Integrated Energy Scheduling and Routing for a Network of Mobile Prosumers

BY

Mohammed Alqahtani
B.S., King Khalid University, Abha 2010
M.S., University of Arizona, Tucson, 2015

Dissertation

Submitted as partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Industrial Engineering and
Operations Research in the Graduate College of
The University of Illinois at Chicago, 2021

Chicago, Illinois

Defense Committee:

Prof. Michael J. Scott, Chair and Advisor
Prof. Houshang Darabi
Prof. David He
Prof. Mohammad Shadmand, Electrical and Computer Engineering Department
Dr. Yang Chen, Oak Ridge National Laboratory

This dissertation is dedicated to my parents, who encouraged me to become a better person each and every single day. And taught me how to be passionate and resilient in life. To my wife, and all my family for their eternal love and limitless support. To my friends and colleagues who have been there for me whenever I needed them along my Ph.D. journey.

ACKNOWLEDGEMENTS

First of all, I would like to express my ultimate gratitude to my loving parents, and my beloved wife Manar, for their unconditional love and support throughout this journey. Thank you for believing in me and for being there when I needed you. You are a blessing in my life. I love you all.

Next, I would like to thank my Ph.D. advisor, Dr. Michael Scott for his support, guidance and advice. Additionally, I would like to thank my Ph.D. committee members, Dr. Houshang Darabi, Dr. David He, Dr. Mohammad Shadmand, and Dr. Yang Chen for their support, constructive feedback, and overwhelming encouragements.

My next special thanks go to my UIC family, the rest of the faculty and staff members at the Mechanical and Industrial Engineering Department.

Furthermore, I would like to express my sincere gratitude to Dr. Mengqi Hu, my former academic advisor for helping me grow as a researcher and becoming the person I am today. I will carry the lessons I have learned from you for the rest of my life. I would like to extend my thanks to my past and current colleagues in the SMART Lab, Jinghang Li, Di Wang, and Zishuan Yu, for their patience, discussions, and suggestions throughout on my research.

I would also like to thank the government of Saudi Arabia, represented by King Khalid University for granting me this scholarship to continue my graduate studies and maximize my potentials and their guidance and support throughout my education journey. In addition, I would like to thank the National Science Foundation for funding this research under grant number CMMI 1634738.

CONTRIBUTION OF AUTHORS

Chapter 1 is an introduction of the dissertation, including background on smart grid and network of mobile prosumers (MPN), Deep Reinforcement Learning (DRL), and Multi-agent Reinforcement Learning (MARL). It also includes literature review, research motivation, objectives, and contributions.

Chapter 2 introduces the concept of a mobile prosumer network, which is one of my published papers (Alqahtani, M., & Hu, M. (2020). Integrated energy scheduling and routing for a network of mobile prosumers. *Energy*, 117451.(My contributions to the paper include conceptualization, data curation, investigation, methodology, simulation, validation, visualization, writing, and original draft. Dr. Mengqi Hu (my former advisor) contributed to the supervision, revision writing, and editing.

Chapter 3 discusses the uncertainties and dynamics in smart grid energy systems, which represents my second paper (submitted for publication in Renewable and Sustainable Energy Reviews journal). My contributions to the paper include conceptualization, data curation, investigation, methodology, simulation, validation, visualization, writing, and original draft. Dr. Mengqi Hu (my former advisor) contributed to the manuscript editing, revision, and general supervision of the research.

Chapter 4 investigates scalability issues for large fleets of EV's as energy prosumers in a smart grid, which is my third paper (submitted for publication in Applied Energy journal). I have been the major driver of the research. Dr. Mengqi Hu (my former advisor) contributed to manuscript editing, revision, and general supervision of the research. In addition, Dr. Michael Scott (my current advisor) contributed to manuscript editing, revision, and general supervision of the research.

Chapter 5 concludes the dissertation and suggests future work.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF ABBREVIATIONS

OSG	On-Site Generation
DER	Distributed Energy Resources
PV	Photovoltaic Panels
MPN	Mobile Prosumer Network
DC	Direct Current
AC	Alternative Current
EV	Electric Vehicle
V2G	Vehicle-to-Grid
CO2	Carbon Dioxide
VR	Vehicle Routing
ES/ED	Energy Scheduling/Dispatch
SP	Stochastic Programming
MDP	Markov Decision Process
QP	Quadratic Programming
HP	Hydropower Plant
DEC-MDP	Decentralized Markov Decision Process
MARL	Multi-Agent Reinforcement Learning
DRL	Deep Reinforcement Learning
GA	Genetic Algorithm
PSO	Particle Swarm Optimization
AFSA	Artificial Fish Swarm Algorithm
PHEV	Plug-in Hybrid Electric Vehicles
MIP	Mixed-Integer Programming
MILP	Mixed-Integer Linear Programming
RES	Renewable Energy Sources
AMPSO	Angle Modulated Particle Swarm Optimization

LIST OF ABBREVIATIONS (continued)

MIQP	Mixed-integer quadratic programming
DP	Dynamic Programming
E-PGVR	Ecofriendly Pheromone-based Green Vehicle Routing
DIRS	A distributed intelligent resource scheduling
TG	Thermoelectric Generator
DM	Disjoint Model
IM	Integrated Model
DQN	Deep Q-Network
CTDE	Centralized Training and Distributed Execution.
ReLU	Rectified Linear Unit
MAE	Mean Absolute Error

SUMMARY

Due to the spatio-temporal complexity in energy consumption profiles of multiple consumers at different regions, it is expected that more cost savings and higher resiliency to power disruptions can be achieved using a network of mobile prosumers, defined as entities that can both provide and consume energy. In this dissertation, the prosumer is assumed to be an autonomous vehicle equipped with different distributed energy resources (e.g., solar panel, battery). In addition, the demand on energy is uncertain and subject to change with time due to several factors including the emergence of new technology, entertainment, divergence of people's consumption habits, changing weather conditions, etc. Moreover, increases in energy demand are growing every day due to increases in world population and growth of the global economy, which substantially increase the chances of disruptions in power supply. This makes the security of power supply a more challenging task especially during peak usage seasons (e.g., summer and winter). Furthermore, the majority of the applications of smart grids in the existing literature include a large number of agents (e.g., electric vehicles), and each agent interacts with other agents and affects their decisions (i.e., mobility and scheduling). This causes the operational complexity of smart grids to increase exponentially.

To date, most of the existing research in energy management in smart grids aims to find efficient controlling schemes that seek renewable energy resources and dispatch energy to shift energy peaks between different daily periods. However, there are limitations and gaps in the current state-of-art in the field of energy management in smart grid including: (i) most of the literature focuses on energy transactions between EV's and power grids while the benefits of mobility in EV to mitigate the impact of spatio-temporal complexity in energy load and production are not maximally exploited; (ii) the existing solution approaches (e.g., mixed integer programming, dynamic programming, etc.) are not computationally efficient to handle system dynamics with the uncertainty inherent in energy problems; (iii) traditional/exact optimization tools are computationally expensive and are not efficient in solving high dimensional energy problems.

Thus, the aim of this dissertation is to thus advance the state-of-the-art in energy management of smart grids to contribute to reducing energy costs, dependency of the main power grid, computational time, and carbon emissions. This is done in several stages. First, an integrated vehicle routing and energy scheduling decision model is proposed to adaptively dispatch vehicles to balance the temporally and spatially distributed energy requests subject to vehicle mobility constraints, and thus to maximally exploit the potential of a mobile prosumer network for cost savings and carbon emission reductions. Second, a reinforcement learning model is developed to address the uncertainties in power supply and demand by dispatching a set of electric vehicles to supply energy to different consumers at different locations. Finally, a multiagent reinforcement learning (MARL) algorithm is introduced to address the scalability issues of large-scale smart grid systems. This algorithm uses centralized training and distributed execution, where all agents are trained using an actor network for each agent and sharing the same critic network, and then executed to make actions independent of other agents to reduce computation time.

1. INTRODUCTION

1.1 Background

Since the middle of the twentieth century, the population around the world is in continuous increase due to several factors including the end of World War II, industrial revolution, medical advancements, agricultural productivity, etc. According to United Nations statistics [116], the world's population is projected to reach 8.6 billion by 2030 and 9.8 billion by 2050.

With this increase in population along with the growth in world's economy, the electrical industry is going through major changes due to the rapid increases in loads and needs for security of power supply to end consumers. Figure 1 illustrates the world energy demand according to Energy Information Administration, since it is expected that the demands on energy are projected to increase 28% by 2040 [117].

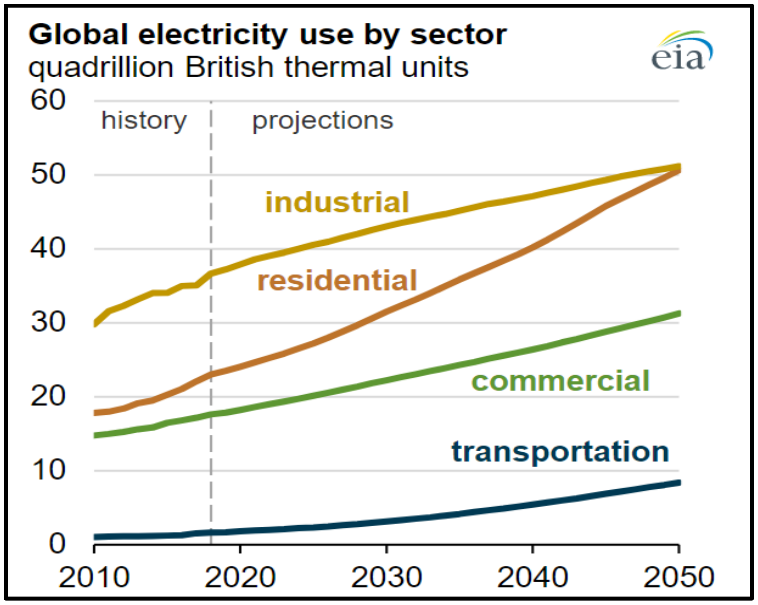

Figure 1: World energy consumption [117]

In addition, the fact that the primary source of energy is fossil fuel resources such as natural gas, coal, etc. that are limited and nonrenewable adds to the challenge of the energy supply to meet continuous increases in energy demands. Figure 2 shows the world's energy consumption by energy source, and indicates the dominance of the nonrenewable sources over the renewables.

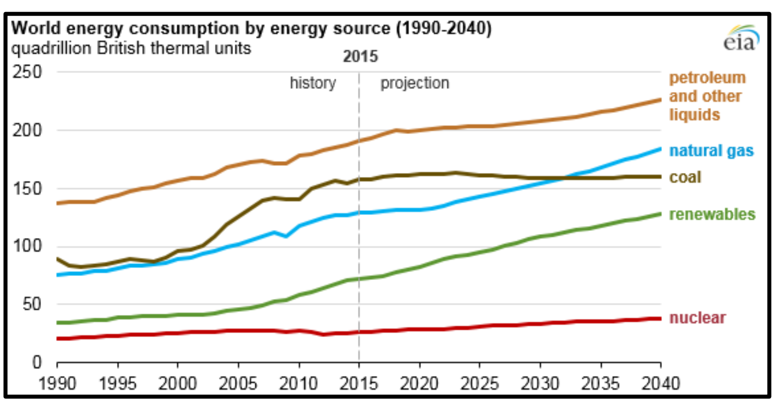

Figure 2: World energy consumption by source [118]

To address these challenges in the energy industry, it is imperative to seek alternative sources of energy that are clean, reliable, and sustainable to ensure security and stability in energy supply, improved customer satisfaction, and reduced carbon footprint. Distributed generation, also known as distributed energy, on-site generation (OSG), or district/decentralized energy, is electrical generation and storage performed by a variety of small, grid-connected or distribution system connected devices referred to as distributed energy resources (DER) (illustrated in Figure 3). Nowadays, DER is one of the most important research interests in the domain of energy due to its role in the shift to green and sustainable energy. It helps communities to use their own resources and facilities to power themselves and become independent from main power grids.

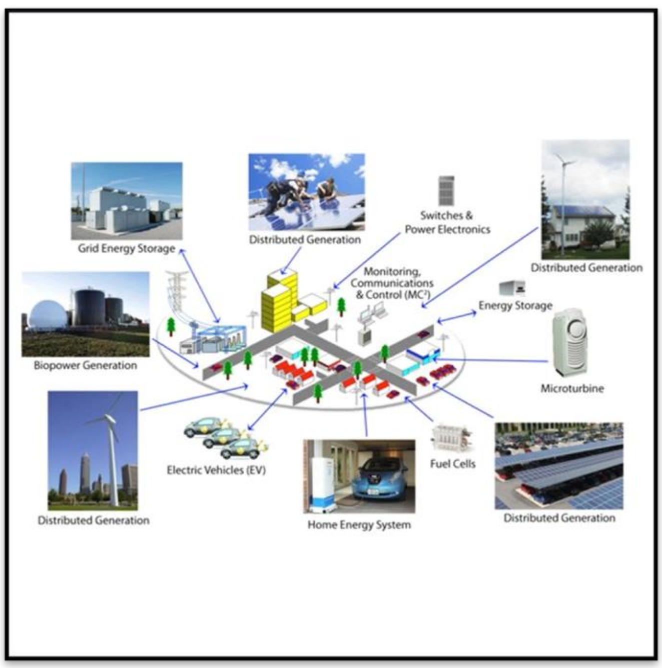

Figure 3: Typical smart grid (DER) [119]

It is well known that energy loads are varied at both spatial and temporal scales, so it will be more cost effective if the DERs can be adaptively dispatched to balance the energy load at different regions at different time periods. On the other hand, the amount of energy can be produced by DERs is also varied at both spatial and temporal scales. For example, the energy cost may be significantly reduced if the PV panels are dispatched to harvest more energy at higher solar irradiance regions and serve the closest high load consumers. To this end, it is expected that more cost savings and higher resiliency to power disruptions can be achieved if multiple mobile DERs (e.g., trucks with solar panels), termed a mobile prosumer network (MPN), can be dynamically dispatched to different regions. For example, mobile PV systems are autonomic systems which can be easily mounted on a tow trailer and transported from one location to another to supply green energy to homes, businesses, hospitals or schools in the aftermath of disasters. Mobile PV systems are composed of PV panels, voltage regulators, lead-acid batteries to store the electricity, inverters to change DC into AC, and wiring to connect various components (see Figure 4).

Figure 4: fixed or temporary Hybrid power systems [120]

1.1.1 Electric Vehicles as Energy Suppliers

In today's world, the electrical industry is going through major changes due to rapid increases in loads and needs for security of power supply to end consumers. Several emerging distributed energy resource

(DER) technologies, such as internal combustion engine, gas turbine, microturbine, photovoltaic (PV), fuel cell, wind turbine, and electric vehicle (EV) [1] have been developed to mitigate the impacts of increased energy consumption and global warming effects. These technologies can act as local energy service providers to address the issue of power outages and environmental sustainability, and motivate the traditional energy consumers to received/supply provide energy at the same time, i.e., prosumer [2].

As an emerging technology in the transportation sector, EVs can be operated in the vehicle-to-grid (V2G) mode where the energy flow between power grids and EVs is bidirectional [3], and thus can be used as energy storage to shift energy from peak periods to off-peak periods and thus to balance the energy request to power grids. V2G technology allows EVs to receive energy from renewable sources such as solar and store it for future consumption or supply it to power grids [4]. In addition, EVs can ensure sustainability of energy supply, especially for sensitive facilities such as hospitals, since they can act as a response to grid outages due to equipment failures [5]. An energy dispatch model of the Scandinavian and German electricity systems presented in [6] demonstrates that EVs can minimize energy costs and reduce CO2 missions to 15g/km by year 2030. An integrated control strategy proposed in [7] can minimize the operation costs for EVs in the distribution grid with renewable energy sources. Additionally, vehicle-to-building is another concept of energy integration that has been introduced to satisfy the energy needs of buildings using vehicle storage. The energy costs can be reduced by 14% in a collaborative network of commercial buildings, EVs, and power grids using a two-stage stochastic programming model [8].

Considering the mobility of EVs and spatio-temporal complexity of energy behaviors of DERs, it is hypothesized that a group of EVs with DERs (e.g., solar powered EVs), called a mobile prosumer network (MPN), will result in significant cost savings and be more resilient to power disruption. It is well known that energy loads are varied at both spatial and temporal scales, so it will be more cost effective

if the DERs are allowed to be adaptively dispatched to balance the energy load at different regions subject to mobility constraints of EVs. On the other hand, the amount of energy can be produced by DERs also varies at both spatial and temporal scales. For example, the energy cost may be significantly reduced if the PV panels are dispatched to harvest more energy at higher solar irradiance regions and serve the closest high load consumers. To this end, the energy scheduling and vehicle routing decisions should be made jointly to maximally exploit the benefits of MPN for cost savings, which has been less studied in the existing literature.

One of the promising applications of MPN is disaster response. For example, when a disaster strikes, there will be a high chance of power outages in disaster areas. The mobile prosumers can be dispatched to provide power service for these areas until power is restored from main grids. The mobile PV system has already been used to provide power service for (i) a community center after Hurricane Hugo in 1989, (ii) some communication centers and Southern California residents after the Northridge earthquake in 1994, and (iii) some medical clinics, shelters, street lights, and communication systems after Hurricane Andrew in 1992 [9].

With this as motivation, an integrated energy scheduling and vehicle routing decision model is developed to enable energy sharing in the MPN and thus to ensure energy security, cost effectiveness and CO2 reductions. Specifically, this study will investigate several cases to evaluate the performance of MPN under various DER configurations, energy pricing, energy consumption due to vehicle mobility, consumer distances, and solar irradiance profiles. The simulation results demonstrate that the proposed integrated decision model can achieve more cost savings compared to a disjoint energy scheduling and routing decision process. This study makes three main contributions to the existing research:

- An integrated decision model is developed to enable more flexible energy sharing among mobile prosumers at both spatial and temporal scales.
- The tradeoff between vehicle mobility and energy service quality is explored and the performance of MPN is evaluated using multiple metrics including energy cost, energy requested from the main grid, and carbon emissions.
- A systematic study is conducted to evaluate the cost effectiveness of the MPN concept, which can provide valuable insights for the energy industry to realize this novel concept.

1.1.2 Dynamics and Uncertainties in Energy Systems

Power outage is one of the major power issues of the twenty-first century due to changes in the world's population, weather conditions, people's lifestyles, and technology advancement. The U.S.-Canadian blackout of August 14, 2003 is one real world example of power issues caused by an overload of the main power grid and lack of supporting units, which affected nearly 50 million people in both the U.S. and Canada [10]. These power issues affect people's security, quality of life, and health [11]. Hence, the need of an immediate response to these power disruptions is imperative to ensure the sustainability of power supply via providing temporary backup energy solutions until the power is restored from the power grid.

To address or mitigate the impact of power supply problems, many studies investigate the causes of energy supply problems and develop plans to increase the resilience of the power grid to ensure sustainability of power supply. Different solution approaches and new technologies, such as electric vehicle (EV), smart grid, microgrid, and wide area monitoring applications are developed to provide temporal solutions and enable faster restoration of power services [12]. An EV can charge its battery with energy received from solar cells and supply it to power grids later [4]. Due to their mobility, EVs can be dispatched to disrupted areas to supply energy.

With this as a motivation, a study will be conducted to investigate how to optimally dispatch multiple EVs equipped with energy storage and photovoltaic (PV) panel to provide energy services for spatially and temporally distributed consumers. To ensure efficient energy dispatching and supply, a simultaneous vehicle routing (VR) and Energy Scheduling (ES) problem will be studied. In addition, the energy demands of the consumers may be uncertain and dynamically changing, and the energy production from renewable energy are highly uncertain. Stochastic Programming (SP) is one of the commonly used algorithms for handling problems with uncertainties. For example, an SP model for emergency response has been developed in [13] to dispatch first-aid commodities to affected locations during disasters. A multi-objective, robust, SP model is developed in [14] to optimize the preparedness and response of relief operations. The study in [15] employs an SP model to handle the uncertainties in solar energy, electricity prices, and energy load in a cluster of energy sharing providers. SP is used in [16] to find the optimal energy pricing for EV to increase both the profitability of a power aggregation business and customer satisfaction. To improve the efficiency of wind farms, a dynamic economic dispatching model using SP and wind speed forecasting is developed in [17].

The above publications indicate that SP is capable of dealing with uncertainties in optimization problems and yielding solutions for energy problems that have multiple sources of uncertainties. However, it may not be an efficient approach to handle the dynamics in the ES and VR problems due to its high computational cost [18, 19]. To enable fast decision-making for the dynamic ES and VR problems, first this problem is reformulated using Markov decision processes (MDP). A large volume of historical data including energy demand and solar irradiance are collected to train a deep reinforcement learning agent to map the system states (e.g., energy load, solar irradiance, battery energy level) to optimal actions (e.g., VR, ES). Once the training process is complete, the agent can quickly generate near-optimal solutions to handle system dynamics and uncertainties in energy load and radiant energy without solving the problems from scratch. The performance of the developed reinforcement learning algorithm is compared with exact optimization and heuristic algorithms to demonstrate its effectiveness for fast decision making.

1.1.3 Scalability in Energy Systems

Sustainability of power delivery to end customers is a main essential in modern life as it directly affects people's security, quality of life, and health [11]. Security of power supply is considered one of the major challenges in the modern power industry, and is greater in the case of delivering power to largescale urban areas due to the continuous changes in the demographic areas. These changes result in overloading the power infrastructure in large cities, which poses many challenges for electric utilities and city planners [20]. Hence, it is important to seek alternative power resources as an additional layer to power security to reduce the dependency on the main power grid, ensure the sustainability of power supply, and preserve the environment.

To increase the resilience of power grids and energy supply infrastructure in urban areas, many approaches and technologies to provide alternative power solutions and diversify sources of power services have emerged recently including electric vehicle (EV), smart grid, microgrid, and wide area monitoring applications [12]. Batteries of EVs can be used as a sustainable source of energy, since they can store energy from solar and discharge it to power grids later to reduce energy costs [4]. Further, given the fact that EVs can move among different places, they can be sent to supply electricity to different customers at different locations.

With this as a motivation, a study is carried out to investigate how to efficiently dispatch large numbers of EVs which are mounted with energy storage and photovoltaic (PV) panel to deliver power services for spatially and temporally distributed customers. To guarantee optimal energy dispatching and delivery, a simultaneous vehicle routing (VR) and Energy dispatching (ED) problem should be investigated. In addition, the large numbers of customers and EVs makes solving high-dimensional power problems to be a hard task and computationally expensive. There are studies that use traditional optimization tools to solve large-scale problems in the energy management field. For instance, Ref. [21] formulate a multi-objective mixed integer nonlinear programming model with economic and environmental aspects to optimize the operational efficiency of large-scale combined cooling, heat, and power system. A mathematical model to optimize the operation of a Brazilian hydrothermal power plant is proposed in [22], which uses linear and non-linear programming algorithms to solve the model. Ref. [23] presented a novel quadratic programming (QP) model to reduce energy waste by optimizing hourly operations in a group of hydropower (HP) plants.

The above studies show that conventional optimization algorithms can solve optimization problems with high dimensions and output solutions for energy problems that have a large number of agents. However, it takes a long time to produce solutions, which makes it inefficient approach to manage the dynamics in ED and VR problems [18,19]. Further, besides that this problem is a large-dimensional problem, it includes uncertainties and system dynamics, which makes a decomposition approach inefficient.

As an illustration of increasing the efficiency of decision-making for the dynamics in ED and VR problems, Multi-Agent Reinforcement Learning (MARL) is used to address the scalability issues in a network of mobile prosumers (EV's). MARL is a framework in which a group of self-ruling, communicating entities share the same environment, which they observe with sensors and upon which they act with actuators [123]. This research makes the following contributions:

- For the first time, a Decentralized Markov Decision Processes (DEC-MDP) is applied to a mobile prosumer network.
- Agents are trained using a large set of historical data for energy loads and solar irradiance in a centralized way, where each agent has its own actor network, and all agents share the same critic network. Each agent then executes its actions in a decentralized way, taking actions independently from other agents and generating near-optimal solutions using the sequence of actions learned from the training process. This approach reduces the interactions and dependencies among agents, which leads to reducing computation time and makes it more robust when handling system dynamics in large-scale problems. This is the first application of this framework to this problem.
- The performance of a number of different solution methods are compared: MARL, deep reinforcement learning (DRL) and three other heuristic algorithms (genetic algorithm (GA), particle swarm optimization (PSO), and artificial fish swarm algorithm (AFSA)) in terms of solution quality and effectiveness for fast decision-making at large-scale energy problems.

