk power transactions, though. Instead, it prefers to follow principles, like multi-options, peer-to-peer, sharing economy friendliness, negotiability, and so on, that are utilized successfully in the customer-centric IT industry. This characteristic is also the reason for popular proposals such as the energy Internet [6] and digital grid [7] in many references.

Around the world, many countries are also pushing the reform of the electricity power

[1] This chapter was previously published as an article in a peer-reviewed journal: [1] **T. Chen**, Q. Alsafasfeh, H. Pourbabak, and W. Su, ”The Next-generation Retail Electricity Market with Customers and Prosumers - A Bibliographical Survey”, Energies, vol.11, no.1 2017.

sector very positively. Chile pioneered in the 1980s the deregulation of the electric power industry [8]. The European Union had taken steps to liberalize its electricity industries in the 1980s and, late in the 2000s, allowed all customers to choose their electricity suppliers [9]. The restructuring and regulatory reforms in the PR China's power sector happened in the 2000s along with other Asian countries [10]. The electricity retailing in Japan was fully deregulated with fierce competition in April 2016 [11]. In today's U.S. retail electricity market, 14 states have already adequate retail competition with Texas, Illinois and Ohio having 100%, 60% and 50% of their residential customers receiving service from electricity suppliers [12]. However, many customers still have very limited "energy choice" or direct participation in the existing retail electricity market.

The key to open innovation in the power sector has been believed to be the development of consumer-centric business models and well-designed demand side management (DSM) programs [13] [14]. Following these ideas, the recent work in [15] looks even further forward to more subtle modeling of customer behavior, with considerations of their willingness to participate and even emotional or irrational features. With these prevailing ideas in the research community, the next-generation retail electricity market infrastructure will be a level playing field, where all energy end-users and customers have equal opportunity to play the role of active participants rather than pure passive price-takers [16] [17]. Fortunately, the recent development of the functionalities of the energy service companies (EsCos) and the distribution system operator (DSO) has opened many new possibility for monitoring, coordinating and controlling short-term or real-time delivery of electricity at the distribution level [18]. Especially with the further development of the concept of the DSO, deregulation of the electricity market has been spreading out from wholesale market design into retail market design, as shown in Figures 1.1 and 1.2. In the new paradigm for energy transactions, different customers or customer groups (e.g., energy communities) are free to choose their service provider, either a distribution company or utility company, including even pure energy retailers, periodically.

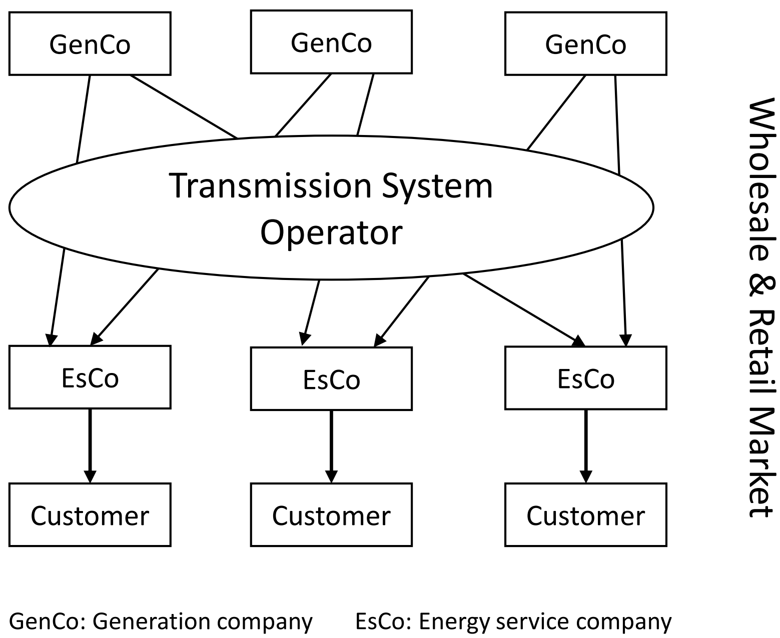


Figure 1.1: The deregulation of the wholesale electricity market.

Moreover, in smart grids, more and more customers will be able to have local generation capability, i.e., distributed energy resources (DERs), along with various flexible controllable loads, such as thermostatically-controlled loads (TCLs), distributed energy storage devices (DESDs) and washing machines [19] [20]. Electric vehicles (EVs) and plug-in electric vehicles (PEVs) are also appealing as the most controllable loads because they can be curtailed for significant periods of time (e.g., several hours) without impact on end-use function [21] [22]. These kinds of customers are encouraged to actively participate in the retail market to provide demand response or localized power balance between energy surplus and energy deficit.

Some existing survey papers focus mainly on the decision-making process of retailers in the wholesale market and somehow ignore the significant effect that various types of future energy end-users will have on the whole electricity market landscape [17]. The entire

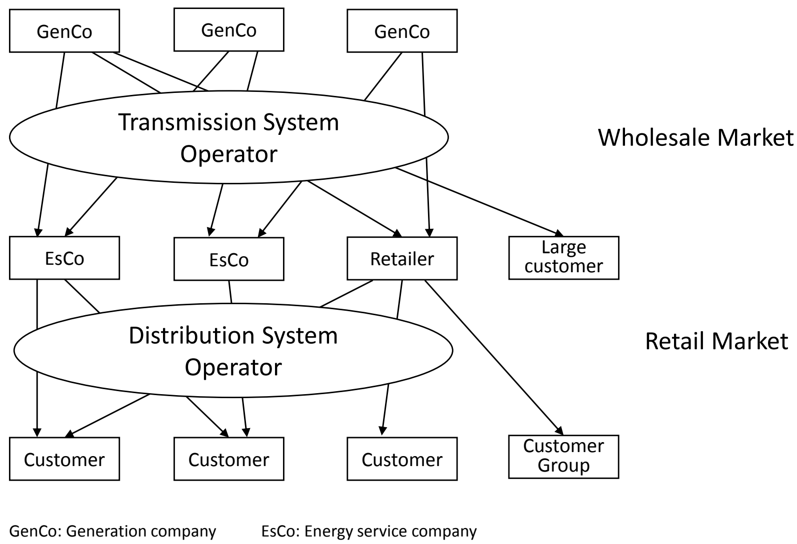


Figure 1.2: The further deregulation of the retail electricity market.

energy business ecosystem will be re-formed if the most recent research trends and principles, such as transactive energy [23] [24], transactive control [25] (Transactive energy and transactive control are explained further in Section 2.4), an energy sharing economy [26], and so on, are adopted.

1.2 Retail Electricity Market with Pure Consumers

In most scenarios, customers play a passive role as price-takers in retail electricity, purely serving as consumers of energy at different locations. Those who have the capability to generate power locally with the help of microgrids and are able to supply electricity to other customers are called prosumers at the distribution level. We will leave the discussion of the retail electricity market that includes prosumers for the next section.

1.2.1 DSO with Distribution Level Pricing

As a result of the distribution grid's increasing number of roles and functionalities, the deployment of a DSO is becoming a necessity to ensure efficient and reliable delivery of electricity to emerging proactive customers. Customers now have more willingness to control their energy use and transactions with the utility grid, as their energy preferences have evolved. In parallel, there is a potential need for an intermediate entity between the regional transmission operators (RTOs) or independent system operators (ISOs) and energy end-users due to the limited visibility and control over the meter resources (e.g., advanced metering infrastructure) at the customer side [27]. A DSO in the future energy system and energy market design may be considered the evolution of a distribution management system, with, however, more functionality at different layers (Figure 1.3).

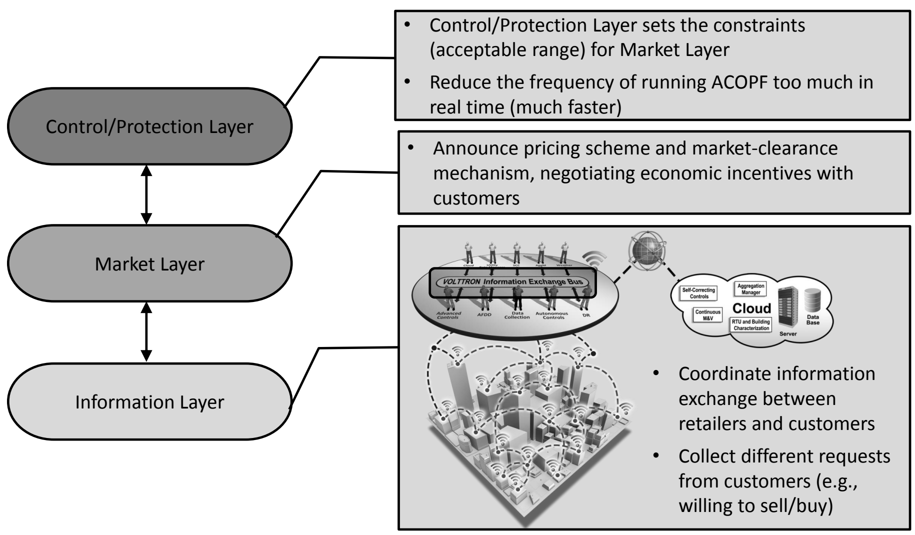


Figure 1.3: The new role design of the DSO. ACOPF, alternating current optimal power flow.

In addition to the traditional mission to operate, maintain and develop an efficient electricity distribution system, the DSO possesses more functionality rather than only mimicking the ISO's pure responsibility of electricity pricing and independent market-clearance at the transmission level as a non-profit entity. In the wholesale market, many ISOs nationwide implement the locational marginal pricing (LMP) strategy either in the form of ex ante LMP, for example New York ISO (NYISO), or ex post LMP, for example ISO-New England (ISO-NE), PJM and Midcontinent-ISO (MISO) [28]. Based on the fact that LMP has been widely adopted to compute electricity prices in the wholesale electricity market [29], some scholars have begun to downscale the LMP schema for distribution networks by proposing its counterpart, distribution locational marginal pricing (DLMP) [30], which can directly work for individual energy end-users without referring to a load serving entity (LSE) or other demand bidding aggregators. It has been applied to several scenarios, such as the congestion management problem and the electric vehicle charging problem [31].

However, as shown in Figure 1.3, the DSO may not only play the role of an ISO at the distribution level since there is a huge difference between a distribution network and a transmission network, such as three-phase imbalance, radial system topology, high ratio of power loss, numerous low-voltage buses, and so on. To some extent, DLMP is hardly effectively obtained through running alternating current optimal power flow (ACOPF) for a distribution system. A very recent three-phase ACOPF-based approach has been developed to define and calculate DLMP accurately [32].

1.2.2 Decision Making of Retailers

Retailers in the electricity market are supposed to purchase electricity in the wholesale market and sell electricity to their subscribed end-user customers through assigning appropriate tariffs, either in a temporal variance way or at a flat rate. Currently, the electricity retail company is usually operated as an entity that is independent of any generation or distribution company [17]. The decision-making process involved in buying and selling strategies usually contains some volatile market risks that are similar to the ones in any other market, such as the stock market and oil market. Especially with the further deregulation of the electricity market, along with the development of DSM and the proliferation of DERs, retailers participating in both the wholesale market and the retail market should carefully design their buying-selling trade-off and electricity portfolio optimization [33]. In the future, many innovative pricing schemes will be necessary, taking into account emerging factors such as the increasing penetration of renewable energy, wide deployment of storage devices, adoption of advanced information and communication technologies

(ICT) and rising customer awareness of switching among electricity suppliers. These new challenges also require retailers to incorporate some typical operations into their decision-making processes, which include retail energy forecasting, portfolio evaluation and risk management.

Due to the page limit and many mature approaches that already exist for residential load forecasting [34] [35] and portfolio evaluation [36] [37], risk management will be the focus of discussion here, along with many recent advances in the research community. In a typical example such as [38], the author utilizes stochastic programming techniques to determine the day-ahead market bidding strategies for retailers with flexible demands to maximize their short-term profit, specifically including a case study based on Sweden's electricity market and consideration of the demand uncertainty of retail customers. In most studies of retail electricity market operation, with risk from either real-time price or demand uncertainty, conditional value-at-risk (CVaR) is widely used to consider risk management [39]. CVaR is a risk assessment technique often used to reduce the probability that a portfolio will incur large losses, which is performed by taking a weighted average between the value at risk and losses exceeding the value at risk [40].

1.2.3 Price Scheme and Demand Response

On the customer side, energy end-users do respond to the long-term electricity contract and price schemes offered by the utilities; however, they are usually insensitive and uncomfortable with respect to the highly dynamic or real-time pricing, due to the lack of competence to immediately respond to the price signal or little awareness of instantaneous opportunity [41] [42]. However, electricity prices that describe marginal costs can vary substantially over time. Fixed rates may ignore continuous changes in the electricity system conditions. Setting prices that differ for certain periods is an approach to approximating the real-time price. If rates are set much in advance and fixed over periods of time, they miss the majority of the potential gain as measured by the variance index [43]. Both time-of-use (TOU) and critical-peak pricing (CPP) play crucial roles in providing load flexibility and tariff design in the retail electricity market [44]. Based on the similar idea of rationalizing energy consumption behavior for the whole system cost, a prediction-of-use (POU) tariff is proposed and believed to better reflect the predicted cost for a customer [45]. The possible combination of POU with the more widely-known TOU tariffs is also considered, which allows customers to fully benefit from meaningfully managing their consumption, as well as from their contribution to the system's delivering energy-efficient solutions.

Using TOU, CPP or other price schemes as baselines, some additional incentive mechanisms are also proposed on top of them to reflect the demand response from customers with energy awareness, which are aware of the electricity price elasticity and reasonable energy saving. Energy tokens, coupons and eVouchers, similar to their literal meaning in daily commercial activities, are proposed in [46] [47] [18] to encourage voluntary energy demand adjustment based on the negotiation principle. These kinds of negotiation-based demand response programs can be categorized as incentive mechanisms [48] that provide an additional economic management tool for the power system and market efficiency.

1.2.4 Transactive Energy and Transactive Control

In order to combine power systems tightly with economic or market-based operation, the term "transactive energy" has begun to be used to refer to techniques for managing the generation, consumption or flow of electric power within an electric power system through the use of economic or market-based constructs, while considering grid reliability constraints [49]. In fact, transactive energy (TE), one of many promising solutions to electrical grid restructuring issues, has gradually become a more and more concrete concept among many discussions [50]. Some experts give it the official definition of "a set of economic and control mechanisms that allows the dynamic balance of supply and demand across the entire electrical infrastructure using value as a key operational parameter" [49]. Specifically, transactive energy mainly focuses on the value or economic operation of a modernized electrical grid, primarily from an economic perspective. It emphasizes the innovative business models and new consumption patterns in electric markets, along with taking some social impact into consideration. Some researchers [51] also believe TE is a potential framework to close the gap between wholesale and retail markets. Most DERs and demand resources can be aggregated as virtual power plants (VPPs) to provide bulk power adjustment capacity in different markets.

In practice, analogous to the price reaction approach, a concept named transactive control [25] allows the operation of flexible devices to be optimized economically by a local intelligent controller (or agent) under the control of the end-user and follows the principles of TE. In a way, a society of intelligent devices is formed to allow for market bids to be sent by a particular group of devices (e.g., hot water buffer, dishwasher, air conditioner, etc.) [24]. The local-level bidding process, or laminar market architecture, is extremely suitable for thermostatically-controlled loads [52]. Furthermore, both transactive energy and transactive control concepts are becoming more and more widely adopted by many recent pilot projects [24, 50].

1.3 Retail Electricity Market with Prosumers

In this section, the retail electricity market with prosumer participation in the local power supply will be discussed. Prosumers are defined as agents that both consume and produce energy [53]. With the growth in the number of small- and medium-sized energy entities using solar photovoltaic panels, small wind turbines, vehicle-to-grid EV/PEVs, home storage energy systems, smart meters and other smart devices, prosuming offers the potential for consumers and smart device owners to re-evaluate the energy practices in their daily lives. As the number of prosumers increases, the retail electricity market of today is likely to undergo significant changes over the coming decades, not only offering possibilities for localized inter-network energy trading and balance, but also introducing many challenges and risks that need to be identified and managed. To develop strategies for the future, policymakers and planners need knowledge of how prosumers can be integrated effectively and efficiently into a competitive retail electricity market. Some promising potential market mechanisms, such as prosumer grid integration, peer-to-peer models, indirect customer-to-customer trading and prosumer community groups, along with their implementation approaches, are identified and discussed below.

1.3.1 Prosumer Grid Integration

Most prosumer integration problems can be incorporated by extending the conventional optimization model to solve the pure energy consumption and energy management problems. However, some characteristics of two-way power flow need to be carefully considered for various types of challenges and optimization constraints, such as inverse current fault detection, distribution topology estimation, power surplus balance, and so on. When leveraged by an energy storage system (ESS), including vehicle-to-grid (V2G) technology, distribution network operation with a high penetration of prosumers needs to make sure that prosumers' benefits are aligned with the regulator/DSO's concerns, thus satisfying the requirements of both sides. The authors in [54] propose a market-based control to solve this issue. The complexity in the environment and in the interactions among players prompts techniques derived from complex systems theory. The work in [55] analyzes the optimal planning and operation of aggregated DERs with participation in the electricity market. In most cases, the aggregator of a large amount of DERs can operate as a virtual power plant (VPP) [56], which is connected as part of the main grid and participates directly in the wholesale electricity market. Many similar ideas based on aggregation frameworks have been frequently employed in solving the prosumer grid integration problem. However, from an energy utilization and market efficiency point of view, localized integration at low voltage levels with direct delivery to end-users is still highly encouraged. More and more decentralized decision-making frameworks without the necessity of aggregation are welcome nowadays.

The integration of various DERs and EVs also provides a new chance for building an innovative business model and a new energy ecosystem. There is a plethora of research and development areas related to prosumer grid integration that can be exploited for new business opportunities, thus spawning another branch of the so-called "green economy" focused on turning smart energy usage into a profitable business [57].

1.3.2 Inter-Network Trading with Peer-To-Peer Models

The encouragement of localized energy trading within a distribution network at low voltage levels promotes an eBay-like market platform and peer-to-peer models. Additionally, a high penetration of distributed energy resources raises operational and market challenges such that existing incentives and tariff support cannot be sustained with penetration growth at the microgrid level. As a result, some competitive market mechanisms or peer-to-peer models are required at the local distribution level. In [58], a matching mechanism is proposed to allow individual generators and load units to meet to conduct a bilateral trade. Each unit interested in maximizing its benefit adopts its own bid strategy. Trade between a randomly matched generation and load unit is established if their bids are compatible, which does not require the units to share their private cost or value information.

Sometimes, this type of peer-to-peer energy sharing is described as "Energy AirBnB" for future electricity retail markets. Furthermore, some peer coalition might be allowed in the electricity retail market and work as a new business model for a very short term. Some possible strategic coalitions among independent electricity retailers or prosumers may happen under the designed distributed framework [59] to maximize profits, which implies that electricity retailers may solely compete with each other, while some of them may cooperate with others to form a coalition in the economic operation of future electricity retail markets. In [60], a scalable and modular system is proposed and demonstrated for energy trading between prosumers. Even a novel decentralized digital currency, named after NRGcoins, is proposed by the same group of researchers to encourage prosumers to locally trade their excess energy while payments are carried out using NRGcoins [61].

The driving force of such a peer-to-peer mode becoming welcome in the retail electricity market is mainly due to two facts. The first one is frequently discussed: that the rapid adoption of DERs enhances people's willingness to trade in a decentralized way. The centralized operation will put too much burden on the central controller when all the individual customers send the trading requests at the same time. The other fact seems not so explicit: that the rise of Internet-connected devices (e.g., Internet of Things) has led to a wide energy connection, which is also strengthened by the disappearance of the conceptual gap between energy as a physical supply service and energy as an information service. The behavior of trading energy among peers more or less carries some meaning of social interactions. It is also another important source for the proposed idea: the energy Internet.

1.3.3 Indirect Customer-To-Customer Trading

Although peer-to-peer models are very attractive for a highly decentralized energy supply, some customers or prosumers can find it difficult and time-consuming to search for suitable partners. They may feel more comfortable and find more convenience trading through an intermediate trader, like an agent or broker in real estate business. This role of intermediate trader particularly in a local retail electricity market allows them to keep additional energy transaction options besides only selling back or buying from utility companies. There are already several pilot projects and demonstration projects underway, verifying the possibility of monetized local energy exchange. For instance, since 2010, Pecan Street Inc. (Austin, TX, USA) has been collecting high-resolution data on how and when homes and small businesses in the United States use and, when PV is present, generate energy [62]. Then, this temporal and spatial information of energy usage/generation can be used to test potential energy trading programs along with predicting the market capacity. On the other hand, in order to reduce the energy transaction cost and search friction in such an indirect customer-to-customer trading paradigm, a local energy market is proposed in [63] to accommodate localized energy trading and exchange for communities, buildings and campuses, which may own surplus local energy produced by on-site DERs.

In this framework, as shown in Figure 1.4, an important new role, named the energy broker (EB) and working as a middleman or trader in this localized retail electricity market, is introduced to get buyers and sellers together, serving as a solution to search friction [64]. Both the buyers and sellers who would like to participate in this local energy market will provide bid/offer information of price-quantity pairs (price (PC) and amount (AM) in Figure 1.4) in each open time interval. The trader itself will also choose to maximize transaction efficiency or revenue with consideration for search cost in each open market time interval accordingly. The index of historical credit for energy transactions, sellers' commitment probabilities (SCPs), is also proposed for power allocation of different trading peers. In this way, the proposed market structure can be modeled based on search theory and an optimal stopping problem (OSP) [63]. Some other similar works about this topic can also be found in [65].

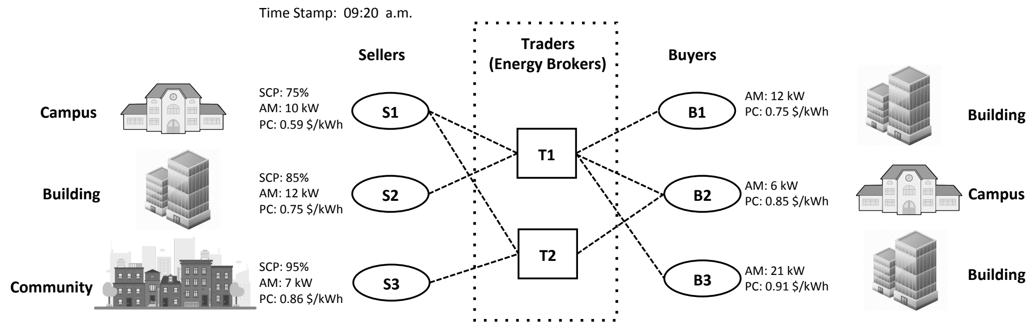


Figure 1.4: The localized retail market with energy broker and search theory. SCP, sellers' commitment probability.

1.3.4 Prosumer Community Groups

The prosumer community group is another typical prosumer paradigm that aims to provide common platforms for coordinating neighboring or local prosumers for exchanging energy and information within the local community or interfacing with outside energy entities as a whole. The authors in [26] argue that energy sharing among neighboring PV prosumers in the microgrid could be more economical than the independent operation of prosumers. They propose an energy-sharing model with price-based demand response (DR) for microgrids of peer-to-peer (P2P) PV prosumers to validate the benefit of forming prosumer communities. In [66], a new vision for local distribution systems is proposed, in which prosumers are encouraged to better balance their electricity usage in a local community through psychological balancing premiums. Even the social behaviors and some quantitative psychological characteristics of self-interested prosumers are considered in modeling the energy exchange and transactions. Price-responsive generation and demand of an individual prosumer are affected by his/her inherent characteristics and the individual's attitudes toward benefit and comfort, which evolve during social interactions. The authors in [67] also introduce a novel concept to manage prosumers in the form of goal-oriented virtual communities. They meanwhile discussed different aspects of the formation, growth and overall management of a prosumer-community. The main significant implication of this approach is that the prosumer-communities are able to facilitate the joining together of prosumers with similar interests. In this way, the quantity of energy to be auctioned to the smart grid can be increased accordingly, and furthermore, the prosumers' bargaining power is increased in the energy market. In a smart community, the benefits of DERs can also be considered in an energy management scheme, where a large number of residential units can participate and a shared facility controller (SFC) can be introduced [68]. The SFC is defined as a public controller that exclusively controls electricity for those publicly sharing used equipment, devices and machines (e.g., water pumps, lifts, parking lot gates, lights, etc.) by the residents of the community. Therefore, the SFC needs to afford all its energy cost either buying from the main grid or buying from the residential units with DERs due to its lack of electricity generation capability.

1.4 Methodology

The methodology used in the study of the retail electricity varies greatly according to particular application scenarios, including making market rules, predicting customer behavior, reducing system operation cost, and so on. On the other hand, all the methods applied in different projects also depend on how to describe the dynamics of the market mechanism in a quantitative way, namely system modeling. In this section, some common methodologies are discussed. However, those methodologies are usually not applied very independently and have the trend of being combined in a hybrid framework to make the system modeling more accurate and efficient.

1.4.1 Optimization, Distributed Optimization and Blockchain

Optimization methods are still dominant in most decision-making problems pertaining to system and market operation. Stochastic optimization, robust optimization, multi-objective optimization and mathematical programming have been widely adopted for research on the wholesale market for market-clearance, and most of this research takes into consideration various types of uncertainties resulting from variable demand or renewable energy supply [36] [30] [59]. Retailers in the retail electricity market often refer to these optimization methods to guarantee their revenue through deterministic analysis. However, since there are numerous decision variables at the distribution level associated with frequent monitoring activities and a large number of customers, especially given more and more small-sized local generation units, global optimization has become rarely implemented due to its increasing computational complexity. Consequently, the state-of-the-art strategy has begun to shift to distributed optimization with necessary decomposition, such as the alternating direction method of multipliers (ADMM), consensus-based algorithms, proximal message passing (PMP), and so on [20] [69].

Blockchain technology, borrowed from the IT industry, has also attracted much attention due to the prevailing distributed optimization implementation in practice. It has been suggested as promising and suitable for such a decentralized decision-making process [70]. The authors in [71] present an architecture for peer-to-peer energy markets that can guarantee that operational constraints are respected and payments are fairly rendered, without relying on a centralized utility or microgrid aggregator. They demonstrate how to address trust, security and transparency issues by using blockchains and smart contracts, two emerging technologies that can facilitate decentralized coordination between non-trusting agents. While blockchains are receiving considerable interest as a platform for distributed computation and data management, this work may be the first one to examine their use to facilitate distributed optimization and control of DERs. Some other works also introduce the utilization of blockchains in local energy transactions between DERs, including a custom-designed blockchain mechanism designed to maintain a distributed database trusted by all DERs and to stipulate and store a smart contract that enforces proportional fairness [72].

1.4.2 Game Theoretic Method and Prospect Theory

In prosumer-centric energy trading, since most interconnected microgrids or DERs operate autonomously and have their own goals of optimizing performance and maximizing benefits through energy trading, the selfish nature of players participating in local energy transactions is inclined to be described by game theoretic methods. A Nash bargaining theory-based incentive mechanism is proposed and designed in [73] to encourage proactive energy trading and fair benefit sharing. It takes autonomous microgrids independent self-interested entities, without assuming that all the microgrids are coordinated by a common grid operator or controlled following a hierarchical structure. In [74], game-theoretic day-ahead energy scheduling in a residential distribution system is proposed, in which the distributed electricity prosumers may only compete with each other while some of them may cooperate with others to form a coalition. A similar noncooperative Stackelberg game between the residential units and the shared facility controller is proposed in [68] in order to explore how both entities can benefit, in terms of achieved utility and minimizing total cost, respectively, from trading energy with each other and with the grid.

It is noteworthy that the proposed game in [75] accounts for each prosumer's subjective decision when faced with the uncertainty of profits, induced by the random future price. In particular, the framing effect from the framework of prospect theory is used to account for each prosumer's valuation of its gains and losses with respect to an individual utility reference point. Prospect theory (PT) is mainly an interpretative theory that considers weighting effect to transform the objective probabilities into subjective probabilities, which was proposed to explain the fact that people usually over-weigh the low probability bad outcomes and under-weigh their favorite outcomes with high probabilities [75]. PT is helpful for modifying conventional game-theoretic methods because it relaxes the assumption of rationality in most game frameworks by taking into account subjective irrational decision behavior [15, 76]. It is not at the same level as game theory, but possible to be combined into building utility functions in game models. Even so, most game-theoretic methods still possess too much simplification, making it hard to find the equilibrium solution, especially for large-scale systems with high computational complexity.

1.4.3 Agent-Based Simulation

Agent-based simulation (ABS) has been another popular tool, since at least the early 1990s, to model the dynamics of the electricity market, including both the wholesale market and the retail market [77]. ABS is particularly suitable for large-scale systems involving various types of interacting system participants. These participants are usually assigned distinct roles, functionalities, behaviors and decisions, which depend on different objective design and interactions with other system participants [78]. In an agent-based system, an agent can be as simple as a single variable (e.g., energy price-amount pair) within a computer program or as complex as an intelligent object, such as a human being (e.g., speculator), involving possibly an infinite number of states, decisions and actions/reactions. However, most ABS are mainly designed for the electricity wholesale market, neglecting transmission/distribution grid constraints [79] [80]. The difficulty of validating an ABS model's outcomes against empirical data is also one of the weaknesses of the ABS methodology.

In recent years, many agent-based systems have become popular again for electricity market simulation, due to the further development of reinforcement learning and the other computational resources available. The Power Trading Agent Competition (Power TAC) is an influential event and simulator that allows rich competitive simulation of future retail power markets and helps with understanding the dynamics of customer and retailer decision-making and the robustness of market designs. Power TAC models a liberalized retail electricity market, where competing business entities or brokers offer energy services to customers through tariff contracts [81]. On the other hand, some researchers also mimic the wholesale market mechanism to study the behavior of a day-ahead retail electrical energy market with price-based demand response from air conditioning loads through a hierarchical multi-agent framework [82]. Meanwhile, ABS is also frequently used as a validation tool for testing certain market rules for policy makers. For instance, an agent-based simulation of the liberalization of a retail electricity market has been developed to introduce competition into a sector historically characterized by the regional monopoly of retail electricity [83]. It is worth mentioning that most existing ABS usually assign some learning capability to intelligent agents and often leverage the Q-learning algorithm from the machine learning field [82].

1.4.4 Machine Learning Techniques

Machine learning has become the status quo for most intelligent systems, including power systems and the electricity market. Utilizing machine learning techniques to detect distinct energy consumption patterns of customers and select high-quality customers for energy programs (e.g., demand response programs) is becoming more and more popular in addressing competitive utility companies and the future energy business ecosystem [84] [85] [86] [87].

Together with various types of machine learning techniques, including successful application of supervised learning in demand response targeting [85] and unsupervised learning in individualized pricing design [87], reinforcement learning (RL) is also believed to have the potential to deal with the energy trading problem and guide energy entities to interact with the market environment. The most important feature distinguishing RL from other types of learning is that it uses training information that evaluates the actions taken rather than instructions by giving correct actions [88]. This is very suitable for economic activities (e.g., energy transactions) that are based on the voluntary principle and associated with privacy issues regarding consumption information. On the other hand, the online nature of RL makes it possible to approximate the best decision-making strategy or optimal policies in ways that put more effort into learning to make good decisions for frequently-encountered states (e.g., high energy demand during the daytime), at the expense of less effort for infrequently-encountered states (e.g., peak load happening at night). The project in [89] demonstrates a data-driven control approach for demand response in real-life residential buildings, in which the objective is to optimally schedule the heating cycles of the domestic hot water (DHW) buffer.

However, most RL applications in power systems depend heavily on Q-learning or other tableau methods, which are based on look-up tables and afford low computational efficiency with increasingly big datasets. In recent research advancements, combining deep learning (DL) and RL to form a deep Q network (DQN) is suggested as a good approach for value function approximation and improving algorithm performance [90]. It can also conquer many of the weaknesses (e.g., feedback delay, partially-observable environments, numerous enumerations) in the energy system decision-making process for the retail electricity market.

These methods, including the optimization model, game-theoretic model, agent-based simulation and machine learning, are usually correlated with each other and often combined together as hybrid frameworks. For instance, the game-theoretic method that finds the equilibrium point is easily transformed to an optimization problem that solves for equivalent optimal results [74]. Agent-based simulation is mostly combined with machine learning techniques to facilitate the interaction dynamics among different agent entities [82]. The machine learning technology also frequently uses optimization methods to train its model parameters and hyper-parameters [85]. Last but not least, the summary of each individual modeling method is shown in Table 1.1.

Table 1.1: Solution methods for the new paradigm of the retail electricity market.

Solution methods	**Advantage**	**Disadvantage**	**Prosumer easily considered**	**Computational complexity**
(Distributed) optimization	Accurate analytical solution result with clear interpretation; Easily consider power flow constraint and network operation conditions; Deterministic conclusion;	Hard to describe every trading features in constraints; Central or regional controllers are needed; Usually need high computational resources;	Yes	Medium
Game theoretic method	Intuitive description about different market participants; Suitable for distributed control; Good economic interpretation;	Convergence is not guaranteed and hard to find the equilibrium point; Limited to stylized trading situations involving few actors;	No	High
Agent-based modeling	Highly adaptive to market and trading environment; Heterogeneity of different types of market participants; Easily incorporate social abilities to exchange information;	Most neglect transmission/distribution grid constraints; Results are mostly non-deterministic with poor interpretation; Not reliable due to external conditions and for policy makers;	Yes	Low
Machine learning techniques	Very autonomous decision-making process; Insensitive to market structure and large data sources;	Data-driven and need realistic experiments; Usually need high computational resources;	Yes	Medium

1.5 Discussion and Policy Issues

Based on the various aforementioned studies of the retail electricity market in recent years, some trends can be easily observed that: (1) the system or market operation is more fine-grained from different perspectives, trying to balance credits' assignment and benefit sharing among many types of market entities, including suppliers, speculative retailers, utilities, service providers, customers and other new parties introduced by new business models; (2) more and more consideration is given for economic operation on top of pure system requirement satisfaction, and a certain degree of risk is acceptable given the improving uncertainty of the whole system; (3) customers are expected to be more active in this market-loop instead of passive participants, which are allowed to directly interact with other market participants and exercise negotiation power.

The study of the electricity market is more or less not a pure technique problem, especially considering the fairness rule (e.g., non-discrimination), data privacy and renewable energy subsidy policy in the retail electricity market close to the customer side. In North America, the U.S. electricity ownership structure is actually quite complex. The U.S. electric power industry consists of approximately 3300 publicly-owned, investor-owned and cooperative utilities; more than 1000 independent power generators; 3 regional synchronized power grids; 8 electric reliability councils; about 150 control-area operators; and thousands of separate engineering, economic, environmental and land use regulatory authorities [16]. We provide a retrospect of the history of U.S. electricity deregulation in Table 1.2 based on our previous work in [16] and hope to remind that electricity deregulation should keep track of the development of emerging technologies, especially considering the manipulative market power brought by these technologies and new business models. Further deregulation of the retail electricity market definitely requires cooperation and technical support from the wholesale market, which is still under intense discussion across the industry and the academic community [91].

Table 1.2: History of U.S. electricity deregulation.

Year	**Effect**
1935	Congress passes the Public Utility Holding Company Act of 1935 (PUHCA) to require the breakup and the stringent federal oversight of large utility holding companies.
1978	Congress passed the Public Utility Regulatory Policies Act (PURPA) which initiated the first step toward deregulation and competition by opening power markets to non-utility electricity producers.
1992	Congress passed the Energy Policy Act of 1992 (EPACT), which promoted greater competition in the bulk power market. The Act chipped away at utilities monopolies.
1996	FERC implemented the intent of the Act in 1996 with Orders 888 and 889, with the stated objective to remove impediments to competition in wholesale trade and to bring more efficient, lower cost power to the nations electricity customers.
2005	Congress passed the Energy Policy Act of 2005, a major energy law to repeal PUHCA and decrease limitations on utility companies ability to merge or be owned by financial holding/non-utility companies.
2007	FERC issued Order 890, reforming the open-access regulations for electricity transmission, in order to strengthen non-discrimination services.
2008	FERC issued Order 719 to improve the competitiveness of the wholesale electricity markets in various ways, and to enhance the role of RTOs.
2012	FERC issued Order 768 to facilitate price transparency in markets for the sale and strengthen the Commissions ability monitor its retail markets for anti-competitive and manipulative behavior.

Although our discussion is mainly focused on the U.S. electricity market, it is worth mentioning that many countries in Europe and on other continents also meet similar challenges with the retail electricity market, such as electricity buy-back volatility, cross-subsidies, distribution cost allocation, and so on [92]. Take for instance the residential U.K. electricity market: it was opened for the first time in 1999, introducing the choice of supplier, and about 40% of households changed supplier in the first four years. After three years, price caps were removed. The work in [93] reviews this process and assesses the competitiveness of the market by examining how the charges levied by suppliers depend on cost and demand factors for three different payment methods and consumption levels, whose experience may be helpful for U.S. retail electricity market development. However, the market deregulation process is not always so successful and full of various kinds of challenges that are far beyond our expectation. Some researchers have summarized two main negative phenomena that could reduce the impact of introducing competition in the retail electricity market: cognitive bias affecting consumers' decisions to switch and a technological paradigm reducing innovation opportunities in commercialization [94]. These discussions can go on and on due to the many research perspectives involved in this field. Research on the electricity market is always hungry for more interdisciplinary study from other fields, such as economics, computer science and operational research.

1.6 Conclusions

To help researchers have an overall understanding of the recent research work on the retail electricity market, different sub-topics with/without prosumers and commonly-used methodologies are surveyed and discussed in this paper. The state-of-the-art, emerging new market functionalities (e.g., DSO's new role, incentive mechanisms, transactive energy, prosumer community groups) and recent innovative techniques (e.g., prospect theory, blockchain, reinforcement learning) have been discussed, covering the entire landscape of the retail electricity market.

In the survey of more than 90 papers published within the last five years that study the retail electricity market, the phenomenon can be observed that more and more intelligent system technology, like machine learning and the Internet of Things, is coming into play in this field. These new automation methods and autonomous systems or controllers allow customers to easily coordinate with each other and actively participate in the electricity market, instead of only passively accepting what they are provided. Another observation is that innovative business model design remains the key driving force behind the reform of traditional energy exchange and transactions.

We intentionally skip some common topics, such as load forecasting and demand response, covered by many existing survey papers, and mainly focus on the most recent developments in the area of innovative conceptual frameworks in the study of the retail electricity market. In the future, more localized energy market models under incubation will come into practice and revolutionize the whole energy ecosystem.

CHAPTER II

DSO service

2.1 Introduction

[1] Although electricity market deregulation has been under way since the United Kingdom opened a Power Pool in April 1990 [2], competitive forces in the U.S. retail electricity market have been largely silent since the early-2000s California electricity crisis. In today's retail electricity market, customers have very limited ”energy choice” or participation, i.e., ability to choose their supplier from competing electricity retailers. The key to open innovation in the retail electricity market is the development of a consumer-centric market and well designed demand side management (DSM) programs [13][14]. The work in [15] looks even further forward to more subtle modeling of customer behavior with consideration for emotional or irrational features and their willingness to participate. It is believed that the next-generation retail electricity market infrastructure will be a level playing field, where all customers have an equal opportunity to actively participate [16][17]. Fortunately, the recent development of functionalities of the distribution system operator (DSO) and the electricity distribution company (EDC) has already opened new possibilities of coordinating, monitoring, and controlling short-term or real-time delivery of electricity at the distribution level.

[1]This chapter was previously published as an article in a peer-reviewed journal: [95] **T. Chen**, H. Pourbabak, Z. Liang, and W. Su, ”An Integrated eVoucher Mechanism for Flexible Loads in Real-Time Retail Electricity Market”, IEEE Access, vol. 5, pp. 2101-2110, 2017.

Moreover, in the smart grid, more and more customers will be able to have local generation capability, i.e., distributed generation (DG), along with various flexible controllable loads, such as thermostatically controlled load (TCL), distributed energy storage device (DESD) and washing machine [19][20]. Plug-in electric vehicles (PEVs) are also appealing as the most controllable loads because they can be curtailed for significant periods of time (e.g., several hours) without impact on end-use function [21][22]. Sooner or later, PEVs will shift the traditional energy demand from crude fossil energy to electricity for the personal transportation sector and also heavily impact power system operation [96]. In the future, through aggregators or parking lots, PEVs will play a much more proactive role in the retail electricity market as very flexible loads with the capability to provide ancillary services, such as power system frequency support [97], and participate in many DSM programs [98][99].

In this paper,In this paper, we aim to study how incentive-based demand response program can benefit both utility company and energy end-users. We proposed an integrated eVoucher mechanism for customer participation in the real-time retail electricity market based on a new designed voluntary demand response program. This integrated eVoucher mechanism also has the potential to support reducing the frequency-deviation in power system, with the idea that demand response can be utilized as a load control tool to realize reliability goals [100]. Conventional demand response programs are easily implemented and straightforward for customers understanding, however typically assume a predefined price scheme or mandatory instruction acceptance after a customer has signed a contract. In addition, most DSM programs are not implemented in a real-time manner, which depends heavily on day-ahead forecasting. However, both the real-time price in the wholesale electricity market and the loads at the distribution level are volatile and unpredictable in a long time interval [101]. Retail customers are also usually insensitive to pre-defined multi-phase price schemes (e.g. time-of-use (TOU) price) or distribution locational marginal price (DLMP), and not used to making day-ahead decisions [102]. In this way, the electricity retailer (e.g. EDC) may be exposed to the risks involved in the difference between the real-time wholesale electricity market and the retail electricity market. Some economic mechanisms are needed to deal with such risks and improve market diversity. In [102] and [103], the authors propose a coupon-incentive-demand-response program to solve the problem of the high risk of the wholesale electricity market. However, they did not consider the demand adjustment in an increasing way and required a large amount of aggregated load demand to bid directly in the wholesale market, which may make small customer participation hard. Some works [104] also estimate the impact of different distribution tariff structures on residential customers, which even include a power related component to reduce the peak load. Other works use dynamic subsidy [92], customer reward [105] and monetary incentives [106] to better schedule the demand response or load management on top of the usual electricity price schemes. The authors in [107] design a non-discriminatory individualized demand aware price scheme to optimize the economic benefit for every customer, and claim to reduce the ”rebound peak”. In contrast, our proposed eVoucher consists of several different components that target both engineering and economic advantages. More importantly, the implementation of the eVoucher is based on a voluntary principle and can also avoid the effect of rebound peak through managing the individualized incentives. Additionally, it can include the frequency-deviation triggered component in eVoucher price signal to benefit the power system operation. Some authors, such as in [108][109] have tried some demand side management frameworks to promote the efficient frequency control and make demand response from customers contribute more to the system stability. Similar to [110], the embedded frequency-deviation triggered component in the eVoucher program is reflected through a dynamic pricing process and, as an engineering signal, coupled with economic incentive. We study this problem mainly because the conventional demand response programs have aforementioned disadvantages compared with the new incentive-based demand response program, and the concept of energy coupon should be fully explored.

In this paper, our contribution is to: (1) explore the potential coordination among DSO, EDC and other participants in the retail electricity market; (2) propose an eVoucher mechanism to provide opportunities for all customers at distribution levels to actively participate and improve economic efficiency; (3) verify various applications of the proposed eVoucher program in different scenarios from both engineering and economic perspectives; (4) discuss the extension of the high-level framework for more business models in an efficient and fair retail market.

The remainder of this paper is organized as follows: Section II introduces the basic features of the eVoucher program and its components. Section III gives the mathematical description of the eVoucher mechanism for several parking lots with a high penetration of electric vehicles. Section IV presents several case studies for different applications. Section V summarizes the major findings of this paper.

2.2 eVoucher program

In the electricity retail market, energy users usually passively accept offers from energy providers without much economic consideration or negotiation. We hope to design a kind of energy token similar to a *Voucher* or *Coupon* used in daily commercial activities, which aims to guide the consumption behavior by economic incentives. The energy token, named after *eVoucher*, can encourage energy customers at the distribution level to actively participate in the energy transaction and load management based on discount or reward. The eVoucher based customer participation program for parking lot owners is different from

Table 2.1: The Innovation and Uniqueness of the eVoucher Program

	Price-based DR	Incentive-based DR	The Proposed eVoucher Program
Examples	Time-of-Use (TOU) Real-time Pricing (RTP) Critical Peak Pricing (CPP)	Emergency Demand Response Program (EDRP) Direct Load Control (DLC)	New Concept Energy Users and Coalitions (consumers and prosumers), EDC, and DSO
Participants	Consumers; EDC	Consumers; EDC	Active energy users and energy community, EDC, and DSO
Flexibility	Mandatory	Mandatory after signing contract	Voluntary at any time and any location
Control Strategy	EDC reshapes load using pre-defined price over a long time scale (hours)	EDC applies the direct load control at peak periods (minutes)	Self-guided control of DG, DESD, dispatchable loads through near-real-time eVoucher negotiation (seconds-minutes)
Dynamics	Pre-defined electricity price	Pre-defined rate and one-time incentive	Reflect the temporal and spatial values of electricity in a competitive market

conventional demand response (DR) programs that are either incentive or price based. It can reflect more dynamic features from both market and system operation perspectives. A comparison between the eVoucher and other kind of demand response programs is presented in Table 2.1. The features of the eVoucher are shown in Figure 2.1 and summarized as follows:

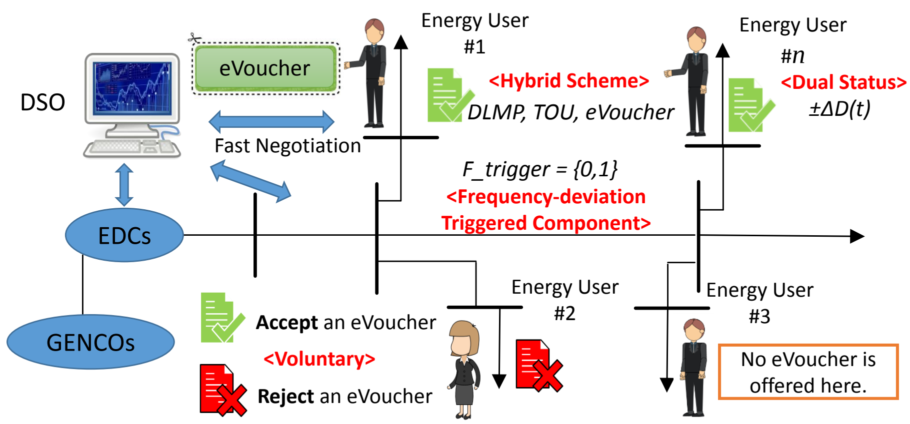


Figure 2.1: The eVoucher program features

Voluntary: Customers can freely decide whether and when to accept the eVoucher.

There are no specific contracts for customers to obey the instructions or accept the predefined price scheme mandatorily. The customers can even take into account their own attitude toward the extra financial benefit. Customers may reject the eVoucher without any penalty. For example, a savings of $10 per month may not seem like a significant gain for a relatively wealthy customer; however, a poor customer might view this amount as a highly significant reduction. Clearly, the objective measure of $10 can be viewed differently by different customers. Additionally, DSO facilitates a fast negotiation between customers and EDC, which is simply not possible in the existing DR programs.

Frequency-deviation event triggered: The integrated eVoucher program does not advocate real-time frequency-based price that may expose customers to the risk of price volatility. Instead, it will use the grid frequency deviation as an event-trigged (e.g., a preset deadband frequency threshold) and location-based signal. Customers can read this signal locally and respond to the needs of EDC directly (e.g., providing frequency regulation as an ancillary service). With minimal communication between customers and DSO, the frequency- deviation-based signal is able to incentivize customers to limit frequency deviation at local levels, ultimately creating a more stable and reliable grid as a whole while allowing increased customer participation.

Dual status: While traditional DR compensates end-users for reducing their electricity use (load) or shifting the peak periods, the proposed eVoucher program provides customers with temporal-spatial-dependent incentives for changing power in the positive or negative direction. In other words, some customers may increase their demand if needed.

Hybrid scheme: It is also worth noting that either the flat rate or the day-ahead price (e.g. TOU) can still serve as a baseline price and reflect the regular retail price. In this way, the eVoucher program can provide real-time-like flexibility with a hybrid electricity price scheme, and meanwhile keep the price clearing process easy and the system operation efficient.

In our assumption, the economic token named after the *eVoucher*, as a signal, can be broadcast with the help of DSO to energy users. However, the benefit is mainly shared by the EDC and energy users (e.g. parking lots), since the DSO is non-profit seeking. Sometimes, an EDC may also sacrifice some economic benefit by adjusting the eVoucher price to achieve the goal of maintaining system reliability.

2.3 Problem formulation

The economic benefits achieved through the eVoucher program are considered as follows for both EDC and parking lots. In the whole system, DSO is assumed to be similar to an independent system operator (ISO) at the transmission level with non-profit characteristics, playing a role as a third party to coordinate the system integration. Thus, we will model the two profit-seeking entities involved in the negotiation process of energy consumption. The variables used are defined in the following table 2.2. Figure 5.2 shows how parking lots participate in the real-time retail electricity market with EVs as flexible loads. It is noteworthy that the eVoucher mechanism can generally be utilized for the management of various types of flexible loads. Here, EV is used as a typical example to demonstrate some basic dynamic features of the integrated eVoucher program.

2.3.1 EDC

The EDC participates in bidding and the day-ahead energy commitment in the wholesale market and purchases a few part in the real-time wholesale balance market, possibly with some forecasting capability. It then resells energy to customers or end-users in the retail electricity market. Due to the highly stochastic and dynamic real-time pricing scheme for power balance in the wholesale electricity market, the EDC has to face some unpredictable risks when a real-time price spike appears. On the other hand, it also has some

Table 2.2: Parameters of different parking lots

D_{RT}	Real time power demand of EDC
D_{DA}	Day-ahead power commitment of EDC
$D_{i,0}$	Original power demand for ith parking lot
ΔD_i	Demand adjustment from ith parking lot
Rev	Revenue for EDC
Rev_i	Revenue from ith parking lot
π_{RTP}	Real-time energy price
π_{DA}	Day-ahead energy price
$\pi_{TOU,i}$	Time-of-use energy price for ith parking lot
π_{PEN}	Penalty for day-ahead energy commitment
$\pi_{V,i}$	eVoucher price assigned for ith parking lot
$\pi_{r,i}$	Rebate rate in ith parking lot
$SoC_{i,j}$	State-of-charge of jth EV in ith parking lot
$R_{pa,i}$	Parking revenue for ith parking lot
$R_{ch,i}$	Charging revenue for ith parking lot
$r_{pa,i}$	Parking rate in ith parking lot
$r_{ch,i}$	Charging rate in ith parking lot
$C_{op,i}$	Operation cost for ith parking lot
$T_{i,j}^{pa}$	Parking time for jth EV in ith parking lot
$T_{i,j}^{out}$	Departure time for jth EV in ith parking lot
$T_{i,j}^{in}$	Arrival time for jth EV in ith parking lot
$P_{i,j}(t)$	Charging power for jth EV in ith parking lot at time t
P_{max}	Maximum charging power
$\mu_{i,j}(t)$	Binary variable to control charging status
$B_{i,j}$	Battery capacity for jth EV in ith parking lot
T	The number of time intervals

responsibility to maintain the system reliability of the distribution network, especially the nominal frequency in the power system. In this way, the EDC has strong motivation to accepted some well-designed incentive mechanism to deal with these situations. Parking lots with a high penetration of EVs, as a typical kind of flexible loads, happen to be able to provide the control flexibility necessary for an EDC to consider utilizing such demand response capacity with our proposed eVoucher program. According to different day-ahead power commitment and real-time price scenarios, the EDC can decide the eVoucher price as an incentive to induce the voluntary demand adjustment in the retail electricity market.

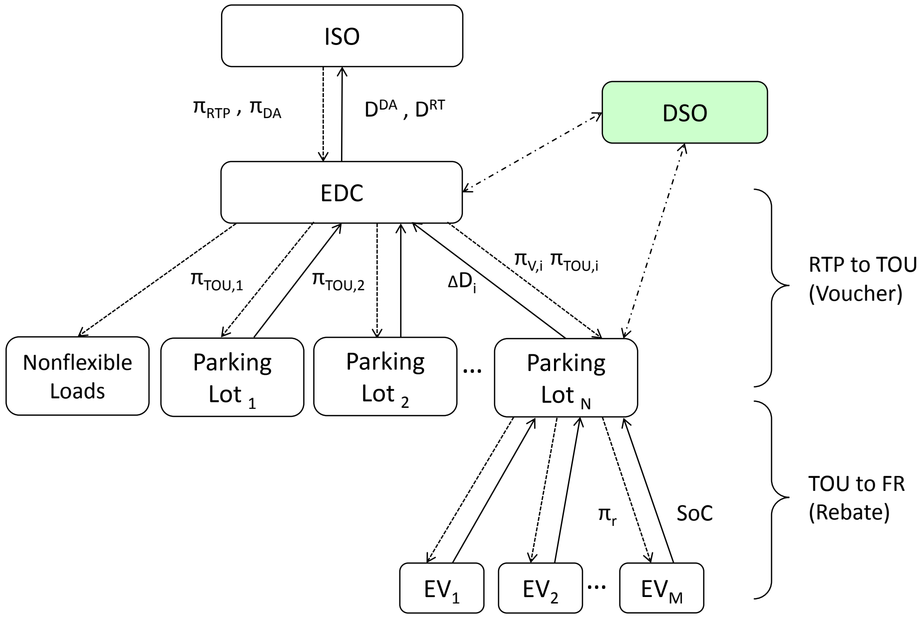


Figure 2.2: Retail electricity market structure for EV parking lots

$D_{RT} > D_{DA}$

In scenario $D_{RT} > D_{DA}$, the EDC can increase its profit or reduce its economic loss by assigning eVoucher price $\pi_{V,i} \in \{0, \pi_{RTP} - \pi_{TOU,i}\}$, as economic incentive to encourage parking lots to adjust their energy demand $D_{i,0} \pm \Delta D_i$, according to different real-time prices in (2.2).

$$Rev = \sum_{i=1}^{N} Rev_i - \sum_{i=1}^{N} \Delta D_i \pi_{V,i} \tag{2.1}$$

$$Rev_i = \begin{cases} \left(D_{i,0} - \Delta D_i\right)\left(\pi_{TOU,i} - \pi_{RTP}\right) \\ \qquad \pi_{RTP} > \pi_{TOU,i} \\ \left(D_{i,0} + \Delta D_i\right)\left(\pi_{TOU,i} - \pi_{RTP}\right) \\ \qquad \pi_{RTP} < \pi_{TOU,i} \end{cases} \tag{2.2}$$

$$0 < \pi_{V,i} < \pi_{RTP} - \pi_{TOU,i} \tag{2.3}$$

$D_{RT} < D_{DA}$

In scenario $D_{RT} < D_{DA}$, the real-time demand is less than the day-ahead energy commitment, therefore EDC should consider economic effect of penalty (π_{PEN}) in wholesale electricity market when making decisions about negotiated eVoucher price. It can only increase its profit by assigning eVoucher price $\pi_{V,i} \in \{0, \pi_{TOU,i} - \pi_{DA} + \pi_{PEN}\}$, as economic incentive to encourage parking lots to adjust their energy demand, $D_{i,0} + \Delta D_i$, as in (2.5).

$$Rev = \sum_{i=1}^{N} Rev_i - \sum_{i=1}^{N} \Delta D_i \pi_{V,i} \tag{2.4}$$

$$Rev_i = (D_{i,0} + \Delta D_i)\left(\pi_{TOU,i} - \pi_{RTP}\right) \tag{2.5}$$

$$0 < \pi_{V,i} < \pi_{TOU,i} - \pi_{DA} + \pi_{PEN} \tag{2.6}$$

The eVoucher price should be within a certain range according to cost-revenue consideration to satisfy the requirement of maximizing the EDC's economic benefit. However, the specific value of the eVoucher accepted by both sides (i.e., EDC and parking lots) can be determined and negotiated based on different conditions. These conditions will be discussed later in this section.

2.3.2 Parking lot

Similar to electric vehicle charging operators [111], every parking lot aims to maximize its own revenue through controlling the EV charging load and responding to the eVoucher price signal from the EDC. A parking lot also provides a rebate $\pi_{r,i}$, similar to [112], to discount EV parking time while an EV is under charging service. The following optimization model (2.7)-(2.14) is solved in every negotiation process for the rest of the time intervals to satisfy EV charging requirements.

$$\max \quad R_{pa,i} + R_{ch,i} + \pi_{V,i}\Delta D_i - C_{op,i} \tag{2.7}$$

$s.t.$

$$R_{pa,i} = r_{pa,i} \sum_{j=1}^{M} T_{i,j}^{pa} \tag{2.8}$$

$$\sum_{j=1}^{M} T_{i,j}^{pa} = \sum_{j=1}^{M} \left(T_{i,j}^{out} - T_{i,j}^{in}\right) - \pi_{r,i}\Delta t \sum_{j=1}^{M} \sum_{t=t_0}^{t_0+T} \mu_{i,j}(t) \tag{2.9}$$

$$R_{ch,i} = \sum_{j=1}^{M} \sum_{t=t_0}^{t_0+T} \left[r_{ch,i} - \pi_{TOU}(t)\right] P_{i,j}(t) \tag{2.10}$$

$$SoC_{i,j}\left(T_{i,j}^{in}\right) + \sum_{t=t_0}^{t_0+T} \frac{P_{i,j}(t)\Delta t}{B_{i,j}} = 1 \tag{2.11}$$

$$0 \le P_{i,j}(t) \le P_{max}\mu_{i,j}(t) \tag{2.12}$$

$$P_{i,j}(t) = 0, \quad \forall t \notin \left[T_{i,j}^{in}, T_{i,j}^{out}\right] \tag{2.13}$$

$$\sum_{j=1}^{M} P_{i,j}(t) \le D_{i,0} - \Delta D_i(t_0), \quad \forall t \in \left[t_0, t_0 + T\right] \tag{2.14}$$

The eVoucher price signal $\pi_{V,i}$ in (2.7) will seduce and compensate the adjusted demand ΔD_i in the current time interval. Meanwhile, a parking lot that accepts such eVoucher compensation must guarantee all the state-of-charge (SoC) statuses (i.e. the percentage of charged battery capacity) are satisfied before the EVs' departure.

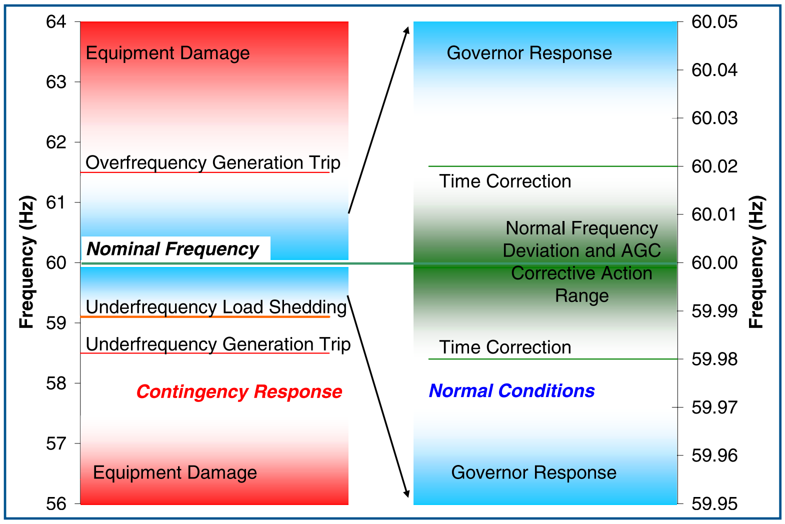


Figure 2.3: Normal and abnormal frequency ranges

2.3.3 Frequency-deviation-based price component

In the supply-side paradigm, many generators nowadays are equipped with frequency responsive governors that produce an output change proportional to the frequency deviation. If system frequency deviates sufficiently far from its setpoint (e.g., 35 mHz in Figure 2.3 [113]), governor response is activated to prevent further growth of the deviation.

There are also some early stage efforts to support maintaining frequency with a frequency-responsive load by providing the equivalent of generator droop [114] [115] in the demand side. Furthermore, one of the most recent interesting works in [110] designs a pricing mechanism for power system frequency and utilizes economic incentive to encourage demand adjustment for frequency-deviation or potential contingency. In [110], the assumed electricity price changes continuously with the frequency and requires frequent responsive load control. However, here we just consider the very high frequency-deviation as a triggered event to modify the corresponding eVoucher price and focus on demand adjustment as potential protective actions. In addition, the triggering cannot be too sensitive and must follow the detection band when the frequency goes above or below some threshold value

[116]. In other words, we do not only consider the frequency deviation f, but also the evolution over time of f. The f-τ characteristics of frequency-deviation triggering are presented in Figure 2.4. A detection method similar to the one in [108] is used.