1.2 Literature Review

This section provides a brief review of state-of-the-art in operational modeling of individual DERs and a system of multiple DERs and existing algorithms to handle system dynamics and scalability issues in power systems.

1.2.1 Electric Vehicles and Distributed Energy Resources

During the past few decades, there is a rich literature focusing on energy management of DERs to maintain sustainable energy supply. A study in [24] demonstrates that EVs can be efficiently integrated with power grids by providing real-time frequency regulation. A two-stage model is proposed in [25] to manage energy in office buildings by adopting EVs as flexible energy resources. A novel EV fleet aggregator model is developed in [26] to study the DER investment and scheduling problems under uncertainty and assess the impact of EV interconnections on optimal DER solutions. A mixed integer linear programming model is constructed in [27] to optimize energy supply to buildings in a district-scale DER through the energy distribution networks to minimize the total annual costs. A multi-objective mixed integer linear programming model is developed in [28] to optimize a DER system at a community level containing residential and office buildings, leading to minimized total costs and reduced carbon emissions. A short-term model with a two-phase approach is developed in [29] to study the schedule operations of renewable energy resources. It is also demonstrated that renewable energy resources can reduce environmental pollution [30]. A mixed integer programming model is developed to obtain a cost effective and environmentally friendly operational strategy for DERs [31]. A conceptual model for a demand response management system is presented in [32] to improve the efficiency of DER operations. Energy storage is an essential component in DER to maintain stability and reliability of system operations [33]. A smart energy management system is proposed in [34] to optimize the operational strategies of energy storages in a microgrid. Multiple scale operation models are proposed in [35] to explore the performance of energy storage operations at different time scales, such as daily, weekly, etc.

The performance of DER integration to power grids can be measured by operational costs and reliability [36]. Many studies focus on assessing the operational performance of distributed generators in terms of cost savings, emission reduction, and service quality improvement [37]. The performance of DER in terms of capabilities of energy supply, sustainability, and reliability is also investigated [38].

Other than operational modeling of individual DERs, it is also demonstrated that more cost savings can be achieved if multiple DERs with complementary capabilities can be integrated. For example, an integrated renewable energy system can satisfy the energy request of seven unelectrified villages in India [39]. Through coordinating multiple energy devices including energy storages, an efficient energy supply can be ensured [40]. A net zero-energy residential building target can be achieved using a system of building-integrated PV panels and hydrogen fuel cell EVs where the EVs can help to reduce the energy purchased from power grids by approximately 71% over one year [41].

A transactive energy operation model is developed in [42] to enable efficient energy sharing among a cluster of microgrids with various DERs. A better energy management system for grid-connected DERs can be achieved using a distributed solution based on the paradigm of multi-agent systems [43]. An operational model for aggregated DER systems is developed in [44] to ensure efficient energy transaction between market and aggregator according to the size of the aggregated resources. A particle swarm optimization algorithm is employed to find near optimal operation strategies for a system of residential level DERs to maximize the net benefits of residents [45]. Taking dynamic load and benefits that end-user can derive from services into account, an optimization model is developed in [46] to optimize energy supply to end-users and increase their satisfaction.

An operation model is developed to manage a residential microgrid including a charging spot for EVs and renewable energy sources where different load profiles, owner behaviors, and energy consumption due to mobility are considered [47]. The EVs can be optimally charged to achieve peak load shaving and reduction of load variability of households connected to a local distribution grid [48]. A two-layer evolution strategy particle swarm optimization algorithm is developed to optimize the control strategies of EVs in a residential distribution grid to shave peak loads [49]. A linear programming model is developed in [50] to evaluate the benefits of the interconnection and joint management of a district with residential loads, DERs, EVs, and an electrical metro substation. A similar study to shave peak load using a two-stage optimization model is conducted in [51]. A practical model to assess the contribution of V2G with energy management support of small electric energy systems is designed in [52]. A real-time digital simulator for a fleet of EVs to implement the V2G power transactions is developed in [53].

There are challenges facing the adoption of the V2G concept including a lack of frameworks to apply this concept in the real world, efficiency and stability of power grids, and charging/discharging strategies between vehicles and power grids. Optimization techniques of V2G operation [54] have recently been developed to overcome the possible challenges [55] of the plug-in hybrid electric vehicles (PHEV) and V2G concept and improve the performance of power grids in terms of efficiency, stability, and reliability [56]. Managing charging/discharging of large numbers of PHEV/EVs and their capabilities is discussed in [57].

While promising, most of the existing research focus on energy transactions between EVs and power grids while the benefits of EV mobility to mitigate the impact of spatio-temporal complexity in energy load and production are not maximally exploited. For instance, deploying an EV with higher stored energy to serve consumers with higher energy loads may be efficient at the current time period, but it may not be efficient to serve higher energy load users at future time periods due to vehicle mobility constraints. This poses critical challenges to the existing disjoint vehicle routing and energy scheduling decision process. To bridge these research gaps, an integrated decision model is proposed in this thesis to study the tradeoff between vehicle mobility and energy service quality.

1.2.2 Addressing Uncertainty Issues in Smart Grids

EV is one of the most commonly used alternatives to provide temporal solutions for power disruptions and work as a secondary power supplying unit as they have the feature of vehicle-to-grid (V2G) where EVs can exchange energy with power grids [3], which enables EVs to act as mobile energy storage to balance energy requested from power grids by shifting energy from peak periods to off-peak periods. EVs can supply energy to compensate energy shortages from the main power grid and maintain energy supply to users [5]. Ref. [58] discusses the integration of the EV-grid including power interaction mode and mainstream dispatching method. Ref. [7] presents a control scheme with renewable energy to reduce the operational expenses for EVs in the distribution network. Ref. [59] discusses the capabilities of V2G to secure power and to provide ancillary services using renewable energy resources.

A two-stage SP model is developed in [8] to reduce energy costs in a community of buildings, EVs, and power networks. A novel management system is developed in [60] to manage and coordinate microgrid which uses plug-in EVs as distributed energy resource (DER). EVs can be efficiently incorporated with an energy distribution network via providing instantaneous frequency arrangements [24]. A two-stage model is presented in [25] to dispatch energy in office buildings by implementing EVs as flexible sources of energy. Ref. [26] presents a novel EV squadron aggregator model to discuss problems in the implementation and scheduling of DER under stochastic conditions and evaluate the contribution of EVs on optimal solutions. Ref. [29] investigates the scheduling operations of RES and develops a model for short time decisions operated in two phases. Renewable energy resources are considered a reliable source of clean energy that can be used to supply energy to communities while preserving the environment [30]. Ref. [35] presents operation models that operate under multiple scales to evaluate the performance of power storage at different time periods. Ref. [61] proposes a MIP model to develop an operational strategy for DERs to reduce energy costs and sustain the environment. A MIP model is presented in [27] to reduce the annual energy expenses in a district-scale DER by optimizing the operation of power grid. Ref. [34] presents a smart energy management system to improve the performance of energy storage units in a microgrid. Ref. [28] presents a multi-objective MILP model to reduce total energy costs and environmental pollution in a community of residential and office buildings.

There is literature discussing the uncertainties that stem from unpredictability of energy demand and supply. For example, a risk-constrained SP model is developed in [62] to address the energy procurement issue for large number of users. While SP and heuristic algorithms can handle the uncertainties in power supply and demand, they are computationally expensive [18, 63] and cannot provide real-time decisions in a power supply network [19]. The dynamic energy system requires instant decision-making processes to maintain its operations under various dynamics and uncertainties.

Traditional optimization tools (e.g., SP) can be used to handle the uncertainties in energy systems and yield optimal solutions for ES. However, these tools may not be efficient to handle system dynamics for real-time decision-making. Since these tools may take longer time to generate solutions and so may not response to the systems dynamics faster enough compared to reinforcement learning approach, especially in the case of routing problems [64]. Hence, reinforcement learning may be a better alternative than traditional optimization approaches, since it can be trained using real world data to interact more efficiently with system dynamics and uncertainties in energy supply and demand to yield faster real-time decisions for the routing and energy dispatching.

Reinforcement learning is widely used to address dynamic decision-making problems under uncertainties. For example, a stochastic optimization framework based on reinforcement learning is developed in [65] to address the uncertainty in energy prices and available wind energy. Uncertainties in traffic demand and supply in a stochastic traffic network are studied using a MDP formulation, and the formulated dynamic speed limit problem is solved by a real-time control mechanism [66]. A reinforcement learning algorithm is developed in [67] to solve the VR problem which can provide fast routing decisions without solving the new problem from scratch. Ref. [68] proposes a model and learning procedure to solve the problem of routing multiple vehicles with heterogeneous capacities. A real time delivery of products is modeled in [69] in the framework of average-reward reinforcement learning with the context of stochastic demands and multiple vehicles. Ref. [70] demonstrates that a reinforcement learning algorithm is able to yield optimal routing actions for autonomous taxis in the city of Singapore. A deep reinforcement learning-based neural combinatorial optimization strategy is presented in [71] to solve VR problems with minimal computation time.

Ref. [72] presents a deep reinforcement learning-based method to address the stochastic and dynamic nature of RES and power electronic devices in modern power systems without system modeling. Ref. [73] proposes a partially observable MDP model to address the stochastic nature of future electricity loads and renewable power generation. A reinforcement learning algorithm for hierarchical automatic generation control is proposed in [74] to achieve multi-objective dynamic optimal allocation of a microgrid in island mode. Distributed multi-energy systems optimization problem is addressed in [75] using an alternative approach based on multi-agent reinforcement learning. A reinforcement learning model is proposed in [76] for flexible and cost-effective dispatch of the electricity used by buildings in a smart grid. Ref. [77] proposes a novel energy dispatching approach for real-time scheduling of a microgrid considering the stochastic nature of energy demand, renewable energy, and energy price. A fuzzy Q-learning model is proposed in [78] with the use of renewable resources for modeling an hour-ahead electricity market. The dimensionality issues in the dynamic optimization of generation command dispatch for automatic generation control is addressed in [79] using an improved hierarchical reinforcement learning approach. Ref. [80] presents a deep reinforcement learning model using the deep deterministic policy gradient algorithm to control energy transactions of shared energy storage assets within building clusters. An economic dispatch problem is formulated as a multi-stage decision making problem in [81] and is solved using reinforcement learning.

1.2.3 *Addressing Scalability Issue in Smart Grids*

The power supply in large urban areas is one of the main targets and challenges in the energy industry, since the energy system in large cities is highly sophisticated in nature and interconnected, posing many challenges for electric utilities and city planners [20]. Hence, the need for an efficient energy management is essential to satisfy large energy demands and ensure the sustainability in power supply.

The optimization of large-scale problems has been an interest for the current state-of-the-art in the optimization field. For instance, a novel min-max dynamic programming (DP) model is presented in [82] to manage operations of HP system during peak times. Ref. [83] proposed an Angle Modulated Particle Swarm Optimization (AMPSO) to optimize the operations of a large-scale power systems. A mixed-integer quadratic programming (MIQP) model is used in [84] to optimize the operations of a large-scale wind power plant with respect to coordination of steady state and transient state operations.

Decomposition methods are a commonly used approach to handle large-scale optimization problems, aiming to partition the problems into sub-problems and solve them separately and then combine them again to reduce computational time. For example, a study in [85] adopts the concept of decomposition to solve a large-scale vehicle routing problem to maintain acceptable run time. Ref. [86] investigates the use of decomposition techniques to reduce computational time for a large-scale vehicle routing problem with time windows. A solution method for a large scale mixed-integer problem with a specific structure is proposed in [87], which decomposes the primal problem using duality while guaranteeing that the solutions produced are feasible for the original unmodified primal problem. A study in [88] uses the Dantzig-Wolfe decomposition method to address a power system problem with a 16-bus system and a modified IEEE 30-bus system. Ref. [89] adopts a Benders decomposition based framework for addressing a high-dimensional energy dispatching problem. However, the decomposition approach assumes a natural decomposition of the problem, since it is very subjective in drawing boundaries around physical components and subassemblies in large, highly dependent systems. Consequently, it may fail to account for interdependencies in dynamic and integrated systems. Given the stochastic dynamic nature of the problem in this thesis, the decomposition method may not be an efficient approach to solve it [90].

There are well-known optimization tools (e.g., MIP, DP, etc.) that can be used to handle the large-scale power systems and generate optimal solutions for ED. Nevertheless, these tools may take considerable time to produce solutions, especially in large-scale decision-making problems [91, 92]. Furthermore, a reinforcement learning approach may not be a good alternative to these traditional optimization tools when it comes to large-scale problems, since it takes more time to converge and output solutions [93]. On the other hand, multi-agent reinforcement learning (MARL) is more efficient than reinforcement learning and other optimization algorithms when it comes to high-dimensional problems [94, since it is able to do the training in a centralized way and execute in decentralized way, imposing fewer restrictions on the agents and resulting in less execution time. This feature enables MARL to handle large scale energy problems that have a large number of variables that are dynamic in nature and require faster decision-making processes [95].

The Decentralized Markov decision process (DEC-MDP) is commonly used to handle the multi-agent decision problems [96]. It has been used to solve energy management problems with distributed decisions. Ref. [97] utilizes a decentralized partially observable Markov decision process (DEC-POMDP) to reduce delay for impedance networks with environmentally friendly power sources. A study in [98] uses DEC-MDP framework to formulate and solve a problem of finding decentralized transmission policies in a wireless communication network. A wireless sensor network for an energy harvesting communication network is investigated using the DEC-MDP framework in [99]. Ref. [100] presents a study that uses DEC-POMDP to formulate and solve a decentralized partially observable offloading problem to maximize the network performance with energy harvesting and Internet of Things (EH-enabled IoT).

Multi-agent reinforcement learning is widely used to address challenges of scalability in DEC-MDP problems. For example, Ref. [101] presents a novel reinforcement learning algorithm that generates fast routing solutions for a largescale logistics system. A multi-agent routing based on learned value iteration is introduced in [102], which aims to output routing solutions for multiple agents in a dynamically changing traffic condition. Ref. [103] proposes a proactive Ecofriendly Pheromone-based Green Vehicle Routing (E-PGVR) scheme to minimize fuel consumption in traffic congestions via finding optimal routing solutions. A multi-agent based model is developed in [104] to optimize energy scheduling in a large-scale microgrid with modern homes interacting together to minimize energy costs. A distributed intelligent resource scheduling (DIRS) system is introduced in [105], which performs centralized training and decentralized execution by each agent dispatched in each mobile edge computing server. Ref. [106] presents a study that implements a multi-agent reinforcement learning based model to reduce energy costs and achieve the energy balance in a microgrid with grid-connected mode. A study in [75] discusses the implementation of a multi-agent reinforcement learning for optimal energy dispatch in distributed multi-energy systems. A multi-agent reinforcement learning approach for residential buildings is presented in [107] to address microgrid energy scheduling problem and guarantee fairness in a microgrid market. Ref. [108] discusses battery scheduling problems and proposes a reinforcement learning framework to increase the lifetime of batteries used in EVs.

1.3 Motivation

Sustainability of the power supply to end clients is a basic need in the present day as it impacts individuals' security, personal satisfaction, and wellbeing [1]. Security of the power supply is considered one of the significant difficulties in the advanced power industry, particularly when providing capacity to largescale metropolitan zones because of persistent changes in demographics in these areas. These changes provoke an overburdening of the power infrastructure in huge urban areas, presenting numerous difficulties for electric utilities and city organizers [2].

Existing literature explores the reasons for energy supply issues and creates plans to build the strength of the power grid to guarantee maintainability of the power supply. Different solution approaches and new advancements, for example, electric vehicle (EV), smart grid, microgrid, and wide territory checking applications are created to give temporal solutions and enable faster restoration of electricity services [3]. An EV can charge its battery with solar energy and supply it to the power grid later [4]. Because of their portability, EVs can be dispatched to the affected territories to supply energy. However, following the above literature review, some of the existing limitations in the literature on smart grid and energy management considering this emerging technology can be summarized as follows:

1. The energy scheduling and vehicle routing decisions should be jointly made which is less studied in the existing research to maximally exploit the benefits of MPN for cost savings.
2. The literature indicates that traditional optimization tools (e.g., Stochastic Programming) are capable of dealing with uncertainties in optimization problems. However, it may not be an efficient approach to handle the dynamics in energy scheduling and vehicle routing problem due to its high computational cost [10, 11].
3. The above studies show that conventional optimization algorithms take a long time to produce solutions, which makes them inefficient approaches to manage the dynamics in the joint energy scheduling and vehicle routing problem [8, 9].

Therefore, advancing the state-of-the-art in energy management and sustainability of the power supply in smart grids is necessary.

1.4 Research Scope, Objectives, Assumptions, and Framework

Based on the mentioned motivations and literature review, the following research scope, objectives, and framework are proposed.

1.4.1 Research Scope

This thesis is composed of three main stages. In the first stage, MIP is used to model and optimize MPN to form a benchmark for future comparisons of MPN performance. The performance of MPN is evaluated under different operating conditions. In the first case, the impact of different pricing mechanisms on the performance of MPN is investigated. In the second case, the performance of MPN under different solar irradiance profiles corresponding to different weather conditions (e.g., sunny, cloudy, snowy) is evaluated. In the third case, the performance of MPN under various consumer patterns is studied. Two consumer patterns are analyzed: (i) the "similar" pattern where each consumer has only one peak load and (ii) the "different" pattern where some consumers have one peak load, while others may have two or multiple peaks. In the fourth case, the performance of MPN under various energy consumption rates due to mobility is assessed. For demonstration, four different consumption rates are considered: 40, 70, 96, and 121 kWh per 100 miles. In the last case, the impact of distances between consumers on the performance of the mobile prosumer network is analyzed. For demonstration, four different distances will be considered: 1, 2, 4, and 8 miles.

In the second stage, the focus is on addressing the uncertainty issues for a network of mobile prosumers using a DRL approach, since traditional optimization tools (e.g., SP) may not be an efficient approach to handle the dynamics in ES and VR problem due to its high computational cost [10, 11]. The dynamic decision problem is reformulated using the Markov decision processes (MDP) framework. Four state variables ($s \in S$) are defined for each prosumer including location, battery level, consumer's load, and solar irradiance at consumer's location. Also, the action space ($a \in A$) is defined with a total of fifteen combinations of actions, five of which are mobility actions (up, down, left, right, and no move) and three of which are battery transaction actions (charge, discharge, and idle). Then the reward function ($r(s, a)$) is defined as the energy cost for each combination of prosumers' actions, which represents the amount of energy requested from the power grid multiplied by the energy price. After that, the model is solved using deep reinforcement learning algorithm (DRL). The decision model is trained to learn the dynamics and changes in energy loads and solar irradiance from historical data and make real-time decisions for the future. The performance of the proposed machine learning algorithm is compared with an exact solution approach and heuristic algorithms including genetic algorithm (GA), particle swarm optimization (PSO), and artificial fish swarm algorithm (AFSA) using different scales of problems.

In the third stage, a multi-agent reinforcement learning model is developed to address the scalability challenges for a network of mobile prosumers and operate it with larger number of consumers and prosumers. Conventional optimization tools (e.g., SP and DP) take a long time to produce solutions, which makes them inefficient approaches to manage the scalability complexities in the ED and VR problem [8, 9]. The centralized training and decentralized execution mechanism is adopted to train a decision model in a centralized way to enable independent decisions for each prosumer. The performance of the proposed algorithm is evaluated using different scales of problems including (i) One hundred consumers and sixteen prosumers, (ii) One hundred consumers and twenty-five prosumers, (iii) One hundred consumers and twenty-nine prosumers, (iv) One hundred consumers and thirty-five prosumers, and (v) One hundred consumers and forty prosumers. Heuristic models including GA, PSO, and AFSA are used to compare their results with the decentralized ML model.

1.4.2 *Research Objectives*

Based on these motivations, literature review, and research scope, the goal of this thesis is to integrate the vehicle routing and energy scheduling operations to enable energy sharing among consumers at both spatial and temporal scales. In addition, this research aims to address the system's dynamics due to sources of uncertainty in both energy load and solar irradiance by implementing the Artificial Intelligence concept of Deep Reinforcement Learning Algorithm. Moreover, this thesis investigates the scalability issues in implementing the concept of a mobile prosumer network on large scales and uses Multi-agent Reinforcement Learning algorithm to enable distributed decision-making among all agents in the system, hence increasing the computational efficiency of the model.

1.4.3 *Research Assumptions and Limitations*

There are some assumptions that have been made in this research for the sake of simplicity of the models' computations. The models in this thesis focus on the systems perspective and energy costs of the consumers only. Assumptions in this research can be summarized as follows:

- In the future there will be variable energy prices in metropolitan cities since they will have independent generation units. The goal is to optimally dispatch EVs to different locations and make energy transactions to minimize the overall costs.
- The cost is calculated from the perspective of the consumers only. EVs are assumed to be owned by the government to ensure sustainable power supply to the community.
- The grid services, consumer comfort level, and grid stability are beyond the scope of this thesis and will be considered as future work.
- The routes between each pair of consumers are available all the time.
- Vehicles are assumed to operate all the time.
- Vehicle maintenance, equipment degradation, design, and investment costs will be addressed in the future stages of this research.

1.4.4 *Research Framework*

In this research, the focus is on developing models and algorithms to study operations of a network of mobile prosumers and evaluate its performance under different scenarios. The overall vision for this research is described in Figure 5, and more detailed descriptions of the stages and the research challenges in each stage are in the following chapters.

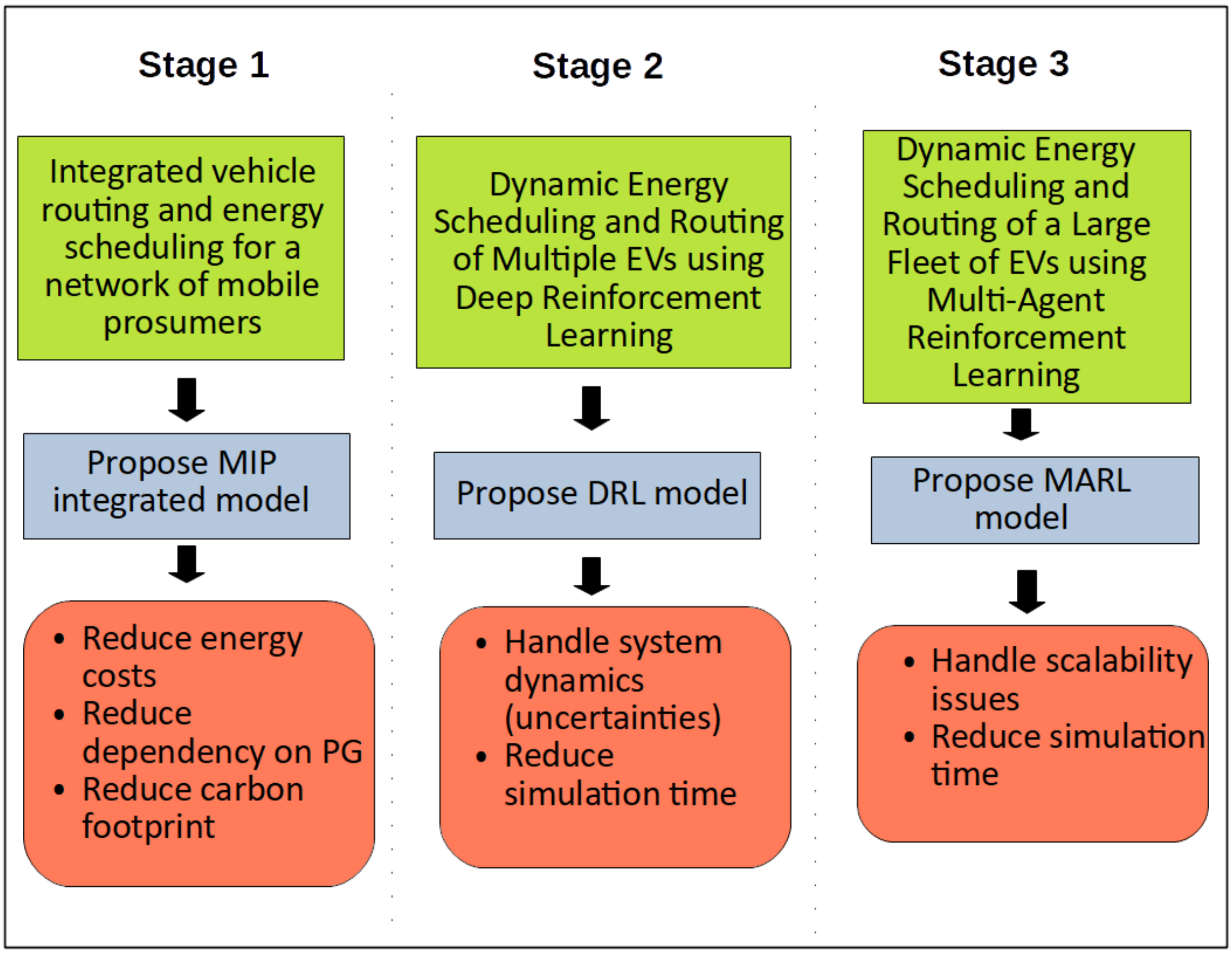

Figure 5: Overall framework of this research

Note that most of existing literature in DER focuses on energy scheduling while the benefits of mobile DERs to mitigate the impact of spatial-temporal complexity in energy load and production are not maximally exploited. Therefore, this research proposes to develop models and algorithms for integrated vehicle routing and energy scheduling for a mobile prosumer network which can find optimal solutions for energy transactions between prosumers and consumers along with optimal routing decisions of mobile prosumers.