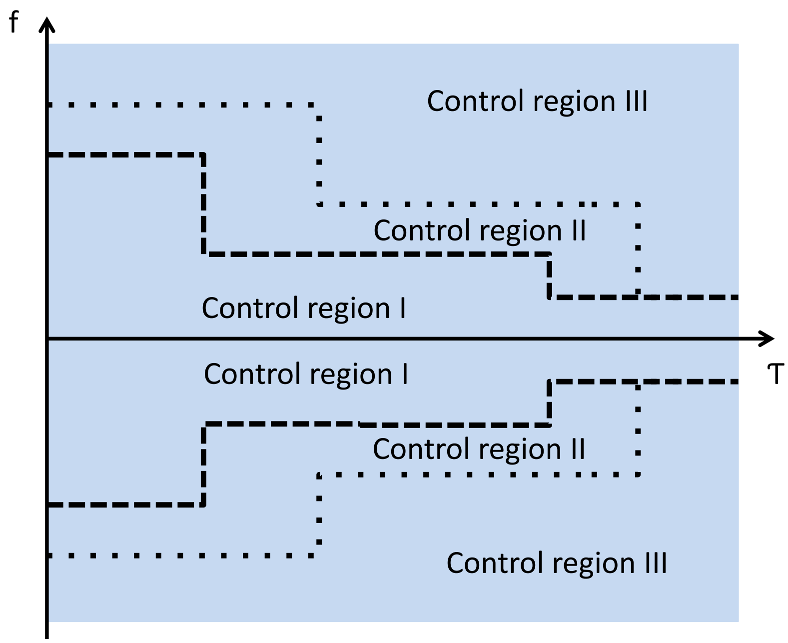


Figure 2.4: f-τ characteristics of frequency-deviation triggering

The different control regions in Figure 2.4 represent different triggering sensitivities according to the eVoucher implementation in different scenarios. Once the frequency deviation goes out of range n_f times , where n_f is larger than threshold value N_{th}, a binary variable $\xi \in \{0,1\}$ is set to assign $\xi \times \frac{\sqrt{n_f}}{\beta}$ additional value to the eVoucher price. Finally, the frequency deviation will be reflected as economic incentive to facilitate load adjustment.

2.3.4 eVoucher price

The status and price decision process of the eVoucher program can be summarized generally as in Figure 2.5 and described in detail as follows:

Step 1: EDC forecasts a potential price spike $\pi_{RT}(t)$ in the wholesale market and receives a frequency-deviation triggered signal through an informational message sent by

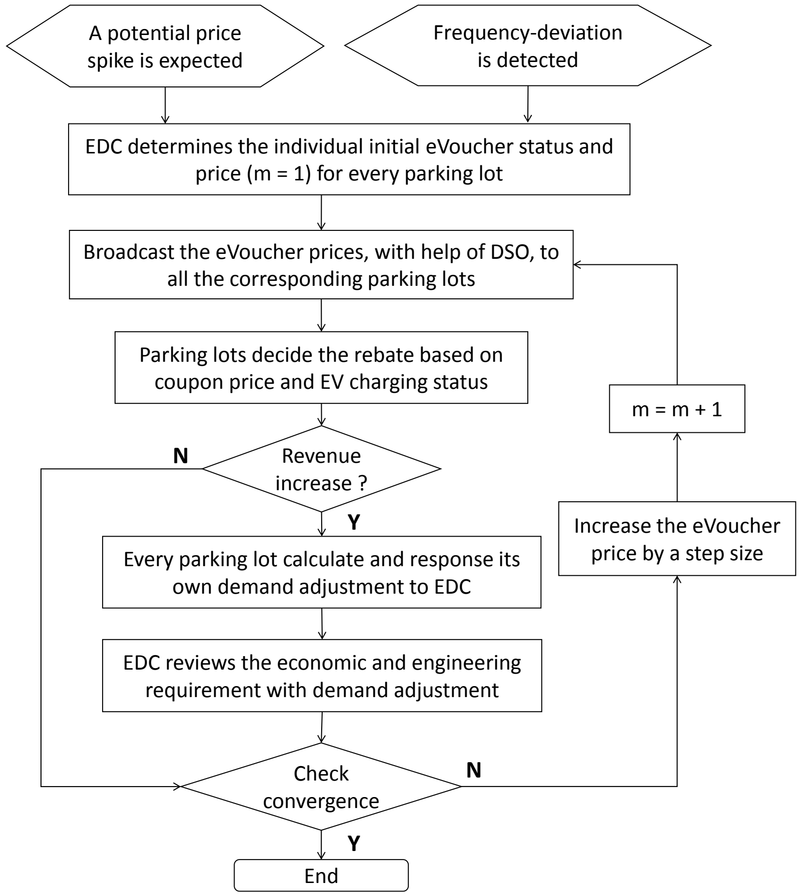


Figure 2.5: Decision process of eVoucher price

the network operation monitoring system.

Step 2: Both the price gap $\pi_{RT}(t) - \pi_{TOU}(t)$ and the severity of the frequency-deviation are estimated, and one of two modes (i.e., demand increase or decrease) of the eVoucher will be determined to be implemented in the next step. In some rare extreme situation, the consideration for system reliability may dominate the economic benefits.

Step 3: Based on the chosen mode and locational information (e.g., number of EVs) of every parking lot, the EDC will set the eVoucher status, $flag = 1$ *when* $\Delta D(t) > 0$ or $flag = -1$ *when* $\Delta D(t) < 0$, as well as the initial individual eVoucher price, $\pi^{m}_{V,i}(t)$ with $m = 1$.

Step 4: The required demand adjustment direction, $flag \in \{-1, 1\}$, and individualized eVoucher price information for the current time interval will be broadcast to each parking lot with the help of DSO.

Step 5: After receiving the eVoucher information from the DSO, every parking lot will determine whether to participate or not according to the EV charging status (e.g., SoC). Each of them estimates their accepted demand adjustment, $\Delta D_i^m(t)$, and solves an optimization problem (2.7)-(2.14). Then, the available decreased or increased demand will be reported back to the EDC.

Step 6: The EDC decides whether or not to increase the eVoucher price for each parking lot in the next negotiation iteration, according to its satisfaction and the total demand response from all the parking lots. If the EDC still would like to encourage more demand adjustment and the parking lots still have motivation to increase their revenue, the eVoucher price will be increased by a fixed step size, $\pi_{V,i}^{m+1}(t) = \pi_{V,i}^{m}(t) + k^m \times \pi_{V,i}^{m}(t)$. Then, *Step 3* is repeated with $m = m+1$, and the negotiation process continues.

Step 7: The DSO checks the convergence based on a limited number of negotiation rounds and revenue increase (loss reduction), making sure the EDC and parking lots are no longer negotiating after the convergence. The EDC finishes the implementation of the eVoucher program in the current time interval and prepares for the next time interval, $t = t+1$.

2.3.5 Implementation issues

In order to implement the eVoucher program, it was assumed that every targeted parking lot is able to be controlled automatically using an energy management system (EMS), which is similar to the home energy management system used by most residential customers [117]. As long as the economic preference setting (e.g., willingness or sensitivity level) is determined beforehand while participating in the eVoucher program, the EMS will take over the specific response process without referring to the decision maker every time. In short, the proposed mechanism is likely to be more effective in cases where there is an intelligent energy management system available, which can automatically respond to external signals and make decisions, rather than depending on an indirect or manual user response to the price signal [107]. The communication hardware cost of implementing such a program is expected to be low and without too much hardware modification, since the maximum number of iterations and the number of variables in the negotiation process are limited to only a few in a single time interval. Since it is a voluntary demand response program, some target customers could be selected according to their daily consumption flexibility reflected in their historical records. Meanwhile, some energy aggregators, like parking lot operators, will be motivated to participate if the increased economic profit are considerable. For residential customers, some potential lottery-based incentives can also be designed to encourage the small energy users to participate in the promoted demand response program with low-probability but high-gain bonus. With increasing awareness of energy saving and possible contribution to non-dispatchable renewable energy, it is believed that more and more customers would like to actively participate in the carefully designed demand response programs.

2.4 Numerical case studies

The demonstrated system contains four parking lots at different locations with different EV population sizes and various parameters. They charge different service fees and have different electricity pricing schemes, according to the locational traffic flow information and regional distribution network conditions. The owners of these parking lots try to increase their own profit by responding to the eVoucher price signal. We use 15 *min* as the time interval. The parameters of the parking lots are given in Table 2.3. Their location related TOU pricing schemes serve as a baseline price as shown in Figure 3.5. In the following test cases, we will show that the eVoucher price can be combined with TOU to create a dynamic hybrid pricing scheme to benefit all the participants in the retail electricity market.

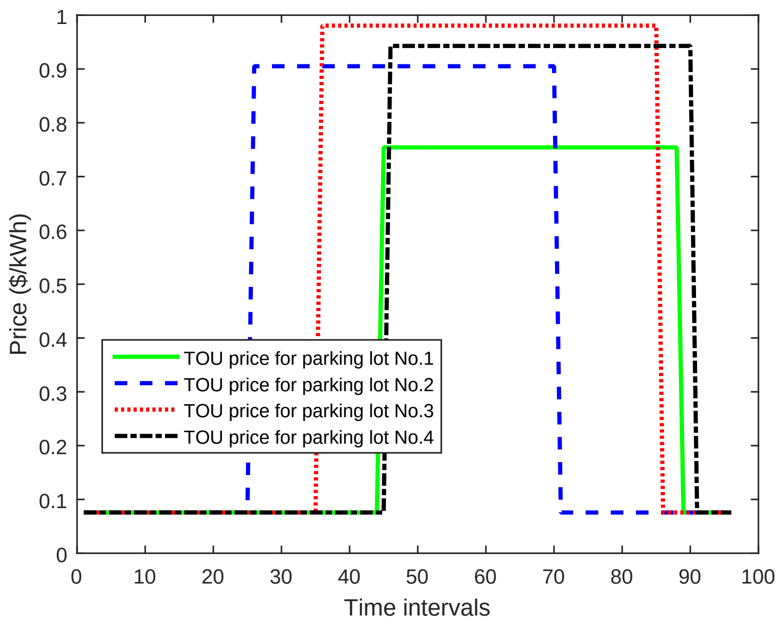


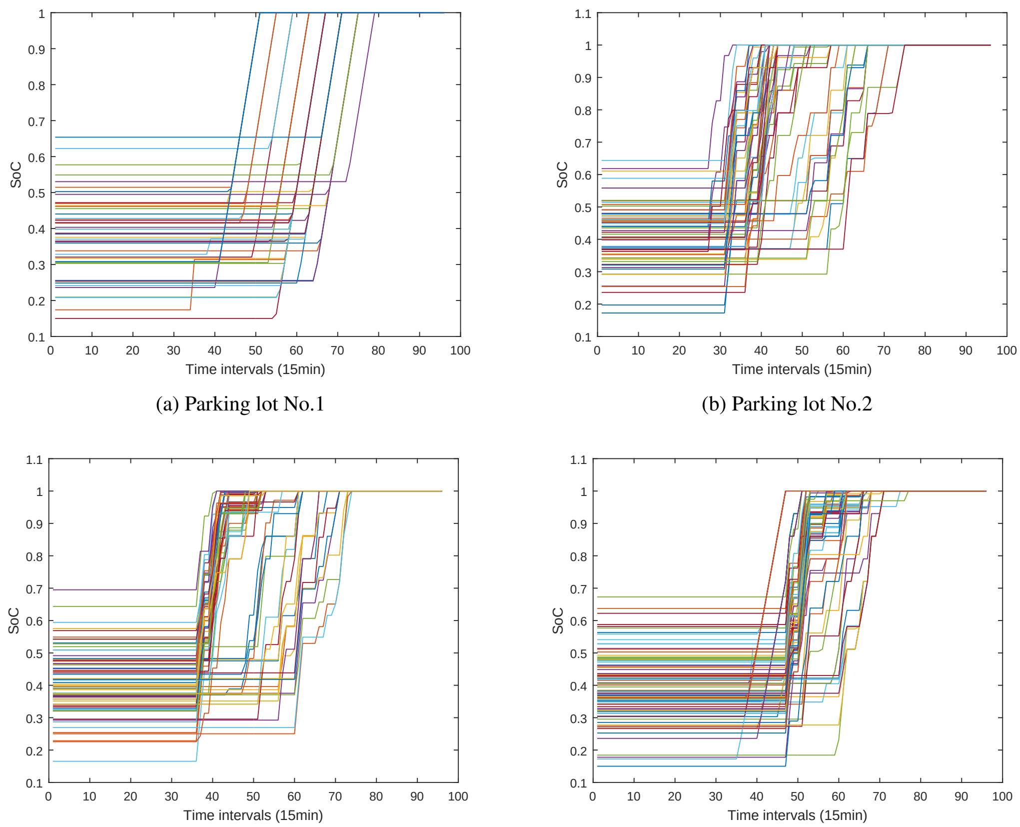
Figure 2.6: Time-of-use price scheme for different parking lots

Table 2.3: Parameters of different parking lots

#Parking lot	No.1	No.2	No.3	No.4
Number of EVs	50	70	80	100
Charging fee	1.80 $/h	1.60 $/h	2.00 $/h	1.70 $/h
Parking fee	1.20 $/h	1.50 $/h	1.80 $/h	1.60 $/h
Parking rebate	0.40	0.35	0.45	0.40

2.4.1 eVoucher program

With the integrated eVoucher program, all four of these parking lots decrease or increase their demand when the EDC experiences a real-time price spike in the wholesale market, there is a huge gap between the real-time load and the day-ahead committed power consumption, or some serious frequency-deviation events are detected. Meanwhile, the charging requirements of any parking EV should always be satisfied, as shown in 2.7, no matter how parking lots negotiate with EDC. The real-time price and mismatched day-ahead committed power are shown in Figure 2.8. Figure 2.9 shows the aggregated demand adjustment of all the four parking lots at different time intervals.



(a) Parking lot No.1

(b) Parking lot No.2

(c) Parking lot No.3

(d) Parking lot No.4

Figure 2.7: SoC status of all the EVs in different parking lots

It is interesting to observe the real-time price and corresponding aggregated demand adjustment during time intervals 60-70. An extremely high real-time price appears in this range, and, due to a response induced by the negotiated eVoucher, the aggregated demand is able to decrease significantly (the positive direction is defined as reduction in Figure 2.9) due to the flexible charging loads of EVs. However, during time intervals 80-90, when another price spike appears, the eVoucher negotiation process will terminate quickly since few EVs are parking there to provide demand response capability.

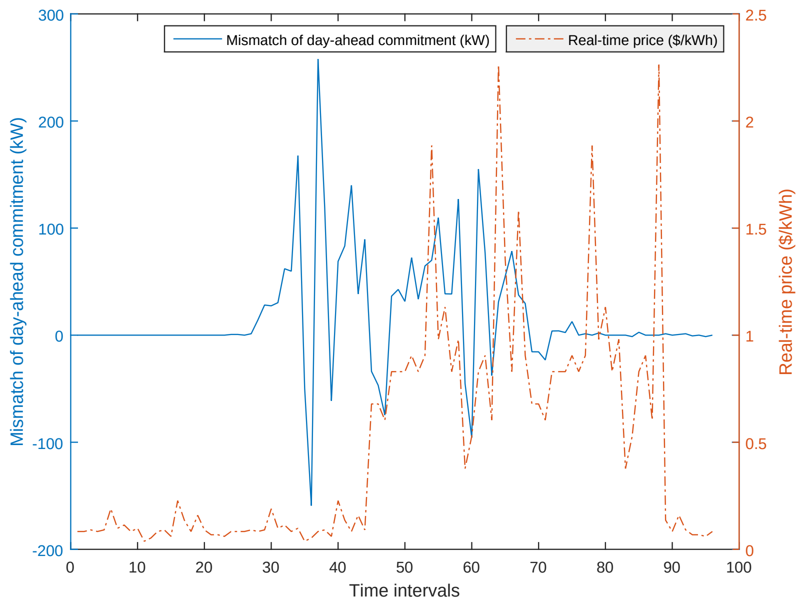


Figure 2.8: Real time price spike and mismatch of day-ahead power commitment

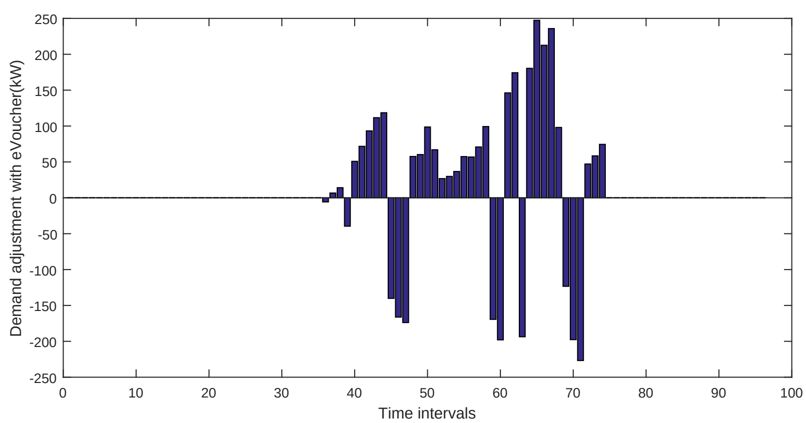


Figure 2.9: Aggregated demand adjustment from all the parking lots

2.4.2 Frequency-deviation triggered component in eVoucher

The integrated eVoucher program considers frequency deviation in its operation, which makes customers unconsciously contribute to improving the system reliability based on economic incentives. However, due to the uncertainty of demand response rate, automatic generation control (AGC) still works as a dominant tool for frequency regulation by frequent adjustments to the output of generators and change in the load [118]. The eVoucher program brings extra adjustment flexibility on top of AGC. Based on the method discussed in section 3.3, we use Figure 2.10 as a detection window for frequency deviation triggering. The fact that it is the frequency deviation evolution rather than individual anomaly that can affect the final eVoucher price, should be emphasized. We chose $N_f = 50$, $\beta = 100$, and used modified *FNET* frequency data [119], obtained through CURENT, the University of Tennessee, Knoxville, and Oak Ridge National Laboratory, to show the application of the eVhoucher program.

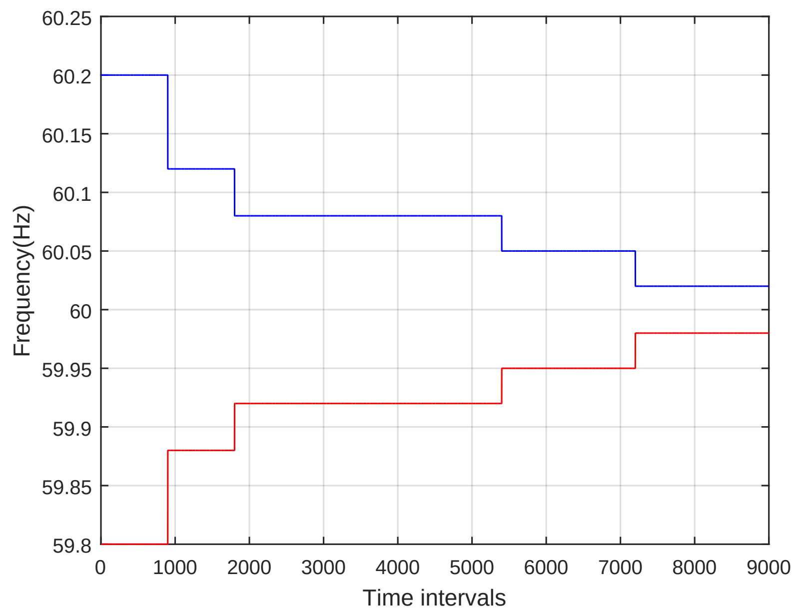


Figure 2.10: Detection window for frequency deviation triggering

In Figure 2.11, some under-frequency can be observed during roughly the time intervals $4.9 \times 10^5 \sim 6.1 \times 10^5$ in $0.1s$. If we use $1s$ moving average data to detect the frequency deviation and trigger the corresponding eVoucher price component, some very high eVoucher price signal after the final negotiation can be observed from 50 to 70 time intervals in $15min$ for parking lot No.1 (Figure 2.12) and all the other parking lots. In other words, there will be more economic incentive caused by frequency-deviation to encourage demand response from rational customers to help system operation. However, as

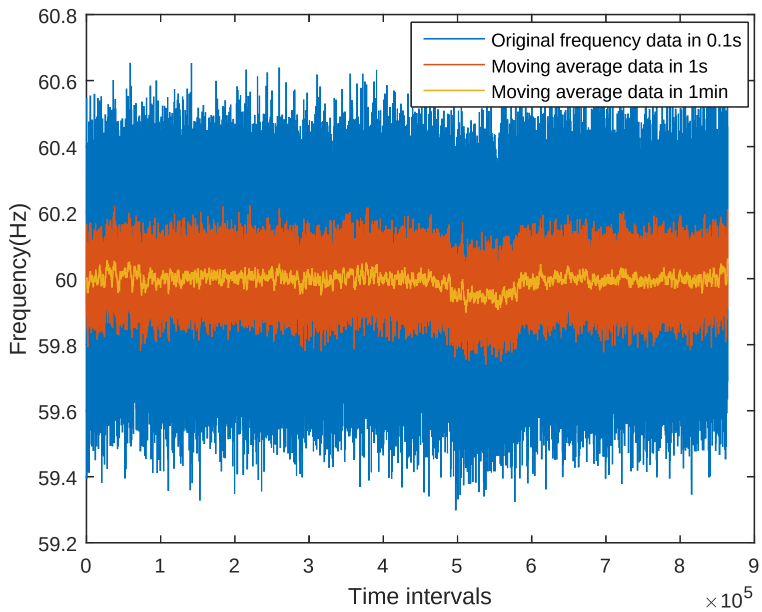


Figure 2.11: One day modified FNET frequency data

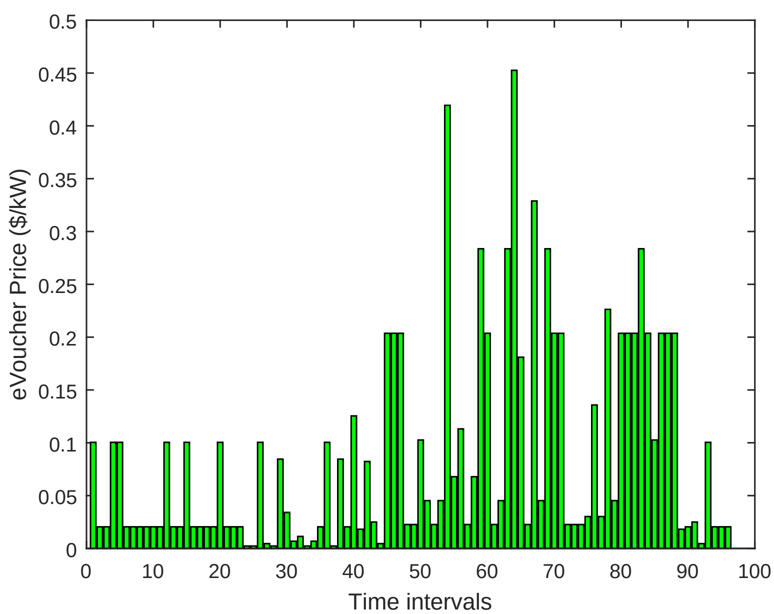


Figure 2.12: Final negotiated eVoucher price for parking lot No.1

shown in Figure 2.9, the eventual load adjustment is usually not just in proportion to the eVoucher price, since both the economic incentive (e.g., TOU, eVoucher price) and the physical constraint (e.g., SoC, congestion) should be considered.

2.4.3 Economic analysis

The integrated eVoucher program will entail reducing loss or improving the part of the revenue associated with available flexible loads for the EDC when it meets some real-time price spikes like in Figure 2.8. It should be pointed out that actually both very high real-time price spikes and frequency deviation can be considered some type of extreme events, which implies the eVoucher program can also be used for many other extreme events, such as natural disasters. As long as the cost of the extreme events can be quantified and amortized with economic analysis, the eVoucher mechanism will be utilized to benefit both market and system participants.

Table 2.4: Economic benefit of EDC and parking lots

Participants	Without eVoucher	With eVoucher
EDC	\$1087.6	\$3174.8
PL No.1	\$1988.5	\$2067.7
PL No.2	\$3107.4	\$3189.2
PL No.3	\$4055.3	\$4099.6
PL No.4	\$6154.8	\$6206.3

In Table 2.4, it is observed that all the participants in the retail market benefit themselves by improving the revenues. One important phenomenon is that the EDC improves its revenue dramatically compared with that of the parking lots because only the EDC is exposed to volatile real-time price as its main risk source. In contrast, the buying price in the retail electricity market is more stable and just a small part of the whole operation cost for parking lots.

2.5 Conclusion

This paper proposes an integrated eVoucher mechanism to encourage typical flexible loads, such as parking lots with a high penetration of electric vehicles, to participate in real-time retail electricity market. This eVoucher program can work for various scenarios involving economic or physical extreme events (e.g. frequency-deviation). The ultimate goal is to explore the possibility of introducing a more flexible hybrid price scheme and operation reliability into the redefined distribution network, meanwhile evaluating the potential of innovative business models in the power industry. In future work, we may combine the eVoucher program with some behavioral economic models, like prospect theory, to analyze more characteristics of human-in-the-loop energy systems.

CHAPTER III

Utility service

3.1 Introduction

[1] Smart grids have been revolutionizing electrical generation and consumption through a two-way flow of power and information. As an important information source from the demand side, advanced metering infrastructure (AMI) has gained increasing popularity worldwide. For example, in Nordic countries, the Finnish government passed a new act, which states that at least 80% of the customers of each distribution system operator (DSO) must have a smart meter by December 2013, and nowadays in 2017 almost every customer (98%) in Finland is supplied with a smart meter [121]. The abundant data set of electricity consumption of residential customers enables accurate load profiling and data analytic application [85][86]. Usually, the load profiles refer to electricity consumption behaviors of customers over a specific period, e.g., one day, and can help utility companies understand how electricity is actually used for different customers and obtain the load patterns to provide better customized service.

Traditionally, the information or data set about an individual energy customer's load profile has been unavailable or incomplete. Consequently, research on retail electricity price design usually assumes that all energy customers have very similar electricity usage

[1]This chapter was previously published as an article in a peer-reviewed journal: [120] **T. Chen**, K. Qian, A. Mutanen, B. Schuller, P. Jarventausta, and W. Su, "Classification of Electricity Customer Groups Towards Individualized Price Scheme Design", 2017 North American Power Symposium (NAPS), Morgantown, WV, USA, September 17-19, 2017.

patterns [84]. Thus, the implemented retail price schemes are designed independent of energy customers' load profiles, even for some demand response (DR) projects. Nowadays, however, with the comprehensive data sets of individual load profiles having been made available, many researchers have found remarkable heterogeneity in energy customers' load profiles [122][123].

In this paper, we aim to study how energy customers are divided into different customer groups and provided different price schemes. Instead of focusing on the clustering of different load curves, we mainly focus on the design of individualized electricity price schemes, such as time-of-use (TOU) [44], for different types of residential electric customers based on their clustering results. Most similar works only consider extracting typical load profiles from clustering results, and do not study the further utilization of these obtained load profiles for end-users, such as customized pricing or individualized demand response program design. However, the clustering of load profiles still plays a vital role in the reasonable price design. So far, a large number of clustering techniques, including K-means [124], hierarchical clustering [125], self-organizing maps (SOM) [126] and support vector machines (SVM) [127], have already been widely applied in power systems. Most of these do not provide a concrete description of how to utilize the clustering results towards improving electricity services though. Here, we use the simple K-means method combined with a dimensionality reduction technique, principal component analysis (PCA), and a fast efficient supervised learning method, extreme learning machine (ELM) [128][129], to make the classification of load profiles more reasonable. Then, the achieved typical daily load profiles in every group can serve better for the design of an individualized electricity price scheme, with the help of the symbolic aggregate approximation (SAX) method. At the high level of a distribution network, the proposed method aims to provide a potential tool for price-based coordinated control and future DR programs in a

Smart Grid.

3.2 Framework and methods

The proposed price scheme design mechanism can be separated into two parts: achieving the typical daily load profiles of every assigned group after classification, and matching a suitable price structure to every typical daily load profile.

3.2.1 Classification

Data normalization

Data preparation including data cleaning is not the subject of this paper, and will not be discussed. In order to focus on the relative consumption level of specific energy customers and make the load profiles comparable, the normalization process transforms the AMI data, y_{ij}, as shown in (3.1).

$$y_{ij} = \frac{y_{ij}^{old} - y_{i,min}}{y_{i,max} - y_{i,min}}, \tag{3.1}$$

where, y_{ij}^{old} denotes the actual electricity consumption for customer i at time j, and $y_{i,max}$ and $y_{i,min}$ denote the minimum and maximum consumption over T periods, respectively.

Principal component analysis

PCA is one of the most widely used dimension reduction techniques available. It aims to find a small set of orthogonal variables with manageable reduced dimensionality. These principal components are actually linear combinations of original variables, which represent the variance of the original data set in a low dimensional subspace [130]. The purpose of utilizing PCA is to speed up the convergence of the following clustering algorithm and make the result more robust [131]. More specifically, lower dimensional data will rationalize the clustering of time series based on the Euclidean distance.

K-means

The aim of classification is to divide a set of objects into different groups such that objects in the same group are more similar to each other than to those in other groups. K-means, as the most widely used and easily implemented clustering algorithm, will divide the input data set into K groups by their similarity [132]. Consider a data set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ consisting of N independent input vectors with D-dimension. The goal of the algorithm is to partition the data set into K groups. In order to obtain those groups, a set of vectors μ_k, with D-dimensionallity and $k = 1, \ldots, K$, is introduced to indicate centers (centroids) of K clusters. In other words, an assignment of data points to clusters is found, along with a set of vectors $\{\mu_i\}$, to ensure that the sum of the squares of the distances of each data point to its closest vector μ_k is minimized as in (3.2). In this paper, $\mathbf{x}_i$ stands for the PCA components of normalized AMI measurements, and $\mathbf{y}_i$, for an input vector.

$$J = \sum_{j=1}^{N} \sum_{i=1}^{K} r_{ij} \parallel \mathbf{x}_j - \mu_i \parallel^2, \tag{3.2}$$

$$r_{ij} = \begin{cases} 1 & \text{if } i = \operatorname{argmin}_i \parallel \mathbf{x}_j - \mu_i \parallel^2 \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

Extreme learning machine

The ELM is a comparably novel learning technology for working with generalized single hidden layer feed-forward neural networks (SLFN) [133]. An SLFN usually includes three layers, which are the input layer, hidden layer, and output layer, as shown in Figure 3.1. Given a training data set with N samples, the output function of the SLFN with L hidden nodes and the activation function θ is as shown in (3.4).

$$f_L(x_j) = \sum_{i=1}^{L} \beta_i \theta \left(\omega_i x_j + b_i \right) = t_j, \quad j = 1, 2, \ldots, N \tag{3.4}$$

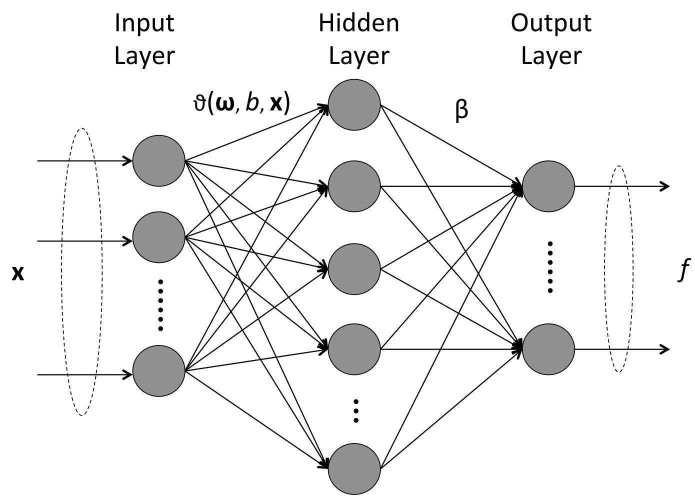


Figure 3.1: Structure of an SLFN

ELM distinguishes itself from other conventional iterative learning algorithms because it randomly selects the biases and input weights for hidden nodes, ω and b. Besides, it usually calculates the output weights, θ, analytically by finding a least-square solution. In [133] and [134], the authors theoretically prove that the training error are usually minimized with better generalization performance and higher accuracy. The reason of why ELM can be calculated quicky is because that it uses only single hidden layer and calculate the weighting terms analytically by satisfying ridge regress theory and neural network generalization theory [128].

3.2.2 Price Scheme design

Symbolic aggregate approximation

SAX mainly works as a powerful technique for the representation of time series data with lower bounding of the Euclidean distance [135]. Through the following two steps – transforming the load data into a piecewise aggregate approximation (PAA) representation and then symbolizing the PAA representation into a discrete string –, SAX can discretize a numeric time series into symbolic strings. As shown in (3.5), the intuitive idea of PAA is to use the mean values to represent the amplitude values that fall into the same time interval.

$$\bar{x}'_i = \frac{1}{k_i - k_{i-1}} \sum_{j=k_{i-1}+1}^{k_i} x'_j, \tag{3.5}$$

where j is the index of the normalized load data; i is the index of the transformed PAA load data; k_i is the ith time domain breakpoint; and $\bar{x}'_i$ is the average value of the ith segment [136]. In many applications, the averaging feature of the PAA can be utilized to smooth out short-duration, sudden and large 'spikes' of time series [85]. PAA has been proven to have all the pruning power of the Haar-based discrete wavelet transform (DWT) and can be defined with lower computation cost for arbitrary length queries [136].

Flowchart

The complete framework and method will follow the general principles of data analytics-type processing, including normalization, feature extraction (dimensionality reduction) and data post-processing of the clustering results.

Step 1: Pre-process the collected AMI data of regional energy customers, which includes removing of invalid data sets and normalizing of customers' daily load profiles.

Step 2: Implement the dimensionality reduction with the PCA technique to make daily load profiles more suitable and easier for classification.

Step 3: Cluster the PCA components of the analyzed daily load profiles into initial K typical groups of energy customers with the K-means classification algorithm.

Step 4: Check the clustering index and accuracy with ELM. If the training and testing accuracy are below a chosen threshold Th, the number of clusters will be decreased by $K' = K - N_c$, and *Step* 3 will be repeated again.

Step 5: Obtain the typical daily load profiles for every energy customer group by averaging the grouped daily load profiles based on the clustering index.

Step 6: Use SAX to assign symbols to the segmentations of the obtained typical daily load profiles in every customer group.

Step 7: Match the symbols associated with every typical daily load profile to suitable energy price levels.

Step 8: Analyze the economic effect and explore the different potential demand response programs for all the energy customer groups.

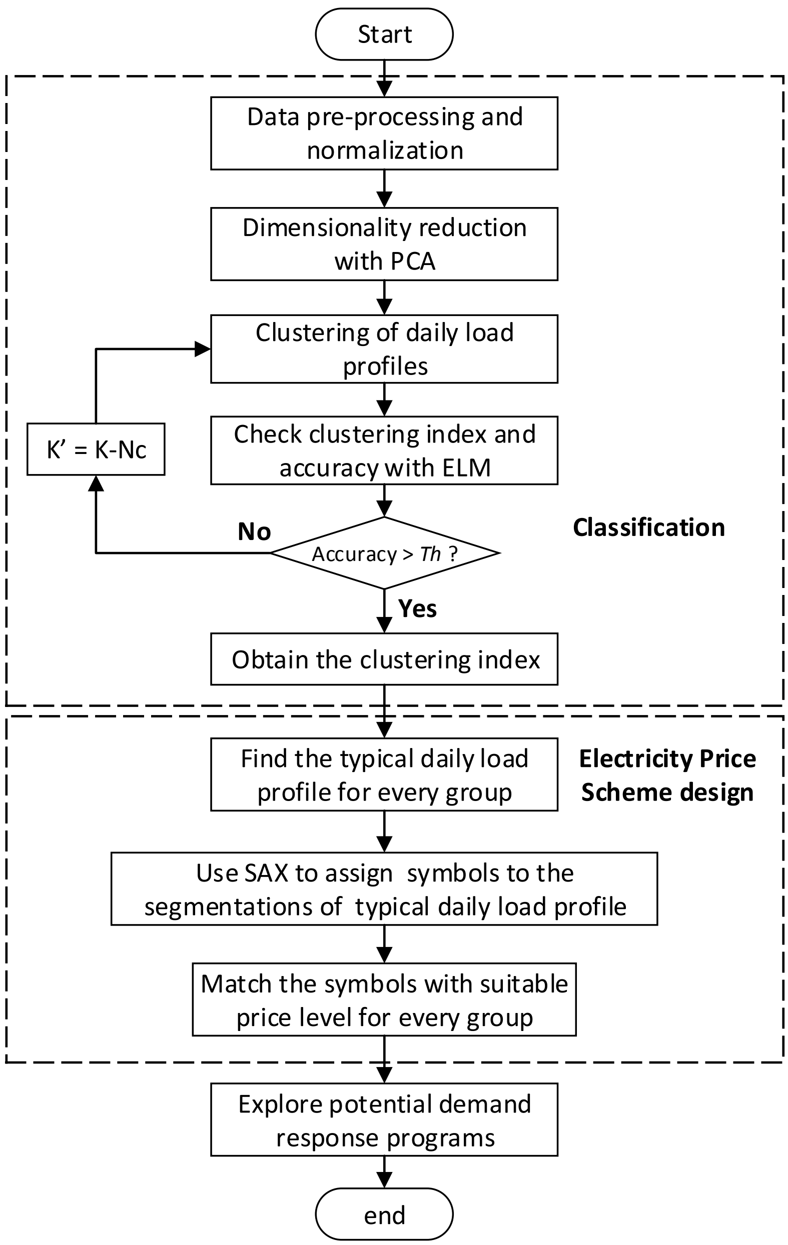


Figure 3.2: Flowchart of the price scheme design process

3.3 Results and discussion

The following test cases include an AMI data set collected from a realistic Finnish distribution system operator (DSO), which includes 3,398 non-empty low voltage customers in a small region. We randomly picked 1,500 customers from them, and chose several typical normal dates (without special national holidays) to demonstrate the proposed framework.

3.3.1 Individualized price scheme design

In the classification stage, 90% is chosen as a criteria for the explained variance in the PCA and Th. 16 energy customer groups are obtained (as shown in Figure 3.3) to stand for the typical energy consumption patterns extracted from the chosen 1,500 customers. In most groups, one or two peaks can be observed during a typical 24-hour time interval.

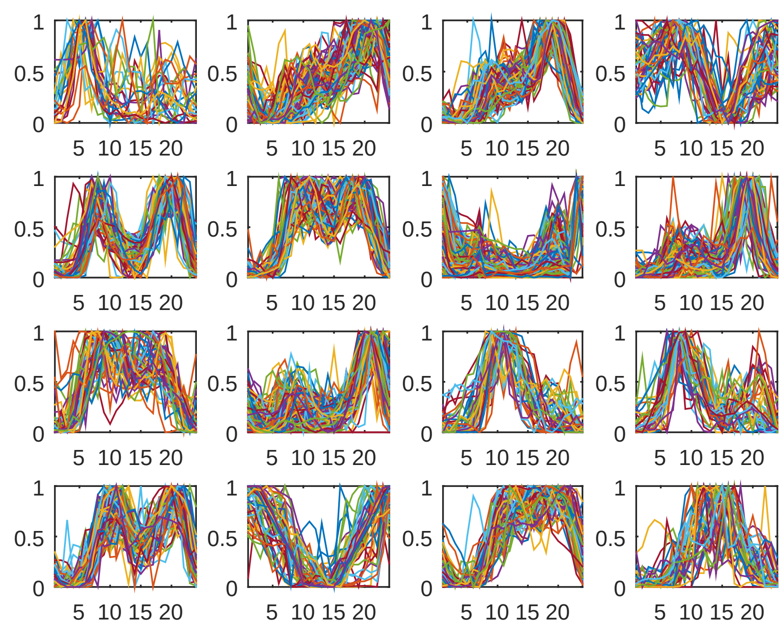


Figure 3.3: Clustering of 1500 customers into 16 groups

The dynamic behavior of energy consumption for the whole group can be represented by SAX symbols as shown in Figure 3.4. It is noteworthy that the number of symbols forming a string can be very flexible. For simplification purposes and implementation easiness of the utility company (retailer), we just use three different symbols "**a**, **b**, **c**" in this test case. While, a larger number of symbol types will produce more accurate pricing for electricity products of different energy customers, it will also produce more complexity in terms of utility operation. The mapping between these SAX symbols and specific energy price levels should depend on a utility's historical operation experience and market analysis. A typical example of an electricity pricing level based on an existing global TOU pricing scheme is presented in Table 3.1. Accordingly, the individualized price scheme designs for all 16 energy customer groups are shown in Figure 3.5.

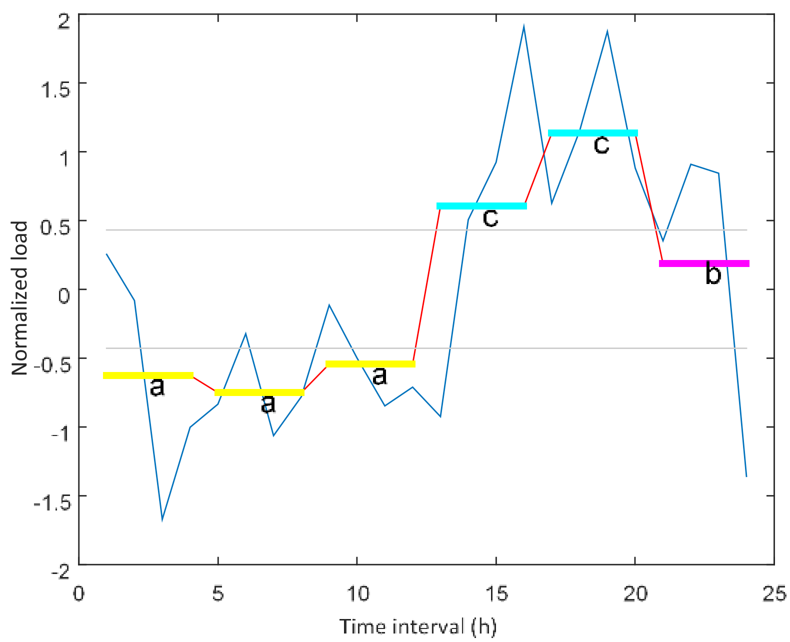


Figure 3.4: The electricity price scheme design for a typical load profile with SAX

Table 3.1: Price scheme design with symbols in SAX

Symbols in SAX	Pricing Level	Price
a	Low level pricing	0.013 $ / kWh
b	Intermediate level pricing	0.075 $ / kWh
c	High level pricing	0.180 $ / kWh

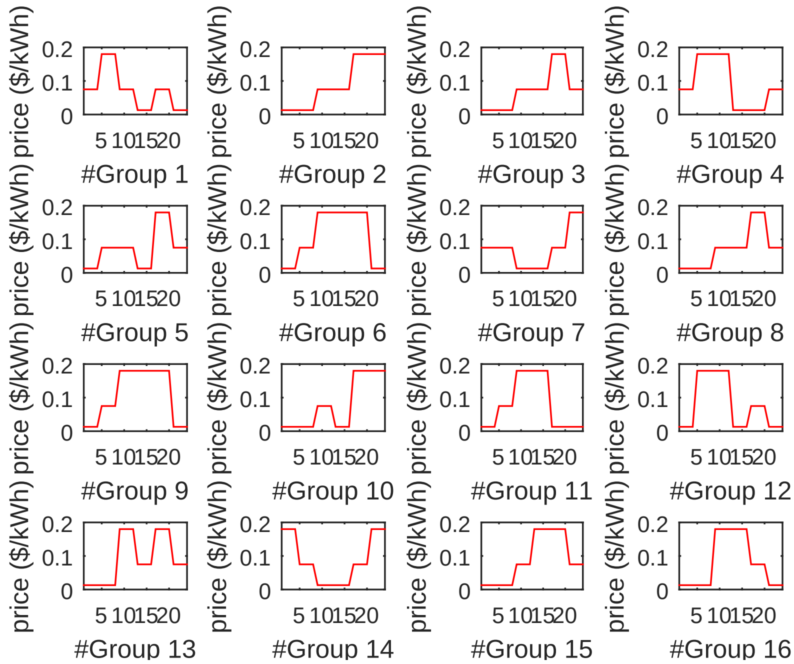


Figure 3.5: The design of an individualized price scheme structure for every group

3.3.2 Economic analysis

In the retail electricity market, different energy customers are usually given different preferences and actually have cross-subsidy with each other [84]. By testing those 1,500 customers, we found that, as shown in Table 3.2, the individualized TOU proposed in this paper mainly benefits retailers rather than energy customers in the short-term. However, if some demand response programs are introduced for customers, and their awareness of DR is reflected by some responsive rates, customers will still be able to achieve smart energy usage and economic benefit. In this way, retailers and customers will interact with each other more actively to commonly reach a better energy management and service.

Table 3.2: Economic analysis of the revenue and payment

Pricing strategy	Customers (payment)	Retailer (cost)	Retailer (revenue)
Global TOU	$1458.33	$1178.12	$280.21
Individualized TOU	$1525.34	$1178.12	$347.22
Individualized TOU with DR	$1429.42	$1148.89	$280.53

3.4 Conclusion

In this paper, we proposed an individualized electricity price scheme design mechanism for various types of customers based on SAX and some combined classification methods, namely K-means and ELM. The final goal is that the utility company can make better use of the collected smart meter data and provide customized service to end-users. The customers can also reach more awareness of the possible energy usage strategy. In the future, some more accurate and computationally efficient classification method should be studied for other related applications involving large-scale distribution networks with an industrialized big-data platform. More innovative business models of demand response programs based on the designed individualized electricity price scheme may also be discussed in the Smart Grid context.

CHAPTER IV

Modeling of energy broker

4.1 Introduction

[1] In today's retail electricity market, customers have very limited ”energy choice,” or freedom to choose different types of energy services. Although the installation of distributed energy resources (DERs) has become prevalent in many regions, most customers and prosumers who have local energy generation and possible surplus can still only choose to trade with utility companies. They either purchase energy from or sell energy surplus back to the utilities directly while suffering from some price gap. The key to providing more energy trading freedom and open innovation in the retail electricity market is to develop consumer-centric business models and possibly a localized energy trading platform. The current research community is pursuing these ideas so that the next-generation retail electricity market infrastructure will be a level playing field, where all customers have an equal opportunity to actively participate directly [16].

Many works have proposed future market mechanisms with promising potential, such as prosumer grid integration [55], peer-to-peer models [58][60], and prosumer community groups [138][139]. Direct integration of DERs in the main grid and participation in the electricity market through aggregators are widely accepted ideas because of their

[1]This chapter was previously published as an article in a peer-reviewed journal: [137] **T. Chen**, and W. Su, ”Indirect Customer-to-Customer Energy Trading with Reinforcement Learning”, IEEE Trans. on Smart Grid, 2018. (accepted)

simple applicability and manageability. The work in [55] describes a typical model that analyzes the optimal planning and operation of aggregated DERs with participation in the electricity market as a whole. However, in order to diversify the customer-side energy ecosystem and promote a deregulated market, a peer-to-peer (P2P) energy trading model with an eBay-like market platform is also frequently discussed and attracting more and more research interest, although still at a very conceptual level [140][141][142]. Admittedly, the e-commerce-like retail activity of energy trading presents only a small ratio of the overall energy business, especially of the wholesale market or bilateral contracts. The trend of gamification of energy activities [143] and novel business model design is still being intensively pursued to drive the new energy ecosystem building and market reformation. For instance, the concept of a prosumer community is utilized in [138][144] for local energy sharing and internal trading. In [58], a matching mechanism based market model is proposed to allow individual prosumers to meet each other to conduct a bilateral trade. In [60], a scalable and modular P2P system is designed, along with a prototype, to demonstrate on-site building energy trading between prosumers. Even a novel decentralized digital currency, named after NRGcoins, is proposed by the same group of researchers to encourage prosumers to locally trade their excess energy while carrying out payments using NRGcoins [61].

In this work, we aim to study how to use reinforcement learning framework to solve a decision-making problem of the local energy market clearance, from the traders perspective. We explore the role of emerging energy brokers (middlemen) in a localized event-driven market (LEM) at the distribution level for facilitating indirect customer-to-customer (iC2C) energy trading. On one hand, many customers are expecting a higher degree of ”energy choice” so that they can locally share or exchange self-generated electricity. As a perfect complement to the existing market mechanism, localized energy trading will also reduce distribution loss due to long-distance exchange and improve the efficiency of resource allocation with consideration for both attitudinal and contextual factors [145]. On the other hand, small-scale electricity consumers and prosumers usually cannot afford a time-consuming search for a particular trading partner, which makes the pure P2P mode unsuitable. In contrast, the specialty associated with a localized adaptive service can help collect, synthesize, and explore distinct renewable energy trading characteristics, as well as size-related regularities [146]. In such a localized electricity market, trade does not involve just buyers and sellers, but also multiple middlemen serving as intermediaries [147]. Classical economic approaches to studying electricity markets, such as competitive equilibrium analysis, largely abstract away the role of such middlemen. That said, there is a need to understand middlemen's optimal strategies when choosing from a series of potential opportunities of random quality, under the assumption that delaying choice is costly.

In this paper, our contribution is to: (1) propose a new paradigm of indirect customer-to-customer energy trading in a localized event-driven market; (2) introduce a new energy trading role called the retail energy broker to facilitate market operation with different well-defined actions; (3) utilize reinforcement learning techniques to benefit all market participants based on a smart strategy with learning capability; and (4) provide some interesting observations for the proposed market mechanism based on various simulation results.

The remainder of this paper is organized as follows: Section II introduces the event-driven market architecture and emphasizes the motivation. Section III introduces the system modeling of the proposed market mechanism with a REB. Section IV gives a mathematical description of a Markov decision process with a modified Q-learning algorithm. Section V presents several numerical results in different scenarios. Section VI summarizes the major findings and potential extensions of this paper.

4.2 Event-driven Market Architecture

A LEM is designed at the distribution level for facilitating retail energy trading aided by the new role of the middleman, called a retail energy broker (REB). This market mechanism will provide additional energy trading opportunities, besides the existing utility service, to allow customers to directly participate in the local energy transaction platform with the help of the REB as a trader. This market is named event-driven, similar to [148], because it is only supposed to work as a back-up trading platform, unnecessary to be open permanently for the whole year, and is affected by local requirements and seasonal events, such as high solar irradiance or energy shortages in regional grids. In this way, the future retail electricity market, including the LEM, can be shown as in Figure 4.1.

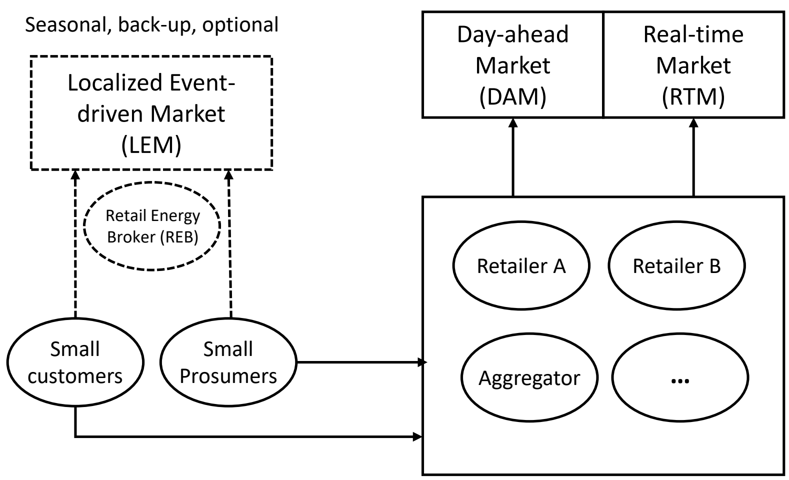


Figure 4.1: The future retail electricity market

All the energy entities, including buildings, campuses, communities, etc., can choose to participate in the local energy trading platform as sellers or buyers according to their own forecasting of their energy surplus or deficit. The LEM model is built as a Markov decision process (MDP) [149] from the REB's perspective. The REB can operate this market while obtaining increasing trading experience and experience-based strategies formulated by reinforcement learning techniques. Given this setup, by applying intelligent methods and taking into account the characteristics of customers' behavior, it can be proven that both customers and traders benefit more at the same time. Some economic concepts, like *search friction* [150], related to this kind of typical search cost involved market model are also discussed in following sections, which offer insight into the improvement of energy market efficiency.

The market model of a LEM is not chosen because of its best performance. The motivation for developing such a localized occasional energy market is twofold: 1) deregulation, may sound a little cliché, is still the driving force for reshaping the legacy energy landscape. The deregulation of the airline industry and telecommunications industry proves how it can benefit consumers. The emergence of *Priceline* and *Uber* are also good examples of two of the few notable industries that contextualize both deregulation and new platforms that boost new business paradigm design and new ecosystem building within the existing market context. Meanwhile, the electricity energy industry is one of the largest in the world, yet it is also considered the least innovative because of its highly regulated market structure; 2) there is a need for a new business model design and even gamification of the energy ecosystem. If successfully implemented, the new model will significantly promote the massive introduction of new players (e.g., REBs) in a more competitive retail market. Rather than passively accepting the trend of the emerging energy trading paradigm, we can help customers attain a better understanding of the future energy landscape. Taking smartphone as an example, before the boom of the smartphone, mobile internet occupied only 5% of the overall internet connections. Nowadays, mobile internet connections have become mainstream and account for half of the overall connections. We should create the same "desire" for local energy trading. In light of these reasons, the LEM seeks to improve the efficiency of regular retail activity mainly through diversifying the existing energy business models (commercial strategies) and providing energy trading opportunities at the edge of the distribution network, near the source of the energy demand.

In the conceptual design of our holistic market model [2] (Figure 4.2), the major functionalities of each entity are highlighted. For example, the proposed localized event-driven market (LEM) will facilitate the short-term or immediate local energy transactions. The Distribution System Operator (DSO) or Distribution Network Operator (DNO) is responsible for market regulation (e.g., reliability and security checks). The electric utility will not only serve customers as per usual but will also offer various long-term retail plans. Meanwhile, customers may develop their own bidding/offering strategies in this local energy trading platform by using various types of energy devices (e.g., batteries, DGs, flexible loads, etc.) that are available to them. A local energy transaction will be physically fulfilled by leveraging the existing distribution line and smart meters (for billing and payment).

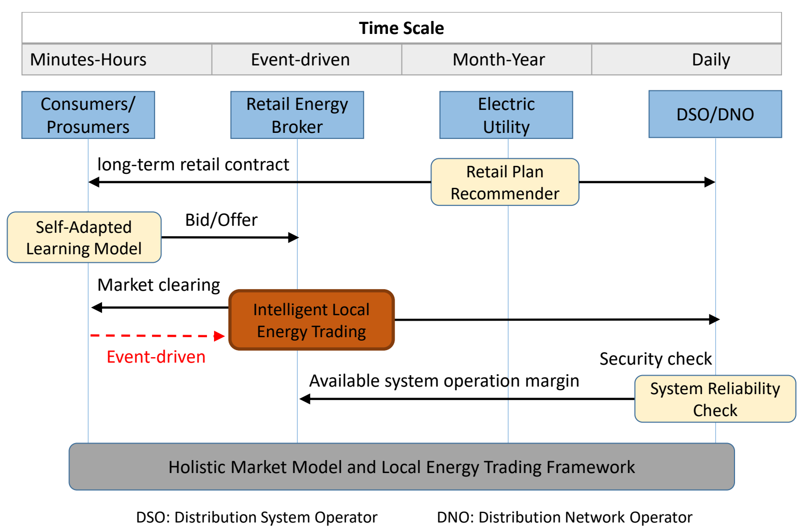


Figure 4.2: Holistic market model design

It is worth noting that it is not our intention to cover every aspect of the aforementioned market model in a single paper. In this paper, we mainly focus on the new role of the retail energy broker (REB), as highlighted in an orange box in Figure 4.2. Without loss of

[2] The comprehensive system regarding holistic market model will be presented in our recent work.

generality, we assume that most local participants are geographically close to each other. In other words, they are most likely connected to the same distribution feeder, as illustrated in Figure 4.3. Therefore, this paper does not explicitly model the AC power flow constraint of distribution networks. In the future, we will further investigate the impacts of physical constraints on the proposed LEM. Distribution substations connect to the transmission system and lower the transmission voltage to medium voltage ranging between 2 kV and 35 kV with the use of transformers. Usually the standard primary distribution voltage levels include 4.16kV, 7.2kV, 13.2kV, 23.9kV, and 34.5kV.

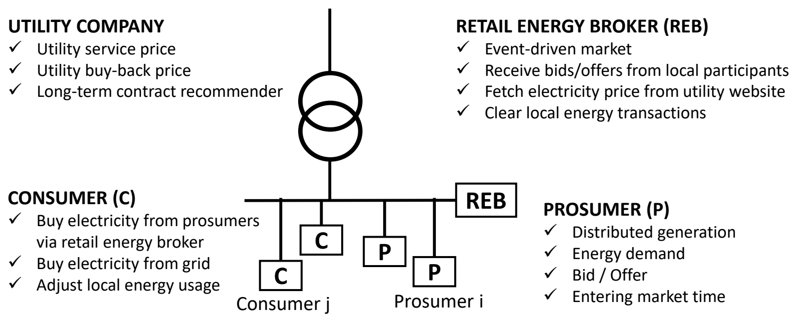


Figure 4.3: Physical connection of consumer/prosumer and REB

4.3 System model

We consider a localized market model that consists of a REB, a set of consumers (buyers) $\mathscr{I}^b$ and a set of prosumers (sellers) $\mathscr{I}^s$. The market is operated in a day-ahead bidding and time-slotted fashion, where each time slot has an equal duration and allows for the REB's decision-making. At each time slot, the REB can choose different actions to clear the market so far and satisfy sellers and buyers on both sides. Hence, the market model will be formulated as an MDP problem, which enables a multi-scenario study of the dynamics and characteristics of the proposed market mechanism.

4.3.1 Market operation

The market mechanism is designed following a day-ahead bidding process similar to the day-ahead market in the wholesale electricity market [151]. Moreover, each day will be divided into n time slots during which the market will be open for sellers' or buyers' participation, as shown in Figure 4.4. Thus, each time interval T will be equal to 24 $hours/n$. No matter when the sellers (buyers) enter the market, there will be initially at most consecutive H time intervals for them to stay in the market after entering. The REB can choose different actions for each time interval according to the current market status. It is also noteworthy that sellers and buyers usually enter the market following heterogeneous time schedules due to the fact that the renewable energy supply (e.g., PV roofing) may not overlap with the demand peak load of energy consumption. Buyers are highly likely to enter the market around the early morning or late afternoon when they experience the twin-peak energy consumption pattern, while sellers prefer to enter the market around noon or midnight, when the solar radiance or wind source is at maximum. Without considering energy storage system (ESS), this temporal heterogeneity requires the REB to use tricky decision-making strategies, depending on whether it is a buyer's market or a seller's market. In particular, taking into account search friction [3], similar to opportunity cost, the decision-making delay or waiting time of the REB can give rise to totally different market-clearance results.

This timeline design allows for future extension of customers' behavioral dynamics, such as withdrawing or modifying their bids (offers) in consecutive time slots. However, in this paper, we mainly consider the market dynamics from the REB's perspective and limit the customers' participation to one-time bidding/offering. Additionally, we introduce a remaining time counter $h^s_{i,t} \in \mathscr{H}^s_t$ and $h^b_{j,t} \in \mathscr{H}^b_t$ associated with each seller and buyer

[3] Search friction will be introduced and discussed in following sections

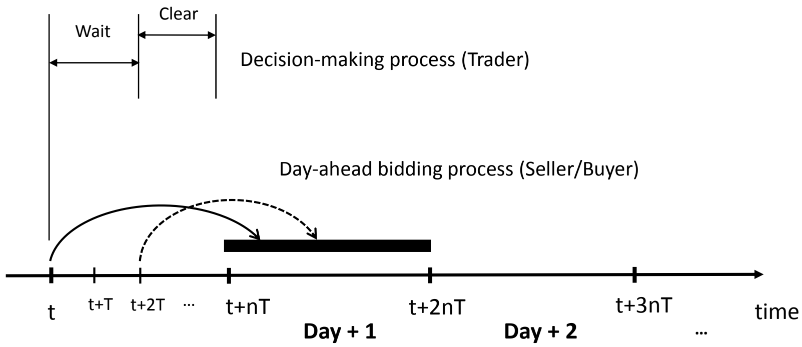


Figure 4.4: Timeline of the proposed market model

to record their staying status in this market with an initial value of $h^s_{i,t^s_i} = H$ and $h^b_{j,t^b_j} = H$, respectively, where the subscript $t^s_i \in \mathscr{T}^s$ ($t^b_j \in \mathscr{T}^b$) indicates the market entering time of each seller (buyer). As the time evolves $t \leftarrow t+1$, the remaining time counter decreases accordingly after each time interval: $h^s_{i,t} \leftarrow h^s_{i,t} - 1$, $h^b_{j,t} \leftarrow h^b_{j,t} - 1$. If the remaining time counter approaches empty, $h^s_{i,t} = 0$ or $h^b_{j,t} = 0$, the seller (buyer) will quit this market and exercise direct transaction with the utility company.