1.5 Contributions and Thesis Outline

The main contributions of this thesis are as follows: First, an integrated decision model is developed to enable more flexible energy sharing among mobile prosumers at both spatial and temporal scales. In addition, a systematic study is conducted to evaluate the cost effectiveness of the MPN concept, which can provide valuable insights for the energy industry to realize this novel concept. Second, this problem is reformulated using Markov decision processes (MDP) to enable fast decision-making for the dynamic ES and VR problems. A large volume of historical data including energy demand and solar irradiance is collected to train a deep reinforcement learning agent to map the system states (e.g., energy load, solar irradiance, battery energy level) to optimal actions (e.g., VR, ES). Once the training process is complete, the agent can generate near-optimal solutions very quickly to handle system dynamics and uncertainties in energy load and radiant energy without solving the problems from scratch. Third, the problem model is reformulated using decentralized Markov decision processes (DEC-MDP) to boost the decision-making for the dynamics in ED and VR problems. A large set of historical data including energy loads and solar irradiance is collected to train a multi-agent reinforcement learning algorithm (MARL). Agents are trained in a centralized way, where each agent has its own actor network and all agents share the same critic network. Then each agent executes the actions in a decentralized way, where each agent takes actions independently from other agents and generates near-optimal solutions using the sequence of actions learned from the training process.

The rest of the thesis is organized as follows.

Chapter 2 introduces the concept of a mobile prosumer network and studies the feasibility of this concept to increase energy savings, reduce dependency on power grid, and reduce carbon emissions due to power generation. An integrated model of energy scheduling and routing for a network of mobile prosumers is presented. Where the integration of scheduling and routing is shown to yield more energy savings, reduce energy requested from main grid, and reduce carbon emissions.

In Chapter 3, the uncertainty issues in the implementation of the MPN model are investigated. Here, the sources of uncertainty are the energy demand and the solar irradiance, which affects the energy supply and demand in the energy system. Hence, models and algorithms are needed to overcome or reduce the impact of these uncertainties.

In Chapter 4, the scalability issues in a smart grid with large fleet of mobile prosumers (EVs) are studied, where a large number of mobile prosumers are dispatched to satisfy more consumers, causing it to take longer to get optimal solutions as the size of the problem increases.

Finally, a summary of the main conclusions and possible future work of this thesis are provided in Chapter 5.

2. INTEGRATED ENERGY SCHEDULING AND ROUTING FOR A NETWORK OF MOBILE PROSUMERS

[The majority of work presented in this chapter of the dissertation is published as: “Alqahtani, M., & Hu, M. (2020). Integrated energy scheduling and routing for a network of mobile prosumers. *Energy*, 117451.”.
For more information, please refer to Appendix B (Copyright Statement).]

This chapter presents the concept of a Mobile Prosumer Network (MPN) and evaluates its performance using three metrics: operational cost, energy requested from power grids, and carbon emissions. It then compares this performance with disjoint models. A *prosumer* is defined as an entity in an energy system that both produce and consume energy. In section 2.1 presents problem definition and model formulation for the proposed integrated decision process. The simulations to evaluate the performance of MPN are presented in Section 2.2 and conclusions are drawn in Section 2.3.

2.1 Mathematical Model for Mobile Prosumer Network

The overall system architecture to model the MPN is presented in Section 2.3.1 and the mathematical model to formulate the operational decision problem of MPN is discussed in Section 2.3.2.

2.1.1. System Architecture for Mobile Prosumer Network

In this section, the studied geographic area is assumed to be divided into mutually exclusive and collectively exhaustive hexagonal regions (see Figure 6) where the distance between two connected regions is one mile. Each region may have one consumer or may not have any consumer. This setting allows us to study the energy consumers with different distances from one to many miles. For demonstration purpose, Figure 5 assumes that there are eight consumers and four prosumers. A network of mobile energy prosumers is assumed to provide energy service for these consumers with various energy and solar irradiance profiles.

Each mobile prosumer can be considered as an autonomous vehicle equipped with multiple DERs, such as thermoelectric generator, PV panel, and energy storage. The decision time horizon (e.g., one day) is assumed to be divided into multiple time steps (e.g., one hour).

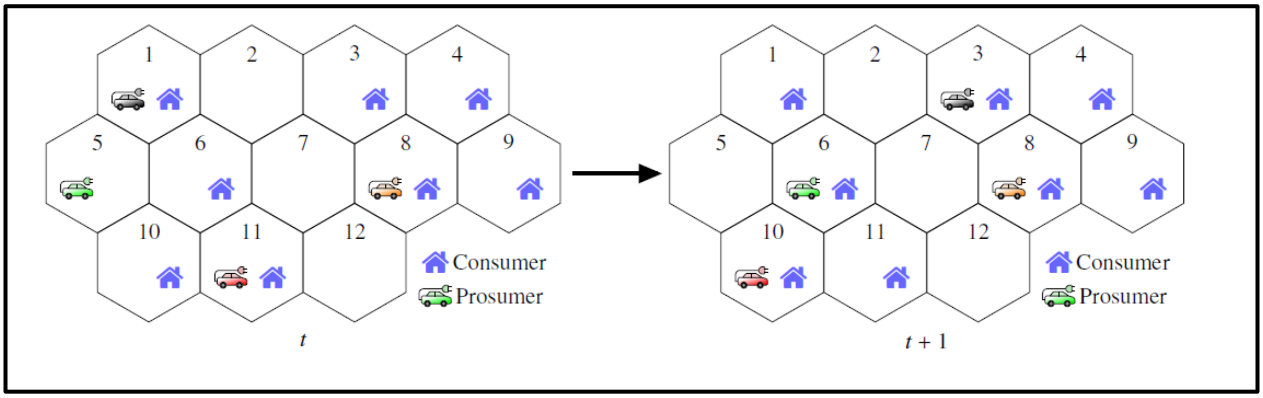

Figure 6: Overall system architecture of mobile prosumer network at two consecutive time steps t and t + 1

Given the energy loads and solar irradiance profiles of each consumer, the mobile prosumers will be dynamically dispatched to serve different consumers at different time periods. In this section, it is assumed that each mobile prosumer can only serve the consumers at its current region. The energy load of each customer will be satisfied using the mobile prosumer if it is available at the customer's region. Otherwise, the power grid will provide energy for that consumer. In addition, each mobile prosumer can also request energy from the power grid at its current region to charge its energy storage.

2.1.2. Operational Decision Model for Mobile Prosumer Network

The integrated vehicle routing and energy scheduling decision process is shown in Figure 7. Given the load and solar irradiance profiles and initial vehicle region, the optimization loop aims to find optimal vehicle routing and energy scheduling decisions for each time step (e.g., hourly) over the given decision time horizon (e.g., one day).

The following symbols are used to illustrate decision variables:

f_i	Energy cost for consumer i (\$)
f	Total energy costs for all consumers (\$)
eGp_i	Electricity purchased from power grids for consumer i at time t (kWh)
$fTG_{n,i,t}$	Fuel used to generate electricity from prosumer *n* to consumer i during time t (l)
$xTG_{n,t}$	ON/OFF state variable of TG for prosumer n at $time$ t
$eTG_{n,t}$	Electricity generated by TG of prosumer n at time *t* (kWh)
$ed_{n,i,t}$	Electricity supplied from prosumer n to consumer i at time *t* (kWh)
$ec_{n,i,t}$	Electricity received by prosumer n from power grid at consumer *i*'s region at time t (kWh)
$xed_{n,i,t}$	Indicates whether prosumer *n* supply energy to consumer i at time *t*
$xec_{n,i,t}$	Indicates whether prosumer n receive energy from power grid at consumer *i*'s region at time *t*
$ePV_{n,t}$	Electricity provided by PV of prosumer n at $time$ t (kWh)
$eB_{n,t}$	Amount of stored energy in storage of prosumer n at $time$ t (kWh)
$eBd_{n,t}$	Discharging rate of storage of prosumer n at time t (kWh)
$eBc_{n,t}$	Charging rate of storage of prosumer n at time *t* (kWh)
$xBd_{n,t}$	Discharging state variable of storage of prosumer n at time *t*
$xBc_{n,t}$	Charging state variable of storage of prosumer n at time *t*
$x_{n,i,t}$	Indicates whether prosumer n is located at consumer *i*'s region at time *t*
$xm_{n,i,t}$	Indicates whether prosumer *n* moves to another consumer's region at time *t*

The following symbols are used to illustrate parameters:

N	Total number of mobile prosumers indexed by n
I	Total number of consumers indexed by i
T	Total number of time steps indexed by t (hr)
PGp_t	Price of electricity purchased from power grid at time t (\$/kWh)
CE	Carbon emission conversion factor for electricity
PF_t	Price of fuel used to generate electricity at time t (\$/kWh)
CF	Carbon emission conversion factor for fuel
η_{TG}	Electricity generating efficiency of TG
$L_{i,t}$	Electricity load of consumer i at time t (kWh)
$Sol_{i,t}$	Solar irradiance at consumer i's region at time t (kWh/m^2)
SPV_n	Size of PV for prosumer n (m^2)
η_{PV}	Electricity generating efficiency of PV
η_{Bd}	Discharging efficiency of energy storage
η_{Bc}	Charging efficiency of energy storage
SBS_n	Size of energy storage of prosumer n (kWh)
α_{Bmin}	Coefficient for minimum storage limit of energy storage
α_{Bdmax}	Coefficient for maximum discharging limit of energy storage
α_{Bdmin}	Coefficient for minimum discharging limit of energy storage
α_{Bcmax}	Coefficient for maximum charging limit of energy storage
α_{Bcmin}	Coefficient for minimum charging limit of energy storage
EOB_n	Initial stored electricity in energy storage of prosumer n (kWh)
$E_{n,t}$	Electricity consumption by prosumer n at time t (kWh)

E_0	Electricity consumption rate due to mobility (kWh)
$X_{n,i,0}$	Indicates whether prosumer *n*'s initial location is *i*

2.1.2.1. Objective Function

In this study, the focus is on system operations (i.e., vehicle routing and energy scheduling) for a given system configuration, so design, investment, and configuration costs are ignored. The objective aims to minimize the total energy costs and carbon emission taxes for all consumers which is computed as

$$min\ f = \sum_i f_i \tag{1}$$

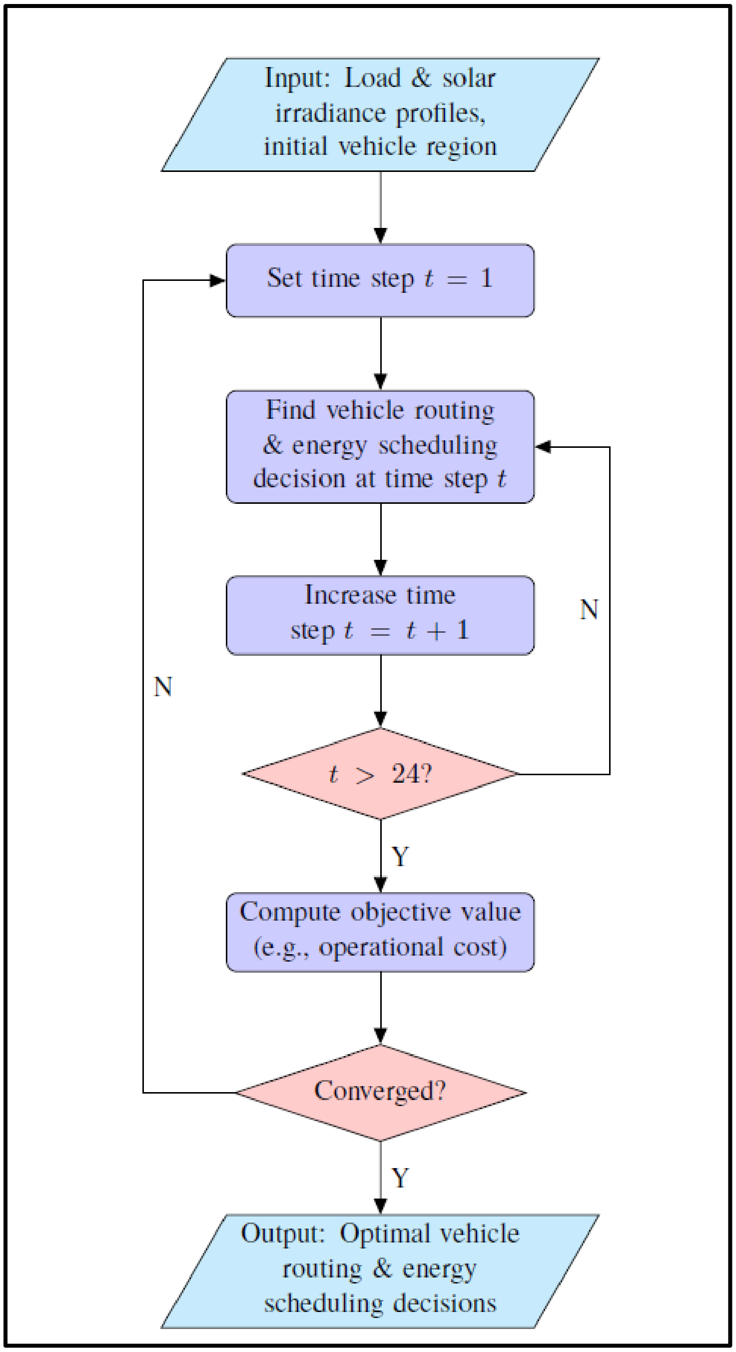

Figure 7: Flowchart of integrated vehicle routing and energy scheduling decision process

where the total cost for each consumer i is calculated as

$$f_i = \sum_t \sum_n fTG_{n,i,t} \times (PF_t + CF \times CT) + \sum_t eGp_{i,t} \times (PGp_t + CE \times CT) \tag{2}$$

and CT denotes the carbon tax.

2.1.2.2. Constraints

Energy balance constraints for consumers

$$eGp_{i,t} + \sum_n ed_{n,i,t} - \sum_n ec_{n,i,t} \geq L_{i,t} \quad \forall i, t \tag{3}$$

The right hand side of this equation is the energy load of each consumer, and the left hand side is the net energy supply from all the mobile prosumers and power grids.

Energy balance constraints for prosumers

$$ePV_{n,t} + eTG_{n,t} + eBd_{n,t} \times \eta_{Bd} + \sum_i ec_{n,i,t} = \frac{eBc_{n,t}}{\eta_{Bc}} + \sum_i ed_{n,i,t} + E_{n,t} \quad \forall n, t \tag{4}$$

The left hand side of this equation is the amount of energy coming from PV, thermoelectric generator, energy storage, and grid at consumer's region. The right hand side of this equation is the amount of energy requested by energy storage, consumers, and prosumer's energy consumption due to mobility.

Constraints for photovoltaic panel

$$ePV_{n,t} \geq \sum_{i} SPV_n \times Sol_{i,t} \times x_{n,i,t} \times \eta_{PV} \quad \forall n,t \tag{5}$$

The energy harvested by the PV panel depends on the size of the PV panel (SPV_n), solar irradiance at consumer i's region ($Sol_{i,t}$), the availability of prosumer at consumer i's region ($x_{n,i,t}$), and the PV efficiency (η_{PV}).

Constraints for thermoelectric generator

$$\sum_{i} fTG_{n,i,t} \leq M \times xTG_{n,t} \quad \forall n,t \tag{6}$$

$$eTG_{n,t} = \sum_{i} fTG_{n,i,t} \times \eta_{TG} \quad \forall n,t \tag{7}$$

$$fTG_{n,i,t} \leq M \times x_{n,i,t} \quad \forall n,i,t \tag{8}$$

Where M is a big number commonly used in mixed-integer programming modeling. The amount of fuel consumed by thermoelectric generator of prosumer n to generate electricity to consumer i is controlled by state variables of the generator ($xTG_{n,t}$) and the availability of prosumer at consumer i ($x_{n,i,t}$).

Constraints for energy storage

$$xBc_{n,t} + xBd_{n,t} \leq 1 \quad \forall n,t \tag{9}$$

$$SBS_n \times \alpha_{Bmin} \leq \ \ eB_{n,t} \leq SBS_n \quad \forall n, t \tag{10}$$

$$eB_{n,1} = EOB + \left(eBc_{n,1} - eBd_{n,1}\right)\Delta t \quad \forall n \tag{11}$$

$$eB_{n,t} = eB_{n,t-1} + \left(eBc_{n,t} - eBd_{n,t}\right)\Delta t \quad \forall n, t \geq 2 \tag{12}$$

$$SBS_n \times \alpha_{Bcmin} \times xBc_{n,t} \leq \ \ eBc_{n,t} \leq SBS_n \times \alpha_{Bcmax} \times xBc_{n,t} \quad \forall n, t \tag{13}$$

$$SBS_n \times \alpha_{Bdmin} \times xBd_{n,t} \leq \ \ eBd_{n,t} \leq SBS_n \times \alpha_{Bdmax} \times xBd_{n,t} \quad \forall n, t \tag{14}$$

Energy storage cannot be charged and discharged at the same time (see Eq. (9)). The amount of electricity stored in energy storage at any time should be kept between the permitted lowest level and its capacity (see Eq. (10)), and depends on charging/discharging activities (see Eqs. (11)-(12)), where Δt is the value of the time step. The charging/discharging rates should be kept within the range of lowest and highest level (see Eqs. (13)-(14)).

Constraints for energy exchanged between prosumers and consumers

$$xec_{n,t} + xed_{n,t} \leq x_{n,i,t} \quad \forall n, i, t \tag{15}$$

$$ec_{n,i,t} \leq M \times xec_{n,i,t} \quad \forall n, i, t \tag{16}$$

$$ed_{n,i,t} \leq M \times xed_{n,i,t} \quad \forall n, i, t \tag{17}$$

A prosumer cannot supply and receive electricity from consumers' regions at the same time (see Eq. (15)). Prosumer *n* is allowed to request electricity from power grid at consumer *i*'s region only if the state variable $xec_{n,i,t} = 1$, and supply electricity for consumer *i* only if the state variable $xed_{n,i,t} = 1$.

Mobility constraints for prosumers

To model prosumer mobility, two sets of decision variables are introduced: 1) $x_{n,i,t}$ indicates whether prosumer *n* is at consumer *i*'s region at time t, and 2) $xm_{n,t}$ indicates whether the prosumer will move at time *t*. It is assumed that the prosumer in consumer *i*'s region can only move one step ahead to its neighborhood $j \in N_i$. Let $\sum_{j \in N_i} x_{n,j,t}$ represent whether the prosumer *n* will move to its neighborhood when it is at consumer *i*'s region, then the relationships among $xm_{n,t}$, $x_{n,i,t}$ and $\sum_{j \in N_i} x_{n,j,t}$ will be modeled using the constraints in Eqs. (21)-(24).

$$\sum_i x_{n,i,t} = 1 \quad \forall n, t \tag{18}$$

$$\sum_i x_{n,i,t} \leq 1 \quad \forall i, t \tag{19}$$

$$E_{n,t} = xm_{n,t} \times E_0 \quad \forall n, t \tag{20}$$

$$x_{n,i,t-1} - x_{n,i,t} \leq xm_{n,t} \qquad \forall n, i, t \geq 2 \tag{21}$$

$$x_{n,i,t-1} - \textstyle\sum_{j \in N_i} x_{n,j,t} \leq 1 - xm_{n,t} \qquad \forall n, i, t \geq 2 \tag{22}$$

$$X_{n,i,0} - x_{n,i,1} \leq xm_{n,1} \qquad \qquad \forall n, i \tag{23}$$

$$X_{n,i,0} - \sum_{j \in N_i} x_{n,j,t} \leq 1 - xm_{n,1} \qquad \qquad \forall n, i \tag{24}$$

Each prosumer cannot be at more than one region at any time step (see Eq. (18)), also each region cannot have more than one prosumer at the same time (see Eq. (19)). If the prosumer moves, it will consume a fixed amount of energy (see Eq. (20)).

2.2 Simulation Results Analysis

In this study, a cluster of twelve energy consumers distributed at different regions in the city of Chicago is studied. One-day data for electricity load profiles for these consumers in the climate zone of Chicago are collected [42]. Solar irradiance for the Chicago area in year 2010 is used [109]. Residential time-of-use rate from [110] is adopted as the electricity purchasing price from the power gird. All the other parameters have the same settings as parameters in [42, 111]. The prosumers' electricity consumption rates for mobility are adopted from [112]. All the parameter settings are provided in Appendix A. All the mathematical decision models implemented in this section are solved by CPLEX with a relative gap of 0.005.

To compare the performance of the proposed integrated model (IM), three disjoint solutions are proposed. In the first model (DM1), the four prosumers are assumed to stay at the four consumers with minimum load, the second model (DM2) assumes they will stay at the four consumers with maximum load, and the third model (DM3) assumes that each prosumer will move to the four neighbor consumers with maximum load at each time step. A baseline model (BM) which assumes no mobile prosumers, and in which consumers always request energy from the power grid, is also studied.

2.2.1 Mobile Prosumer Network Under Different Pricing Mechanisms

Energy price is an important factor that directly affects energy costs. Nowadays, utility companies offer different pricing plans to their clients, and clients can choose the plans that best suit their energy consumption to reduce the costs. In this section, the performance of MPN is evaluated under two different pricing mechanisms: (i) a basic plan where the energy price keeps as a constant for all of the consumers and (ii) a time of use (TOU) plan where the energy price at peak period is much higher compared to the off-peak period. The operational costs, energy requested from the power grid, and carbon emissions for five models (baseline, three disjoint models, integrated model) at two different pricing mechanisms are shown in Figure 8. The operational cost saving, requested energy reduction, and carbon emission reduction for the three disjoint models and integrated model compared to the baseline are recorded in Table III. It is observed that the integrated model is more cost effective, is less dependent on power grids, and produces less carbon emissions compared to disjoint models. The integrated model can achieve more cost savings using TOU pricing plan since the prosumers can be more effectively dispatched to supply lower energy cost to consumers of high prices.