4.3.2 Sellers and buyers

Consumers and prosumers participate as sellers or buyers in the localized retail electricity market according to their power demand and generation surplus. Different types of prosumers are considered uniformly with simplification of their particular generation resources (e.g., wind, PV, biomass, etc.). As long as the prediction for the next hourly time interval is obtained, seller i provides bidding information with selling price p^s_i and supply energy block e^s_i as price-quantity pairs to the REB. Meanwhile, it should be guaranteed that

$$p^{ub} \leq p^s_i \leq p^u - \varepsilon_p, \quad \forall i \in \mathscr{I}^s \tag{4.1}$$

Where p^{ub} is the buy-back price of utility if sellers directly sell their energy surplus to the utility company, and p^u is the regular electricity service price of the utility if buyers directly purchase energy from the utility company. The lower-bound p^{ub} of a seller's bidding price p_i^s is intuitive because the only motivation for sellers' participation in this market is the fact that the seller can sell energy at a higher price than through direct trading with the utility. However, the upper-bound $p^u - \varepsilon_p$ is implicitly established because any bidding price higher than p^u will force buyers to prefer direct trading with the utility, pushing them away from participating in this localized market. Furthermore, item ε_p is a small speculation margin controlled by the REB to guarantee fairness and enhance the seller/buyer's pairing chance (no exactly equivalent price to the utility's is allowed to participate, otherwise all customers will bid the same as the utility's price). Similarly, the same principle also applies to buyers and their offering information, with demand energy block e_j^b and purchasing price p_j^b as

$$p^{ub} + \varepsilon_p \leq p_j^b \leq p^u, \quad \forall j \in \mathscr{I}^b \tag{4.2}$$

Therefore, buyers can benefit from purchasing electricity at a lower price than the utility's regular service price p^u while leaving enough price margin to attract sellers, and vice versa. These price-quantity pairs of both sellers and buyers can be contained in sets, $\mathscr{P}^s = \{p_i^s\}$, $\mathscr{P}^b = \{p_j^b\}$, $\mathscr{E}^s = \{e_i^s\}$ and $\mathscr{E}^b = \{e_j^b\}$. These sets, along with sets of the remaining time counters at time t, will be reformulated again as follows:

$$\mathscr{P} = \mathscr{P}^s \times \mathscr{P}^b, \quad \mathscr{E} = \mathscr{E}^s \times \mathscr{E}^b, \quad \mathscr{H}_t = \mathscr{H}_t^s \times \mathscr{H}_t^b \tag{4.3}$$

We also use $\mathscr{I}_t = \mathscr{I}_t^s \times \mathscr{I}_t^b$ to denote the set of sellers (buyers) that are still present in the market at current time t with $h_{i,t}^s, h_{j,t}^b > 0$. In this way, the market status at current time t will be totally determined by all the current customers' bidding parameters, the current timestamp and the status of the remaining time counters associated with current customers. Thus, the transition probability of the market status, consisting of the status of sellers and buyers, from state $s_t = \left(\mathscr{H}_t, \mathscr{I}_t\right)$ to state $s_{t+1} = \left(\mathscr{H}_{t+1}, \mathscr{I}_{t+1}\right)$ with the REB's given action a_t can be represented as:

$$p_m\big(s_{t+1}\big|s_t,a_t\big) = \prod_{i\in\mathscr{I}_t^s} p_s\big(h_{i,t+1}^s\big|s_t,a_t\big) \times \prod_{j\in\mathscr{I}_t^b} p_b\big(h_{j,t+1}^b\big|s_t,a_t\big) \tag{4.4}$$

Where $\mathscr{I}_t$ implicitly contains all the information associated with parameter sets $\mathscr{P}$ and $\mathscr{E}$, and $\mathscr{H}$ explicitly indicates the possible action candidates available to the energy broker.

4.3.3 Retail energy broker

In the proposed market model, the REB will be responsible for determining the occasional market open rate when any events or requests are detected. It plays a role as a middleman between sellers and buyers to enable local energy transactions in a LEM in an indirect way (i.e. iC2C). Thus, the seller or buyer does not need to search for a particular trading partner during the searching and pairing process.

For the REB, how to choose actions is totally dependent on current time t and current market situations given different price-quantity pairs collected from present seller (buyer) bid (offer) information. In other words, the REB determines mapping function $a_t : \mathscr{R}_+^{|\mathscr{S}|} \to \mathscr{A}$ as follows

$$a_t\Big(\mathscr{H}_t,\mathscr{I}_t,t\Big) = a_i, \quad i = 1,2,...,A \tag{4.5}$$

From there, it decides how to process the market-clearance at current time-slot t, where $|\mathscr{S}|$ denotes the cardinality of market status tuple set $\mathscr{S} = \mathscr{H}_t^s \times \mathscr{H}_t^b \times \mathscr{I}_t^s \times \mathscr{I}_t^b$ and $\mathscr{A}$ denotes finite action set $\{a_1,a_2,...,a_A\}$, as follows (let $A = 4$; four types of actions are designed here):

1) *Waiting*: do not accept any bid or offer price-quantity pairs and delay the market-clearance until the next time slot, $t+1$; expect more sellers or buyers to enter the market with more speculative bidding (offering). However, search friction should be considered

in this action since the REB may lose the opportunity to achieve more or even myopic profit. The seller and buyer's remaining time counters will be updated by $h \leftarrow h-1$ for the next time interval.

2) *Retention*: do not accept any bid or offer price-quantity pairs and delay the market-clearance until the next time slot, $t+1$; however, in contrast to *waiting*, the REB will issue a retention bonus ξ to all the sellers/buyers present in the current market to compensate for the delay of the market-clearance. Moreover, the seller and buyer's remaining time counters will be updated by $h \leftarrow h+1$ for the next time interval.

3) *Partial-clearance*: accepts the highest bid p^s_{high} and lowest offer p^b_{low} associated with a price-quantity pair to lock the instant profit so far, setting the chosen seller/buyer's remaining time counter $h \leftarrow 0$, meanwhile delaying the rest of the market-clearance until the next time-slot, with all the other sellers/buyers' remaining time counters updated by $h \leftarrow h-1$.

4) *Clearance*: accept all the bids and offers belonging to price-quantity pairs to exercise the standard double-auction market (action $a_t = a_4$ in (4.7) and Figure 4.5), and set all the chosen sellers/buyers' remaining time counters $h \leftarrow 0$ to force them to quit the market. The rest of the customers that fail to be selected continue to stay in the market for future chances, with their remaining time counters updated by $h \leftarrow h-1$.

The goal of the REB in this LEM is to choose a series of actions to maximize its profit with constant consideration of some opportunity cost when many a seller/buyer frequently enters and quits the market. The optimal policy π^* can be defined as: 1) the best action is always correctly chosen for maximizing long-term expected profit (reward), and 2) the decision is always made at the most suitable time without any delay. Thus, the expected discounted revenue of the REB is defined as in the following MDP problem **P**:

$$\mathbf{P}: \quad \max_{\pi:\mathscr{S}\to\mathscr{A}} E\left[\sum_{t=0}^{nT}(\gamma)^t r_t\Big(s_t,\pi(s_t)\Big)\right] \tag{4.6}$$

where $0 \leq \gamma \leq 1$ is the discount factor, which represents the relative importance of the future market profit compared with that of the present market profit. The reward function $r_t\Big(s_t,\pi(s_t)\Big)$ is calculated according to different chosen actions, as follows:

$$r_t\Big(s_t,a_t\Big) = \begin{cases} 0 & a_t = a_1 \\ -\xi & a_t = a_2 \\ p^s_{high}e^s_{high} - p^b_{low}e^b_{low} & a_t = a_3 \\ z^\star & a_t = a_4 \end{cases} \tag{4.7}$$

where a_1, a_2, a_3 and a_4 indicate the actions of *waiting, retention, partial-clearance* and *clearance*, respectively, $p^s_{high} = max\{p^s_i | i \in \mathscr{I}^s_t\}$, and $p^b_{low} = max\{p^b_j | j \in \mathscr{I}^b_t\}$. The numerical value $z^\star$ corresponding to $a_t = a_4$ in (4.7) is obtained as an optimal value through solving the following optimization problem, **Pa**, in (4.8)-(4.11), which is also a double-auction problem, as shown in Figure 4.5. It follows a pay-as-bid model, in which all the price-quantity pairs left at the equilibrium price will be cleared.

$$\mathbf{Pa}: \quad \max_{x^s_i, x^b_j} \quad z = \sum_{j\in\mathscr{I}^b_t} p^b_j x^b_j - \sum_{i\in\mathscr{I}^s_t} p^s_i x^s_i \tag{4.8}$$

s.t.

$$\sum_{j\in\mathscr{I}^b_t} x^b_j - \sum_{i\in\mathscr{I}^s_t} x^s_i = 0 \tag{4.9}$$

$$0 \leq x^s_i \leq e^s_i, \quad \forall i \in \mathscr{I}^s_t \tag{4.10}$$

$$0 \leq x^b_j \leq e^b_j, \quad \forall j \in \mathscr{I}^b_t \tag{4.11}$$

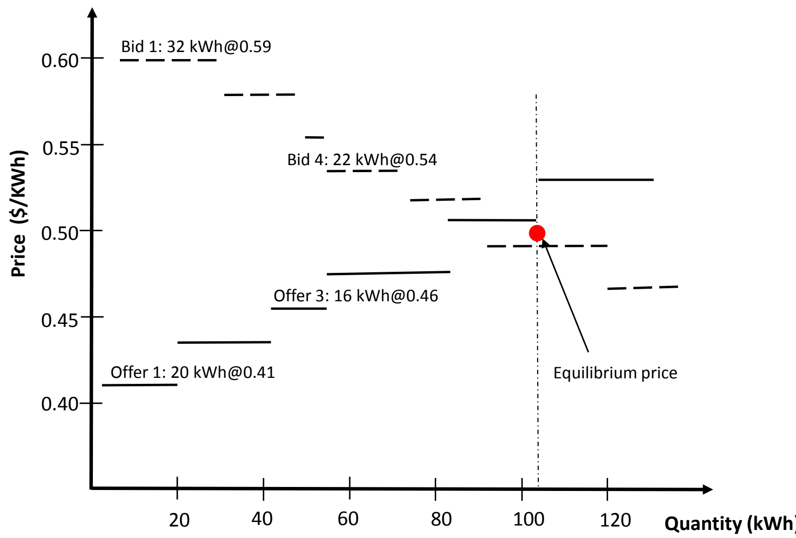


Figure 4.5: Market-clearance of the double-auction model

It is also worth commenting that in a reinforcement learning framework, the defined actions can be low-level controls, such as assigning a particular numerical value to a variable, or high-level decisions, such as whether or not to solve a sub-decision problem [152] like ours. Additionally, retention bonus ξ in *retention* explicitly indicates the fact that decision delay is costly, and some other implicit *search cost* is also contained in *waiting* because delaying the decision-making by waiting may often introduce some regret or opportunity cost, which will be quantitatively measured in Section 4.5.4.

4.3.4 Real-world implementation issues

In order to implement the overall local energy trading process, it is assumed that every interested participant is able to join the LEM in an intelligent way by using a home energy management system (HEMS) [153] with trading functionality embedded. The trader also must have access to a regional central market operation system that will be responsible for collecting the local customers' subscription to this optional energy service. In the long term and as more trading experience is accumulated, both customers and the trader will be able to adapt to the daily, weekly or even seasonal patterns of energy trading activities. The communication hardware cost of implementing such an energy trading service is expected to be low and without too much hardware modification, since no subjective negotiation process is involved in the intelligent trading platform, and localized energy service activities are usually stable when considering only nearby customers in a neighborhood or local community. The current distribution network would be better to support such local energy trading if real-time monitoring, two-way communication and feeder reconfiguration devices are widely available. However, there is no need to increase the distribution line capacity since the congestion can even be relieved to some extent if many energy sources are generated and consumed locally without long distance transmission.

4.4 Markov Decision Process and Modified Q-learning Algorithm

An MDP is usually described by state space $\mathscr{S}$, action space $\mathscr{A}$, reward function set $\mathscr{R}$, state transition probability matrix $\mathscr{P}$ and a discount factor $\gamma \in [0,1]$. It is also assumed that an MDP remains the Markov property, which implies the transition probabilities of a state are only affected by the previous one step with no memory [152]. As our market model and problem **P** is built as an MDP (Figure 4.6), it can be supposed to be solved by any standard reinforcement learning techniques, such as temporal-difference learning.

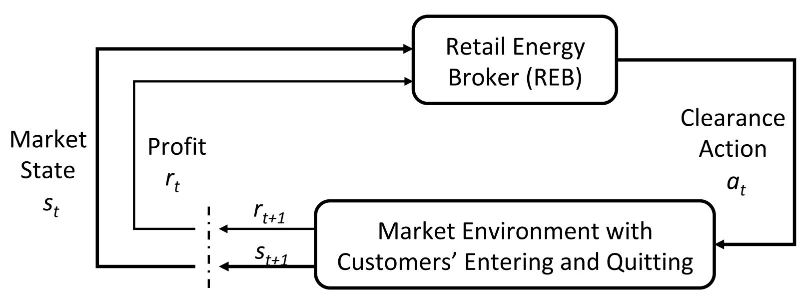


Figure 4.6: The market model as an MDP with agent-environment interaction

However, there will always be some issues associated with most similar practical problems, namely how to deal with the continuous state space. For instance, the market status

$\mathscr{S}$ in this paper consists of price information, energy quantity, customers' entering time, and many possible combinations of customers' staying status, which are all continuous variables or hard to be enumerated. In order to deal with the continuous state space, one possible solution is to take it as a multi-armed bandits problem, especially a contextual bandits problem [154]. We can take timestamp t as the only state space and other information $\{\mathscr{P}, \mathscr{E}, \mathscr{T}^s, \mathscr{T}^b, \mathscr{H}_t, \mathscr{I}_t\}$ as pure reward parameters or external guiding signals in a contextual bandits problem with the assumption of some certain distribution of these parameters. That said, the learning ability of the decision-maker will be very limited since no knowledge of previous experience is utilized when choosing the next actions. In a bandits problem, each action affects only the immediate reward. If actions are allowed to affect the next situation as well as the reward, then we have to take it as a full reinforcement learning problem [152]. In our market model, how to choose the current action (i.e., *waiting, retention, partial-clearance* and *clearance*) determines how much profit will be exploited right now and how much will be left for the next time interval. The chosen action will obviously affect the market situation and profit capacity (reward structure) in the next time interval in a certain episode. Therefore, we introduce an episode-dependent modified Q-learning algorithm, as shown in Algorithm 2, based on the temporal-different learning structure and theoretic analysis in [155] to solve the problem. We named the market operation method based on this modified Q-learning algorithm *smart strategy*, which we use in following sections.

The function $\phi(\tau)$ in Algorithm 2 refers to the tabular non-stationary reward structure described in [155]. Here, in our market model, the term non-stationary is twofold: 1) the reward distribution is affected by the market status, which changes along episodes (days) and may follow some certain probability distribution functions, with one sample generated in each episode, and 2) the action chosen in each step of a given episode will affect the

Algorithm 1 Modified Q-learning algorithm

Initialize $Q(s,a), \forall s \in \mathbb{S}, a \in \mathbb{A}(s)$ arbitrarily
Repeat (for each episode τ):
  Initialize s with episode dependent information
    Repeat (for each step of episode):
      Choose a from s using policy derived from Q (e.g., ε-greedy)
      Take action a, observe $r \in \phi(\tau)$ via (4.7) and s'
      Update the non-stationary reward structure $\phi(\tau)$
      $Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_{a'} Q(s',a') - Q(s,a)\right]$
      $s \leftarrow s'$
    Until s is terminal

potential reward value in following steps. It is these mechanisms that will introduce the many interesting market dynamics demonstrated in the next section through various cases studies. In this paper, we will not go deep into the theoretic fundamental that is still under study and has produced many cutting edge research topics. Interested readers may refer to [155][156][157] for more details.

Similar to the basic Q-learning algorithm, the optimal stationary policy π^* can be well defined by using the optimal *action-value function* $Q^* : \mathscr{S} \times \mathscr{A} \to \mathscr{R}$, which satisfies the following Bellman optimality equation:

$$Q^*(s,a) = r(s,a) + \gamma \sum_{s' \in \mathscr{S}} p\left(s' \middle| s,a\right) V^*\left(s'\right), \tag{4.12}$$

where $V^*\left(s'\right)$ is the optimal *state-value function* [152], which is defined as

$$V^*\left(s'\right) = \min_{a \in \mathscr{A}} Q^*(s',a), \quad \forall s \in \mathscr{S}. \tag{4.13}$$

Since $V^*\left(s'\right)$ is the expected discounted system cost, with action a in state s, we can obtain the optimal stationary policy as

$$\pi^*\left(s\right) = \operatorname{argmin}_{a \in \mathscr{A}} Q^*(s,a). \tag{4.14}$$

In solving problem $\mathbf{P}$, any temporal-different learning based method, unlike dynamic programming methods [158], can provide the solution without acquisition of the state transition probabilities $p_s\left(s_{t+1} \middle| s_t, a_t\right), \forall s \in \mathscr{S}$ a priori [159]. In this way, reinforcement learning techniques based on an MDP will allow the REB, as a trader, to take the LEM as a pure data-driven model to build its strategy intelligently with increasing experience.

4.5 Numerical results

In this section, we provide numerical results to show some interesting observations regarding the proposed LEM model and evaluate the performance of our market learning algorithm. Unless specifically indicated, the timeline, one day, consists of 72 time intervals, each of which lasts for 20 *min*. The simulation environment is Python 3.6 running on a desktop with an Intel i7 and 16.0 GB RAM.

4.5.1 Smart strategy with learning

In this first case study, we run a basic market model to initially compare the results of the smart strategy with learning and a dummy strategy without learning. The dummy strategy means the trader REB does not utilize any knowledge of historical trading experience to operate this market and clears the market immediately, as much as possible, with the myopic action to maximize its current profit. In contrast, the smart strategy allows utilizing learning capability to choose different actions (i.e., waiting, retention, clearance and partial-clearance) to maximize profit along the complete time horizon for the whole day. All the parameters of the market model in this case are defined in Table 5.2.

Table 4.1: Parameters of the market model

Parameter sets	Values
$\mathscr{T}^s$, $\mathscr{T}^b$	$\mathscr{T}^s \sim \mathscr{N}\left(39, 12^2\right)$, $\mathscr{T}^b \sim \mathscr{N}\left(54, 12^2\right)$
$\mathscr{P}^s$, $\mathscr{P}^b$	$\mathscr{U}\left(0.4, 0.6\right)$
$\mathscr{E}^s$, $\mathscr{E}^b$	$\mathscr{U}\left(20, 40\right)$
	$\varepsilon_p = 0$, $\lvert\mathscr{I}^s\rvert = 100$, $\lvert\mathscr{I}^b\rvert = 110$, $H = 3$, $n = 72$, $\xi = 0.5$

It is noteworthy that this market model is totally data-driven, with accumulated trading experience, and independent of the choice of distribution. The parameter sets, $\mathscr{T}^s$ and $\mathscr{T}^b$, are chosen as normal distributions just for convenience and reasonable assumption; however it is not necessary that they are chosen as normal, Weibull, uniform or any particular distribution. Some more sensitivity analysis in following study cases will show similar results with a hybrid distribution or purely messy random data. Some parameters related to the modified Q-learning algorithm are also given as follows: $\varepsilon = 0.1, \alpha = 0.5, \gamma = 1$. The algorithm runs 500 episodes in total and provides the last 400 episodes' results while averaging over 20 independent runs in Figure 4.7. The learned optimal policy with chosen actions is also shown in Figure 4.8. The action code is defined to be the same as in eq. (4.7): $1 \rightarrow waiting, 2 \rightarrow retention, 3 \rightarrow partial - clearance, 4 \rightarrow clearance$.

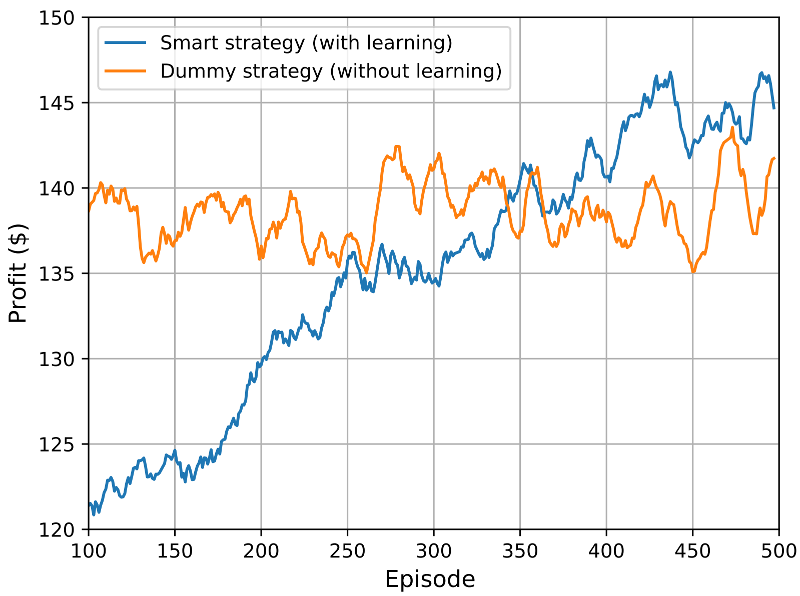


Figure 4.7: Smart strategy (with learning) vs dummy strategy (without learning)

From Figure 4.7, it can be observed that the smart strategy actually does not work well at the beginning since it is still undergoing trial and error as part of its learning process. However, with more experience obtained through running more episodes and for a longer time, the smart strategy starts to adapt to the market characteristics and *exploits* this *knowledge* to adjust its *optimal policy* [152]. In the long term, it will outperform the

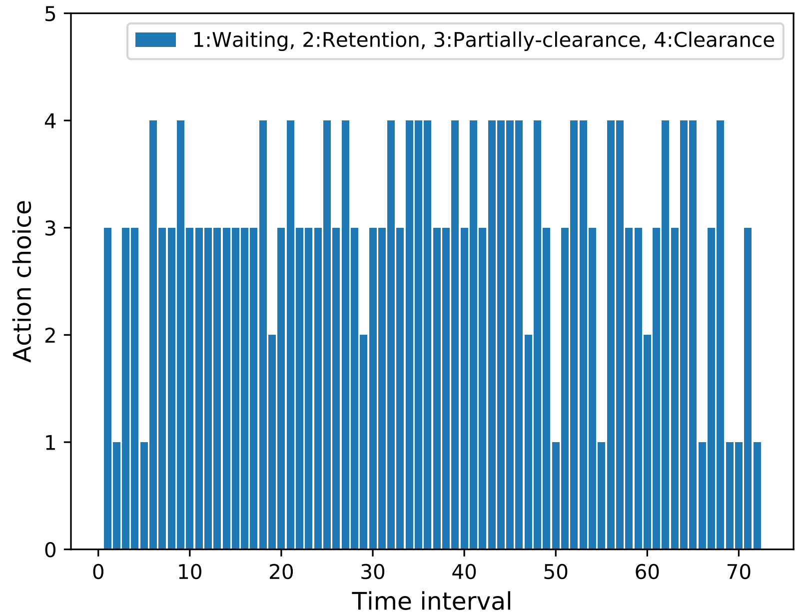
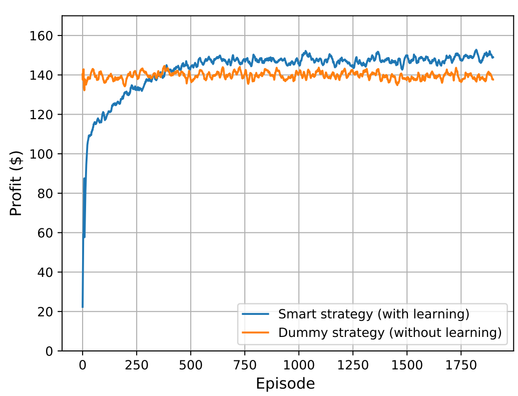
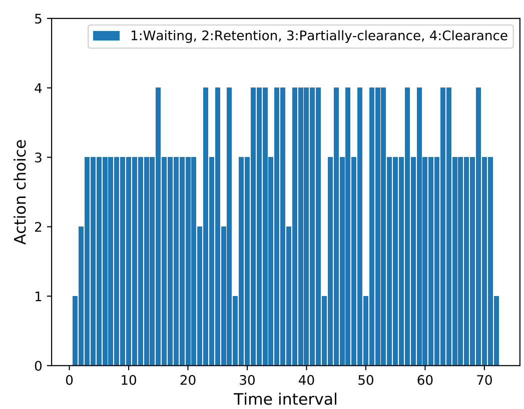


Figure 4.8: Optimal policy with chosen actions

dummy strategy that has no learning capability. In order to see a more obvious comparison in the long term, we run the algorithm again for around 2000 episodes and show the complete results, from the very beginning episode, in Figure 4.9.



(a) Profit with more experience



(b) Optimal policy with more experience

Figure 4.9: Results of running for a longer time with more experience accumulated

It is interesting to see the smart strategy with learning can improve its performance quickly and stably outperform the dummy strategy without learning in the long term. Additionally, the optimal policy will also be adjusted accordingly with more experience obtained. That said, it is noteworthy that Figure 4.8 and Figure 4.9a do not imply the smart strategy will always outperform the dummy strategy strictly in every case every day (episode), which is because the comparison result is not a deterministic one, and occasionally it might be true that the smart strategy is worse (obtains less profit) than the dummy strategy according to a particular market status. We can only conclude that in the long term and in most scenarios, the smart strategy with learning wins in a statistical way.

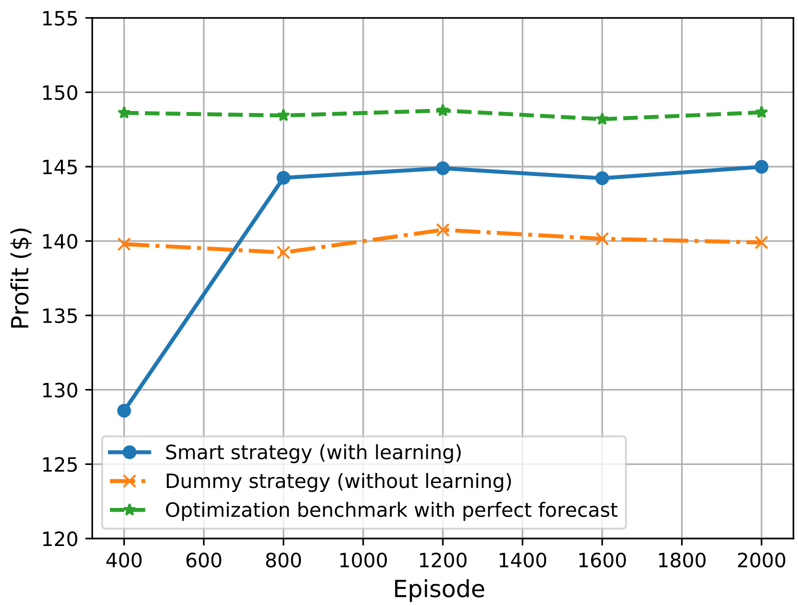


Figure 4.10: Optimization benchmark with perfect forecast

As shown in Figure 4.10, analysis of a benchmark optimization model similar to (4.8)-(4.11) with consideration for all the time intervals and perfect forecast (for future time horizons) is also conducted to compare with the reinforcement learning method. The objective function is given as follows:

$$\max \quad Z_{opt} = \sum_{j \in \mathscr{I}^b} p_j^b x_j^b - \sum_{i \in \mathscr{I}^s} p_i^s x_i^s \tag{4.15}$$

Where $\mathscr{I}^b = \bigcup_{t=1}^{nT} \mathscr{I}_t^b$ and $\mathscr{I}^s = \bigcup_{t=1}^{nT} \mathscr{I}_t^s$. We can see that because of the assumption that perfect forecast is available and given as a priori, the benchmark optimization actually has the best performance with maximum profit. However, this perfect forecast that removes all the uncertainty characteristics of customers' trading behavior is rarely the case in reality. Considering its adaptivity to volatile local market situations, it is still reasonable to accept reinforcement learning as a well-performing method for the proposed market model.

4.5.2 Advantages of LEM

In order to show the advantages the LEM model has over the current existing energy service without a local energy trading platform, we firstly present the total benefit for all the participants at different levels of participation in a one-day market operation in Table 4.2. The sellers' economic benefit achieved thanks to the LEM is described by how much they were able to increase their revenue by bidding, while the buyers' is described by how much they were able to reduce their cost by purchasing for energy demand.

Table 4.2: The benefit ($) of a LEM for all participants

Participation level	**Sellers**		**Buyers**		**Trader**	
	Mean	Std.	Mean	Std.	Mean	Std.
Low (∼200)	240.4	16.8	265.2	12.4	121.1	2.4
Medium (∼400)	501.2	27.8	552.6	21.5	276.8	4.2
High (∼1000)	1307.9	49.0	1441.9	39.2	698.9	9.1

It is interesting to observe that the total benefit at the medium participation level (~ 400 sellers plus buyers) is slightly more than two times of that at the low participation level (~ 200 sellers plus buyers), because more participation of customers (buyers or sellers) usually means more additional energy transaction opportunities if more different types of customers enter the market with a volatile range of bid/offer prices. Secondly, we also reflect the average annual benefit of each participant (per capita) at different participation levels in the LEM with the assumption of occasional opening of the market across the whole year, using a one-time average benefit (4.16).

$$\overline{R_{seller}} = \frac{1}{3N_r}\left[\frac{1}{N^s_{low}}\sum_{k=1}^{N_r}\sum_{i=1}^{N^s_{low}} R^s_{low}(i,k) + \frac{1}{N^s_{med}}\sum_{k=1}^{N_r}\sum_{i=1}^{N^s_{med}} R^s_{med}(i,k) + \frac{1}{N^s_{high}}\sum_{k=1}^{N_r}\sum_{i=1}^{N^s_{high}} R^s_{high}(i,k)\right] \tag{4.16}$$

$\overline{R_{seller}}$ indicates the average revenue of a seller for one-day participation in different scenarios, where $N^s_{low}, N^s_{med}, N^s_{high}$ are the number of customers in different participation level, respectively. N_r is the number of algorithm running times. The annual revenue of a seller in a LEM can be calculated by $365 \times r_o \times \overline{R_{seller}}$, where r_o indicates the occasional market open rate. The average annual cost of a buyer and the average annual profit of a trader can be calculated in a similar way. The overall annual benefit effect (per capita, with $r_o = 20\%$) of introducing a LEM in addition to an existing utility's energy service is shown in Figure 4.11.

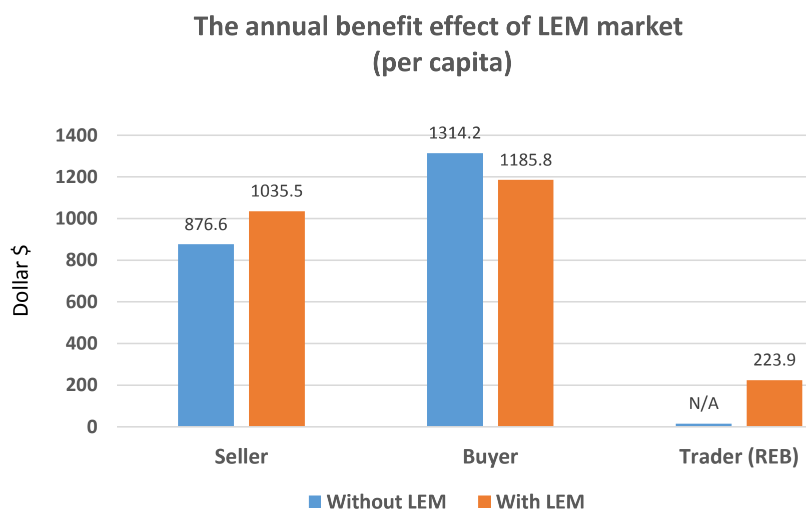


Figure 4.11: The annual benefit effect (per capita) on participants in a LEM (values in dollar are revenue for seller and REB, and cost for buyer)

4.5.3 Sensitivity analysis

In this market model, a lot of market parameters and learning parameters are involved, and many of them are significant in the REB's decision-making process. Here, some important parameters are studied during the sensitivity analysis with many interesting results and phenomenon observed. In the following cases, we use a uniform distribution for market entering time and a bid/offer price with $|\mathscr{I}^s| = 200$, $|\mathscr{I}^b| = 220$ and other parameters the same as in Section 4.5.1 if without explicit illustration.

Time intervals

The number of time intervals determines how many chances customers have to enter the market each day. If we fix the value H of the remaining time counter, fewer time intervals implies a larger ratio of customers' staying time in the market for the whole time horizon. Thus, the sellers and buyers may have a higher possibility of being chosen for market-clearance, which meanwhile provides the trader REB more chances to pursue profit, as shown in Figure 4.12. On the other hand, increasing the number of time intervals is actually also equivalent to reducing the value of the remaining time counters while improving some sort of opportunity cost from the trader's perspective. This behavior is related to a microeconomic concept, *search friction*, which will be discussed in Section 4.5.4.

Exploration ratio

Most reinforcement learning methods always need to trade-off *exploration* and *exploitation*, which means the agent has to exploit what it already knows in order to obtain a reward, but it also has to explore in order to make better action selections in the future [152]. The exploration ratio ε controls how much effort the agent, such as the REB, will exert to select an action with non-maximum reward value in every episode. In order to see the effect of the exploration ratio in the long term, we extend the run of the algorithm to 1000 episodes and show the result in Figure 4.13.

We can observe that a very low exploration ratio $\varepsilon = 0.05$ introduces very low profit

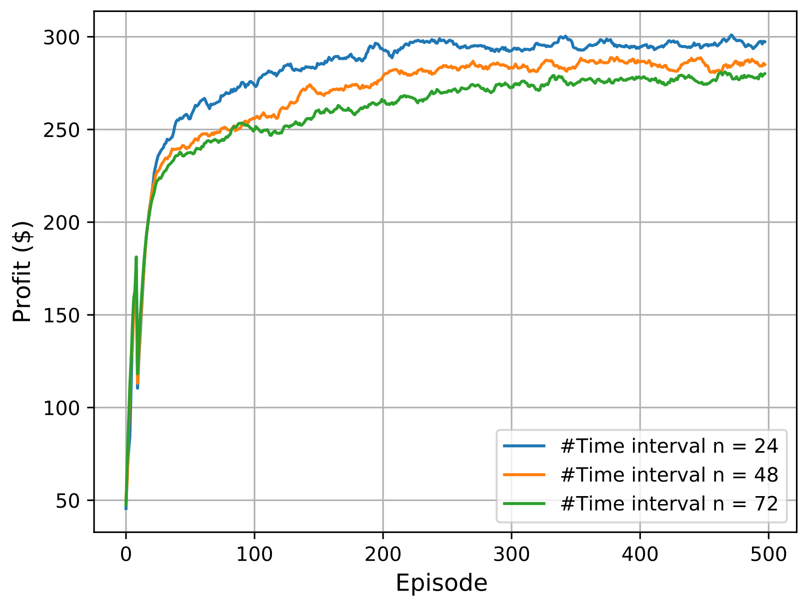


Figure 4.12: The impact of the number of time intervals

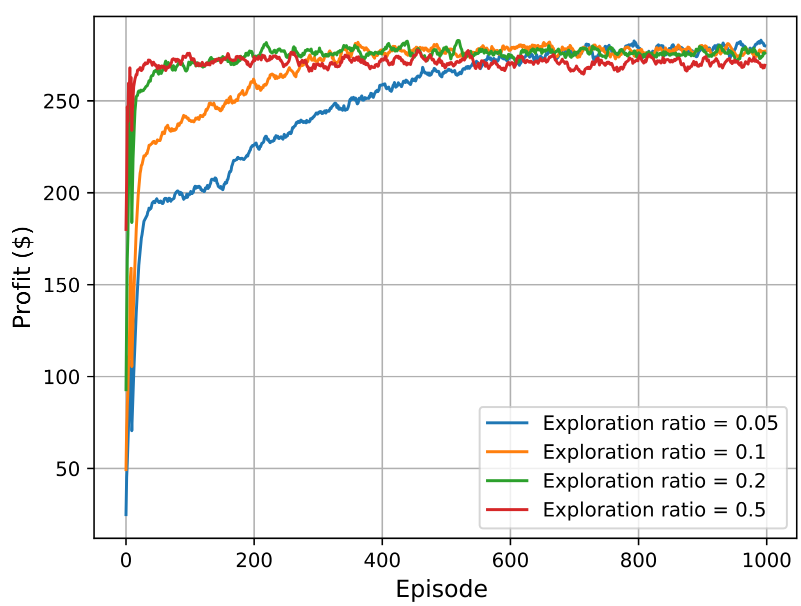


Figure 4.13: The obtained profit with different exploration ratios

at the beginning; however, the profit will gradually converge to a normal level, similar to that of other choices, in the long term. In contrast, a very high exploration ratio $\varepsilon = 0.5$ provides higher profit at the beginning but will suffer some loss in the long term because of the inefficient effort on random exploration.

Price margin

In Section II-B, the price margin ε_p is used to provide the price bar for customers who would like to participate in this market, as they have to guarantee trading partners' benefit for fairness instead of only maximizing their own profit. The sensitivity analysis of parameter epsilon is also provided in Figure 4.14 to demonstrate the effect of different values of epsilon for overall trading profit. On the other hand, the price margin is allowed to be 0 set by the REB, and most customers can bid/offer at the same price as the utility service, though they may suffer the risk of never being selected. However, increasing the price margin to a reasonable level can improve the REB's profit and the market efficiency through facilitating customers' trading as quickly and frequently as possible. Since the price margin can only affect the price samples statistically, we compare the results with different ε_p in a *violin plot*, as shown in Figure 4.14. It can be seen that with a little higher price margin, the total profit distribution from customers' trading will shift right towards a larger number.

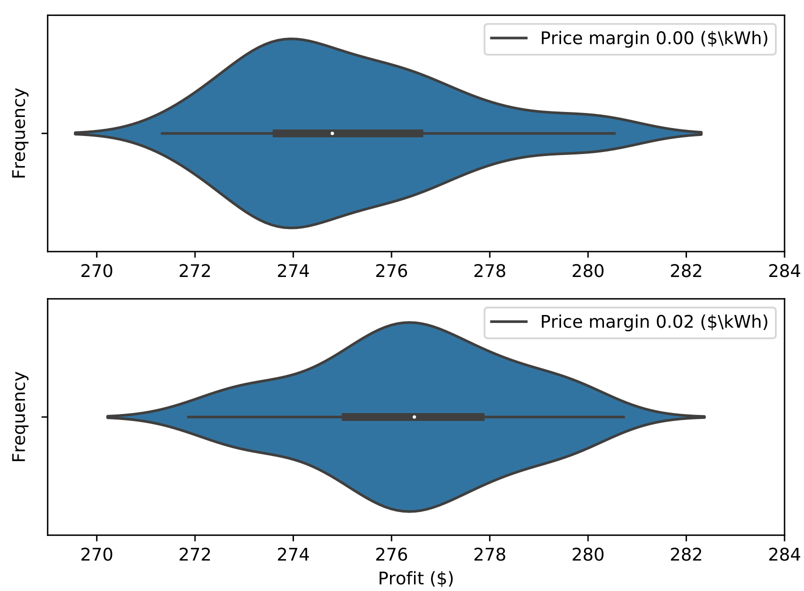


Figure 4.14: The effect of price margin ε_p

4.5.4 The impact of H and search friction

As this market operation is event-driven and only open occasionally to customers when there is a seasonal weather event, such as high solar irradiance or an energy supply shortage from the main grids. How actively participants join this local trading platform during these occasional scenarios will also determine the efficiency of the market operation, since some customers may quit the market if long-time waiting is needed for an occasional trading opportunity with few customers present. The kind of anxiety caused by staying idle in the market and *searching* for a suitable trading opportunity can be described qualitatively and quantitatively.

In this test case, we will try different values of parameter H, which is used to describe the remaining time counter, $h^s_{i,t^s_i}, h^b_{j,t^b_j} = H$, or how many chances each customer (seller or buyer) still has before they must quit the market, to see the impact on the final profit of the REB for the whole day. Some other parameters are the same as in the first case study with more customers, $|\mathscr{I}^s| = 200$ and $|\mathscr{I}^b| = 210$, distributed more intensively over all the time intervals. The result is averaged over 10 independent runs and shown in Figure 4.15. It matches the intuition that the more chances to enter or the longer customers are allowed to stay in the market, the more profit the REB can obtain by delaying its decision-making until the overlap of many a suitable bid and offer.

In microeconomics, the concept of *search friction* that arises in a search process can be used to explain such a phenomenon. A search process is seen as sequential sampling of a population X, in which the sample points x_t can be energy prices, quality of services, bids for an asset, and so on [150]. The 2010 Nobel Prize in Economics was awarded for an analysis of markets with search frictions [160]. Additionally, the optimal stopping problem is a typical application example involving search friction, which tries to solve a problem that involves choosing a suitable time to take a particular action, in order to

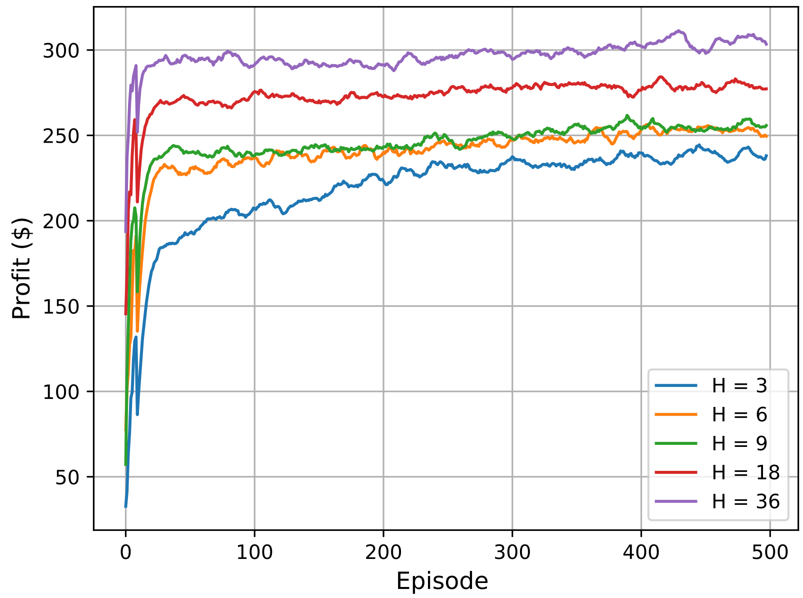


Figure 4.15: The impact of parameter H

maximize an expected reward or minimize an expected cost [161]. In this energy market model, if we go into the details of the single time interval decision problem, at the beginning of every interval t, the REB, as the decision-maker, will look forward to the following time horizon and suffer some type of loss when making its decision too early or too late. The REB cannot make a decision too early because some more suitable customers (i.e., with lower bids or higher offers) may come in the next few time intervals. Similarly, the REB cannot make a decision too late because it may regret the opportunity cost when some potential customers quit the market in the next few time intervals. This problem is also studied quantitatively, with a certain assumption regarding the probability distribution of customers' quality and market entering time, in our previous work [162]. The optimal stopping result can be estimated by eq. (4.17), where P_{dummy} is the baseline profit collected by the dummy strategy, and eq. (4.18) indicates the normalized profit in an optimal stopping problem while assuming an equal chance of profit distribution in each time interval [163]. Additionally, the reduced search friction of the learning method can also be obtained by averaging the profit gap in Figure 4.15. Some results associated with the search friction analysis are summarized in Table 4.3. In this paper, the actions *retention* and *partial-clearance* are actually designed accordingly for the purpose of reducing search friction while locking in a portion of the potential profit (partial-clearance) and increasing the chance of decision-making.

$$C_{friction} = n \times \min_{c=2,\ldots,H} \Big[V_T(c) - V_T(c-1) \Big] \times P_{dummy} \tag{4.17}$$

$$\begin{aligned} V_T(c) = \sum_{t=c}^{H-1} \left[\prod_{w=c}^{t-1} \left(\frac{w-1}{w} \right) \right] \left(\frac{1}{t+1} \right) + \\ 0.5 \left[\prod_{w=c}^{H-1} \left(\frac{w-1}{w} \right) \right] \end{aligned} \tag{4.18}$$

Table 4.3: Reduced search friction by learning method (daily)

Remaining time counter	Reduced cost	Analytical cost estimation
$H = 3$	\$5.96	\$23.33
$H = 6$	\$17.32	\$46.67
$H = 9$	\$19.12	\$62.22
$H = 18$	\$40.25	\$81.67
$H = 36$	\$63.45	\$97.22

4.6 Conclusion

In this paper, we studied how the retail energy broker facilitates indirect customer-to-customer energy trading using reinforcement learning techniques. The localized event-driven market model can provide customers or prosumers additional energy trading options, besides the conventional utility service, and further promote the deregulation of the retail electricity market, perhaps reshaping the future energy business landscape. It should be emphasized that this innovative local energy market does not aim to replace any existing energy service, nor is it proposed as a best market paradigm. Instead, it mainly seeks to diversify or even gamify the energy ecosystem at the edge of distribution networks and near end-users. Without loss of generality, this paper assumes that most local participants are geographically close to each other. In other words, they are most likely connected to the same distribution feeder. Therefore, this paper does not explicitly model the AC power flow constraint of distribution networks. In the future, we will further investigate the conceptual market model design under more realistic operating conditions. We may also try some virtual experience or experience replay based deep reinforcement learning techniques to improve the intelligence and learning capability of the market operator. It is believed that in the next-generation distribution network, more relevant energy business models and local energy trading platforms under incubation will come into practice and revolutionize the overall energy ecosystem.

CHAPTER V

Modeling of prosumer

5.1 Introduction

[1] In today's retail electricity market, customers have very limited ”energy choice” or freedom to choose different types of energy services. Although the installation of distributed energy resources (DERs) has become prevalent in many regions, most customers or prosumers who have local energy generation and possible surplus can still only choose to trade with utility companies. They either purchase energy from or sell energy surplus back to the utilities directly while suffering some price gap. The key to providing more energy trading freedom and open innovation in the retail electricity market is to develop consumer-centric business models and possibly a localized energy trading platform. Following these ideas in the current research community, the next-generation retail electricity market infrastructure will be a level playing field, where all customers have an equal opportunity to actively participate directly [16].

Recently, many a market mechanism has been proposed to support prosumers' participation, such as prosumer grid integration [55], peer-to-peer models [58][60], and prosumer community groups [138][139]. Direct integration of prosumers' DERs in the main grid with or without aggregators is still the most widely accepted idea because of its simple

[1] This chapter was planned to be published as an article in a peer-reviewed journal: [164] **T. Chen**, and W. Su, ”Local Energy Trading Behavior Modeling with Deep Reinforcement Learning”, prepared to be submitted.

applicability and manageability. The work in [55] describes a typical model that encourages the aggregation of prosumers' DERs and participation in the electricity market as a whole. However, in order to diversify the customer-side energy ecosystem and promote a deregulated market, prosumers' participation in a peer-to-peer (P2P) eBay-like market is also frequently discussed. In [58], a matching-mechanism-based market model is proposed to allow individual prosumers to meet each other to conduct bilateral trades. The concept of a prosumer community is studied in [138][144] for local energy sharing and trading. In another example, a new vision of local distribution systems with embedded trading functionality is also proposed [66], in which prosumers are encouraged to better balance their electricity usage in a local community through psychological balancing premiums and internal trading.

In this work, we aim to study how to use deep reinforcement learning to solve a decision-making problem of selling or buying energy in the local market, from the prosumers perspective. We aim to study prosumers' trading behavior in a similar local energy market (LEM), with utilization of new advances in reinforcement learning technologies. The description of the LEM will be given in Section 5.2, and the physical feeder connection of a prosumer is given in Figure 5.1. For simplification and specification, the prosumer is assumed to have only one wind turbine (WT) and only one energy storage system (ESS), which can be easily extended and modified for multiple installations and other types of renewable energy resources.

With increasing installation of distributed energy resources and development of deregulated retail electricity market, the prosumer will be able to make the best use of all the energy resources (e.g., DERs, ESS, flexible loads) available on-site to maximize its own benefit strategically. The trading process of a prosumer will be modeled as a Markov decision process (MDP) to fully account for the volatile market and physical conditions.

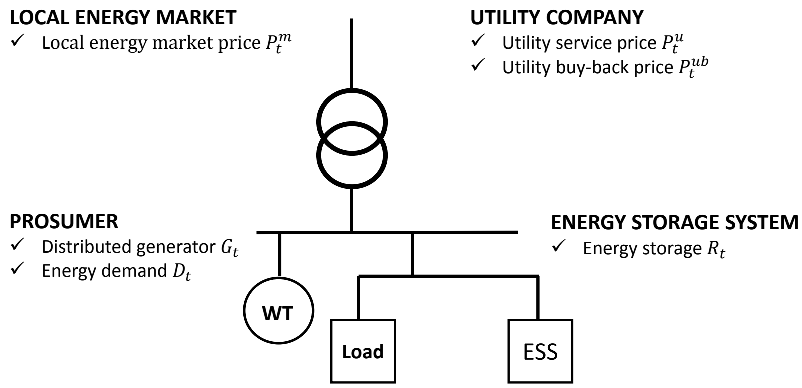


Figure 5.1: Prosumer with energy trading conditions

With the help of the deep reinforcement learning (DRL) method, the prosumer can exercise trading actions without analytical calculation (e.g., optimization) and knowledge of the market model. The use of DRL also abandons many assumptions due to discretization of various continuous variables. Although, the application of reinforcement learning techniques for microgrid energy management is commonly found recently in other recent works [165][166], the application of the DRL technique, which combines deep learning and reinforcement learning, is firstly studied in this paper, as it is especially suitable for such a unique local energy trading problem with many external data collections.

In this paper, our contribution is to: (1) propose a new paradigm of local energy trading in a carefully designed local energy market; (2) study the prosumer's trading strategy with well-defined trading actions; (3) explore the features of the deep reinforcement learning technique for dealing with a data-driven market model; and (4) provide some interesting findings through experiments on prosumers' participation.

5.2 Holistic local energy market model

The LEM is operated at the distribution level for facilitating local retail energy trading aided by a retail energy broker (REB). This market mechanism can provide extra energy transactions options for customers or prosumers, who have the willingness to directly participate in the retail electricity market, on top of the existing utility service. This LEM is event-driven, similar to [148], because it works only as a back-up trading platform, unnecessary to be open permanently for the whole year. It is affected by local requirements and seasonal events, such as high solar irradiance or energy shortages in regional grids. In this way, the future retail electricity market, including the LEM, can be shown as in Figure 5.2.

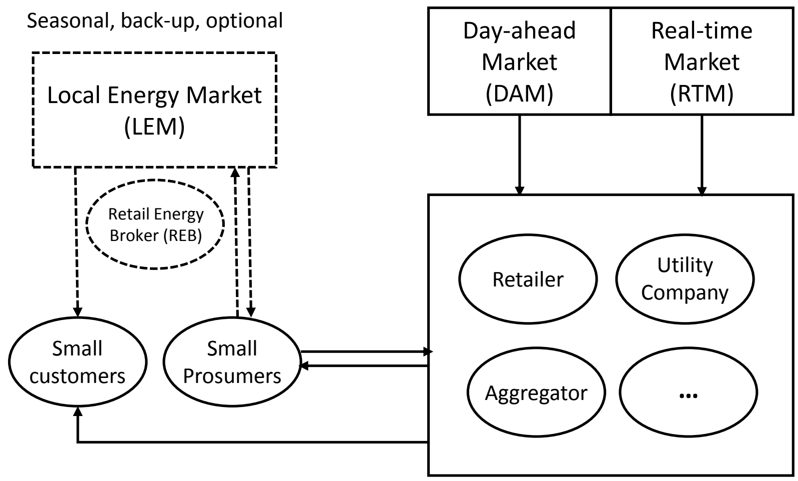


Figure 5.2: Local energy market with prosumer's participation

The motivation of introducing such a localized occasional energy market is twofold: 1) continuous deregulation, which may sound a little cliché, is still the driving force for reshaping the legacy energy landscape. The deregulation of the airline industry and telecommunications industry proves how it can benefit customers. The emergence of *Priceline* and *Uber* are also good examples of two of the few notable industries that contextualize both deregulation and new platforms that boost new business paradigm design and new ecosystem building within the existing market context; 2) there is a need for a new business model design and even gamification of the energy ecosystem. If successfully implemented, the new model will significantly promote the massive introduction of new players (e.g., REBs) in a more competitive retail market. Rather than passively accepting the trend of the emerging energy trading paradigm, we can help customers attain a better understanding of the future energy landscape. Taking the smartphone as an example, before the boom of the smartphone, mobile internet occupied only 5% of the overall internet connections. Nowadays, mobile internet connections have become mainstream and account for half of the overall connections. We should create the same ”desire” for local energy trading. Given these reasons, the LEM seeks to improve the efficiency of regular retail activity mainly through diversifying the existing energy business models (commercial strategies) and providing energy trading opportunities at the edge of the distribution network, near the source of the energy demand. The conceptual design of our holistic market model, as shown in Figure 5.3, and the major functionalities of each entity are presented.

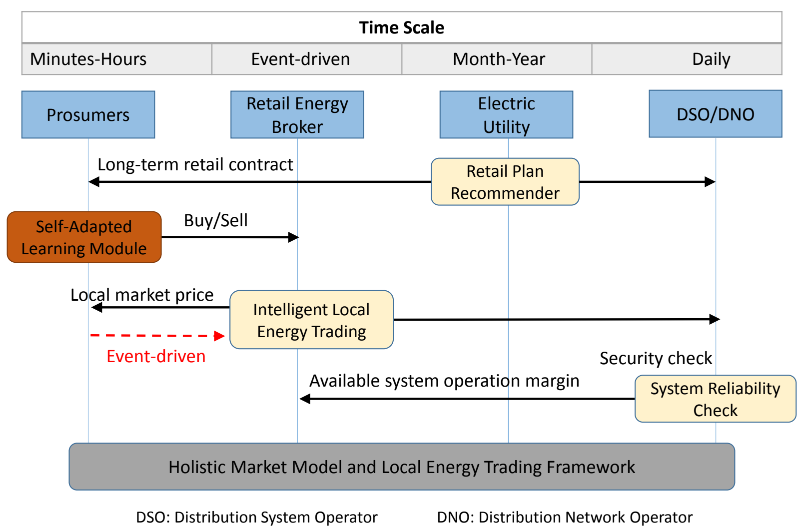


Figure 5.3: Holistic market model design

For example, the proposed LEM will facilitate short-term and immediate local energy transactions. The distribution system operator (DSO) or distribution network operator (DNO) is responsible for market regulation (e.g., reliability and security checks). The electric utility will not only serve customers as per usual but will also offer various long-term retail plans. Meanwhile, prosumers may develop their own trading strategies by using various types of energy devices (e.g., batteries, DERs, flexible loads) that are available. A local energy transaction will be physically fulfilled by leveraging the existing distribution line and smart meters for billing and payment. It is worth noting, however, that it is not our intention to cover every aspect of the aforementioned market model in a single paper. In this paper, we mainly focus on the self-adapted learning module for prosumers, as highlighted in an orange box in Figure 5.3. More description regarding other modules can be found in [137].

5.3 Prosumer model

A prosumer is always trying to benefit by making the best use of the energy resources available, while observing the energy trading conditions of the LEM and the utility company. As shown in Figure 5.1, we explicitly consider a wind turbine and only one ESS for a prosumer. However, the generic model described in this section can be easily applied for prosumers with other renewable energy resources and multiple ESSs.

5.3.1 Wind power generation model

The wind turbine is used as the major power supply unit owned by the prosumer, which is integrated into the system as a renewable source. The generation power output can be related to wind speed, approximately, by using the following function [167],

$$G_t(v) = \begin{cases} 0 & v < v_{ci} \text{ or } v > v_{co} \\ \frac{G_r(v-v_{ci})}{(v_r-v_{ci})} & v_{ci} \leq v \leq v_r \\ G_r & v_r \leq v \leq v_{co} \end{cases} \tag{5.1}$$

where v_{ci} is cut-in speed, v_{co} is cut-off speed, v_r is the rated wind speed, and G_r is the rated output power of the wind turbine.

5.3.2 Energy storage system model

The ESS is another one of the core parts of the energy management system for a prosumer. The strategy of operating the ESS (i.e., charging and discharging) significantly impacts the performance of the overall trading behavior. We consider the SOC at time t with current remaining energy storage R_t with charging and discharging power, P_{ch} and P_{dis},

$$SOC_{t+\Delta t} = SOC_t + \frac{P_{ch}\eta_c \Delta t}{B} - \frac{P_{dis}\Delta t}{B\eta_d} \tag{5.2}$$

$$R_{t+\Delta t} = R_t + E_{ch}\eta_c - E_{dis}/\eta_d \tag{5.3}$$

where B is the capacity (kWh) of the ESS, and η_c (η_d) (%) is the charging (discharging) efficiency.

Additionally, the ESS wear cost will be taken into account in the energy trading decision-making, in which myopic reliance on charging and discharging of the ESS to arbitrage in the local energy market is discouraged, since it may lead to an increased long-term cost caused by ESS degradation. An analysis of the effect of the weighting factor of the SOC and batter wear cost can be found in [168]. The empirical ESS wear cost coefficient κ ($\$/kWh$) is expressed as follows:

$$\kappa = \frac{C_i}{\eta_d B N_c \delta} \tag{5.4}$$

where C_i is the initial investment cost for the ESS, N_c is the corresponding number of life cycles at a rated depth of discharge (DOD), and δ is the DOD of the ESS.

5.3.3 Trading actions and utilities

In this section, we will define the local energy trading actions of the prosumer, which form its trading strategies and behaviors when participating in the local energy market.

The prosumer can choose to *buy* or *sell* energy with consideration for the energy deficit or surplus conditions and ESS operation. Usually, the ESS can have three statuses in each scenario, namely *charging*, *discharging* and *idle*. However, the *idle* status can be easily combined with either charging or discharging status when let $P_{ch} = 0$ or $P_{dis} = 0$. Thus, we define the energy trading actions with four possible options: $(buy, charge)$, $(buy, discharge)$, $(sell, charge)$ and $(sell, discharge)$. They are also indicated by a_{11}, a_{10}, a_{01} and a_{00}, respectively, with detailed explanation and justification given in the following subsections.

Buy and charge, a_{11}

This action is suitable for the scenario where the prosumer has little energy stored and a huge load demand that cannot be covered by its on-site generation. The price signals from the utility company and the LEM are also a very significant factor that will determine the trading benefit with consideration for the wear cost of the ESS. The utility function of this action is as follows:

$$u\Big(G_t, R_t, D_t \;\Big|\; a_{11}\Big) = \Big(P_t^u - P_t^m\Big) \times E_b - \kappa \times \Big(B - R_t\Big) \tag{5.5}$$

where,

$$E_b = max\big\{0, \quad (B - R_t)/\eta_c + D_t - G_t \Delta t\big\} \tag{5.6}$$

The underlying trading benefit mainly comes from the price gap between buying energy from the utility company directly, P_t^u, and purchasing in the LEM, P_t^m. The second term in (5.5) indicates the wear cost of the ESS. Purchasing energy amount E_b in (5.6) should be evaluated strategically according to the estimation of random variables G_t and D_t. The charging action is also assumed to be always fully charge.

Buy and discharge, a_{10}

This action is similar to a_{11}; however, it considers discharging of the ESS in the circumstance where the purchasing price is high and unnecessary energy import should be avoided. Both the energy obtained from purchasing and battery discharging can be used for the energy consumption. In other words, the ESS can be used for arbitrage purpose if the market price is extremely high.

$$u\Big(G_t, R_t, D_t \;\Big|\; a_{10}\Big) = \Big(P_t^u - P_t^m\Big) \times E_b - \kappa \times E_{dis} \tag{5.7}$$

where,

$$E_b = max\Big\{0, \quad D_t - G_t\Delta t - R_t\eta_d\Big\} \tag{5.8}$$

$$E_{dis} = min\Big\{max\Big\{0, D_t - G_t\Delta t\Big\}, R_t\eta_d\Big\} \tag{5.9}$$

The utility function is given in (5.7) with adjusted discharging energy amount E_{dis}.

Sell and charge, a_{01}

This action takes into account the situation where power surplus is huge and can be exported (sold). The trading benefit of a_{01} comes from the price gap between selling energy back to the utility company directly, P_t^{ub} and selling in the LEM, P_t^m. The utility function of this action is as follows:

$$u\Big(G_t, R_t, D_t \;\Big|\; a_{01}\Big) = \Big(P_t^m - P_t^{ub}\Big) \times E_s - \kappa \times E_{ch} \tag{5.10}$$

where,

$$E_s = max\Big\{0, \quad G_t\Delta t - (B - R_t)/\eta_c - D_t\Big\} \tag{5.11}$$

$$E_{ch} = min\Big\{max\Big\{0, G_t\Delta t - D_t\Big\}, (B - R_t)/\eta_c\Big\} \tag{5.12}$$

As presented in (5.11) and (5.12), it is also implicitly assumed that the power generation would give priority to the charging of the ESS. The remaining energy left after charging the ESS will then be sold in the LEM.

Sell and discharge, a_{00}

This action is similar to a_{01}; however, it considers the scenario of an extremely high selling price for a given arbitrage opportunity. The ESS will be fully discharged in order to earn extra revenue, and then charge again in other time intervals with a lower purchasing price. The utility function of doing so is given as follows:

$$u\Big(G_t, R_t, D_t \;\Big|\; a_{00}\Big) = \Big(P_t^m - P_t^{ub}\Big) \times E_s - \kappa \times R_t \tag{5.13}$$

where,

$$E_s = max\Big\{0, \quad R_t\eta_d + G_t\Delta t - D_t\Big\} \tag{5.14}$$

It is worth noting that all these actions cannot be intuitively decided only according to the observation of the current market price and generation-demand balance, because the random variables G_t and D_t possess high stochasticity and can only be realized at the end of every decision-making step. Besides, the prediction and historical knowledge of the market price, as well as its impact on ESS charging and discharging coordination, also contribute to the difficulty of choosing the most suitable action. However, a very intuitive rule-based dummy trading strategy is still provided in Section 5.5.2 to help with the benchmark analysis.

5.3.4 Self-adaptive learning problem

The trading strategy design problem is to maximize the total utility or economic benefit U_t for a prosumer across the overall time horizon, which can be written as:

$$\mathbf{P1} : \max\ U_t = \sum_{t=0}^{T} u\Big(G_t, R_t, D_t \;\Big|\; \pi(G_t, R_t, D_t, \hat{s}_t)\Big) \tag{5.15}$$

where $\hat{s}_t$ contains all the price-related uncertain information that can be realized and obtained at the beginning of each time interval while making decisions; G_t and D_t are random variables that can only be realized at the end of each time interval; $\pi(\cdot)$ is a policy function that determines the current trading action according to the uncertain market information obtained so far and predictions, as well as the estimation of random variables G_t and D_t using accumulated historical knowledges. In this way, the problem **P1** can actually also be described as a self-adaptive learning problem that deals with decision-making under many uncertainties with learning through historical knowledge. Further description of this learning problem and solution method will be provided in the next section.