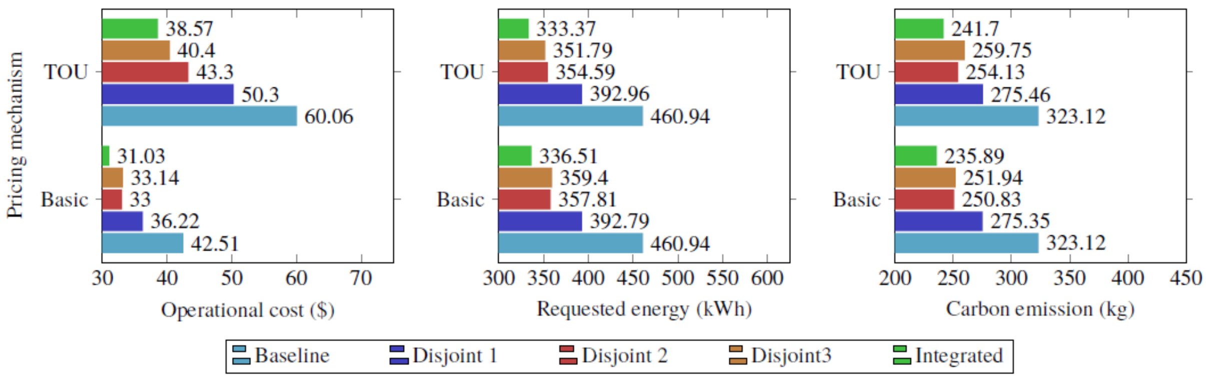

Figure 8: Operational cost, energy requested from the power grid, and carbon emission under different pricing mechanisms

TABLE I. Operational cost saving, requested energy reduction, and carbon emission reduction under different pricing mechanisms (DM1: disjoint model 1; DM2: disjoint model 2; DM3: disjoint model 3; IM: integrated model)

Pricing mechanism	Performance metrics											
	Operational cost saving (%)				Requested energy reduction (%)				Carbon emission reduction (%)			
	DM1	DM2	DM3	IM	DM1	DM2	DM3	IM	DM1	DM2	DM3	IM
Basic	14.78	22.37	22.03	27	14.78	22.37	22.03	27	14.78	22.37	22.03	27
TOU	16.24	27.9	32.73	35.78	14.75	23.07	23.68	27.68	14.75	21.35	19.61	25.2

Simulation results in Figure 8 and Table I show that the integrated model can achieve more cost savings compared to other models (solutions obtained by disjoint models 1, 2, and 3) for the two pricing mechanisms. This is due to the fact that disjoint 1 and 2 models are fixed at certain regions and can only harvest solar energy from those regions, and disjoint 3 model is merely routed based on maximum loads. The integrated model, on the other hand, can be dynamically dispatched to locations depending on their loads, amount of solar irradiation, and cost of travel, thus to satisfy more energy demands with minimum possible cost.

The integrated model under the two pricing mechanisms is less dependent on power grids than the three disjoint models since various distributed energy resources can be more efficiently allocated to different regions to reduce energy waste. In this section, the carbon emissions associated with both energy requested from the power gird and natural gas used for thermoelectric generator are considered.
To this end, it is concluded that the integrated model can produce solutions that are more cost effective, less dependent on power grids, and environmentally friendly since the energy scheduling and routing decisions are considered jointly.

2.2.2 Mobile Prosumer Network Under Different Weather Conditions

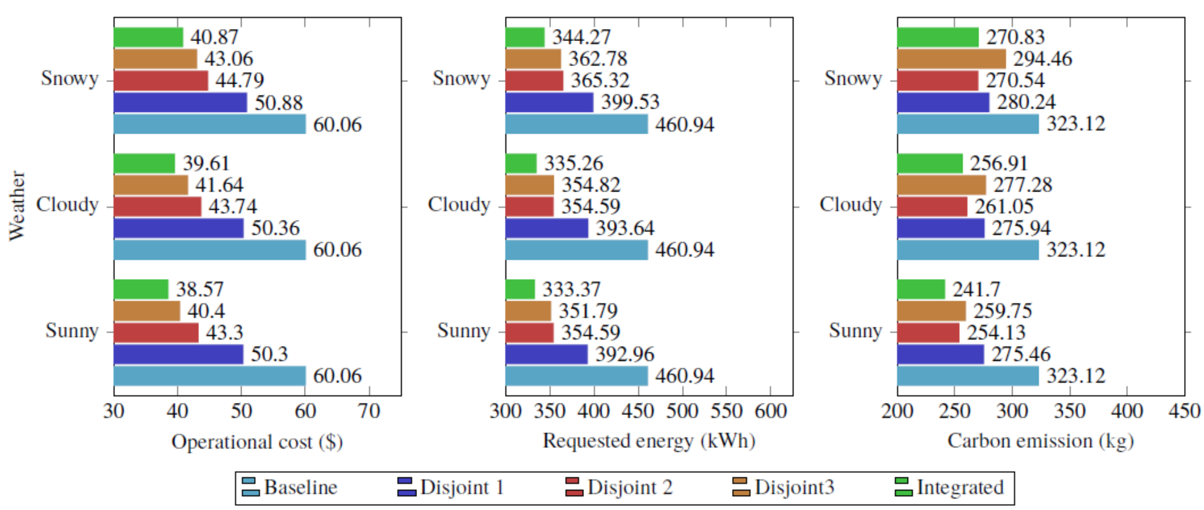

Figure 9: Operational cost, energy requested from the power grid, and carbon emission under

different weather conditions

The mobile prosumer network will be very useful for disaster response when the power supply from the grid is limited. In these scenarios, the mobile prosumers should be dispatched in an efficient way to harvest more renewable energy (e.g, solar energy) and satisfy consumers' demand as much as possible. In this section, the performance of the integrated decision model is evaluated under different solar irradiance profiles corresponding to different weather conditions. Specifically, three representative weather conditions (e.g., sunny, cloudy, snowy) for the Chicago area in year 2010 [44] are chosen. It is expected that the mobile prosumer may harvest more solar energy during sunny days.

The operational costs, energy requested from the power grid, and carbon emissions for the five models (baseline, three disjoint models, integrated model) at three different solar irradiance profiles are shown in Figure 9. The operational cost saving, requested energy reduction, and carbon emission reduction for the three disjoint models and integrated model compared to the baseline are recorded in Table II. It is observed that the integrated model performs better than the three disjoint models under all weather conditions. In addition, the integrated model under cloudy and snowy days can perform better than or comparably to the disjoint models under sunny days. This result further demonstrates the necessity and effectiveness of the integrated model.

TABLE II. Operational cost saving, requested energy reduction, and carbon emission reduction under different weather conditions (DM1: disjoint model 1; DM2: disjoint model 2; DM3: disjoint model 3; IM: integrated model)

Weather	Performance metrics											
	Operational cost saving (%)				Requested energy reduction (%)				Carbon emission reduction (%)			
	DMI	DM2	DM3	IM	DMI	DM2	DM3	IM	DMI	DM2	DM3	IM
Sunny	16.24	27.9	32.73	35.78	14.75	23.07	23.68	27.68	14.75	21.35	19.61	25.2
Cloudy	16.14	27.17	30.67	34.04	14.6	23.07	23.02	27.27	14.6	19.21	14.19	20.49
Snowy	15.27	25.42	28.3	31.94	13.32	20.75	21.3	25.31	13.27	16.27	8.87	16.18

2.2.3 Mobile Prosumer Network Under Different Consumer Patterns

Consumers are expected to have different lifestyles, since they have different financial capabilities, traditions, employment status, etc,, and consequently they will have different energy consumption patterns. In this section, the performance of MPN under various consumer patterns is studied. Two consumer patterns are analyzed: (i) the "similar" pattern where each consumer has only one peak load and (ii) the "different" pattern where some consumers have one peak load, while others may have two or multiple peaks.

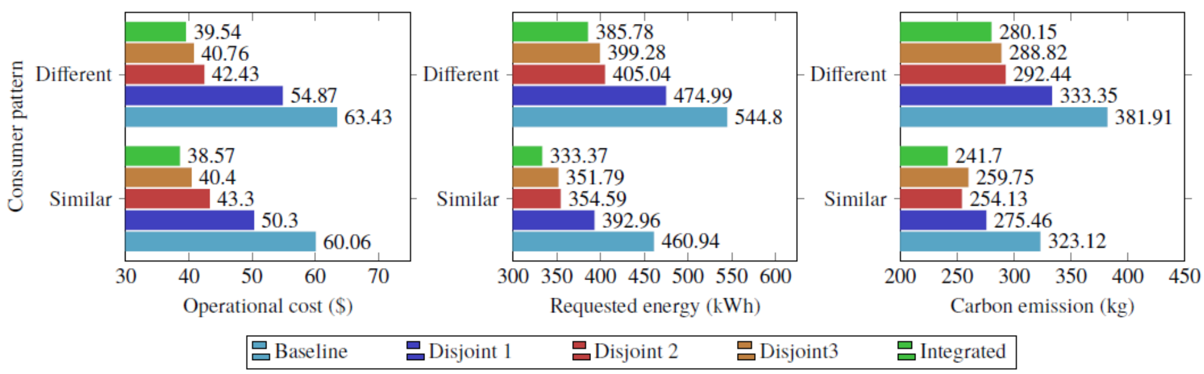

Figure 10: Operational cost, energy requested from the power grid, and carbon emission under different consumer patterns

The operational costs, energy requested from the power grid, and carbon emissions for the five models (baseline, three disjoint models, integrated model) at two different consumer patterns (similar and different patterns) are shown in Figure 10. The operational cost saving, requested energy reduction, and carbon emission reduction for the three disjoint models and integrated model compared to the baseline are recorded in Table III. It is observed that the integrated model performs better than the three disjoint models under the two patterns. The integrated model enables the prosumers move to neighbor regions with maximum loads depending on time periods and satisfy their energy demands with cheaper energy.

TABLE III. Operational cost saving, requested energy reduction, and carbon emission reduction under different consumer patterns (DM1: disjoint model 1; DM2: disjoint model 2; DM3: disjoint model 3; IM: integrated model)

Consumer pattern	Performance metrics											
	Operational cost saving (%)				Requested energy reduction (%)				Carbon emission reduction (%)			
	DMI	DM2	DM3	IM	DMI	DM2	DM3	IM	DMI	DM2	DM3	1M
Similar	16.24	27.9	32.73	35.78	14.75	23.07	23.68	27.68	14.75	21.35	19.61	25.2
Different	13.48	33.1	35.73	37.66	12.81	25.65	26.71	29.19	12.71	23.43	24.37	26.65

2.2.4 Mobile Prosumer Network Under Different Mobility Energy Consumption Rates

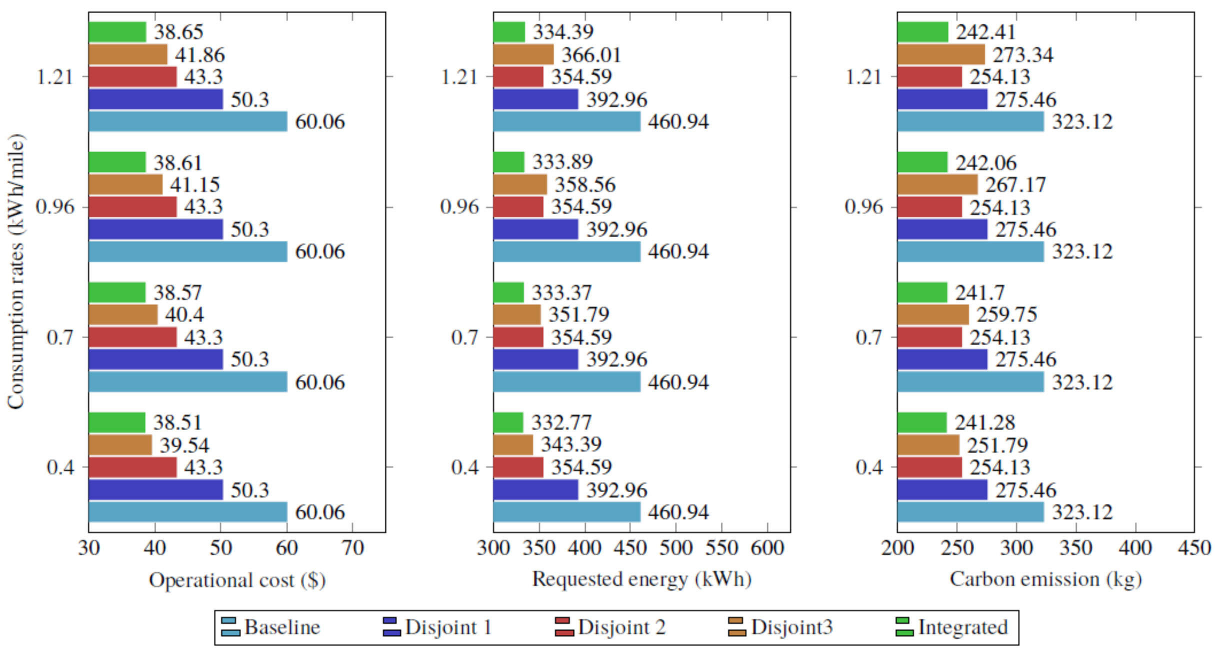

Figure 11. Operational cost, energy requested from the power grid, and carbon emission under different consumption rates

According to the U.S. Department of Energy, EVs consume electricity as they move with rates ranging between 25 and 120 kWh per 100 miles [47]. In this section, the performance of MPN is evaluated under various energy consumption rates due to mobility. For demonstration, four different consumption rates will be considered which are 40, 70, 96, and 121 kWh per 100 miles. The operational costs, energy requested from the power grid, and carbon emissions for these five models (baseline, three disjoint models, integrated model) at four different consumption rates are shown in Figure 11. The operational cost saving, requested energy reduction, and carbon emission reduction for the three disjoint models and integrated model compared to the baseline are recorded in Table IV. It is observed that the integrated model solutions are more cost effective and less dependent on power grids, and produce less carbon emissions compared to the three disjoint models under all four consumption rates. The operational cost, requested energy, and carbon emission will increase when the consumption rates increase.

TABLE IV. Operational cost saving, requested energy reduction, and carbon emission reduction under different consumption rates (DM1: disjoint model 1; DM2: disjoint model 2; DM3: disjoint model 3; IM: integrated model)

Consumption rates (kWh/mile)	Performance metrics											
	Operational cost saving (%)				Requested energy reduction (%)				Carbon emission reduction (%)			
	DM I	DM2	DM3	1M	DMI	DM2	DM3	IM	DM I	DM2	DM3	IM
0.4	16.24	27.9	34.16	35.87	14.75	23.07	25.5	27.81	14.75	21.35	22.08	25.33
0.7	16.24	27.9	32.73	35.78	14.75	23.07	23.68	27.68	14.75	21.35	19.61	25.2
0.96	16.24	27.9	31.47	35.71	14.75	23.07	22.21	27.56	14.75	21.35	17.32	25.09
1.21	16.24	27.9	30.3	35.64	14.75	23.07	20.6	27.46	14.75	21.35	15.41	24.98

2.2.5. Mobile Prosumer Network Under Different Consumer Distances

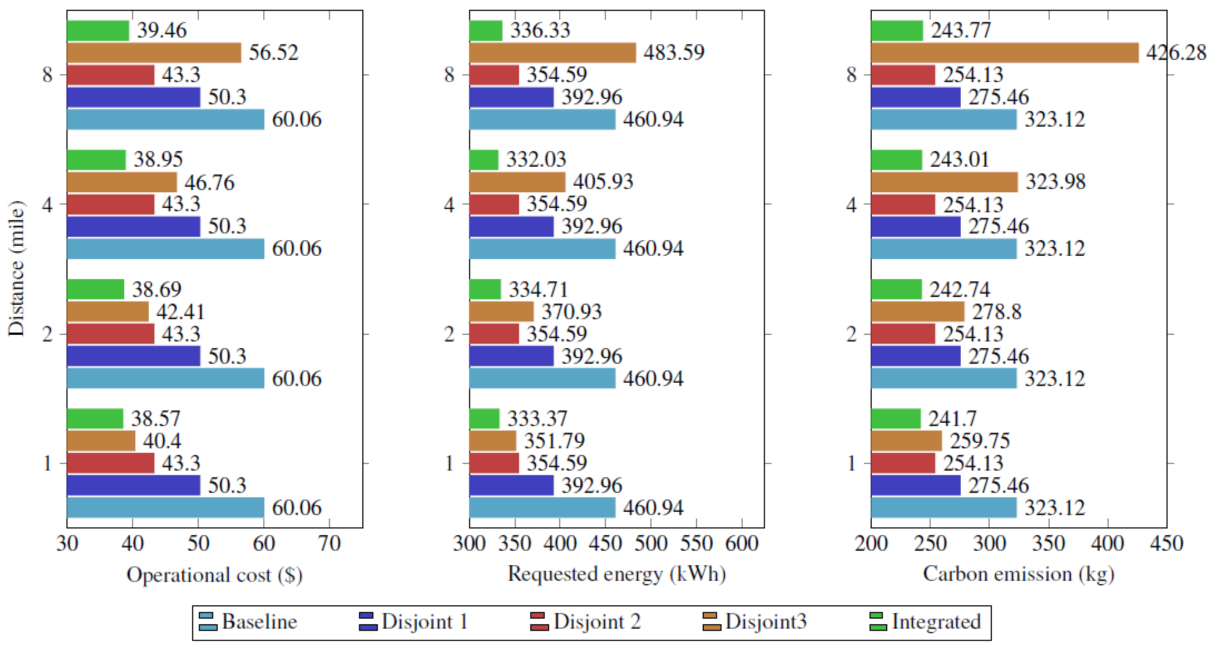

Figure 12: Operational cost, energy requested from the power grid, and carbon emission under different consumer distances

The distances between consumers affect the energy consumption of the mobile prosumers and consequently affect energy supply to consumers. Thus, it is vital to study the impact of distances between consumers on the performance of the mobile prosumer network. For demonstration, four different distances will be considered which are 1, 2, 4, and 8 miles. The operational costs, energy requested from the power grid, and carbon emissions for these five models (baseline, three disjoint models, integrated model) at four different consumer distances are shown in Figure 12. The operational cost saving, requested energy reduction, and carbon emission reduction for the three disjoint models and integrated model compared to the baseline are recorded in Table V. It is observed that the solutions from the integrated model are more cost effective and less dependent on power grids, and produce less carbon emissions compared to the three disjoint models under all the four distances. The operational cost, requested energy, and carbon emission will increase when the distance increases.

TABLE V. Operational cost saving, requested energy reduction, and carbon emission reduction under different consumer distances (DM1: disjoint model 1; DM2: disjoint model 2; DM3: disjoint model 3; IM: integrated model)

Distance (mile)	Performance metrics											
	Operational cost saving (%)				Requested energy reduction (%)				Carbon emission reduction (%)			
	DM1	DM2	DM3	IM	DM1	I)M2	DM3	IM	DM1	DM2	DM3	IM
1	16.24	27.9	32.73	35.78	14.75	23.07	23.68	27.68	14.75	21.35	19.61	25.2
2	16.24	27.9	29.38	35.58	14.75	23.07	19.53	27.39	14.75	21.35	13.72	24.88
4	16.24	27.9	22.14	35.15	14.75	23.07	11.94	27.97	14.75	21.35	-0.27	24.79
8	16.24	27.9	5.88	34.29	14.75	23.07	-4.91	27.03	14.75	21.35	-31.93	24.56

2.3 Conclusions

In this chapter, the performance of a network of mobile prosumers is measured to provide energy to consumers at different regions in terms of cost effectiveness, dependency on the power grid, and environmental friendliness. An integrated vehicle routing and energy scheduling decision model based on mixed integer programming is proposed to study the tradeoff between vehicle mobility and energy service quality. Five case studies designed using the proposed decision model are used to evaluate the impacts of different pricing mechanisms, weather conditions, consumer patterns, mobility energy consumption rates, and consumer distances on the performance of the mobile prosumer network. The simulation results demonstrate that the mobile prosumer network (integrated model) can not only shift energy load from peak to off-peak periods, but also from high demand to low demand consumers to mitigate the impacts of spatio-temporal complexity in energy demand and production. Thus, it can be more cost effective, less dependent on power grids, and more environmentally friendly than disjoint models, as it can save up to 38% of energy costs and reduce up to 29% and 27% of energy requested from main grid and carbon emissions compared to a baseline where consumers always request energy from power grids, respectively. There are a few requirements that should be considered for the implementation of MPN. For example, the vehicles are suggested to be unmanned or autonomous. The charging stations are available at each consumer's region.

3. DYNAMIC ENERGY SCHEDULING AND ROUTING OF MULTIPLE ELECTRIC VEHICLES USING DEEP REINFORCEMENT LEARNING

[The majority of work presented in this chapter of the dissertation is submitted for publication in Renewable and Sustainable Energy Reviews journal. For more information, please refer to Appendix B (Copyright Statement).]

This chapter is arranged as follows: Section 3.1 shows problem statement and mathematical model formulated as mixed integer programming (MIP) optimization problem. Section 3.2 shows the reformulation of the MIP model using MDP. The experiments to assess the performance of MDP are illustrated in section 3.3 and section 3.4 presents the conclusion.

3.1 Mathematical Optimization Model

Section 3.1 discusses the system architecture and components to model the EV network, and Section 3.2 presents the mathematical formulation of the EV network operational decision problem.

3.1.1. System Configuration for the Electric Vehicle Network

In this work, it is assumed that a geographic area is split into mutually exclusive and collectively exhaustive square areas (see Fig. 1) with equal distances of one mile between each pair of connected regions. It is assumed that each area may contain one consumer or may not contain any consumer, hence the distance between each pair of consumers ranges from one to multiple miles. Fig. 1 demonstrates a system of eight users and four EVs. These users have different energy loads and radiant energy and receive energy from a network of EVs. Each EV is mounted with a photovoltaic panel and power storage. Each EV makes decisions at each time step (e.g., one hour), summing up to a total decision time horizon (e.g., one day).

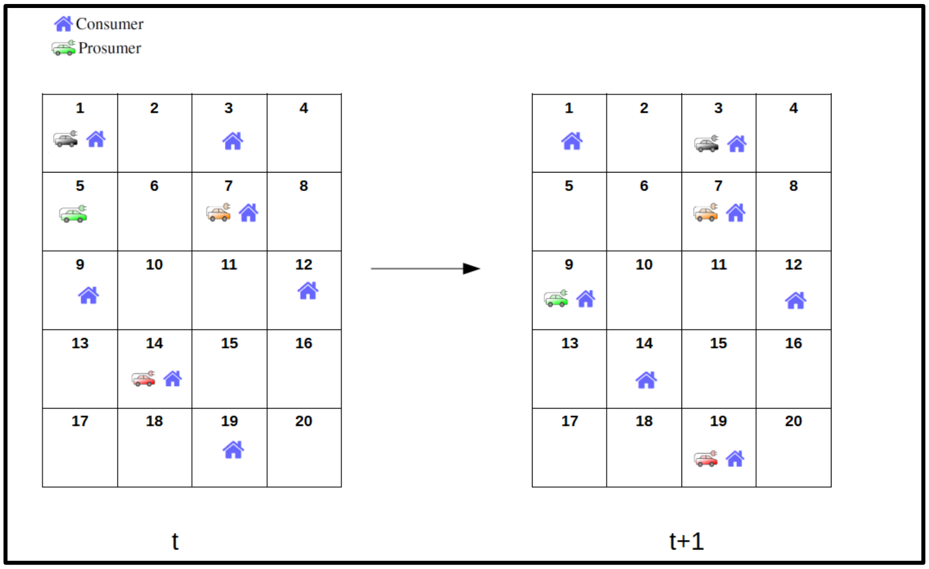

Figure 13: Demonstration of VR decisions at time *t* and *t* + 1

In this research, each EV moves around and supplies energy to the users at its current area only. In other words, each customer's load can be satisfied using an EV only if that EV exists at the user's location. Otherwise, the customer's load will be satisfied from the power grid with an incurred cost. In addition, each EV can charge its energy storage by requesting energy from the power grid at its current location.

3.1.2. Operational Decision Model for Electric Vehicle Network

All the variables and parameters introduced in the model are described in chapter 2 of this thesis.

3.1.2.1 Objective Function

In this study, the main objective is to reduce energy costs for all consumers via reducing the dependence on power purchased from the main power grid. Other system costs including design, investment, and configuration costs are disregarded.