5.4 Deep reinforcement learning and solution algorithms

DRL is a cutting-edge ML technique that has arisen just recently in the research community of intelligence systems, and has been experimented with to be able to achieve human-level or better control performance in various decision-making contexts [169]. DRL is poised to revolutionize the field of intelligence systems with many unpredictable achievements and represents a step towards building autonomous systems, such as prosumer communities or other intelligent energy systems, with a higher level understanding of the decision-making environment. In particular, the leveraging of deep learning is enabling reinforcement learning to scale to problems that were previously computationally intractable [170]. In theory, the DRL technique shares many common characteristics with conventional RL technology and is organized on top of an MDP model, which will be discussed in following subsections. Here, we use Table 5.1 to summarize the comparison between DRL and conventional RL.

The DRL technique is suitable for solving an MDP problem like local energy trading, because the prosumers' decision-making and trading behavior will highly depend on var-

Table 5.1: Reinforcement learning v.s. deep reinforcement learning

Features	Reinforcement learning	Deep reinforcement learning
Typical algorithm	Q-learning	Deep Q-Network
System scale	Usually deal with small or middle-scale problem for computational tractability;	Able to handle extremely large problem with leverage of deep learning techniques for data analytics;
Solution methods	Mostly work with tabular methods or policy-gradient methods;	Mostly use embedded neural network for value function approximation and actor-critic methods;
Function approximation	Inclined to use basic regression model or hand-crafted features;	Usually use deep neural network, such as CNN, RNN with LSTM as default; Ensemble methods are also commonly used;
Experience replay	Almost not used;	Use transition states buffer to realize experience replay with batch training;
Relation to other methods	Close relation to dynamic programing and simulation based methods;	Close relation to deep learning and most data-driven methods;

ious types of continuous variables that are hard to be discretized as state space in conventional Q-learning. Additionally, with the help of an experience replay module embedded in DRL, the best use of the stochastic market information and renewable energy generation, as well as prosumer's load demand pattern, will be possible.

5.4.1 Markov decision process

An MDP usually provides a mathematical description for situations where a system can be partly under the control of decision-making, while also partly random and independent of the control. An MDP model is described by state space $\mathscr{S}$, action space $\mathscr{A}$, reward function set $\mathscr{R}$, state transition probability matrix $\mathscr{P}$ and a discount factor $\gamma \in [0,1]$. It is often assumed that an MDP model keeps the Markov property, which implies the transition probabilities of a state are only affected by the previous one step with no memory. By following the principle of MDP modeling, we can design the reward function to be,

$$r_t = u\Big(G_t, R_t, D_t \;\Big|\; a_t\Big) \tag{5.16}$$

where $u(.|a_t)$ can be calculated according to prosumer's utility estimation (5.5)-(5.14) and affected by various forecasting or historical information organized in the state variable $s_t = \{X_t^T, \hat{s}_t, G_t, D_t, R_t\}$; where X_t^T is the time stamp, $X_t^T \in \{1, 2, 3, ..., T\}$, and $\hat{s}_t$ stores all the price-related uncertain information at time t, $\hat{s}_t = \{P_t^m, P_t^u, P_t^{ub}\}$. In this way, the total utility or economic benefit U_t for a prosumer beginning from time t can be written as:

$$U_t = \sum_{k=0}^{T} \gamma^k r_{t+k+1}\big(\pi(s_{t+k+1}), s_{t+k+1}\big) \tag{5.17}$$

It should be noted that the policy function $a_{t+k+1} = \pi(s_{t+k+1})$ will often affect the next time-step s_t, and most information stored in s_t is highly stochastic with dependent relationships between some of them. To explicitly model the decision-making process under uncertainties with consideration of expectations, the problem **P1** can be re-written as:

$$\mathbf{P2}: \quad \max_{\pi: \mathscr{S} \to \mathscr{A}} E\left[\sum_{t=0}^{T} (\gamma)^t r^t \Big(s^t, \pi(s^t)\Big)\right] \tag{5.18}$$

As our local energy trading model of customer-side learning is built as an MDP model, the problem is supposed to be solved efficiently by dynamic programming techniques, once the transition probability set is given. However, this is not the case, since the transition probability is usually extremely hard to estimate under most circumstances. Therefore, some standard reinforcement learning techniques, such as temporal-difference learning or specifically Q-learning, are supposed to be helpful in dealing with model-free problems, in which no information regarding transition probability is needed. On the other hand, there will also be some issues associated with the model-free techniques in practice, namely how to deal with the continuous state space. For instance, the market status $\mathscr{S}$

in this paper consists of various types of contentious variables, such as price information, energy demand and generation, and ESS status, which are hard to enumerate or even discretize. In order to deal with the continuous state space and overcome the disadvantages of tabular methods [152], we leverage deep neural network (DNN) technology to realize a deep Q-learning algorithm for local energy trading (DQL-LET), similar to another successfully applied algorithm DQN in playing video game Atari [171]. In the next sections, the DQL-LET algorithm and its novel characteristics, especially experience replay, will be described to help solve a continuous, highly stochastic local energy trading problem.

5.4.2 Deep Q-learning algorithm

Similar to the basic Q-learning algorithm, the overall utility of the prosumer can be evaluated by value-function $Q(\cdot)$ or $V(\cdot)$, and the optimal stationary policy π^* can be well defined by using the optimal *action-value function* $Q^* : \mathscr{S} \times \mathscr{A} \to \mathscr{R}$, which satisfies the following Bellman optimality equation:

$$Q^*(s,a) = r(s,a) + \gamma \sum_{s' \in \mathscr{S}} p\left(s' \middle| s,a\right) V^*\left(s'\right), \tag{5.19}$$

where $V^*\left(s'\right)$ is the optimal *state-value function* [152], which is defined as

$$V^*\left(s'\right) = \min_{a \in \mathscr{A}} Q^*(s',a), \quad \forall s \in \mathscr{S}. \tag{5.20}$$

Since $V^*\left(s'\right)$ is the expected discounted system cost with action a in state s, we can obtain the optimal stationary policy:

$$\pi^*\left(s\right) = \mathrm{argmin}_{a \in \mathscr{A}} Q^*(s,a). \tag{5.21}$$

In maximizing the overall utility U_t since time t, any model-free learning based method, unlike dynamic programming methods [158], can provide a solution without acquisition of the state transition probabilities $p_s\left(s_{t+1} \middle| s_t, a_t\right), \forall s \in \mathscr{S}$ a priori [159]. Furthermore, DRL will enable an intelligent agent to make decision without discretization of the continuous state space and make better use of historical data samples, which is called *experience replay*.

In the RL research community, a linear function approximator based on hand-crafted features is often used to estimate the action-value function,

$$Q(s,a;\theta) \approx Q^*(s,a). \tag{5.22}$$

However, a non-linear function approximator, such as DNN, can be used instead with features extracted automatically. In this way, we can refer to this neural network function approximator with weights θ as a Q-network. A Q-network can be trained by minimizing a sequence of loss functions $L_k(\theta_k)$ that changes at each iteration k,

$$L_k(\theta_k) = \mathbb{E}_{s,a\sim\rho(.)}\Big[\big(y_k - Q(s,a;\theta_k)\big)^2\Big] \tag{5.23}$$

where $y_k = \mathbb{E}_{s'\sim\mathscr{S}}\big[r + \gamma \max_{a'} Q(s',a';\theta_{k-1})|s,a\big]$ is the target for iteration k and $\rho(s,a)$ is a probability distribution (i.e., *behavior distribution*) over sequences s and actions a. It is worth mentioning that in contrast with the fixed targets used for supervised learning, in DRL, the targets depend on the network weights. Thus, differentiating the loss function (5.23) with respect to the weights can be written as follows,

$$\begin{aligned}\Delta_{\theta_k} L_k(\theta_k) =& \mathbb{E}_{s,a\sim\rho(.);s'\sim\mathscr{S}}\Big[\big(r + \gamma \max_{a'} Q(s',a';\theta_{k-1}) \\ &- Q(s,a;\theta_k)\big)\Delta_{\theta_k} Q(s,a;\theta_k)\Big]\end{aligned} \tag{5.24}$$

Rather than computing the full expectations of the gradient, it is often computationally expedient to optimize the loss function by stochastic gradient descent. Based on the RL framework and DNN function approximation for evaluating a prosumer's current trading strategy, Algorithm 2 is presented to facilitate local energy trading and is named after the deep Q-learning for local energy trading (DQL-LET) algorithm.

Algorithm 2 Deep Q-Learning for Local Energy Trading (DQL-LET) Algorithm

Initialize trading replay memory $\mathbb{D}$ to capacity N_{max}
Initialize DNN based action-value function Q with random weights
Repeat for each episode $= 1, ..., M$:
Collect the current market, ESS and demand conditions
Forecast the wind power generation via (5.1)
  Initialize sequence $s_1 = \{X_1^T, P_1^m, P_1^u, P_1^{ub}, G_1, D_1, R_1\}$ and preprocessed sequenced $\phi_1 = \phi(s_1)$
  Repeat for each time step of episode, $t = 1, ..., T$:
    With probability ε select a random action $a_t \in \{a_{00}, a_{01}, a_{10}, a_{11}\}$, otherwise select $a_t = \max_a Q^*\big(\phi(s_t), a; \theta\big)$
    Execute trading action a_t in emulator and observe reward $r_t = u(.|a_t)$ via (5.5)-(5.14)
    Set $s_{t+1} = s_t, a_t$ and preprocess $\phi_{t+1} = \phi(s_{t+1})$
    Store transition $(\phi_t, a_t, r_t, \phi_{t+1})$ in trading replay memory $\mathbb{D}$
    If $|\mathbb{D}| > batch\_size$
      Sample random minibatch of transitions $(\phi_j, a_j, r_j, \phi_{j+1})$ from $\mathbb{D}$
      Set $y_j = \begin{cases} r_j & \text{for terminal } \phi_{j+1} \\ r_j + \gamma \max_{a'} Q(\phi_{j+1}, a'; \theta) & \text{for non-terminal } \phi_{j+1} \end{cases}$
      Perform a gradient descent step on $\big(y_j - Q(\phi_j, a_j; \theta)\big)^2$ according to (5.24)
      If exploration rate $\varepsilon > \varepsilon_{min}$
        $\varepsilon \leftarrow \varepsilon \times \varepsilon_{decay}$
  End loop
End loop

In this paper, we will not go deep into the theoretic analysis of the algorithm convergence and boundary conditions, which are still under study with many cutting edge research topics. Interested readers may see [170] and [171] for additional information. That said, the features of DQL-LET with experience replay are further discussed here.

5.4.3 Experience replay

One of the challenges of general DQL is that the neural network used in the algorithm tends to forget its previous experiences as it overwrites them with new experiences, which is similar to updating of the Q table in Q-learning. So we need to maintain a list of previous experiences and observations to re-train the model with its previous experiences. This list of previous experiences forms a dataset of transitions and can be called the replay buffer, as shown in Figure 5.4.

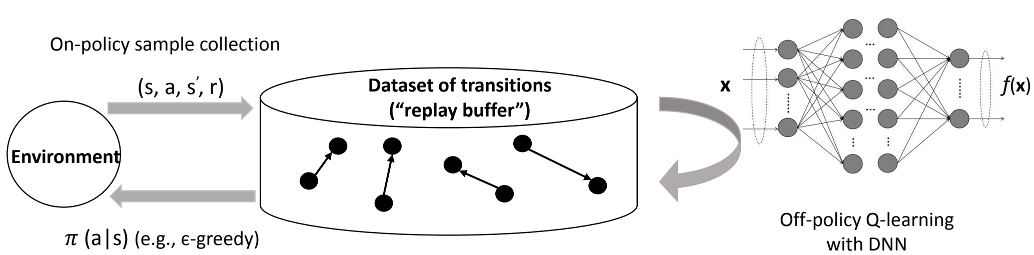


Figure 5.4: Experience replay with transition state buffer

The experience replay procedure stores transition samples and repeatedly presents them to a gradient and DNN-based, incremental RL algorithm. Thus, we can make better use of the computational efficiency of the underlying gradient-based algorithm, while also obtaining high data efficiency by reusing the samples. Meanwhile, the approximate *Q*-function and the greedy policy are able to remain constant during the interval between consecutive parameter updates of the DNN. The term *trajectory* can be used to refer to every sequence of *batch_size* samples (mini-batch) collected between two consecutive updates [172]. If the weights are updated after every time-step, and the expectations are replaced by single samples from the behavior distribution $\rho(\cdot)$ and the emulator $\mathscr{S}$ respectively, then the algorithm will degenerate to the conventional *Q-learning* algorithm.

5.5 Numerical results

In this section, we provide numerical results to show some interesting observations regarding the proposed prosumer energy trading behavior model and evaluate the performance of the algorithm for local energy trading strategies. Unless specifically indicated, the timeline, one day, consists of 24 time intervals, each of which lasts for 1 *hour*. The simulation environment is Python 3.6 running on a desktop with an Intel i7 and 16.0 GB RAM.

5.5.1 Simulation setup

Since the proposed market design of local energy trading is still at a conceptual level without field test data, some distribution generated data are used first to test the validity of the DQL-LET algorithm. The system parameters are given in Table 5.2 with battery model parameters provided in Table 5.3. Some realistic dataset representing the load demand will be assumed for several different prosumers to demonstrate the economic benefit within the proposed framework of local energy trading in section 5.5.3.

Table 5.2: Parameters of the system model

Parameters	Values
P^m	$\mathscr{U}\Big(0.42, 0.48\Big)$
P^u	$\mathscr{U}\Big(0.47, 0.53\Big)$
P^{ub}	$\mathscr{U}\Big(0.37, 0.43\Big)$
G	$G_{noise} \in \mathscr{N}\Big(0, 1.2^2\Big)$, $G_{max} = 12$, $G_{min} = 3$
D	$D_{noise} \in \mathscr{N}\Big(0, 1.1^2\Big)$, $D_{max} = 10$, $D_{min} = 0$
	$\varepsilon = 0.1$, $\varepsilon_{decay} = 0.995$, $\varepsilon_{min} = 0.01$, $\alpha = 0.001$ $N_T = 24$, $\eta_d = \eta_c = 0.90$, $batch\_size = 32$, $\gamma = 0.95$

Table 5.3: Battery parameters

Battery	Lead-Acid
Type	2 V/1000 Ah
Capacity	20 kWh
Minimum limit	3 kWh
Cycle life	1000 @ 30% DOD
Charging and discharging efficiencies	90%
Battery cost	$70 per kWh
Installation cost	$25 per kWh
Transportation cost	$20 per kWh

The stochastic load demand and wind generation are assumed based on typical generation $\widehat{G}_t$ and demand $\widehat{D}_t$ curves (as shown in Figure 5.5) of a typical prosumer with modifications as follows:

$$G_t = min\Big\{max\{\widehat{G}_t + G_{noise}, G_{min}\}, G_{max}\Big\} \tag{5.25}$$

$$D_t = min\Big\{max\{\widehat{D}_t + D_{noise}, D_{min}\}, D_{max}\Big\} \tag{5.26}$$

For the neural network embedded in the DQL-LET algorithm, we use three hidden layers with fully connected architecture and *ReLU* as the activation function. The network is trained with an adaptive stochastic gradient descent based on estimates of lower-order moments (i.e., Adam [173]), which can speed up convergence and is appropriate for problems with very noisy and/or sparse gradients.

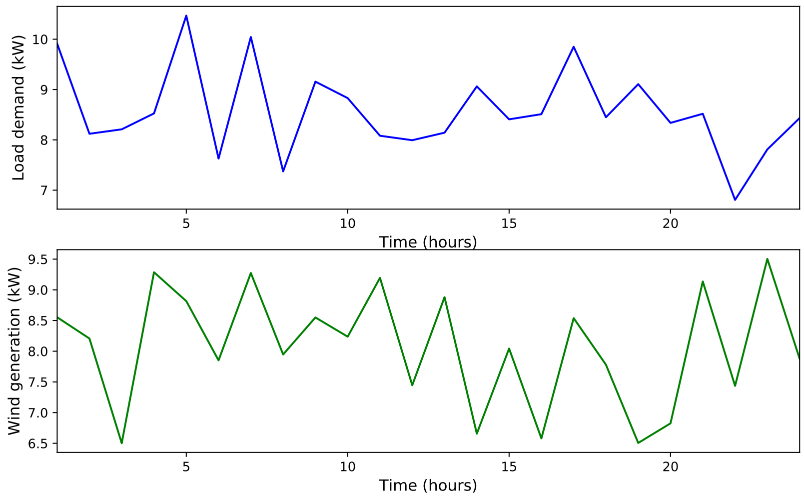


Figure 5.5: Wind generation $\widehat{G}_t$ and typical load demand $\widehat{D}_t$ of a prosumer

5.5.2 Performance evaluation of the proposed method

In this case study, we will present the advantages of using DRL technology for guiding prosumers' trading behaviors in the LEM. We call the trading strategy aided by the DQL-LET algorithm a smart strategy, and in contrast, we call arbitrary trading actions a dummy random strategy, in which the prosumer chooses the trading action randomly without any analysis. In addition, a rule based trading strategy is also designed as a benchmark in Figure 5.6 to reflect prosumers' intuitive trading behaviors with reasonable choices based on myopic analysis.

As shown in Figure 5.6, the prosumer will make a decision mainly depending on the current market information obtained so far. The prosumer will try its best to sell or buy in the LEM immediately if the local market price is significantly too high, $P^m > P^u$, or too low, $P^m < P^{ub}$, using the utility price as a reference. While the local market price is among a certain range, $P^u < P^m < P^{ub}$, the prosumer will have multiple choices according to his/her own estimation of the demand load and generation. However, perfect coordination between the future market price and battery storage, as well as accurate estimation of stochastic demand and generation, is hard to be captured by a prosumer's intuition.

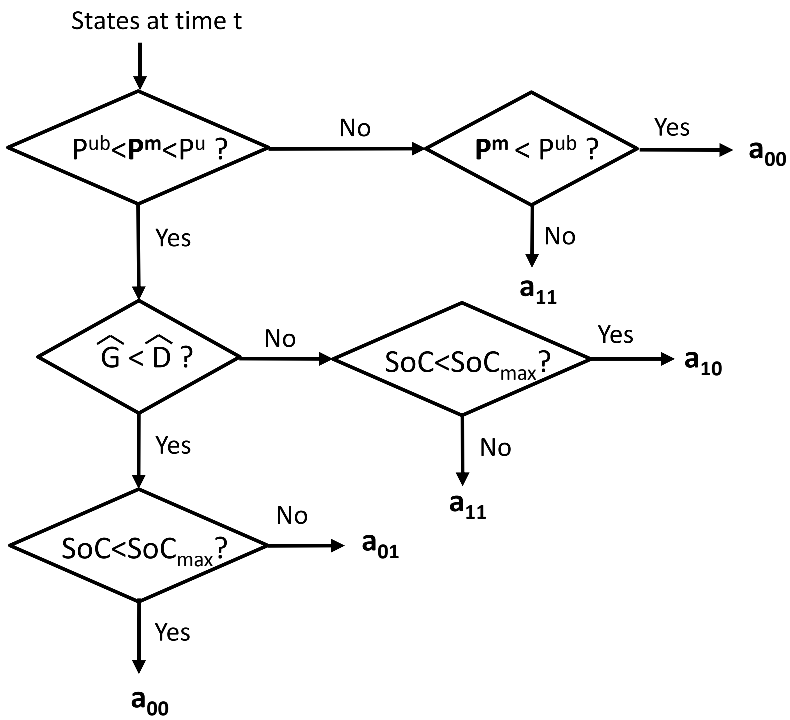


Figure 5.6: Intuitive rule based trading strategy

We can easily observe in Figure 5.7 that the smart strategy with DQL-LET outperforms significantly both the rule-based strategy and dummy random strategy.

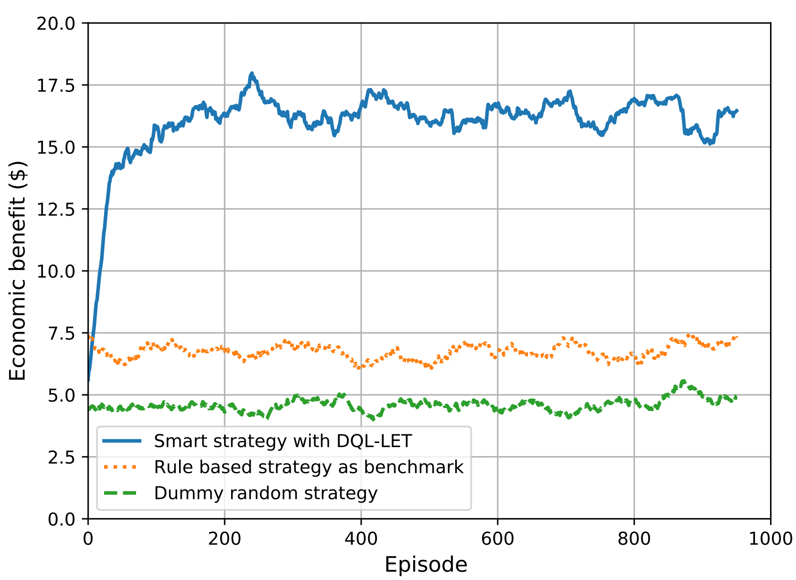


Figure 5.7: Prosumer's daily economic benefit in local energy trading

In addition, it should be emphasized that as deep reinforcement learning technology is usually data-driven and dependent on historical experience replay, it cannot guarantee the learned trading strategy is deterministic and the best every time, every day. We can only conclude that the learning algorithm will perform very well in a statistical way in the long term.

5.5.3 Economic analysis of the prosumer

This case study uses a load demand dataset collected from a realistic Finnish distribution system operator in northern Europe, which includes 3,398 non-empty low voltage customers in a small region [120]. We randomly picked nine customers out of them to analyze the annual economic benefit of these prosumers when participating in local energy trading intelligently. They are assumed to have similar wind turbines and ESSs installed. As shown in Figure 5.8, the load demand data is collected within 20,999 hours from 2010.06.10 to 2012.10.31.

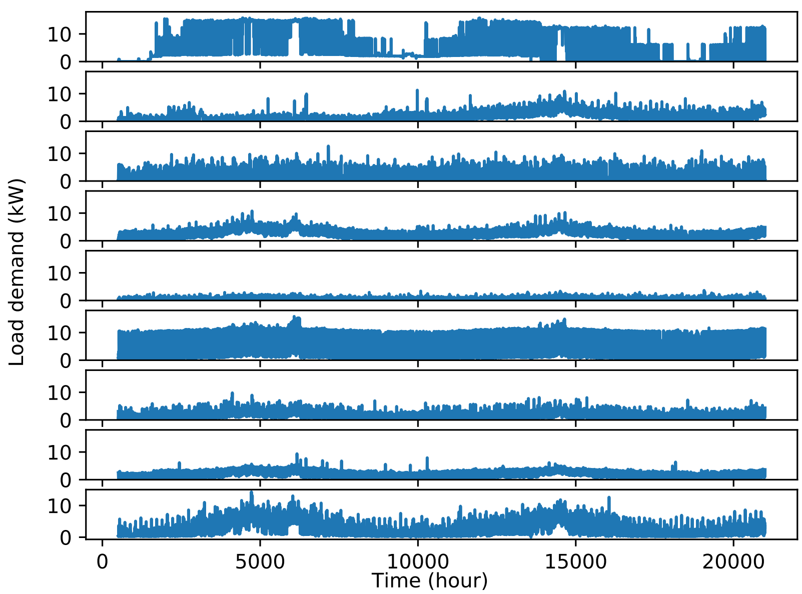


Figure 5.8: The load demand curves of random selected 9 customers , numbering customers from top (No.1) to bottom (No.9)

In order to compare prosumers' economic benefit from trading in the LEM with trading directly with the utility company and fully consider the stochasticity of renewable generation and load demand, as well as uncertain market conditions, under different scenarios, the average annual economic benefit for each prosumer can be calculated as follows:

$$\overline{R_{year}} = 365 \times r_0 \times \frac{1}{N_d N_s} \sum_{k=1}^{N_d} \sum_{i=1}^{N_s} R_{day}(i,k) \tag{5.27}$$

where $\overline{R_{year}}$ and R_{day} indicate the average annual benefit and daily benefit achieved from trading in the LEM, respectively; r_0 indicates the occasional market open rate determined by an event-driven LEM. As mentioned before, since the smart trading strategy outperforms others in a nondeterministic way, the simulation will run for $N_d = 20,999/24 \approx 874$ days (episodes) for different wind generation scenarios and power output (*low*, *medium* and *high*), with $N_s = 3$. The final result with $r_0 = 20\%$ and consideration of whether an ESS is available or not is shown in Figure 5.9.

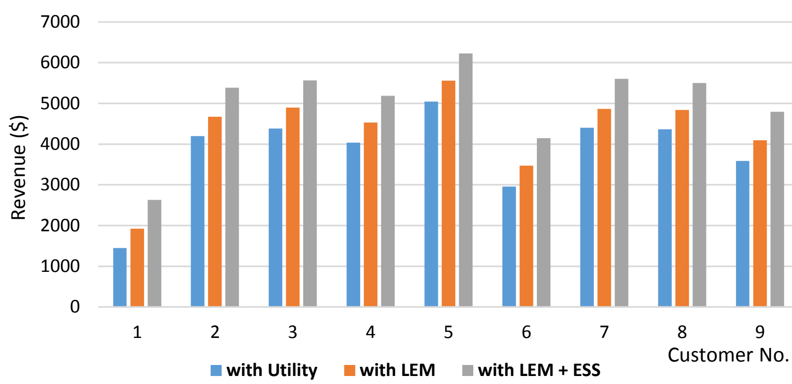


Figure 5.9: The annual economic benefit of each prosumer

5.6 Conclusion

In this paper, we proposed an event-driven local energy market for facilitating energy trading at the distribution level. We also studied how the prosumer chooses its local energy trading strategies given the available energy resources. The problem is built as an MDP and is solved by using new advances in reinforcement learning technology, namely deep reinforcement learning. It should be emphasized that this innovative local energy market does not aim to replace any existing energy service, nor is it proposed as a best market paradigm. Instead, it mainly seeks to diversify or even gamify the energy ecosystem at the edge of distribution networks and near end-users. In the future, some additional energy business models and services provided by the local energy market will be carefully designed. It is also believed that the active participation of end-users will revolutionize the overall energy ecosystem and reshape the energy business landscape.

CHAPTER VI

Conclusion and future work

The massive technology and industry transformation is rapidly reshaping the upgraded power system, also known as Smart Grid, to help improving the energy efficiency. On top of many upgraded energy infrastructures and new market mechanism, the further deregulation of retail electricity market at distribution level will play a important role in promoting such transformation in a socioeconomic way. It is believed that in the next-generation distribution network, more relevant energy business models and local energy trading platforms under incubation will come into practice and revolutionize the energy ecosystem for end-users. Thus, understanding the deregulated retail electricity markets in the future from a perspective of machine learning and optimization is very inspiring for facilitating intelligent operation of the economy-engineering nexus. The global Smart Grid and electricity market development is expected to continue its growth momentum in the next decade. This dissertation proposed a localized electricity market and discussed opportunities of emerging technologies for intelligent energy system operation. Chapter 4 studies a decision-making process aided by reinforcement learning for the retail energy broker, who plays a role of trader or middleman. While chapter 5 considers a decision-making process for an individual prosumer with suitable operation of the energy storage system. This dissertation also suggested many improvement and future work for the proposed framework and methodology.

The main contributions of this dissertation are as follows:

(1) Provided a new holistic localized electricity retail market design for future local energy trading happening at distribution level;

(2) Explored the utilization of machine learning (reinforcement learning) technologies for new paradigm of power system operation and future smart grid;

(3) Combined the adaptive data-driven methods with the static optimization framework to solve a decision-making problem within distribution network;

(4) Suggested new paradigm of research framework based on agent-based modeling for studying interactions among various kinds of energy entities.

In the future, this work can be improved and enhanced from different aspects: 1) for methodology, we can further improve the reinforcement learning efficiency by leverage of actor-critic methods, and combine the data-driven optimization methods, such as distributionally robust optimization, with machine learning methods; 2) for modeling, we may consider multi-type batteries, solar panels, combined-heat-pumps and thermostatically controlled loads; 3) for implementation, we will use some agent-based simulation software packages, such as Repast Symphony and JADE, to consider multi-agents interactions and group behavior in the proposed localized electricity market.

Based on the various aforementioned studies of the retail electricity market in recent years, some trends can be easily observed that: (1) the system or market operation is more fine-grained from different perspectives, trying to balance credits' assignment and benefit sharing among many types of market entities, including suppliers, speculative retailers, utilities, service providers, customers and other new parties introduced by new business models; (2) more and more consideration is given for economic operation on top of pure system requirement satisfaction, and a certain degree of risk is acceptable given the improving uncertainty of the whole system; (3) customers are expected to be more active in this market-loop instead of passive participants, which are allowed to directly interact with other market participants and exercise negotiation power. The study of the electricity market is more or less not a pure technique problem, especially considering the fairness rule (e.g., non-discrimination), data privacy and renewable energy subsidy policy in the retail electricity market close to the customer side. As concluded, this work will also contribute to the socioeconomic development of our society and energy sector reform with consideration for the societal impact on customer-customer interactions, energy communities, and energy policy making. Simply put, the changes brought by the deregulated retail market will be far beyond the scope of merely the engineering and economic fields. Its development will ultimately benefit the value monetization of new technologies and produce ways in which human beings can interact better with each other. Although it is hard to say how long to realize such an open and diverse energy ecosystem. Retail electricity markets are truly reaching something of a breaking point now. This is because the fundamental and underlying architecture of electric power systems are changing in major ways, which will enable a tremendous transformation in just a few decades.

BIBLIOGRAPHY

[1] T. Chen, Q. Alsafasfeh, H. Pourbabak, and W. Su, “The next-generation us retail electricity market with customers and prosumersa bibliographical survey,” *Energies*, vol. 11, no. 1, p. 8, 2017.

[2] “The public utility regulatory policies act. smithsonian museum of american history,” p. http://americanhistory.si.edu/powering/past/history4.htm.

[3] Y. Yan, Y. Qian, H. Sharif, and D. Tipper, “A survey on smart grid communication infrastructures: Motivations, requirements and challenges,” *IEEE communications surveys & tutorials*, vol. 15, no. 1, pp. 5–20, 2013.

[4] M. Prabavathi and R. Gnanadass, “Energy bidding strategies for restructured electricity market,” *International Journal of Electrical Power & Energy Systems*, vol. 64, pp. 956–966, 2015.

[5] W. W. Hogan, “Financial transmission rights: point-to point formulations,” in *Financial Transmission Rights*, pp. 1–48, Springer, 2013.

[6] K. Zhou, S. Yang, and Z. Shao, “Energy internet: the business perspective,” *Applied Energy*, vol. 178, pp. 212–222, 2016.

[7] R. Abe, H. Taoka, and D. McQuilkin, “Digital grid: Communicative electrical grids of the future,” *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 399–410, 2011.