3.1.2.2 Constraints

The constraints of this model are the same as Eqs. (2-24) in Chapter 2.

3.2 Proposed Reinforcement Learning Based Decision Model

In this section, the mathematical model proposed in chapter 2 will be reformulated using Markov decision processes. A reinforcement learning algorithm is adopted to solve the reformulated problem. The agent is modeled in Section 3.2.1.

The environment model is presented in Section 3.2.2. The architecture of the deep Q-network to solve the MDP formulation is discussed in Section 3.2.3.

3.2.1 Electric Vehicle Agent

The EV agent is assumed to be equipped with a PV panel and energy storage unit that move around and supply energy to consumers at various locations to satisfy their energy demand at lower energy costs. The system state $s \in \mathcal{S}$ is a tuple of four variables $s_t = (p_t, eB_t, Sol_t, L_t)$ corresponding to vehicle location p_t, battery state of charge eB_t, solar irradiance Sol_t, and energy load L_t respectively. The action space $a \in \mathcal{A}$ is defined as the mobility actions (moving up, down, left, right, and no move) and energy transaction actions (charging, discharging, and idle). The movement of each vehicle is restricted to the neighboring regions in an orthogonal movement in a 5×4 grid of regions.

3.2.2. Environment Agent

Given the system state s_t at time step t, the system state will transit to s_{t+1} after executing action a_t following the state transition rules. Once the vehicle chooses which mobility action to take, the location of the vehicle p_t is updated to p_{t+1}.

Given the vehicle's location at the next state, the energy load Lt and solar irradiance Sol_t are updated to next state load L_{t+1} and solar irradiance Sol_{t+1}. After that, based on the ES action taken (charge, discharge, idle) at the current state, the battery state of charge variable eB_t is updated to the next state battery level eB_{t+1}, and output from PV will be calculated based on the solar irradiance at the next state Solt+1. The state of charge at the next state (eBt+1) can be computed in Eqs. (8)-(9), given the state of charge at the current state (eB_t). The location of the vehicle at next state (p_{t+1}) can be obtained from routing constraints in Eqs. (15)-(20). The solar irradiance and load at next state (Sol_{t+1} and L_{t+1}) are obtained from the sequence of the solar irradiance and load data.

The reward function is defined as follows:

$$R(s_t, a_t) = \sum_{i \in I} eGp_t^i \times PGp_t \tag{25}$$

Where PGp_t is the price of energy from the power grid. The energy costs can be obtained from Eq. (25), as it represents the amount of demand unsatisfied by EVs but satisfied instead by the main power grid at that region. eGp_t^i represents the energy requested from the power grid and can be computed as follows:

$$eGp_t^i = max\left(L_t^i - \sum_{n \in N} ed_t^{n,i} + \sum_{n \in N} ec_t^{n,i}, 0\right) \quad \forall i \epsilon I, \forall t \tag{26}$$

The EV energy balance constraint in Eq. (4) is automatically satisfied based on the definition of system states and actions. The first energy storage constraint (Eq. (9)) is satisfied based on the definition of the battery actions, so that at each time *t*, the battery can only take one action (charge, discharge, or idle). The battery capacity constraint in Eq. (10) is included in the definition of the storage parameters, so that at each time step, the level in each battery is maintained between certain boundaries. The constraint in Eq. (11) is satisfied during model initialization. For example, at time t = 1, a certain amount of energy is stored in the battery. The constraint in Eq. (12) is satisfied based on the state transition rule. The amount of energy transactions between power grid and battery shown in Eqs. (13)-(14) are maintained between certain boundaries (rates of charging and discharging). These constraints are satisfied by calculating the amount of energy received or supplied by the battery at each time step, such that they do not fall below lower limit or exceed upper limit of charging/discharging rates, respectively.

The energy exchange between EVs and consumers is modeled according to Eqs. (15-17). Eq. (15) is satisfied based on the definition of action space. At each time step, each EV can either receive energy from the power grid or supply energy to the consumer. Constraints in Eqs. (16-17) are handled by setting upper and lower limits for the amount of energy exchange between EVs and consumers to fixed amounts.

Finally, the mobility constraint in Eq. (18) is automatically satisfied. The constraint in Eq. (19) is satisfied by adding a condition for each region to have no more than one EV at each time t. The mobility of vehicles at each time step shown in Eq. (21) and Eq. (23) are considered in the definition of the action space, such that each EV will make decisions about whether to move to surrounding regions or stay at the current region. The locations of EVs at the next time step shown in Eq. (22) and Eq. (24) are satisfied based on the definition of action space, which restricts the mobility to only the neighboring regions of each vehicle.

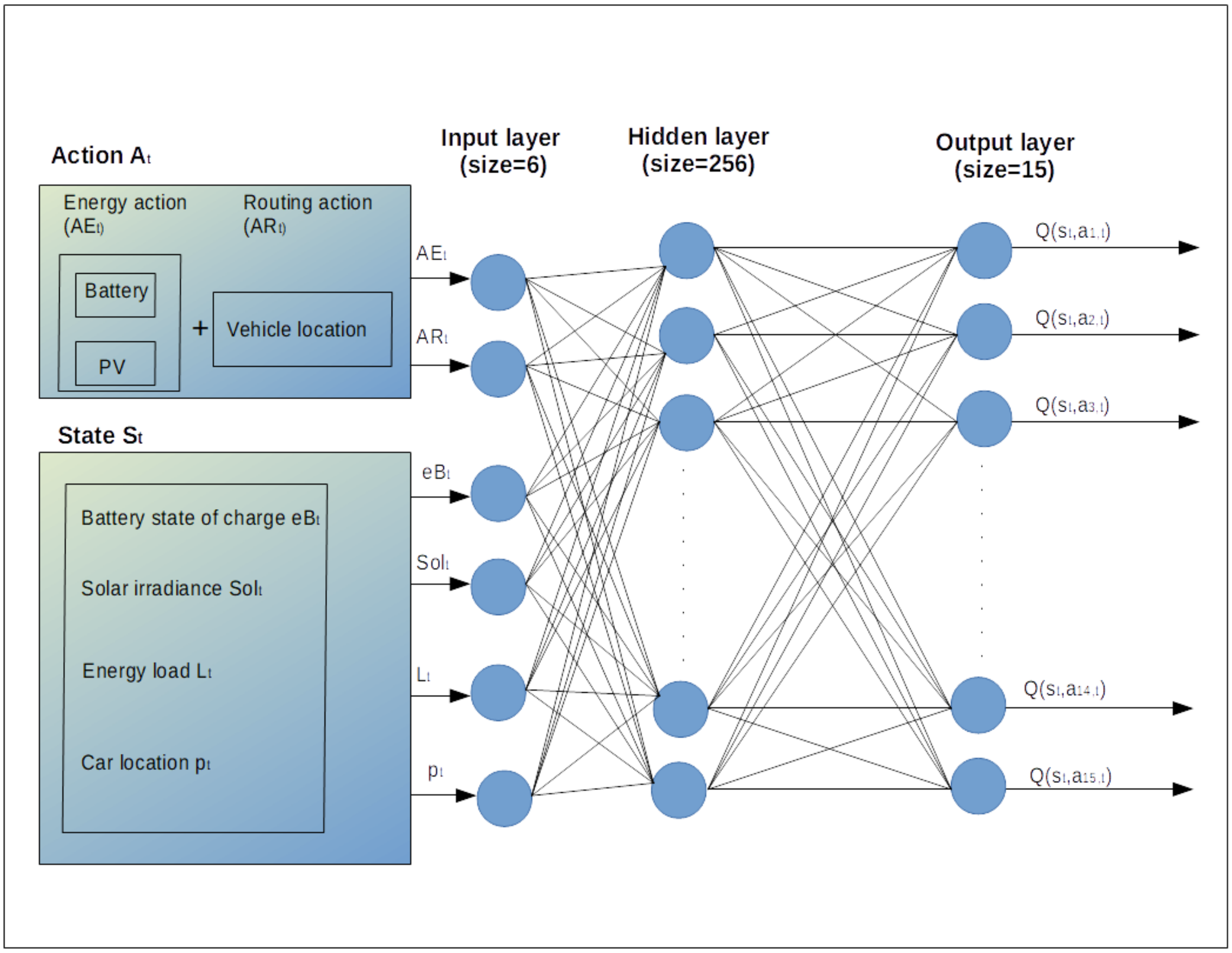

Figure 14: Proposed deep Q-network architecture

3.2.3. Deep Q-Network Architecture

To solve the reformulated decision problem, a deep Q-network (DQN) is adopted to estimate the Q-value for each action by the EV agent. It takes the state variables including the vehicle's location, battery level, consumer load, and solar irradiance at the consumer's location ($s_t = (p_t$,eB, Sol_t , L_t)) and the routing and scheduling actions (a_t) as input, and passes them through the hidden layer and output layer of the network to approximate Q-values corresponding to each combination (routing and scheduling) of the agent's actions. DQN is used because it is more efficient at handling larger numbers of inputs and outputs compared to Q-tables, which can only deal with discrete numbers of inputs and outputs. A neural network with parameters θ, is used to estimate the Q-values, i.e. $Q(s,a;\theta) \approx Q^*(s,a)$. To do this, the following loss function need to be minimized [124]:

$$L_t(\theta_t) = \mathbb{E}_{s,a,r,s'\sim\rho(.)}[(y_t - Q(s,a;\theta_t))^2] \quad (27)$$

Here, ρ represents the behavior distribution over transitions *s,a,r,s′* collected from the environment, and y_t is the temporal difference target which is calculated as:

$$y_t = r + \gamma max_{a'} Q(s', a'; \theta_{t-1}) \quad (28)$$

The architecture of the proposed DQN is illustrated in Fig. 14. In this figure, the DQN receives input from both actions and states at time *t* including state variables and routing and scheduling actions that have been made at time *t*. It then passes the input through a hidden layer with the size of 256 neurons to the output layer. The size of the output layer is 15 neurons corresponding to the discounted rewards (estimated Q-values) for each combination of actions (i.e., routing and scheduling). This process is repeated over many episodes to train the DQN to map each combination of actions along with its approximated rewards. Then, in the execution stage, the agents will use the trained neural network to evaluate the rewards for each action and select the best action.

3.3 Simulation Results Analysis

In this research, it is assumed thar there are twenty energy users located at different locations in the Chicago area. 24hr data for energy demand profiles for these users at the weather zone of Chicago are collected [42]. The solar irradiance profile used in this paper is taken from the city of Chicago in year 2010 [109]. Prices of electricity purchased from the power grid were taken from [110]. All the other parameters used in the experiments in this paper are taken from [42, 111]. The mathematical programming model proposed in Section 3.3 is solved using CPLEX with a relative gap of 0.005. The CPLEX algorithm is considered an exact optimizer that provides an optimal solution for benchmarking.

To compare the performance of the presented DQN model, three heuristic solutions including genetic algorithm, particle swarm, and artificial fish swarm algorithms are studied. Each EV is assumed to go to one of the four neighboring locations (locations of users) at each time step. The DQN is trained using one year energy load and solar irradiance data for 1,000 episodes. Each algorithm is evaluated on 100 problem instances.

3.3.1. Scalability Evaluation

3.3.1.1. Energy Network with Four EVs and Twenty Consumers

Table VI: Simulation results for energy network with four EVs and twenty consumers

Algorithm	Operational cost ($)				Decision time (sec)
	Mean	Std. dev.	Min	Max	
CPLEX	435.44	20.85	231.13	488.37	41
DQN	467.13	10.35	291.46	495.84	2
GA	535.14	19.06	336.86	585.80	68
PSO	537.54	19.06	339.26	588.20	17
AFSA	523.46	17.00	334.69	568.79	215

In this section, the performance of the DQN algorithm is evaluated using a case of four EVs and twenty consumers. The operational costs and decision time per problem instance for five algorithms (CPLEX, DQN, and three heuristic algorithms) are shown in Table VI. It is observed that the DQN algorithm can obtain better solutions compared to the heuristic algorithms, with an average cost reduction of 21.31%. In addition, the DQN algorithm can obtain a near-optimal solution very quickly compared to other algorithms. The convergence curve for DQN during training is demonstrated in Fig. 15.

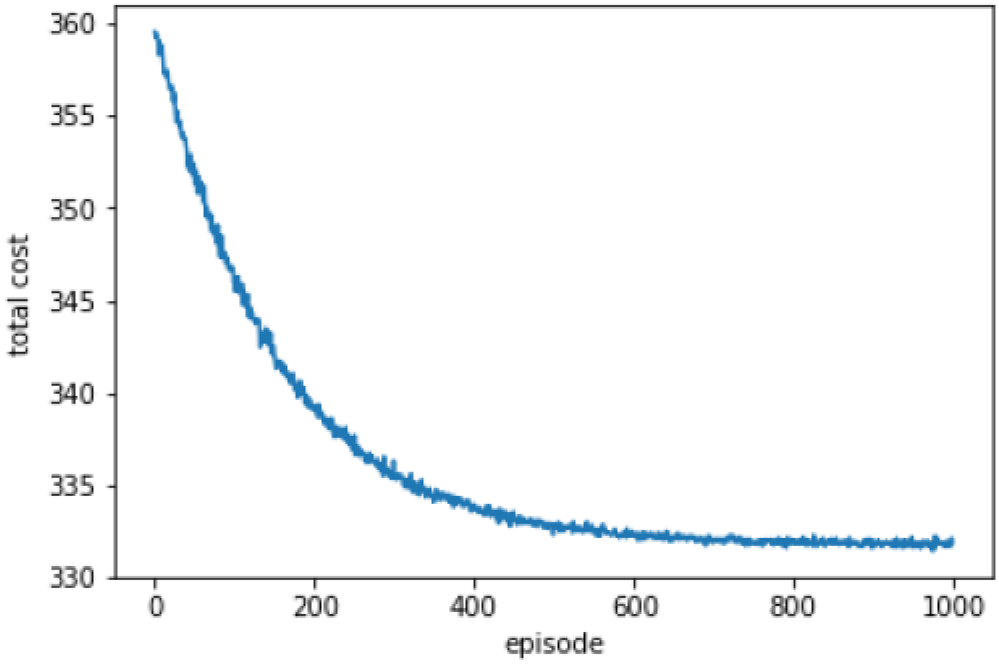

Figure 15: Convergence curve for the DQN during training

3.3.1.2. Energy Network with Eight EVs and Twenty Consumers

In this section, the performance of the DQN algorithm is evaluated using a case of eight EVs and twenty consumers. The operational costs and decision time per problem instance for four algorithms (DQN and three heuristic algorithms) are shown in Table VII. CPLEX is computationally expensive for this case, so results from CPLEX will not be reported here.

TABLE VII. Simulation results for energy network with eight EVs and twenty consumers

Algorithm	Operational cost ($)				Decision time (sec)
	Mean	Std. dev.	Min	Max	
DQN	418.55	7.99	285.85	442.99	6
GA	498.75	25.38	321.81	560.45	81
PSO	501.15	25.38	324.21	562.85	30
AFSA	503.55	25.38	326.61	565.25	230

It is observed that the DQN algorithm can generate better solutions compared to the heuristic algorithms, with average cost reduction of 19.53%. It further demonstrates the computational efficiency of the DQN algorithm for fast decision-making.

3.3.1.3. Energy Network with Twelve EVs and Forty Consumers

In this section, the performance of DQN algorithm is evaluated using a case of twelve EVs and forty consumers. The operational costs and decision time per problem instance for four algorithms (DQN and three heuristic algorithms) are shown in Table VIII.

TABLE VIII. Simulation results for energy network with twelve EVs and forty consumers

Algorithm	Operational cost ($)				Decision time (sec)
	Mean	Std. dev.	Min	Max	
DQN	821.03	18.64	514.47	877.41	13
GA	962.14	44.27	586.29	1074.14	125
PSO	966.93	44.28	591.08	1078.93	44
AFSA	969.33	44.28	593.48	1081.33	246

It is observed that the DQN algorithm can yield better solutions compared to the heuristic algorithms, with average cost reduction of 17.51%. It is observed that the DQN algorithm can output solutions that are better in terms of energy cost compared to the heuristic algorithms. Moreover, the DQN algorithm outperforms the heuristic algorithms in terms of the simulation time.

3.3.1.4. Energy Network with Sixteen EVs and Forty Consumers

In this section, the performance of DQN algorithm is evaluated using a case of sixteen EVs and forty consumers. The operational costs and decision time per problem instance for four algorithms (DQN and three heuristic algorithms) are shown in Table IX.

TABLE IX. Simulation results for energy network with sixteen EVs and forty consumers

Algorithm	Operational cost ($)				Decision time (sec)
	Mean	Std. dev.	Min	Max	
DQN	606.03	14.98	345.50	657.37	23
GA	923.48	49.62	571.23	1048.75	149
PSO	928.25	49.62	576.02	1053.55	55
AFSA	930.65	49.62	578.42	1055.95	279

The DQN algorithm can generate better solutions, with average cost reduction of 52.78%. DQN also outperforms the three heuristic algorithms in terms of computational time.

3.3.1.5. Summary of Scalability Evaluation

Fig. 16 represents the simulation time for these four energy networks. It is observed that GA and AFSA algorithms show quadratic or parabolic patterns, whereas PSO shows a linear pattern as the problem scale increases. The DQN's simulation time increases in a linear pattern with a smaller slope compared to PSO, which makes it more appropriate for fast decision-making to quickly respond to system dynamics.

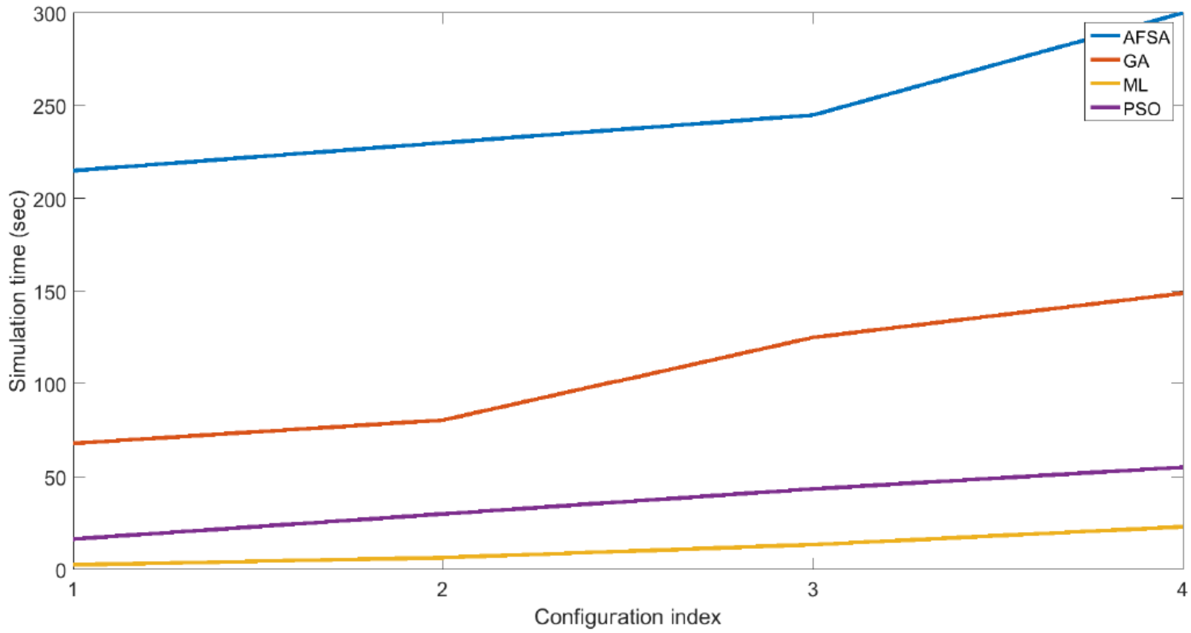

Figure 16: Simulation time for four different energy networks

In summary, a t-test is conducted to compare the five algorithms in terms of solution quality and computational time. Table XIV summarizes the results of the t-test for each case. The first symbol in "+/+" indicates that the DQN algorithm significantly outperforms the compared algorithm in terms of solution quality, and the second symbol indicates DQN algorithm significantly outperforms the compared algorithm in terms of computational time. From Table X, it is concluded that the DQN algorithm significantly outperforms the three heuristic algorithms in terms of both solution quality and computational time.

TABLE X. Summary of t-test results to compare DQN with other algorithms

Algorithm	t-test			
	Case 1	Case 2	Case 3	Case 4
CPLEX	-/+			
GA	+/+	+/+	+/+	+/+
PSO	+/+	+/+	+/+	+/+
AFSA	+/+	+/+	+/+	+/+

Fig. 17 illustrates the routing decisions of EVs that supply energy to consumers in different regions using a DQN algorithm along with other algorithms at critical time steps (hours 11, 12, 13, and 15), respectively. For demonstration purposes, one EV for each algorithm is used to show the routing decisions of an EV at critical time steps for twenty consumers. The total energy costs of each algorithm at each time step are shown to demonstrate how the routing decisions affect the energy costs of multiple consumers in the network. It is observed that the CPLEX and DQN algorithms outperform the heuristic algorithms in terms of total energy costs at critical time steps (hour 11, 12, 13, and 15). Both CPLEX and DQN approaches show better capability to exploit solar energy to satisfy energy demands (better ES decisions) compared to the heuristic algorithms. Further, the CPLEX and DQN algorithms show better routing decisions compared to the heuristic algorithms, since they choose locations where energy loads at these regions and their neighboring regions are higher for consecutive time steps to incur higher net energy savings.

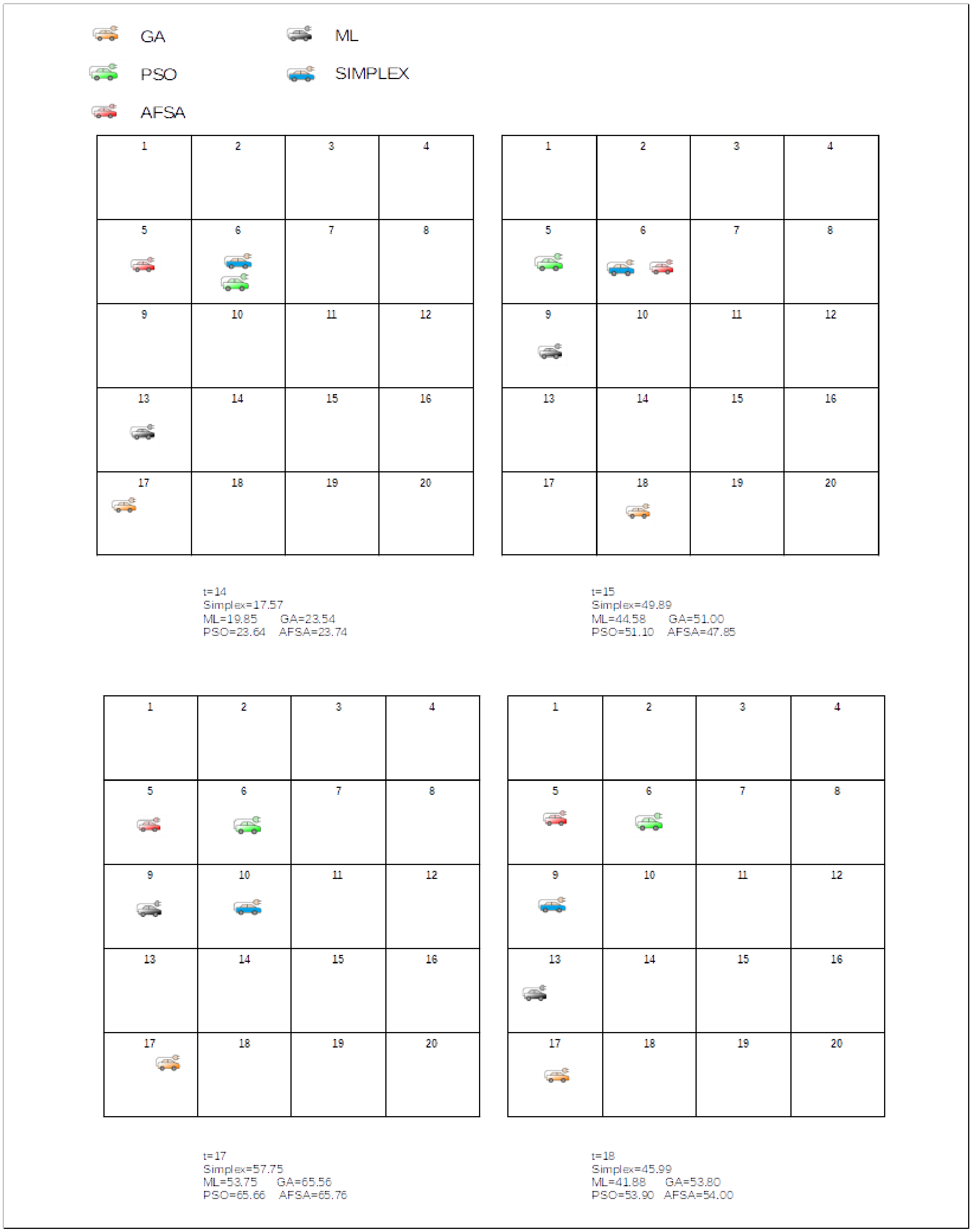

Figure 17: Routing decisions of EVs using different algorithms at critical time steps

3.3.2. Robustness Evaluation

In this section, the DQN algorithm is examined under the case of variations in load and solar irradiance data from one region to another due to different consumption patterns among consumers and different weather conditions among different locations. The DQN algorithm is compared to the CPLEX and heuristic algorithms to determine the impact of variability in input data on solution quality and simulation time.