[8] H. Rudnick, “Chile: Pioneer in deregulation of the electric power sector,” *IEEE Power Engineering Review (Institute of Electrical and Electronics Engineers);(United States)*, vol. 14, no. 6, 1994.

[9] R. Green, “Electricity liberalisation in europehow competitive will it be?,” *Energy Policy*, vol. 34, no. 16, pp. 2532–2541, 2006.

[10] S. Xu and W. Chen, “The reform of electricity power sector in the pr of china,” *Energy Policy*, vol. 34, no. 16, pp. 2455–2465, 2006.

[11] N. Wang and G. Mogi, “Deregulation, market competition, and innovation of utilities: Evidence from japanese electric sector,” *Energy Policy*, vol. 111, pp. 403–413, 2017.

[12] W. Gonzalez, “Restructured states, retail competition, and market-based generation rates,” *Smart Rate Design for a Smart Future*, 2015.

[13] D. S. Kirschen, “Demand-side view of electricity markets,” *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 520–527, 2003.

[14] R. Deng, Z. Yang, M.-Y. Chow, and J. Chen, “A survey on demand response in smart grids: Mathematical models and approaches,” *IEEE Transactions on Industrial Informatics*, vol. 11, no. 3, pp. 570–582, 2015.

[15] W. Saad, A. L. Glass, N. B. Mandayam, and H. V. Poor, “Toward a consumer-centric grid: A behavioral perspective,” *Proceedings of the IEEE*, vol. 104, no. 4, pp. 865–882, 2016.

[16] W. Su, “The role of customers in the us electricity market: Past, present and future,” *The Electricity Journal*, vol. 27, no. 7, pp. 112–125, 2014.

[17] J. Yang, J. Zhao, F. Luo, F. Wen, and Z. Y. Dong, “Decision-making for electricity retailers: A brief survey,” *IEEE Transactions on Smart Grid*, 2017.

[18] T. Chen, H. Pourbabak, Z. Liang, and W. Su, “An integrated evoucher mechanism for flexible loads in real-time retail electricity market,” *IEEE Access*, vol. 5, pp. 2101–2110, 2017.

[19] T. Chen, H. Pourbabak, and W. Su, "A game theoretic approach to analyze the dynamic interactions of multiple residential prosumers considering power flow constraints," in *2016 IEEE Power & Energy Society General Meeting*, IEEE, 2016.

[20] Z. Xu, D. S. Callaway, Z. Hu, and Y. Song, "Hierarchical coordination of heterogeneous flexible loads," *IEEE Transactions on Smart Grid*, vol. 31, no. 6, pp. 4206–4216, 2016.

[21] D. S. Callaway and I. A. Hiskens, "Achieving controllability of electric loads," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 184–199, 2011.

[22] W. Su and M.-Y. Chow, "Performance evaluation of an eda-based large-scale plug-in hybrid electric vehicle charging algorithm," *IEEE transactions on smart grid*, vol. 3, no. 1, pp. 308–315, 2012.

[23] K. Kok and S. Widergren, "A society of devices: Integrating intelligent distributed resources with transactive energy," *IEEE Power and Energy Magazine*, vol. 14, no. 3, pp. 34–45, 2016.

[24] Z. Liu, Q. Wu, S. Huang, and H. Zhao, "Transactive energy: A review of state of the art and implementation," in *12th IEEE Power and Energy Society PowerTech Conference*, IEEE, 2017.

[25] H. Hao, C. D. Corbin, K. Kalsi, and R. G. Pratt, "Transactive control of commercial buildings for demand response," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 774–783, 2017.

[26] N. Liu, X. Yu, C. Wang, C. Li, L. Ma, and J. Lei, "An energy sharing model with price-based demand response for microgrids of peer-to-peer prosumers," *IEEE Transactions on Power Systems*, 2017.

[27] D. Apostolopoulou, S. Bahramirad, and A. Khodaei, "The interface of power: Moving toward distribution system operators," *IEEE Power and Energy Magazine*, vol. 14, no. 3, pp. 46–51, 2016.

[28] T. Zheng and E. Litvinov, "On ex post pricing in the real-time electricity market," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 153–164, 2011.

[29] F. Li and R. Bo, "Dcopf-based lmp simulation: algorithm, comparison with acopf, and sensitivity," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1475–1485, 2007.

[30] M. Caramanis, E. Ntakou, W. W. Hogan, A. Chakrabortty, and J. Schoene, "Co-optimization of power and reserves in dynamic t&d power markets with nondispatchable renewable generation and distributed energy resources," *Proceedings of the IEEE*, vol. 104, no. 4, pp. 807–836, 2016.

[31] R. Li, Q. Wu, and S. S. Oren, "Distribution locational marginal pricing for optimal electric vehicle charging management," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 203–211, 2014.

[32] R. Yang and Y. Zhang, "Three-phase ac optimal power flow based distribution locational marginal price: Preprint," tech. rep., NREL (National Renewable Energy Laboratory (NREL), Golden, CO (United States)), 2017.

[33] A. J. Conejo, M. Carrión, and J. M. Morales, *Decision making under uncertainty in electricity markets*, vol. 1. Springer, 2010.

[34] W. Kong, Z. Y. Dong, D. J. Hill, F. Luo, and Y. Xu, "Short-term residential load forecasting based on resident behaviour learning," *IEEE Transactions on Power Systems*, 2017.

[35] B. Stephen, X. Tang, P. R. Harvey, S. Galloway, and K. I. Jennett, "Incorporating practice theory in sub-profile models for short term aggregated residential load forecasting," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1591–1598, 2017.

[36] A. A. S. de la Nieta, V. González, and J. Contreras, "Portfolio decision of short-term electricity forecasted prices through stochastic programming," *Forecasting Models of Electricity Prices*, p. 203, 2017.

[37] A. Calvo-Silvosa, S. I. Antelo, I. Soares, *et al.*, "Energy planning and modern portfolio theory: A review," *Renewable and Sustainable Energy Reviews*, vol. 77, pp. 636–651, 2017.

[38] M. Song and M. Amelin, "Purchase bidding strategy for a retailer with flexible demands in day-ahead electricity market," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 1839–1850, 2017.

[39] A. Hatami, H. Seifi, and M. K. Sheikh-El-Eslami, "A stochastic-based decision-making framework for an electricity retailer: time-of-use pricing and electricity portfolio optimization," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 1808–1816, 2011.

[40] R. T. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of banking & finance*, vol. 26, no. 7, pp. 1443–1471, 2002.

[41] N. Nezamoddini and Y. Wang, "Real-time electricity pricing for industrial customers: Survey and case studies in the united states," *Applied Energy*, vol. 195, pp. 1023–1037, 2017.

[42] S. Gyamfi, S. Krumdieck, and T. Urmee, "Residential peak electricity demand responsehighlights of some behavioural issues," *Renewable and Sustainable Energy Reviews*, vol. 25, pp. 71–77, 2013.

[43] W. W. Hogan, "Time-of-use rates and real-time prices," *Online at http://www. hks. harvard. edu/fs/whogan/Hogan_TOU_RTP_Newar k_082314. pdf*, 2014.

[44] E. Celebi and J. D. Fuller, "Time-of-use pricing in electricity markets under different market structures," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1170–1181, 2012.

[45] V. Robu, M. Vinyals, A. Rogers, and N. Jennings, "Efficient buyer groups with prediction-of-use electricity tariffs," *IEEE Transactions on Smart Grid*, 2017.

[46] H. Zhong, L. Xie, and Q. Xia, "Coupon incentive-based demand response: Theory and case study," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1266–1276, 2013.

[47] B. Xia, H. Ming, K.-Y. Lee, Y. Li, Y. Zhou, S. Bansal, S. Shakkottai, and L. Xie, "Energycoupon: A case study on incentive-based demand response in smart grid," in *Proceedings of the Eighth International Conference on Future Energy Systems*, pp. 80–90, ACM, 2017.

[48] M. Behrangrad, "A review of demand side management business models in the electricity market," *Renewable and Sustainable Energy Reviews*, vol. 47, pp. 270–283, 2015.

[49] R. B. Melton, "Gridwise transactive energy framework (draft version)," tech. rep., Pacific Northwest National Laboratory (PNNL), Richland, WA (US), 2013.

[50] R. Masiello, "Transactive energy: The hot topic in the industry [guest editorial]," *IEEE Power and Energy Magazine*, vol. 14, no. 3, pp. 14–16, 2016.

[51] F. A. Rahimi and A. Ipakchi, "Transactive energy techniques: closing the gap between wholesale and retail markets," *The Electricity Journal*, vol. 25, no. 8, pp. 29–35, 2012.

[52] S. Behboodi, D. P. Chassin, N. Djilali, and C. Crawford, "Transactive control of fast-acting demand response based on thermostatic loads in real-time retail electricity markets," *Applied Energy*, 2017.

[53] Y. Parag and B. K. Sovacool, "Electricity market design for the prosumer era," *Nature Energy*, vol. 1, p. 16032, 2016.

[54] E. F. Bompard and B. Han, "Market-based control in emerging distribution system operation," *IEEE Transactions on Power Delivery*, vol. 28, no. 4, pp. 2373–2382, 2013.

[55] C. Calvillo, A. Sanchez-Miralles, J. Villar, and F. Martin, "Optimal planning and operation of aggregated distributed energy resources with market participation," *Applied Energy*, vol. 182, pp. 340–357, 2016.

[56] A. Clausen, A. Umair, Y. Demazeau, and B. N. Jørgensen, "Agent-based integration of complex and heterogeneous distributed energy resources in virtual power plants," in *International Conference on Practical Applications of Agents and Multi-Agent Systems*, pp. 43–55, Springer, 2017.

[57] Rodríguez-Molina and et al., "Business models in the smart grid: challenges, opportunities and proposals for prosumer profitability," *Energies*, vol. 7, no. 9, pp. 6142–6171, 2014.

[58] S. Sikdar and K. Rudie, "Microgrid level competitive market using dynamic matching," in *Electrical Power & Energy Conference (EPEC), 2013 IEEE*, pp. 1–6, IEEE, 2013.

[59] J. Lee, J. Guo, J. K. Choi, and M. Zukerman, "Distributed energy trading in microgrids: A game-theoretic model and its equilibrium analysis," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3524–3533, 2015.

[60] M. Mihaylov, S. Jurado, N. Avellana, I. Razo-Zapata, K. Van Moffaert, L. Arco, M. Bezunartea, I. Grau, A. Cañadas, and A. Nowé, "Scanergy: a scalable and modular system for energy trading between prosumers," in *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pp. 1917–1918, International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[61] M. Mihaylov, S. Jurado, N. Avellana, K. Van Moffaert, I. M. de Abril, and A. Nowé, "Nrgcoin: Virtual currency for trading of renewable energy in smart grids," in *the 49th North American Power Symposium*, pp. 1–6, IEEE, 2017.

[62] "Pecan street inc. real energy. real customers. in real time.," p. http://www.pecanstreet.org/energy/ (accessed on 10.24.2017).

[63] T. Chen, Y.-S. Chen, and W. Su, "An innovative localized retail electricity market based on energy broker and search theory," in *European Energy Market (EEM), 2014 11th International Conference on the*, pp. 1–6, IEEE, 2014.

[64] J. Albrecht, "The 2011 nobel memorial prize in search theory," *Department of Economics, Georgetown University*, 2011.

[65] J. Luo, W. Su, and A. Q. Huang, "Bit-energy: An innovative bitcoin-style distributed transactional model for a competitive electricity market," in *Power & Energy Society General Meeting, 2017 IEEE*, pp. 1–5, IEEE, 2017.

[66] Y. Cai, T. Huang, E. Bompard, Y. Cao, and Y. Li, "Self-sustainable community of electricity prosumers in the emerging distribution system," *IEEE Transactions on Smart Grid*, 2016.

[67] A. D. Rathnayaka, V. M. Potdar, and S. J. Kuruppu, "An innovative approach to manage prosumers in smart grid," in *Sustainable Technologies (WCST), 2011 World Congress on*, pp. 141–146, IEEE, 2011.

[68] W. Tushar, B. Chai, C. Yuen, D. B. Smith, K. L. Wood, Z. Yang, and H. V. Poor, "Three-party energy management with distributed energy resources in smart grid," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2487–2498, 2015.

[69] H. Pourbabak, J. Luo, T. Chen, and W. Su, "A novel consensus-based distributed algorithm for economic dispatch based on local estimation of power mismatch," *IEEE Transactions on Smart Grid*, 2017.

[70] M. Swan, *Blockchain: Blueprint for a new economy.* " O'Reilly Media, Inc.", 2015.

[71] E. Munsing, J. Mather, and S. Moura, "Blockchains for decentralized optimization of energy resources in microgrid networks," 2017.

[72] P. Danzi, M. Angjelichinoski, Č. Stefanović, and P. Popovski, "Distributed proportional-fairness control in microgrids via blockchain smart contracts," *arXiv preprint arXiv:1705.01453*, 2017.

[73] H. Wang and J. Huang, "Incentivizing energy trading for interconnected microgrids," *IEEE Transactions on Smart Grid*, 2016.

[74] N. Zhang, Y. Yan, and W. Su, "A game-theoretic economic operation of residential distribution system with high participation of distributed electricity prosumers," *Applied Energy*, vol. 154, pp. 471–479, 2015.

[75] G. El Rahi, S. R. Etesami, W. Saad, N. Mandayam, and H. V. Poor, "Managing price uncertainty in prosumer-centric energy trading: A prospect-theoretic stackelberg game approach," *IEEE Transactions on Smart Grid*, 2017.

[76] Y. Li and L. Yang, "Prospect theory, the disposition effect, and asset prices," *Journal of Financial Economics*, vol. 107, no. 3, pp. 715–739, 2013.

[77] Z. Zhou, W. K. V. Chan, and J. H. Chow, "Agent-based simulation of electricity markets: a survey of tools," *Artificial Intelligence Review*, vol. 28, no. 4, pp. 305–342, 2007.

[78] K. Kahn, "An introduction to agent-based modeling," *Physics Today*, vol. 68, no. 8, p. 55, 2015.

[79] K. Dehghanpour, M. H. Nehrir, J. W. Sheppard, and N. C. Kelly, "Agent-based modeling in electrical energy markets using dynamic bayesian networks," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4744–4754, 2016.

[80] D. Krishnamurthy, W. Li, and L. Tesfatsion, "An 8-zone test system based on iso new england data: Development and application," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 234–246, 2016.

[81] W. Ketter, J. Collins, and M. d. Weerdt, "The 2016 power trading agent competition," 2016.

[82] K. Dehghanpour, H. Nehrir, J. Sheppard, and N. Kelly, "Agent-based modeling of retail electrical energy markets with demand response," *IEEE Transactions on Smart Grid*, 2016.

[83] M. Yadack, B. Vermeulen, and A. Pyka, "Competition in the german market for retail electricity: an agent-based simulation," in *Innovation Networks for Regional Development*, pp. 255–272, Springer, 2017.

[84] Y. Yu, G. Liu, W. Zhu, and et al., "Good consumer or bad consumer: Economic information revealed from demand profiles," *IEEE Transactions on Smart Grid*, 2017.

[85] Y. Wang, Q. Chen, C. Kang, and Q. Xia, "Clustering of electricity consumption behavior dynamics toward big data applications," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2437– 2447, 2016.

[86] T. Chen, A. Mutanen, P. Järventausta, and H. Koivisto, "Change detection of electric customer behavior based on amr measurements," in *PowerTech, 2015 IEEE Eindhoven*, pp. 1–6, IEEE, 2015.

[87] F. Luo, G. Ranzi, X. Wang, and Z. Y. Dong, "Social information filtering based electricity retail plan recommender system for smart grid end users," *IEEE Transactions on Smart Grid*, 2017.

[88] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*, vol. 1. MIT press Cambridge, 1998.

[89] O. De Somer, A. Soares, T. Kuijpers, K. Vossen, K. Vanthournout, and F. Spiessens, "Using reinforcement learning for demand response of domestic hot water buffers: a real-life demonstration," *arXiv preprint arXiv:1703.05486*, 2017.

[90] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[91] L. Hirth and F. Ueckerdt, "Redistribution effects of energy and climate policy: The electricity market," *Energy Policy*, vol. 62, pp. 934–947, 2013.

[92] D. M. Newbery, “Towards a green energy economy? the eu energy unions transition to a low-carbon zero subsidy electricity system–lessons from the uks electricity market reform,” *Applied Energy*, vol. 179, pp. 1321–1330, 2016.

[93] E. Salies and C. W. Price, “Charges, costs and market power: the deregulated uk electricity retail market,” *The Energy Journal*, pp. 19–35, 2004.

[94] C. Defeuilley, “Retail competition in electricity markets,” *Energy Policy*, vol. 37, no. 2, pp. 377–386, 2009.

[95] T. Chen, H. Pourbabak, Z. Liang, and W. Su, “An integrated evoucher mechanism for flexible loads in real-time retail electricity market,” *IEEE Access*, vol. 5, pp. 2101–2110, 2017.

[96] W. Su, H. Eichi, W. Zeng, and M.-Y. Chow, “A survey on the electrification of transportation in a smart grid environment,” *IEEE Transactions on Industrial Informatics*, vol. 8, no. 1, pp. 1–10, 2012.

[97] S. Huang, L. Wu, D. Infield, and T. Zhang, “Using electric vehicle fleet as responsive demand for power system frequency support,” in *2013 IEEE vehicle power and propulsion conference (VPPC)*, pp. 1–5, IEEE, 2013.

[98] S. Shao, M. Pipattanasomporn, and S. Rahman, “Grid integration of electric vehicles and demand response with customer choice,” *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 543–550, 2012.

[99] Z. Fan, “A distributed demand response algorithm and its application to phev charging in smart grids,” *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1280–1290, 2012.

[100] P. J. Douglass, R. Garcia-Valle, P. Nyeng, J. Østergaard, and M. Togeby, “Smart demand for frequency regulation: Experimental results,” *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1713–1720, 2013.

[101] Z. Chen, L. Wu, and Y. Fu, “Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization,” *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1822–1831, 2012.

[102] H. Zhong, L. Xie, and Q. Xia, “Coupon incentive-based demand response: Theory and case study,” *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1266–1276, 2013.

[103] X. Fang, Q. Hu, F. Li, B. Wang, and Y. Li, “Coupon-based demand response considering wind power uncertainty: A strategic bidding model for load serving entities,” *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1025–1037, 2016.

[104] C. Bartusch, F. Wallin, M. Odlare, I. Vassileva, and L. Wester, “Introducing a demand-based electricity distribution tariff in the residential sector: Demand response and customer perception,” *Energy Policy*, vol. 39, no. 9, pp. 5008–5025, 2011.

[105] C. Vivekananthan, Y. Mishra, G. Ledwich, and F. Li, “Demand response for residential appliances via customer reward scheme,” *IEEE transactions on smart grid*, vol. 5, no. 2, pp. 809–820, 2014.

[106] M. R. Sarker, M. A. Ortega-Vazquez, and D. S. Kirschen, “Optimal coordination and scheduling of demand response via monetary incentives,” *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1341–1352, 2015.

[107] B. Hayes, I. Melatti, T. Mancini, M. Prodanovic, and E. Tronci, “Residential demand management using individualised demand aware price policies,” *IEEE Transactions on Smart Grid*, 2016.

[108] A. Molina-Garcia, F. Bouffard, and D. S. Kirschen, “Decentralized demand-side contribution to primary frequency control,” *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 411–419, 2011.

[109] S. A. Pourmousavi and M. H. Nehrir, “Real-time central demand response for primary frequency regulation in microgrids,” *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1988–1996, 2012.

[110] S. Hambridge, A. Q. Huang, and N. Lu, "Proposing a frequency based real-time energy market and economic dispatch strategy," in *2016 IEEE Power & Energy Society General Meeting*, IEEE, 2016.

[111] W. Su, J. Wang, K. Zhang, and A. Q. Huang, "Model predictive control-based power dispatch for distribution system considering plug-in electric vehicle uncertainty," *Electric Power Systems Research*, vol. 106, pp. 29–35, 2014.

[112] Y. Guo, J. Xiong, S. Xu, and W. Su, "Two-stage economic operation of microgrid-like electric vehicle parking deck," *IEEE Transactions on Smart Grid*, vol. 7, no. 3, pp. 1703–1712, 2016.

[113] B. Liscouski and W. Elliot, "Final report on the august 14, 2003 blackout in the united states and canada: Causes and recommendations," *A report to US Department of Energy*, vol. 40, no. 4, 2004.

[114] J. A. Short, D. G. Infield, and L. L. Freris, "Stabilization of grid frequency through dynamic demand control," *IEEE Transactions on power systems*, vol. 22, no. 3, pp. 1284–1293, 2007.

[115] D. J. Hammerstrom, J. Brous, D. P. Chassin, G. R. Horst, R. Kajfasz, P. Michie, T. V. Oliver, T. A. Carlon, C. Eustis, O. M. Jarvegren, *et al.*, "Pacific northwest gridwise testbed demonstration projects; part ii. grid friendly appliance project," tech. rep., Pacific Northwest National Laboratory (PNNL), Richland, WA (US), 2007.

[116] F. C. Schweppe, R. D. Tabors, J. L. Kirtley, H. R. Outhred, F. H. Pickel, and A. J. Cox, "Homeostatic utility control," *IEEE Transactions on Power Apparatus and Systems*, no. 3, pp. 1151–1163, 1980.

[117] P. García, J. P. Torreglosa, L. M. Fernández, F. Jurado, R. Langella, and A. Testa, "Energy management system based on techno-economic optimization for microgrids," *Electric Power Systems Research*, vol. 131, pp. 49–59, 2016.

[118] N. Jaleeli, L. S. VanSlyck, D. N. Ewart, L. H. Fink, and A. G. Hoffmann, "Understanding automatic generation control," *IEEE transactions on power systems*, vol. 7, no. 3, pp. 1106–1122, 1992.

[119] Y. Zhang, P. Markham, T. Xia, L. Chen, Y. Ye, Z. Wu, Z. Yuan, L. Wang, J. Bank, J. Burgett, *et al.*, "Wide-area frequency monitoring network (fnet) architecture and applications," *IEEE Transactions on smart grid*, vol. 1, no. 2, pp. 159–167, 2010.

[120] T. Chen, K. Qian, A. Mutanen, B. Schuller, P. Järventausta, and W. Su, "Classification of electricity customer groups towards individualized price scheme design," in *Power Symposium (NAPS), 2017 North American*, pp. 1–4, IEEE, 2017.

[121] A. Mutanen, H. Niska, and P. Järventausta, "Ming smart meter data - case finland," in *The 23th International Conference & Exhibition on Electricity Distribution (CIRED) Workshop*, (Helsinki, Finland), pp. 120–124, 2016.

[122] F. Quilumba and et al., "Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 911–918, 2015.

[123] C. Beckel and et al., "Revealing household characteristics from smart meter data," *Energy*, vol. 78, pp. 397–410, 2014.

[124] R. Li and et al., "Development of low voltage network templatespart i: substation clustering and classification," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3036–3044, 2015.

[125] T. Räsänen and et al., "Data-based method for creating electricity use load profiles using large amount of customer-specific hourly measured electricity use data," *Applied Energy*, vol. 87, no. 11, pp. 3538–3545, 2010.

[126] S. V. Verdú and et al., "Classification, filtering, and identification of electrical customer load patterns through the use of self-organizing maps," *IEEE Trans. on Power Systems*, vol. 21, no. 4, pp. 1672–1682, 2006.

[127] G. Chicco and I.-S. Ilie, "Support vector clustering of electrical load pattern data," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1619–1628, 2009.

[128] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 2, pp. 513–529, 2012.

[129] K. Qian, Z. Zhang, F. Ringeval, *et al.*, "Bird sounds classification by large scale acoustic features and extreme learning machine," in *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, (Orlando, FL, USA), pp. 1317–1321, 2015.

[130] R. O. Duda, P. E. Hart, D. G. Stork, *et al.*, *Pattern classification*, vol. 2. Wiley New York, 1973.

[131] C. Ding and X. He, "K-means clustering via principal component analysis," in *Proceedings of the twenty-first international conference on Machine learning*, (Alberta, Canada), p. 29, ACM, 2004.

[132] C. M. Bishop, "Pattern recognition," *Machine Learning*, vol. 128, pp. 1–58, 2006.

[133] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489–501, 2006.

[134] R. Zhang and et al., "Short-term load forecasting of australian national electricity market by an ensemble model of extreme learning machine," *IET Gen., Trans. & Distribution*, vol. 7, no. 4, pp. 391–397, 2013.

[135] J. Lin and et al., "A symbolic representation of time series, with implications for streaming algorithms," in *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, (San Diego, CA, USA), pp. 2–11, 2003.

[136] J. Sevcech and M. Bielikova, "Symbolic time series representation for stream data processing," in *Trustcom/BigDataSE/ISPA, 2015 IEEE*, vol. 2, (Helsinki, Finland), pp. 217–222, 2015.

[137] T. Chen and W. Su, "Indirect customer-to-customer energy trading with reinforcement learning," *IEEE Transactions on Smart Grid*, 2018.

[138] N. Liu, X. Yu, C. Wang, C. Li, L. Ma, and J. Lei, "An energy sharing model with price-based demand response for microgrids of peer-to-peer prosumers," *IEEE Transactions on Power Systems*, 2017.

[139] C. Yuan, M. Illindala, and A. Khalsa, "Co-optimization scheme for distributed energy resource planning in community microgrids," *IEEE Transactions on Sustainable Energy*, 2017.

[140] T. Morstyn, A. Teytelboym, and M. D. McCulloch, "Bilateral contract networks for peer-to-peer energy trading," *IEEE Transactions on Smart Grid*, 2018.

[141] C. Zhang, J. Wu, C. Long, and M. Cheng, "Review of existing peer-to-peer energy trading projects," *Energy Procedia*, vol. 105, pp. 2563–2568, 2017.

[142] T. Chen, Q. Alsafasfeh, H. Pourbabak, and W. Su, "The next-generation us retail electricity market with customers and prosumersa bibliographical survey," *Energies*, vol. 11, no. 1, p. 8, 2017.

[143] D. Johnson, E. Horton, R. Mulcahy, and M. Foth, "Gamification and serious games within the domain of domestic energy consumption: A systematic review," *Renewable and Sustainable Energy Reviews*, vol. 73, pp. 249–264, 2017.

[144] R. Zafar, A. Mahmood, S. Razzaq, W. Ali, U. Naeem, and K. Shehzad, "Prosumer based energy management and sharing in smart grid," *Renewable and Sustainable Energy Reviews*, 2017.

[145] X. Xu, A. Maki, C.-f. Chen, B. Dong, and J. K. Day, "Investigating willingness to save energy and communication about energy use in the american workplace with the attitude-behavior-context model," *Energy Research & Social Science*, 2017.

[146] W. A. Brock and D. S. Evans, "Small business economics," *Small business economics*, vol. 1, no. 1, pp. 7–20, 1989.

[147] A. Rubinstein and A. Wolinsky, "Middlemen," *The Quarterly Journal of Economics*, vol. 102, no. 3, pp. 581–593, 1987.

[148] S. Park, J. Lee, G. Hwang, and J. K. Choi, "Event-driven energy trading system in microgrids: Aperiodic market model analysis with a game theoretic approach," *IEEE Access*, vol. 5, pp. 26291–26302, 2017.

[149] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of markov decision processes," *Mathematics of operations research*, vol. 12, no. 3, pp. 441–450, 1987.

[150] M. G. Kohn and S. Shavell, "The theory of search," *Journal of Economic Theory*, vol. 9, no. 2, pp. 93–123, 1974.

[151] J. Contreras, O. Candiles, J. I. De La Fuente, and T. Gomez, "Auction design in day-ahead electricity markets," *IEEE Transactions on power Systems*, vol. 16, no. 3, pp. 409–417, 2001.

[152] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*, vol. 1. MIT press Cambridge, 1998.

[153] D.-M. Han and J.-H. Lim, "Design and implementation of smart home energy management systems based on zigbee," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 3, 2010.

[154] Y. Yue, J. Broder, R. Kleinberg, and T. Joachims, "The k-armed dueling bandits problem," *Journal of Computer and System Sciences*, vol. 78, no. 5, pp. 1538–1556, 2012.

[155] T. Shibuya and S. Yasunobu, "Reinforcement learning with nonstationary reward depending on the episode," in *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on*, pp. 2145–2150, IEEE, 2011.

[156] R. Rana and F. S. Oliveira, "Real-time dynamic pricing in a non-stationary environment using model-free reinforcement learning," *Omega*, vol. 47, pp. 116–126, 2014.

[157] E. A. Feinberg and A. Shwartz, *Handbook of Markov decision processes: methods and applications*, vol. 40. Springer Science & Business Media, 2012.

[158] D. P. Bertsekas, D. P. Bertsekas, D. P. Bertsekas, and D. P. Bertsekas, *Dynamic programming and optimal control*, vol. 1. Athena scientific Belmont, MA, 1995.

[159] C. J. C. H. Watkins, *Learning from delayed rewards*. PhD thesis, King's College, Cambridge, 1989.

[160] J. Albrecht, "The 2010 nobel memorial prize in search theory," 2011.

[161] S. D. Jacka, "Optimal stopping and the american put," *Mathematical Finance*, vol. 1, no. 2, pp. 1–14, 1991.

[162] T. Chen, W. Su, and Y.-S. Chen, "An innovative localized retail electricity market based on energy broker and search theory," in *Power Symposium (NAPS), 2017 North American*, pp. 1–5, IEEE, 2017.

[163] J. N. Bearden, "A new secretary problem with rank-based selection and cardinal payoffs," *Journal of Mathematical Psychology*, vol. 50, no. 1, pp. 58–59, 2006.

[164] T. Chen and W. Su, "Local energy trading behavior modeling with deep reinforcement learning," *IEEE Transactions on Smart Grid*, 2018 (under review).

[165] E. Kuznetsova, Y.-F. Li, C. Ruiz, E. Zio, G. Ault, and K. Bell, "Reinforcement learning for microgrid energy management," *Energy*, vol. 59, pp. 133–146, 2013.

[166] X. Xiao, C. Dai, Y. Li, C. Zhou, and L. Xiao, "Energy trading game for microgrids using reinforcement learning," in *International Conference on Game Theory for Networks*, pp. 131–140, Springer, 2017.

[167] B. Zhao, X. Zhang, J. Chen, C. Wang, and L. Guo, "Operation optimization of standalone microgrids considering lifetime characteristics of battery energy storage system," *IEEE transactions on sustainable energy*, vol. 4, no. 4, pp. 934–943, 2013.

[168] A. Das and Z. Ni, "A computationally efficient optimization approach for battery systems in islanded microgrid," *IEEE Transactions on Smart Grid*, 2017.

[169] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, p. 529, 2015.

[170] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017.

[171] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," *arXiv preprint arXiv:1312.5602*, 2013.

[172] S. Adam, L. Busoniu, and R. Babuska, "Experience replay for real-time reinforcement learning control," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 2, pp. 201–212, 2012.

[173] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.