3.3.2.1. Variability in Energy Load

In this section, the performance of different algorithms is compared under different patterns of energy load. The operational costs for five algorithms (CPLEX, DQN, and three heuristic algorithms) are shown in Table XI.

Table XI: Simulation results for energy network with different load patterns

Algorithm	Operational cost ($)			
	Mean	Std. dev.	Min	Max
CPLEX	661.41	45.77	327.63	714.80
DQN	687.53	32.44	416.85	721.61
GA	729.41	44.78	401.37	782.18
PSO	741.41	44.78	413.37	794.18
AFSA	711.65	43.17	413.42	781.32

The results from Table XV show that the DQN algorithm is more robust considering variability in energy load compared to the CPLEX and heuristic algorithms (GA, PSO, and AFSA) since it has the lowest standard deviation.

3.3.2.2. Variability in Solar Irradiance

In this section, the performance of five algorithms is compared under different patterns of solar irradiance. The operational costs for five algorithms (CPLEX, DQN, and three heuristic algorithms) are shown in Table XII.

Table XII: Simulation results for energy network with different solar irradiance patterns

Algorithm	Operational cost ($)			
	Mean	Std. dev.	Min	Max
CPLEX	415.26	11.04	400.09	451.14
DQN	441.90	9.58	429.82	465.11
GA	444.48	10.80	430.08	476.33
PSO	450.48	10.80	436.08	482.33
AFSA	508.39	34.47	435.20	556.16

The results from Table XVI show that the DQN algorithm is more robust considering variability in solar irradiance compared to the CPLEX and heuristic algorithms (GA, PSO, and AFSA) since it has the lowest standard deviation.

3.4 Conclusions

In this chapter, a network of EVs is studied which can be dispatched to supply energy to a set of consumers in different locations. A reinforcement learning model is proposed to allow multiple EVs to operate under uncertainties of energy loads and solar irradiance at different locations. Two case studies are designed to evaluate scalability and robustness of the proposed reinforcement learning model. The results show that the reinforcement learning model can better handle changes and problem scales compared to the heuristic algorithms since it can reduce the simulation time by 93.87%, 96.30%, 84.77%, and 98.83% compared to the CPLEX, GA, PSO, and AFSA, respectively. Furthermore, the reinforcement learning algorithm outperforms the three heuristic algorithms (GA, PSO, and AFSA) and can achieve a reduction in energy costs up to 22.05%, 22.57%, and 19.33% compared to GA, PSO, and AFSA, respectively. Moreover, the reinforcement learning algorithm shows better robustness to variability in load and solar irradiance compared to the CPLEX and heuristic algorithms. In the future, a multi-agent reinforcement learning algorithm will be proposed where decisions can be distributed among EVs so that they can take actions of mobility and energy transaction in a way that is independent ofother EVs.

4. DYNAMIC ENERGY SCHEDULING AND ROUTING OF A LARGE FLEET OF ELECTRIC VEHICLES USING MULTI-AGENT REINFORCEMENT LEARNING

[The majority of work presented in this chapter of the dissertation is submitted for publication in Applied Energy journal. For more information, please refer to Appendix B (Copyright Statement).]

This work is arranged as follows: Section 4.1 shows problem description and mathematical model formulated as mixed integer programming (MIP) optimization problem. Section 4.2 shows the reformulation of the MIP model as DEC-MDP. The experiments to assess the performance of DEC-MDP are illustrated in section 4.3 and section 4.4 presents the conclusion.

4.1 Mathematical Optimization Model

This section discusses the system architecture and components to model the smart grid with EVs, residential buildings, and a main power grid. The objective function and constraints of the mathematical model for the operational decision problem of EV network are described in Chapter 2.

In this study, it is assumed that a geographic area is broken down into mutually exclusive and collectively exhaustive locations (see Fig. 13) with homogenous distances (one mile) between each pair of connected locations. It is assumed that each location may include one customer or may not include any customer, hence the distance between each pair of customer ranges from one to multiple miles. Fig. 13 demonstrates a system of eight customers and four EVs. These customers have different energy demand and radiant energy and receive energy from a network of EVs. Each EV is mounted with a photovoltaic panel and power storage. Each EV takes actions at each time step (e.g., one hour), summing up to a total decision time horizon (e.g., one day).

In this study, each vehicle moves around and supplies energy to the customers at its current location only. In other words, each customer's demand can be supplied by the EV only if it exists at the customer's location. Otherwise, customer demand will be supplied from the power grid with an incurred cost. In addition, each vehicle can charge its energy storage by requesting energy from the power grid at its current location.

4.2 Proposed Multi-agent Decision Model

There are several challenges to address when reformulating the mathematical model into DEC-MDP. First, from a modeling perspective, the challenge is to map actions to collective reward. To overcome this challenge, the states are defined as tuples of joint state variables including vehicle locations, battery states of charge, solar irradiance, and energy loads. Joint action spaces are defined as tuples including four disposition actions for each vehicle (moving up, down, left, right, and no move) and energy dispatch decisions (charging, discharging, and idle). Then, the joint reward function is defined, which is a function of state and joint actions, as the cost of energy requested by all regions from the power grid.

From an algorithmic perspective, the main challenge is how to apply the MARL algorithm with a large number of agents. This challenge is addressed by a centralized training and decentralized execution framework. Since this problem is large-scale and highly dynamic due to the presence of uncertainties in energy loads and solar irradiance and so, a decomposition approach or traditional optimization algorithm cannot be used to solve this problem.

4.2.1 Electric Vehicle Agent

In this problem, multiple agents are considered, each having a "local" decision. Each agent collaborates with others by means of a number of organized transition dependencies. That is, activities taken by one agent may influence the transition functions of the other agents. Each agent interacts with other agents in the same environment and its decisions are impacted by other agents. The EV agent is assumed to be mounted with a PV panel and power storage unit that move around and discharge electricity to customers at various locations to satisfy their power demand with minimum energy costs. The DEC-MDP problem is defined by a tuple (S, A, P, R), where

- S is a finite set of the system states, which is a tuple of four variables $s_t = (p_{1t}, \ldots, p_{nt}, eB_{1t}, \ldots, eB_{nt}, Sol_t, L_t)$ corresponding to vehicle i's position p_{it}, vehicle i's state of charge eB_{it}, solar irradiance Sol_t, and power load L_t, respectively.
- $A = A_1 \times \ldots \times A_n$ is a finite set of actions. A_i indicates the set of actions taken by agent i which is described as the mobility decisions (moving up, down, left, right, and no move) and energy dispatch decisions (charging, discharging, and idle). The movement of each vehicle is restricted to the neighboring locations in an orthogonal movement in a 5×4 grid of locations. Each agent's action depends on the actions of other agents. Namely, if an agent wants to take a disposition action, then it can only move to a nearby location that is empty of other agents, or it remains at its current location.
- P is a transition function. The closed form of the transition function for this problem cannot be obtained, but simulation is used to approximate this function.
- R is a reward function. $R(s; a_1, \ldots, a_n, s')$ is the reward obtained from taking actions $a_1, \ldots, a_n$ in state s and transitioning to state s'.

4.2.2 Environment Agent

The system state s_t at time step t will transit to the next state s_{t+1} after making decisions $(a_{1t,\ldots,}\ a_{n,t})$ following the state transition rules: once the vehicle chooses which mobility decision to take, the positions of the vehicles $(p_{1t,\ldots,}\ p_{n,t})$ are changed to $(p_{1t+1,\ldots,}\ p_{n,t+1})$. Given the vehicle's position at the next state, the power demand L_t and solar irradiance Sol_t are updated to next state demand L_{t+1} and solar irradiance Sol_{t+1}. After that, based on the ED decision taken (charge, discharge, idle) at the current state, the battery state of charge variables $(eB_{1t,\ldots,}\ eB_{n,t})$ are changed to the next state battery level $(eB_{1t+1,\ldots,}\ eB_{n,t+1})$, and electricity generated from PV will be calculated based on the solar irradiance at the next state Sol_{t+1}. The state of charge at the next state $(eB_{1t+1,\ldots,}\ eB_{n,t+1})$ can be computed in Eqs. (11)-(12), given the state of charge at the current state ($eB_{1t,...,}$ $eB_{n,t}$). The position of the vehicle at the next state ($p_{1t+1,...,}$ $p_{n,t+1}$) can be determined from routing constraints in Eqs. (18)-(24). The solar irradiance and load at the next state (Sol_{t+1} and L_{t+1}) are obtained from real-world solar irradiance and power demand data at a resolution of one hour.

The reward function is defined as follows:

$$R(s_t, a_t) = \sum_{i \in I} eGp_t^i \times PGp_t \tag{29}$$

Where PGp_t is the price of electricity purchased from the power grid. The energy costs can be obtained from Eq. (29), as it indicates the amount of demand not satisfied by the main grid rather than by EVs at that location. eGp_t^i indicates the electricity purchased from the power grid, which can be computed as follows:

$$eGp_t^i = max\left(L_t^i - \sum_{n \in N} ed_t^{n,i} + \sum_{n \in N} ec_t^{n,i}, 0\right) \quad \forall i \epsilon I, \forall t \tag{30}$$

The EV power balance constraint in Eq. (4) is automatically satisfied based on the definition of system states and actions. The first power storage constraint (see Eq. (9)) is satisfied based on the definition of the battery decisions, so that at each time *t*, the battery can only take one decision (charge, discharge, or idle). The battery capacity constraint in Eq. (10) is included in the definition of the storage parameters, so that at each time step, the level in each battery is maintained within certain limits. The constraint in Eq. (11) is satisfied during model initialization. For example, at time $t = 1$, an initial amount of electricity is stored in the battery. The constraint in Eq. (12) is satisfied based on the state transition rule. The amount of electricity exchange between power grid and battery shown in Eqs. (13)-(14) is maintained within certain limits (rates of charging and discharging). These constraints are satisfied by calculating the amount of electricity received or supplied by the battery at each time step, such that they do not fall below the lower limit or exceed the upper limit of charging/discharging rates, respectively.

The energy exchange between EVs and customers is formulated as Eqs. (15-17). Eq. (15) is satisfied based on the definition of action space. At each time step, each EV can either charge electricity from the power grid or discharge electricity to customer. Constraints in Eqs. (16-17) are handled by setting upper and lower bounds for the amount of electricity exchange between EVs and customers to fixed amounts.

Finally, the movement constraint in Eq. (18) is automatically satisfied. The constraint in Eq. (19) is satisfied by adding a condition for each location to have no more than one EV at each time t. The movement of vehicles at each time step shown in Eq. (21) and Eq. (23) is considered in the definition of the action space, such that each EV will make decisions about whether to move to surrounding locations or stay at the current location. The positions of the EV at the next time step shown in Eq. (22) and Eq. (24) are satisfied based on the definition of action space, which restricts the movement to only the neighboring locations of each vehicle.

4.2.3 Centralized Training and Decentralized Execution Framework

The proposed framework uses centralized training and distributed execution (CTDE). This is a commonly used framework to model and solve MARL models, and many recent works have adopted a CTDE framework to solve distributed decision problems [113]. Each agent is trained using an actor network, which is a unique network for each agent, and all agents have access to global information and share the same critic network (which is used to evaluate the value function for agents' decisions) to foster coordinated and effective learning. Then, the model is locally executed in a distributed environment such that each agent takes actions independently from other agents, thus reducing computational time.

4.2.4. Actor Network Architecture

The actor network is used to map the neural commands to each combination (routing and scheduling) of agent's actions and select an action. The gradient of the performance with respect to the actor parameters is directly estimated by simulation, and the parameters are updated in a direction of improvement [125,126,127,128]. Given that the actor parameter vector is θ, the purpose of the critic is to compute an estimate of the Q-value, which is then used by the actor to update its policy in an approximate gradient direction [129].

In this research, the actor network is trained to approximate the control policy of the agents. It takes the state variables including vehicle location, battery level, consumer load, and solar irradiance at consumer locations $s_t = (p_t, eB_t, \ldots, eB_t, Sol_t, L_t)$ and the routing and scheduling actions (a_t) as input and maps the neural commands to each combination (routing and scheduling) of agent's actions and selects an action.

4.2.5. Critic Network Architecture

The critic network depends exclusively on the estimates of the value function and targeting at learning an approximate solution to the Bellman equation in order to generate a near-optimal policy. It uses the approach of indirectly optimizing the performance over a policy space. The critic is a temporal difference algorithm with a linearly parameterized approximation architecture for the Q-value function. The Q-value is the expected discounted reward for executing action *a* at a given state *s* and following policy π [121,129]. It is calculated as:

$$Q_\theta^r(s,a) = \sum_{i=1}^{m} r^j \phi_\theta^j(s,a) \tag{31}$$

where $r = (r^1, \ldots, r^m) \epsilon R^m$ represents the parameter vector of the critic. The features $\phi_\theta^j, j = 1, \ldots, m$, used by the critic rely on the actor parameter vector θ.

In this thesis, the critic network is trained for value function approximation to estimate the Q-values of actions ($a_{1t}, \ldots, a_{n,t}$) and estimate the reward for the decision chosen by the agents. Based on this assessment the weights of the actors are updated. The framework of the proposed MARL model is illustrated in Fig. 18.

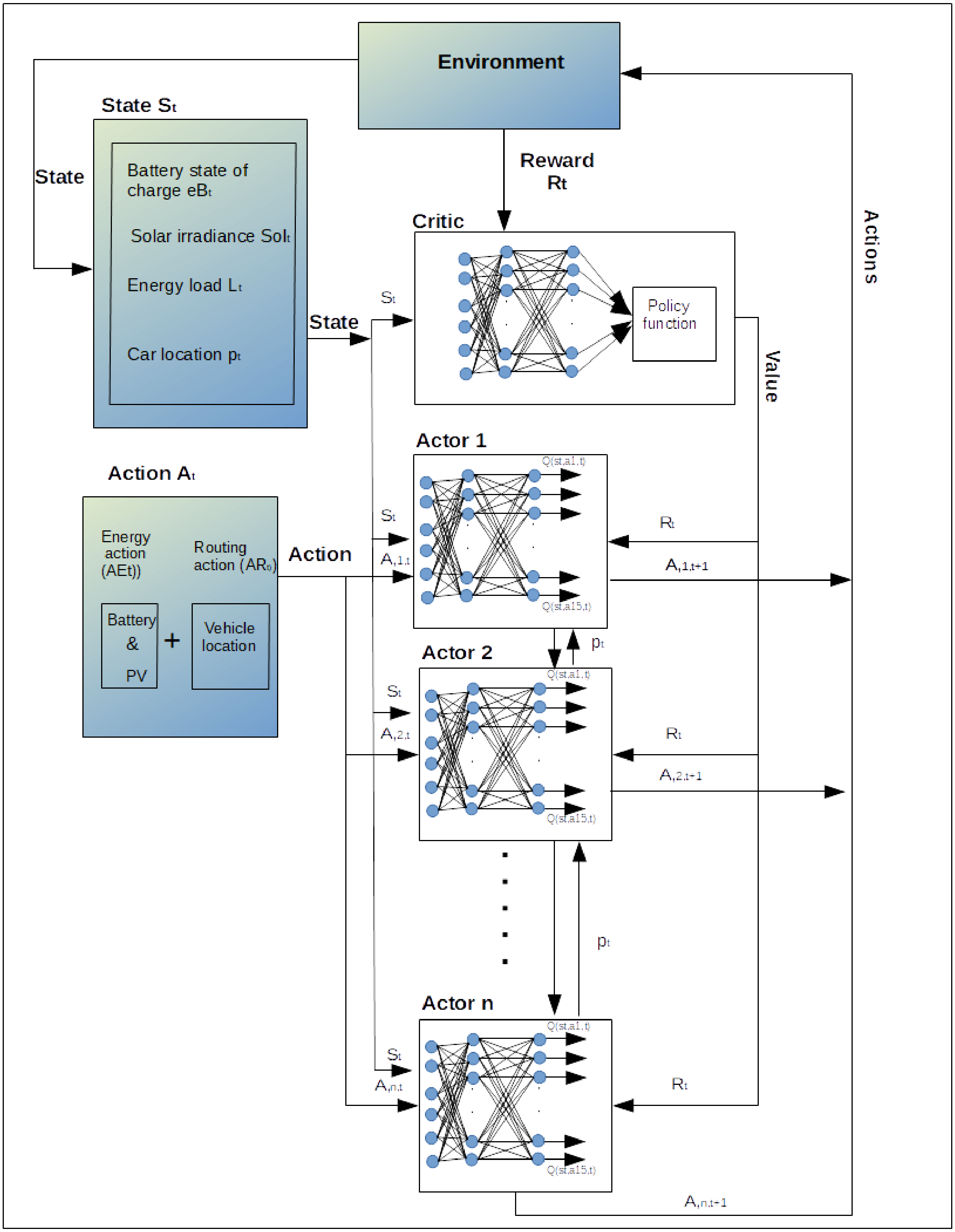

Figure 18: Proposed actor-critic network architecture

4.3 Simulation Results Analysis

To illustrate the application of the methods proposed in this chapter, an example with twenty energy users located at different locations in the Chicago area is considered. Energy demand profile data for one day for these users are collected at the weather zone of Chicago [42]. The solar irradiance profile used is taken from the city of Chicago in year 2010 [109]. Prices of electricity purchased from the power grid are taken from [110]. All the other parameters used in the experiments in this paper are taken from [42, 111]. The mathematical programming model proposed in Section 3 is solved using CPLEX with a relative gap of 0.005. The CPLEX algorithm is considered as an exact optimizer to provide an optimal solution for benchmarking.

To compare the performance of the presented MARL model, three heuristic solutions including genetic algorithm, particle swarm, and artificial fish swarm algorithms are studied. Each EV is permitted to go to one of the four neighboring locations (users) at each time step. The MARL is trained using one year energy load and solar irradiance data for 5,000 episodes. Each algorithm is evaluated on 100 problem instances. For actor and critic networks, one hidden layer with 256 hidden units and one output layer with 15 hidden units for each network are used. A rectified linear unit (ReLU) is used as activation function for each network. Adam optimizer is used to optimize the performance of the algorithm.

In DRL, the agent interacts with the environment. At each time step t, the agent gets a state s_t in a state space S and chooses an action a_t from an action space A, following a policy $\pi(a_t|s_t)$. This policy helps the agent to map states to actions to maximize a cumulative reward R_t, and select changes to the following state s_{t+1} [114]. Neural networks work as approximators, which are especially valuable in reinforcement learning when the state space or action space is large. Neural nets can be trained to map states to actions. Instead of utilizing a query table to store, index, and update every single state and itsvalues, which is not computationally feasible for high-dimensional problems, a neural network can be trained to sample from the state or action space to maximize total accumulated reward [115].

For DRL, the agent component is the same as the agent in MARL model with system state $s \in S$ and action spaces defined as follows:

- System state $s \in \mathcal{S}$ set as a tuple of four variables $s_t = (p_t, eB_t, Sol_t, L_t)$ corresponding to vehicle location p_t, battery state of charge eB_t, solar irradiance Sol_t, and energy load L_t respectively.
- The action space $a \in \mathcal{A}$ is defined as the mobility actions (moving up, down, left, right, and no move) and energy transaction actions (charging, discharging, and idle)

The DRL model used a centralized training and execution framework, where agents are trained for 5000 episodes and executed in a centralized way, and each agent has access to global information about the system state in both training and execution processes.

4.3.1. Optimality Evaluation

In this section, the performance of all five algorithms is evaluated using a case of 40 EV's and 100 customers, and the DRL algorithm is shown to provide the best performance of the five. The Mean Absolute Errors (MAE) for the MARL algorithm and three heuristic algorithms (GA, PSO, AFSA) compared to the DRL results are shown in Table XIII.

The MAE is calculated as [123]:

$$\mathrm{MAE} = \frac{\sum_{e \in E} |x_e - y_e|}{E} \tag{29}$$

Where x represents the observation from MARL, GA, PSO, or AFSA, whereas y indicates the corresponding observation from DRL, respectively, and e is the index of instances with a total of one hundred instances in the set E.

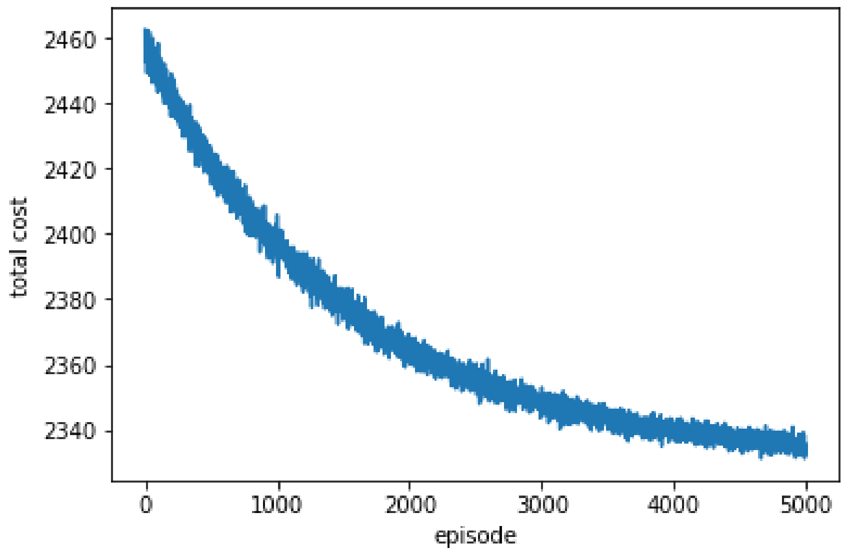

Figure 19: Convergence curve for the DRL during training

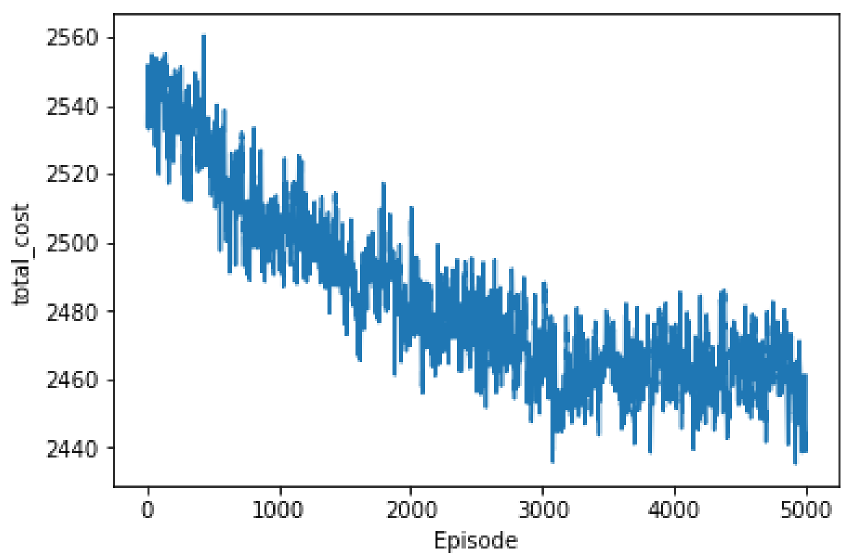

Figure 20: Convergence curve for the MARL during training

Table XIII: Mean absolute error for energy network with 40 EVs and 100 customers

Algorithm	Mean absolute error ($)			
	Mean	Std. dev.	Min	Max
MARL	74.55	36.09	0.33	145.85
GA	162.35	75.28	1.93	401.61
PSO	87.71	42.66	1.09	165.53
AFSA	89.04	43.52	3.49	167.92

It is observed from table XIV that MARL model has the lowest MSE compared to the other models, with reductions of 54.32%,14.94%, and 16.85% compared to GA, PSO, and AFSA, respectively.

Table XIV: Simulation results for energy network with 40 EVs and 100 customers

Algorithm	Operational cost ($)				Decision time (sec)
	Mean	Std. dev.	Min	Max	
DRL	2141.48	130.98	1062.60	2273.73	126
MARL	2181.18	171.76	1208.45	2387.57	16
GA	2301.40	166.19	1302.11	2503.43	164
PSO	2205.45	173.39	1228.12	2422.69	17
AFSA	2207.85	173.39	1230.53	2425.09	160

In addition, Table XVIII shows the total costs for DRL, MARL, GA, PSO, and AFSA. It is observed that the MARL algorithm can obtain better solutions compared to the heuristic models, but underperforms the DRL. The convergence curves for DRL and MARL during training are demonstrated in Figs. (19) and (20), respectively.

From the results in tables XVIII, it is noted that DRL model outperforms all the other models in terms of operational costs, since it can achieve reductions of 1.80%, 6.95%, 2.90%, and 2.99% compared to MARL, GA, PSO, and AFSA, respectively. Nevertheless, MARL can output results faster than all models with reductions in run time of 95.23%, 90.24%, 5.88%, and 90% compared to DRL, GA, PSO, and AFSA, respectively.

The computer used to conduct the training and execution processes had 8 GB RAM, with an Intel(R) Core(TM) i7-2600 CPU 3.40GHz processor. The times consumed to train the DRL model of

different configurations for 5,000 episodes were 32, 83, 103, 139, and 175 hours for configurations of one hundred customers and sixteen, twenty five, twenty nine, thirty five, and forty EV's, respectively.
The training times of the MARL model for the same configurations took 60,90, 103, 122, and 138 hours, respectively.

4.3.2. Scalability Evaluation

In this section, the performance and computational cost of the five algorithms are compared for five different configurations with varying numbers of EV's and 100 customers. CPLEX could theoretically be used to calculate an optimal solution, but this was not computationally tractable for any of the cases reported here. In all cases, the DRL algorithm produced the highest quality solution, followed closely by MARL, and MARL provided the best computational efficiency. The details of the five simulations are presented below.

4.3.2.1. Energy Network with Sixteen EVs and One Hundred Customers

Table XV: Simulation results for energy network with 16 EVs and 100 customers

Algorithm	Operational cost ($)				Decision time (sec)
	Mean	Std. dev.	Min	Max	
DRL	2317.88	76.86	1222.98	2496.26	18
MARL	2319.29	147.82	1248.10	2498.29	8
GA	2389.09	148.31	1314.84	2570.25	100
PSO	2393.77	148.29	1319.48	2574.99	10
AFSA	2396.17	148.29	1321.88	2577.40	148

The operational costs and decision time per problem instance for the five algorithms are shown in Table XV. It is observed that the DRL algorithm can generate better solutions compared to the MARL algorithm, with an average cost reduction of 0.06%. However, the MARL algorithm demonstrates better computational efficiency compared to the DRL algorithm with a 55.60% reduction in run time. MARL outperforms the heuristic methods with average cost reduction of 2.93%,3.19%, and 3.32% compared to GA, PSO, and AFSA, respectively. Moreover, MARL reduces the simulation time by 92%, 20%, and 94.59% compared to GA, PSO, and AFSA, respectively.

4.3.2.2. Energy Network with Twenty Five EVs and One Hundred Customers

The operational costs and decision time per problem instance for two algorithms (DRL and MARL) are shown in Table XVI.

Table XVI: Simulation results for energy network with 25 EVs and 100 customers

Algorithm	Operational cost ($)				Decision time (sec)
	Mean	Std. dev.	Min	Max	
DRL	2198.78	132.86	1105.08	2349.00	60
MARL	2295.39	150.72	1254.68	2476.46	10
GA	2325.06	157.88	1285.01	2537.09	120
PSO	2323.10	156.80	1285.37	2517.87	13
AFSA	2325.50	156.80	1287.77	2520.27	144

The results show that the DRL algorithm can output better solutions compared to the MARL algorithm, with average cost reduction of 4.2%. Nevertheless, the MARL algorithm performs better in terms of computational efficiency, since it can generate solutions with a reduction of 82.33% in simulation time compared to the DRL algorithm. On the other hand, MARL can achieve better solutions compared to the heuristic models with average cost reduction of 1.29%, 1.20%, and 1.30% compared to

GA, PSO, and AFSA, respectively. Moreover, MARL reduces the simulation time by 91.66%, 23.07%, and 93.05% compared to GA, PSO, and AFSA, respectively.

4.3.2.3. Energy Network with Twenty Nine EVs and One Hundred Customers

The operational costs and decision time per problem instance for two algorithms (DRL and MARL) are shown in Table XVII.

Table XVII: Simulation results for energy network with 29 EVs and 100 customers

Algorithm	Operational cost ($)				Decision time (sec)
	Mean	Std. dev.	Min	Max	
DRL	2183.25	132.92	1087.26	2336.69	74
MARL	2280.79	157.69	1255.56	2465.60	12
GA	2306.58	164.53	1278.53	2555.75	135
PSO	2291.74	160.93	1270.11	2492.49	14
AFSA	2294.14	160.93	1272.51	2494.89	150

It is observed that the DRL algorithm can generate better solutions compared to the MARL algorithm, with average cost reduction of 4.25%. However, the MARL algorithm demonstrates better computational efficiency compared to the DRL algorithm for fast decision-making with 84.02% reduction in run time. At the same time, MARL outperforms the heuristic models with average cost reduction of 1.14%, 0.48%, and 0.61% compared to GA, PSO, and AFSA, respectively. Moreover, MARL reduces the simulation time by 91.11%, 14.28%, and 92% compared to GA, PSO, and AFSA, respectively.

4.3.2.4. Energy Network with Thirty Five EVs and One Hundred Customers

The operational costs and decision time per problem instance for two algorithms (DRL and MARL) are shown in Table XVIII.

Table XVIII: Simulation results for energy network with 35 EVs and 100 customers

Algorithm	Operational cost ($)				Decision time (sec)
	Mean	Std. dev.	Min	Max	
DRL	2159.64	133.74	1052.99	2302.60	100
MARL	2227.91	162.37	1234.04	2419.91	15
GA	2301.40	166.19	1302.11	2503.43	144
PSO	2244.71	167.52	1247.10	2454.42	16
AFSA	2247.11	167.52	1249.50	2456.82	155

The results show that the DRL algorithm can output better solutions compared to MARL algorithm, with average cost reduction of 3.06%. Nevertheless, the MARL algorithm performs better in terms of computational efficiency, since it can generate solutions with a reduction of 85.03% reduction in simulation time compared to the DRL algorithm. On the other hand, MARL can achieve better solutions compared to the heuristic models with average cost reduction of 3.22%, 0.75%, and 0.89% compared to GA, PSO, and AFSA, respectively. Moreover, MARL reduces the simulation time by 89.58%, 6.25%, and 90.32% compared to GA, PSO, and AFSA, respectively.

4.3.2.5. Energy Network with Forty EVs and One Hundred Customers

The operational costs and decision time per problem instance for two algorithms (DRL and MARL) are shown in Table XIX.

Table XIX: Simulation results for energy network with 40 EVs and 100 customers

Algorithm	Operational cost ($)				Decision time (sec)
	Mean	Std. dev.	Min	Max	
DRL	2141.48	130.98	1062.60	2273.73	126
MARL	2181.18	171.76	1208.45	2387.57	16
GA	2301.40	166.19	1302.11	2503.43	164
PSO	2205.45	173.39	1228.12	2422.69	17
AFSA	2207.85	173.39	1230.53	2425.09	160

It is observed that the DRL algorithm can generate better solutions compared to the MARL algorithm, with average cost reduction of 1.82%. However, the MARL algorithm demonstrates better computational efficiency compared to the DRL algorithm for fast decision-making with 87.30% reduction in run time. At the same time, MARL outperforms the heuristic models with average cost reduction of 5.21%, 1.08%, and 1.17% compared to GA, PSO, and AFSA, respectively. Moreover, MARL reduces the simulation time by 90.24%, 5.88%, and 90% compared to GA, PSO, and AFSA, respectively.

4.3.2.6. Summary of comparison results

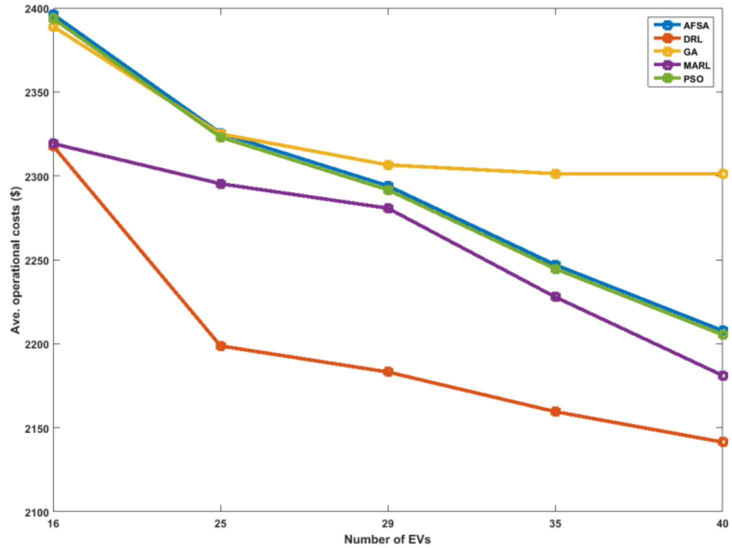

Figure 21: Average energy costs across different numbers of EVs

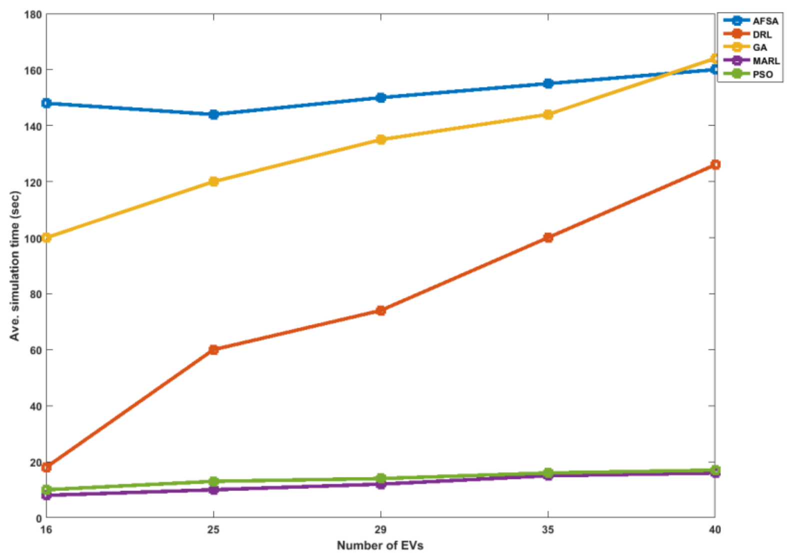

Figure 22: Average simulation time across different numbers of EVs

Figs. 21 and 22 summarize the average energy costs and average simulation times for different models under different configurations presented in the preceding sections. Fig. 21 shows that energy costs for all models decrease as the number of EV's increases. This is expected, since EV's contribute to the reduction of energy costs by supplying free energy to consumers. In addition, the figure shows that the DRL model outperforms all other models under all configurations, followed by the MARL model. However, MARL outperforms all other models when it comes to the simulation time as illustrated by Fig. 22, with a smaller increase in run time as the problem scale increases compared to the other algorithms. This demonstrates the power of the MARL model in handling large scale problems and generating solutions in less time.

4.4 Conclusions

In this chapter, a large fleet of EV's is discussed which are routed to provide electricity to a set of customers in different locations. A DEC-MDP reformulation model is introduced (MARL) to enable a large fleet of EV's to operate in large urban areas to supply energy to a large number of customers at different locations. Two example problems are designed to assess optimality and scalability of the proposed model. The results show that the MARL model outperforms the heuristic models (GA, PSO, and AFSA) and achieves results in MAE with average reductions of 54.32%,14.94%, and 16.85% compared to GA, PSO, and AFSA, respectively. Further, the MARL model can better handle changes and problem scales compared to deep Q-learning reinforcement learning

and heuristic models since it can reduce the simulation time significantly compared to the DRL reinforcement learning model, GA, PSO, and AFSA. In addition, the MARL model demonstrates a better performance compared to the heuristic models in both energy costs and run time with a cost reduction up to 5.21%, 3.19%, and 3.32% and reduction in run time up to 92%, 20%, and 94.59% compared to GA, PSO, and AFSA, respectively.

5. SUMMARY AND FUTURE WORK

This thesis discusses the optimal operation of a fleet of EV's which are routed to provide electricity to a set of customers in different locations. The research project is meant to support EV networks for the grid. This thesis is carried out in three stages: in the first stage, a MIP integrated model is introduced to integrate vehicle routing and energy scheduling to enable energy sharing at both spatial and temporal scales. The performance of the integrated model is evaluated under different scenarios using energy costs, energy requested from the power grid, and carbon emissions as performance measures. Then the performance of the integrated model is compared to disjoint models. The results showed that the proposed model outperforms all disjoint models at all different conditions, as it can save up to 38% of energy costs and reduce up to 29% and 27% of energy requested from main grid and carbon emissions compared to a baseline where consumers always request energy from power grids, respectively. The uncertainty issue in the MPN is investigated in the second stage. The MIP model proposed in the first stage is reformulated as an MDP model, and DRL is used to solve it and output solutions for the routing and scheduling. The results demonstrate that the DRL model is computationally efficient compared to the exact optimization and heuristic algorithms, since it can reduce simulation time by 93.87%, 96.30%, 84.77%, and 98.83% compared to the CPLEX, GA, PSO, and AFSA, respectively. Furthermore, DRL outperforms the three heuristic algorithms (GA, PSO, and AFSA) and can achieve a reduction in energy costs up to 22.05%, 22.57%, and 19.33% compared to GA, PSO, and AFSA, respectively. This makes it a more efficient method for handling uncertainty issues in the system. In the third stage, the scalability issue is studied, and a DEC-MDP reformulation model using multi-agent reinforcement learning (MARL) is introduced to enable a fleet of EV's to operate efficiently to supply energy to a large number of customers at different locations. A sample problem is designed and simulations are run to assess the optimality and scalability of the proposed model. The results show that the MARL model outperforms three heuristic models (GA, PSO, and AFSA) and achieves a reduction in mean absolute error compared with benchmark solutions from deep reinforcement learning (DRL) with averages of 54.32%,14.94%, and

16.85% compared to GA, PSO, and AFSA, respectively. Furthermore, the MARL model reduces the simulation time significantly compared to the DRL reinforcement learning model, GA, PSO, and AFSA. In addition, the MARL model demonstrates better performance than the heuristic models in both energy costs and run time with cost reduction up to 5.21%, 3.19%, and 3.32% and reduction in run time up to 92%, 20%, and 94.59% compared to GA, PSO, and AFSA, respectively.

Future work built on this thesis can be either from the problem perspective or the methodology perspective. From the problem perspective, we can model more realistic situations (e.g., disaster response, energy network disruptions) where routes between regions might not be available at certain periods of time. In addition, grid services will be incorporated into the model as more comprehensive multi-objective optimization in the scheduling of EV charging points, since the charging of EVs at different locations with faster charging times (e.g., 10-15 minutes) might cause instability and stress the power grid due to high fluctuations in the voltage consequently affecting the energy prices at those locations. Furthermore, the power losses will be considered in the calculation of the total costs (i.e., included in the energy prices). Moreover, consumer comfort level, reliability and maintenance of EVs, equipment degradation, design, and investment costs will be included in the model. From the methodology perspective, we can combine MARL and heuristic algorithms such that heuristics can be used to provide initial solutions for MARL to make good starting points to achieve better solution quality and efficiency than starting from scratch.

References

[1] R. H. Lasseter, P. Paigi, Microgrid: Aconceptual solution, in: Power Electronics Specialists Conference, 2004. PESC 04. 2004 IEEE 35th Annual, Vol. 6,IEEE, 2004, pp. 4285–4290.

[2] K. Zhou, S. Yang, Z. Shao, Energy internet: The business perspective,Applied Energy 178 (2016) 212 – 222. doi:https://doi:org/10:1016/j:apenergy:2016:06:052.URL http://www:sciencedirect:com/science/article/pii/S0306261916308273

[3] C. Guille, G. Gross, A conceptual framework for the vehicle-to-grid (v2g)implementation, Energy policy 37 (11) (2009) 4379–4390.

[4] H. Lund,W. Kempton, Integration of renewable energy into the transport and electricity sectors through v2g, Energy policy 36 (9) (2008) 3578–3587.

[5] W. Kempton, J. Tomic, Vehicle-to-grid power implementation: From stabilizing the grid to supporting large-scale renewable energy, Journal of power sources 144 (1) (2005) 280–294.

[6] M. Taljegard, L. Göransson, M. Odenberger, F. Johnsson, Impacts of electric vehicles on the electricity generation portfolio–a scandinavian-german case study, Applied Energy 235 (2019) 1637–1650.

[7] S. Gao, K. Chau, C. Liu, D. Wu, C. C. Chan, Integrated energy management of plug-in electric vehicles in power grid with renewables, IEEE Transactions on Vehicular Technology 63 (7) (2014) 3019–3027.

[8] M. A. Quddus, O. Shahvari, M. Marufuzzaman, J. M. Usher, R. Jaradat, A collaborative energy sharing optimization model among electric vehicle charging stations, commercial buildings, and power grid, Applied energy 229 (2018) 841–857.

[9] M. F. Qazi, F. Qazi, Green technology for disaster relief and remote areas, age 24 (2014) 1.

[10] W.Kempton, V. Udo, K. Huber, K.Komara, S. Letendre, S. Baker, D. Brunner, N. Pearre, A test of vehicle-to-grid (v2g) for energy storage and frequency regulation in the pjm system, Results from an Industry-University Research Partnership 32.

[11] X. Jin, J. Wu, Y. Mu, M. Wang, X. Xu, H. Jia, Hierarchical microgrid energy management in an office building, Applied Energy 208 (2017) 480–494.

[12] G. Cardoso, M. Stadler, M. C. Bozchalui, R. Sharma, C. Marnay, A. Barbosa- Póvoa, P. Ferrão, Optimal investment and scheduling of distributed energy resources with uncertainty in electric vehicle driving schedules, Energy 64 (2014) 17–30.

[13] Y. Yang, S. Zhang, Y. Xiao, An milp (mixed integer linear programming) model for optimal design of district-scale distributed energy resource systems, Energy 90 (2015) 1901–1915.

[14] L. Li, H. Mu, N. Li, M. Li, Economic and environmental optimization for distributed energy resource systems coupled with district energy networks, Energy 109 (2016) 947–960.

[15] G. Gutiérrez-Alcaraz, E. Galván, N. González-Cabrera, M. Javadi, Renewable energy resources short-term scheduling and dynamic network reconfiguration for sustainable energy consumption, Renewable and Sustainable Energy Reviews 52 (2015) 256–264.

[16] A. Tsikalakis, N. Hatziargyriou, Environmental benefits of distributed generation with and without emissions trading, Energy Policy 35 (6) (2007) 3395–3409.

[17] J. J. Ming Kwok, N. Yu, I. A. Karimi, D.-Y. Lee, Microgrid scheduling for reliable, cost-effective, and environmentally friendly energy management, Industrial & Engineering Chemistry Research 52 (1) (2012) 142–151.

[18] A. Zakariazadeh, S. Jadid, P. Siano, Stochastic multi-objective operational planning of smart distribution systems considering demand response programs, Electric Power Systems Research 111 (2014) 156–168.

[19] C. A. Hill, M. C. Such, D. Chen, J. Gonzalez, W. M. Grady, Battery energy storage for enabling integration of distributed solar power generation, IEEE Transactions on smart grid 3 (2) (2012) 850–857.

[20] C. Chen, S. Duan, T. Cai, B. Liu, G. Hu, Smart energy management system for optimal microgrid economic operation, IET renewable power generation 5 (3) (2011) 258–267.

[21] W. S. Ho, S. Macchietto, J. S. Lim, H. Hashim, Z. A. Muis, W. H. Liu, Optimal scheduling of energy storage for renewable energy distributed energy generation system, Renewable and Sustainable Energy Reviews 58 (2016) 1100–1107.

[22] M. Moradi, A. Khandani, Evaluation economic and reliability issues for an autonomous independent network of distributed energy resources, International Journal of Electrical Power & Energy Systems 56 (2014) 75–82.

[23] L. Zhang, N. Gari, L. V. Hmurcik, Energy management in a microgrid with distributed energy resources, Energy Conversion and Management 78 (2014) 297–305.

[24] C. L. Prete, B. F. Hobbs, C. S. Norman, S. Cano-Andrade, A. Fuentes, M. R. von Spakovsky, L. Mili, Sustainability and reliability assessment of microgrids in a regional electricity market, Energy 41 (1) (2012) 192–202.

[25] A. Kanase-Patil, R. Saini, M. Sharma, Integrated renewable energy systems for off grid rural electrification of remote area, Renewable Energy 35 (6) (2010) 1342–1349.

[26] C. Wang, C. Lv, P. Li, G. Song, S. Li, X. Xu, J. Wu, Modeling and optimal operation of community integrated energy systems: A case study from china, Applied Energy 230 (2018) 1242–1254.

[27] C. B. Robledo, V. Oldenbroek, F.Abbruzzese, A. J. vanWijk, Integrating a hydrogen fuel cell electric vehicle with vehicle-to-grid technology, photovoltaic power and a residential building, Applied Energy 215 (2018) 615–629.

[28] Y. Chen, M. Hu, Balancing collective and individual interests in transactive energy management of interconnected micro-grid clusters, Energy 109 (2016) 1075–1085.

[29] J. Lagorse, D. Paire, A. Miraoui, Amulti-agent system for energy management of distributed power sources, Renewable energy 35 (1) (2010) 174–182.

[30] C. Calvillo, A. Sánchez-Miralles, J. Villar, F. Martín, Optimal planning and operation of aggregated distributed energy resources with market participation, Applied Energy 182 (2016) 340–357.

[31] M. A. A. Pedrasa, T. D. Spooner, I. F. MacGill, Coordinated scheduling of residential distributed energy resources to optimize smart home energy services, IEEE Transactions on Smart Grid 1 (2) (2010) 134–143.

[32] M. A. A. Pedrasa, T. D. Spooner, I. F. MacGill, A novel energy service model and optimal scheduling algorithm for residential distributed energy resources, Electric Power Systems Research 81 (12) (2011) 2155–2163.

[33] C. Corchero, M. Cruz-Zambrano, F.-J. Heredia, et al., Optimal energy management for a residential microgrid including a vehicle-to-grid system, IEEE transactions on smart grid 5 (4) (2014) 2163–2172.

[34] K. Mets, T. Verschueren, F. De Turck, C. Develder, Exploiting v2g to optimize residential energy consumption with electrical vehicle (dis) charging, in: Smart Grid Modeling and Simulation (SGMS), 2011 IEEE First International Workshop on, IEEE, 2011, pp. 7–12.

[35] J. Tan, L.Wang, Integration of plug-in hybrid electric vehicles into residential distribution grid based on two-layer intelligent optimization, IEEE Transactions on Smart Grid 5 (4) (2014) 1774–1784.

[36] C. Calvillo, A. Sánchez-Miralles, J. Villar, F. Martín, Impact of ev penetration in the interconnected urban environment of a smart city, Energy 141 (2017) 2218–2233.

[37] A. Rabiee, M. Sadeghi, J. Aghaeic, A. Heidari, Optimal operation of microgrids through simultaneous scheduling of electrical vehicles and responsive loads considering wind and pv units uncertainties, Renewable and Sustainable Energy Reviews 57 (2016) 721–739.

[38] C. Battistelli, L. Baringo, A. Conejo, Optimal energy management of small electric energy systems including v2g facilities and renewable energy sources, Electric Power Systems Research 92 (2012) 50–59.

[39] G. K. Venayagamoorthy, P. Mitra, K. Corzine, C. Huston, Real-time modeling of distributed plug-in vehicles for v2g transactions, in: Energy Conversion Congress and Exposition, 2009. ECCE 2009. IEEE, IEEE, 2009, pp. 3937– 3941.

[40] K. M. Tan, V. K. Ramachandaramurthy, J. Y. Yong, Integration of electric vehicles in smart grid: A review on vehicle to grid technologies and optimization techniques, Renewable and Sustainable Energy Reviews 53 (2016) 720–732.

[41] B. K. Sovacool, R. F. Hirsh, Beyond batteries: An examination of the benefits and barriers to plug-in hybrid electric vehicles (phevs) and a vehicle-to-grid (v2g) transition, Energy Policy 37 (3) (2009) 1095–1103.

[42] M. Yilmaz, P. T. Krein, Review of the impact of vehicle-to-grid technologies on distribution systems and utility interfaces, IEEE Transactions on power electronics 28 (12) (2013) 5673–5689.

[43] J. Aghaei, A. E. Nezhad, A. Rabiee, E. Rahimi, Contribution of plug-in hybrid electric vehicles in power system uncertainty management, Renewable and Sustainable Energy Reviews 59 (2016) 450–458.

[44] National Renewable Energy Laboratory.. National solar radiation data base, http://rredc:nrel:gov/solar/old_data/nsrdb/, accessed: 2018-09-06.

[45] Arizona.. Salt river project utility company, https://www:srpnet:com/prices/ home/tou:aspx, accessed: 2018-09-06.

[46] R. Dai, M. Hu, D. Yang, Y. Chen, A collaborative operation decision model for distributed building clusters, Energy 84 (2015) 759–773.

[47] U.S. Department of Energy. The official U.S. government source for fuel economy information, https://www:fueleconomy:gov/ feg/PowerSearch:do?action=noform&path=1&year1=1984&year2= 2019&vtype=Electric&pageno=1&sortBy=Comb&tabView=0&rowLimit= 10, accessed: 2018-09-06.

[48] G¨oran Andersson, Peter Donalek, Richard Farmer, Nikos Hatziargyriou, Innocent Kamwa, Prabhashankar Kundur, Nelson Martins, John Paserba,Pouyan Pourbeik, Juan Sanchez-Gasca, et al. Causes of the 2003 major grid blackouts in north America and Europe, and recommended means to improve system dynamic performance. IEEE transactions on Power Systems, 20(4):1922–1928, 2005.

[49] Chaamala Klinger and Virginia Murray Owen Landeg. Power outages, extreme events and health: a systematic review of the literature from2011-2012.PLoS currents, 6, 2014.

[50] Yezhou Wang, Chen Chen, Jianhui Wang, and Ross Baldick. Research on resilience of power systems under natural disasters—a review. IEEE Transactions on Power Systems, 31(2):1604–1613, 2015.

[51] Gulay Barbaroso˘glu and Yasemin Arda. A two-stage stochastic programming framework for transportation planning in disaster response. Journalof the operational research society, 55(1):43–53, 2004.

[52] Ali Bozorgi-Amiri, Mohamad Saeed Jabalameli, and SMJ Mirzapour Al-e Hashem. A multi-objective robust stochastic programming model for disaster relief logistics under uncertainty. OR spectrum, 35(4):905–933,2013.

[53] Nian Liu, Minyang Cheng, Xinghuo Yu, Jiangxia Zhong, and Jinyong Lei.Energy-sharing provider for pv prosumer clusters: A hybrid approachusing stochastic programming and stackelberg game. IEEE Transactions on Industrial Electronics, 65(8):6740–6750, 2018.

[54] João Soares, Mohammad Ali Fotouhi Ghazvini, Nuno Borges, and ZitaVale. Dynamic electricity pricing for electric vehicles using stochastic programming. Energy, 122:111–127, 2017.

[55] YZ Sun, Jun Wu, GJ Li, and Jian He. Dynamic economic dispatch considering wind power penetration based on wind speed forecasting and stochastic programming. Proceedings of the CSEE, 29(4):41–47, 2009.

[56] David Abramson and J Abela. A parallel genetic algorithm for solving the school timetabling problem. Citeseer, 1991.

[57] Suvrajeet Sen. Stochastic programming: computational issues and challenges. Encyclopedia of OR/MS, pages 1–11, 2001.

[58] Yanchong Zheng, Songyan Niu, Yitong Shang, Ziyun Shao, and LinniJian. Integrating plug-in electric vehicles into power grids: A comprehensive review on power interaction mode, scheduling methodology and mathematical foundation. Renewable and Sustainable Energy Reviews,112:424–439, 2019.

[59] Sayed Saeed Hosseini, Ali Badri, and Masood Parvania. A survey on mo-bile energy storage systems (mess): Applications, challenges and solutions. Renewable and Sustainable Energy Reviews, 40:161–170, 2014.

[60] Taha Selim Ustun, Umit Cali, and Mithat C Kisacikoglu. Energizing microgrids with electric vehicles during emergencies—natural disasters, sabotage and warfare. In2015 IEEE International Telecommunications Energy Conference (INTELEC), pages 1–6. IEEE, 2015.

[61] Jeremy Jie Ming Kwok, Nan Yu, Iftekhar A Karimi, and Dong-YupLee. Microgrid scheduling for reliable, cost-effective, and environmentally friendly energy management. Industrial & Engineering Chemistry Research, 52(1):142–151, 2013.

[62] Miguel Carrión, Andy B Philpott, Antonio J Conejo, and Jos M Arroyo.A stochastic programming approach to electric energy procurement for large consumers. IEEE Transactions on Power Systems, 22(2):744–754,2007.

[63] Dian Palupi Rini, Siti Mariyam Shamsuddin, and Siti Sophiyati Yuhaniz.Particle swarm optimization: technique, system and challenges. International journal of computer applications, 14(1):19–26, 2011.

[64] Penghao Sun, Yuxiang Hu, Julong Lan, Le Tian, and Min Chen. Tide:Time-relevant deep reinforcement learning for routing optimization. Future Generation Computer Systems, 99:401–409, 2019.

[65] Jose L Crespo-Vazquez, C Carrillo, E Diaz-Dorado, Jose A Martinez-Lorenzo, and Md Noor-E-Alam. A machine learning based stochastic optimization framework for a wind and storage power plant participating in energy pool market. Applied energy, 232:341–357, 2018.

[66] Feng Zhu and Satish V Ukkusuri. Accounting for dynamic speed limit control in a stochastic traffic environment: A reinforcement learning approach. Transportation research part C: emerging technologies, 41:30–47,2014.

[67] Mohammadreza Nazari, Afshin Oroojlooy, Lawrence Snyder, and Martin Takáac. Reinforcement learning for solving the vehicle routing problem. In Advances in Neural Information Processing Systems, pages 9839–9849,2018.

[68] Joseé Manuel Vera and Andres G Abad. Deep reinforcement learning for routing a heterogeneous fleet of vehicles. In2019 IEEE Latin American Conference on Computational Intelligence (LA-CCI), pages 1–6. IEEE,2019.

[69] Scott Proper, Prasad Tadepalli, Hong Tang, and Rasaratnam Logendran.A reinforcement learning approach for product delivery by multiple vehicles. In IIE Annual Conference. Proceedings, page 1. Institute of Industrial and Systems Engineers (IISE), 2003.

[70] Miyoung Han, Pierre Senellart, St´ephane Bressan, and Huayu Wu. Routing an autonomous taxi with reinforcement learning. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, pages 2421–2424, 2016.

[71] JQ James, Wen Yu, and Jiatao Gu. Online vehicle routing with neural combinatorial optimization and deep reinforcement learning. IEEE Trans-actions on Intelligent Transportation Systems, 20(10):3806–3817, 2019.

[72] Jiajun Duan, Haifeng Li, Xiaohu Zhang, Ruisheng Diao, Bei Zhang,Di Shi, Xiao Lu, Zhiwei Wang, and Siqi Wang. A deep reinforcement learning based approach for optimal active power dispatch. In2019 IEEE Sustainable Power and Energy Conference (iSPEC), pages 263–267. IEEE,2019.

[73] Lei Lei, Yue Tan, Glenn Dahlenburg, Wei Xiang, and Kan Zheng. Dynamic energy dispatch in isolated microgrids based on deep reinforcement learning. arXiv preprint arXiv:2002.02581, 2020.

[74] Lingxiao Gan, Tao Yu, Jing Li, and Jie Tang. Smart scheduling strategy for islanded microgrid based on reinforcement learning algorithm. Int. J.Smart Grid and Clean Energy, 1(1):122–128, 2012.

[75] Lynn A Bollinger and Ralph Evins. Multi-agent reinforcement learning for optimizing technology deployment in distributed multi-energy systems.In23rd International Workshop of the European Group for Intelligent Computing in Engineering. Krakow, Poland, 2016.

[76] Sunyong Kim and Hyuk Lim. Reinforcement learning based energy management algorithm for smart energy buildings. Energies, 11(8):2010, 2018.

[77] Ying Ji, Jianhui Wang, Jiacan Xu, Xiaoke Fang, and Huaguang Zhang.Real-time energy management of a microgrid using deep reinforcement learning. Energies, 12(12):2291, 2019.

[78] Mohammad Reza Salehizadeh and Salman Soltaniyan. Application of fuzzy q-learning for electricity market modeling by considering renewable power penetration. Renewable and Sustainable Energy Reviews, 56:1172–1181, 2016.

[79] T Yu, YM Wang, WJ Ye, B Zhou, and KW Chan. Stochastic optimal generation command dispatch based on improved hierarchical reinforcement learning approach. IET generation, transmission & distribution,5(8):789–797, 2011.

[80] Philip Odonkor and Kemper Lewis. Control of shared energy storage assets within building clusters using reinforcement learning. In International Design Engineering Technical Conferences and Computers and In-formation in Engineering Conference, volume 51753, page V02AT03A028.American Society of Mechanical Engineers, 2018.

[81] EA Jasmin, TP Imthias Ahamed, and VP Jagathiraj. A reinforcement learning algorithm to economic dispatch considering transmission losses. In TENCON 2008-2008 IEEE Region 10 Conference, pages 1–6. IEEE,2008.

[82] Zafar I Choudhry. Electric power delivery to big cities. In2008 IEEE/PES Transmission and Distribution Conference and Exposition, pages 1–6.IEEE, 2008.

[83] Qiannan Zhu, Xianglong Luo, Bingjian Zhang, and Ying Chen. Mathematical modelling and optimization of a large-scale combined cooling, heat, and power system that incorporates unit changeover and time-of-use electricity price. Energy Conversion and Management, 133:385–398,2017.

[84] Renato C Zambon, Mario TL Barros, Joao Eduardo G Lopes, Paulo SFBarbosa, Alberto L Francato, and William W-G Yeh. Optimization of large-scale hydrothermal system operation. Journal of Water Resources Planning and Management, 138(2):135–143, 2012.

[85] Zhong-kai Feng, Wen-jing Niu, and Chun-tian Cheng. A quadratic programming approach for fixed head hydropower system operation optimization considering power shortage aspect. Journal of Water Resources Planning and Management, 143(10):06017005, 2017.

[86] Zhong-kai Feng, Wen-jing Niu, Chun-tian Cheng, and Xin-yu Wu. Optimization of large-scale hydropower system peak operation with hybrid dynamic programming and domain knowledge. Journal of Cleaner Pro-duction, 171:390–402, 2018.

[87] Li Liu, Wenxin Liu, David A Cartes, and Il-Yop Chung. Slow coherency and angle modulated particle swarm optimization based islanding of large-scale power systems. Advanced Engineering Informatics, 23(1):45–56,2009.

[88] Lili Hao, Jing Ji, Dongliang Xie, Haohao Wang, Wei Li, and Philip Asaah. Scenario-based unit commitment optimization for power system with large-scale wind power participating in primary frequency regulation. Journal of Modern Power Systems and Clean Energy, 2020.

[89] Ying Zhou, Lingjing Kong, Yiqiao Cai, Ziyan Wu, Shaopeng Liu, Jiaming Hong, and Keke Wu. A decomposition-based local search for large-scalemany-objective vehicle routing problems with simultaneous delivery and pickup and time windows. IEEE Systems Journal, 2020.

[90] Russell Bent and Pascal Van Hentenryck. Spatial, temporal, and hybrid decompositions for large-scale vehicle routing with time windows. In International Conference on Principles and Practice of Constraint Programming, pages 99–113. Springer, 2010.

[91] Robin Vujanic, Peyman Mohajerin Esfahani, Paul J Goulart, S´ebastienMari´ethoz, and Manfred Morari. A decomposition method for large scale milps, with performance guarantees and a power system application. Automatica, 67:144–156, 2016.

[92] Nedal Deeb and SM Shahidehpour. Linear reactive power optimization ina large power network using the decomposition approach. IEEE Transactions on power systems, 5(2):428–438, 1990.

[93] Richard Lusby, Laurent Flindt Muller, and Bjørn Petersen. A solution approach based on benders decomposition for the preventive maintenance scheduling problem of a stochastic large-scale energy system. Journal of Scheduling, 16(6):605–628, 2013.

[94] Nestor F Michelena and Panos Y Papalambros. A hypergraph framework for optimal model-based decomposition of design problems. Computational optimization and applications, 8(2):173–196, 1997.

[95] Ralf Niemann and Peter Marwedel. An algorithm for hardware/software partitioning using mixed integer linear programming. Design Automation for Embedded Systems, 2(2):165–193, 1997.

[96] David Q Mayne, Eric C Kerrigan, and Paola Falugi. Robust model predictive control: advantages and disadvantages of tube-based methods. IFAC Proceedings Volumes, 44(1):191–196, 2011.

[97] Yang Yang, Li Juntao, and Peng Lingling. Multi-robot path planning based on a deep reinforcement learning dqn algorithm. CAAI Transactions on Intelligence Technology, 5(3):177–183, 2020.

[98] Kaixiang Lin, Renyu Zhao, Zhe Xu, and Jiayu Zhou. Efficient large-scale fleet management via multi-agent deep reinforcement learning. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pages 1774–1783, 2018.

[99] Bolong Cheng and Warren B Powell. Co-optimizing battery storage for the frequency regulation and energy arbitrage using multi-scale dynamic programming. IEEE Transactions on Smart Grid, 9(3):1997–2005, 2016.

[100] Aur´elie Beynier, Fran¸cois Charpillet, Daniel Szer, and AbdelIllah Mouad-dib. Dec-mdp/pomdp.Markov Decision Processes in Artificial Intelligence, pages 277–318, 2013.

[101] Huang Huang and Vincent KN Lau. Decentralized delay optimal control for interference networks with limited renewable energy storage. IEEE Transactions on Signal Processing, 60(5):2552–2561, 2012.

[102] Alessandro Biason, Subhrakanti Dey, and Michele Zorzi. Decentralized transmission policies for energy harvesting devices. In2017 IEEE Wire-less Communications and Networking Conference Workshops (WCNCW),pages 1–6. IEEE, 2017.

[103] Alessandro Biason, Subhrakanti Dey, and Michele Zorzi. A decentralized optimization framework for energy harvesting devices. IEEE Transactions on Mobile Computing, 17(11):2483–2496, 2018.

[104] Qinqin Tang, Renchao Xie, F Richard Yu, Tao Huang, and Yunjie Liu. Decentralized computation offloading in iot fog computing system withenergy harvesting: A dec-pomdp approach. IEEE Internet of Things Journal, 2020.

[105] Ke Zhang, Meng Li, Zhengchao Zhang, Xi Lin, and Fang He. Multi-vehicle routing problems with soft time windows: A multi-agent reinforcement learning approach. arXiv preprint arXiv:2002.05513, 2020.

[106] Quinlan Sykora, Mengye Ren, and Raquel Urtasun. Multi-agent routing value iteration network. arXiv preprint arXiv:2007.05096, 2020.

[107] Kian Lun Soon, Joanne Mun-Yee Lim, Rajendran Parthiban, and Mun Chon Ho. Proactive eco-friendly pheromone-based green vehicle routing for multi-agent systems. Expert Systems With Applications, 121:324–337, 2019.

[108] Yaodong Yang, Jianye Hao, Yan Zheng, and Chao Yu. Large-scale home energy management using entropy-based collective multiagent deep reinforcement learning framework. In IJCAI, pages 630–636, 2019.

[109] Feibo Jiang, Li Dong, Kezhi Wang, Kun Yang, and Cunhua Pan. Dis-tributed resource scheduling for large-scale mec systems: A multi-agent ensemble deep reinforcement learning with imitation acceleration. arXiv preprint arXiv:2005.12364, 2020.

[110] Fu-Dong Li, Min Wu, Yong He, and Xin Chen. Optimal control in microgrid using multi-agent reinforcement learning. ISA transactions,51(6):743–751, 2012.

[111] Xiaohan Fang, Jinkuan Wang, Guanru Song, Yinghua Han, Qiang Zhao,and Zhiao Cao. Multi-agent reinforcement learning approach for residential microgrid energy scheduling. Energies, 13(1):123, 2020.

[112] Yu Sui and Shiming Song. A multi-agent reinforcement learning frame-work for lithium-ion battery scheduling problems. Energies, 13(8):1982,2020.

[113] Kevin Corder, Manuel M Vindiola, and Keith Decker. Decentral-ized multi-agent actor-critic with generative inference. arXiv preprintarXiv:1910.03058, 2019.

[114] Yuxi Li. Deep reinforcement learning: An overview. arXiv preprintarXiv:1701.07274, 2017.

[115] A Beginner's Guide to Deep Reinforcement Learning. Neural net-works and deep reinforcement learning. https://wiki.pathmind.com/deep-reinforcement-learningneural. Accessed: 2021-01-02

[116] World's population statistics. (2017, June 21). Retrieved January 18, 2020, from https://www.un.org/development/desa/en/news/population/world-population-prospects-2017.html

[117] Global energy consumption. (2017, September 14). Retrieved January 18, 2020, from https://www.eia.gov/todayinenergy/detail.php?id=32912

[118] Global energy consumption. (2019, September 24). Retrieved January 18, 2020, from https://www.eia.gov/todayinenergy/detail.php?id=41433

[119] Smart grid. (2017, February 26). Retrieved December 1, 2019, from https://www.govtech.com/blogs/lohrmann-on-cybersecurity/how-secure-is-our-smart-grid.html.

[120] Solae Rover. Hybrid power systems, https://www.solarover.com/product_pv_hybrids.html, accessed: 2020-01-30.

[121] Christopher JCH Watkins and Peter Dayan. Q-learning. Machine learning, 8(3-4):279–292, 1992.

[122] Lucian Bus ̧oniu, Robert Babuˇska, and Bart De Schutter. Multi-agent reinforcement learning: An overview. Innovations in multi-agent systems andapplications-1, pages 183–221, 2010.

[123] Kotz, S.; et al., eds. (2006), Encyclopedia of Statistical Sciences, Wiley.

[124] Deep Q-Networks. DQN agent, https://www.tensorflow.org/agents/tutorials/0_intro_rl, accessed: 2020-03-10.

[125] P. W. Glynn, Stochastic approximation for Monte Carlo optimization, in Proceedings of the 1986 Winter Simulation Conference, Washington, DC, 1986, pp. 285–289.

[126] T. Jaakkola, S. P. Singh, and M. I. Jordan, Reinforcement learning algorithms for partially observable Markov decision problems, in Advances in Neural Information Processing Systems 7, G. Tesauro and D. Touretzky, eds., Morgan Kaufman, San Francisco, CA, 1995, pp. 345–352.

[127] P. Marbach and J. N. Tsitsiklis, Simulation-based optimization of Markov reward processes, IEEE Trans. Automat. Control, 46 (2001), pp. 191–209.

[128] R. Williams, Simple statistical gradient following algorithms for connectionist reinforcement learning, Machine Learning, 8 (1992), pp. 229–256.

[129] Konda, V. R., & Tsitsiklis, J. N. (2000). Actor-critic algorithms. In *Advances in neural information processing systems* (pp. 1008-1014).

VITA

NAME:	Mohammed Alqahtani (pronouns he/him/his)
EDUCATION:	B.S., Industrial Engineering, King Khalid University, Saudi Arabia, 2010 M.S., Industrial Engineering (Optimization), University of Arizona, USA, 2015 Ph.D., Industrial Engineering and Operations Research, University of Illinois at Chicago, Chicago, Illinois, 2021
HONORS:	Graduate Scholarship, King Khalid University, Saudi Arabia, 2011-2013 (Masters) 2017-2021 (Doctorate)
JOURNAL PUBLICATIONS:	Alqahtani, M., & Hu, M. (2020). Integrated energy scheduling and routing for a network of mobile prosumers. Energy, 200, 117451. Alqahtani, M., & Hu, M. (2020). Dynamic Energy Scheduling and Routing of Multiple Electric Vehicles using Deep Reinforcement Learning. (submitted to Renewable and Sustainable Energy Reviews journal) Alqahtani, M., & Hu, M. (2021). Dynamic Energy Scheduling and Routing of a Large Fleet of Electric Vehicles using Multi-Agent Reinforcement Learning . (submitted to Applied Energy journal)
CONFERENCE PRESENTATIONS:	The Annual Student Water Conference, Stillwater OK (March 26-28, 2015) .
CONFERENCE POSTERS:	The Student Energy Conference , Tucson AZ (April 8-10, 2015). Arizona Annual Water Conference , Pheonix AZ (May 1-2, 2015). First Saudi Environment Conference, Abha Aseer, KSA (March 11th, 2016) Engineer day association, Abha Aseer, KSA (October 22nd , 2016)

SERVICE AND EXPERIENCE:

-Member of Scholarship Committee in Engineering College at King Khalid University
-Member of Graduation Projects Committee in Industrial Engineering Department
-Member of Engineering College Training and Alumni Unit council
-Coordinator of Training and Alumni Unit in Industrial Engineering Department
-Member of Qiyas Examination Committee at King Khalid University
-Member of Saudi Society for Systems and Industrial Engineering (SSSIE)
-Cost Estimator Saudi Basic Industries Company (SABIC), Aljubail, Eastern rigion, Saudi Arabia.

TEACHING EXPERIENCE:

-Lecturer May 2015 – Present
King Khalid University, Abha, Asir province, KSA.
-Teaching Assistant February 2015 – May 2015
The University of Arizona, Tucson, AZ, USA.
-Teaching Assistant January 2011 – May 2012
King Khalid University, Abha, Asier province, KSA.

Appendix A

TABLE XX. Energy load (kWh) profile for twelve consumers (similar behavior)

Consumer	CI	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
1	0.248	1.105	1.433	0.479	0.83	1.753	0.745	0.871	2.82I	0.385	4.588	2.648
2	0.419	1.315	1.836	0.605	0.89	2.022	0.921	1.187	1.3	0.872	5.339	1.797
3	0.294	0.768	1.806	0.849	0.893	2.012	0.818	2.209	1.554	0.869	2.555	1.278
4	0.303	0.776	1.829	0.808	0.805	2.446	0.79	1.827	2.292	0.863	2.004	1.456
5	0.291	0.812	1.799	0.651	0396	1.99	0.938	1.528	0.922	0.882	1.592	0.835
6	0.292	0.783	1.773	0.71	0.784	2.399	1.017	1.793	0.837	0.876	1.348	0.512
7	0.56	0.789	1.845	0.75	0.809	1.886	1.956	1.03	1.062	0.95	1.207	0.414
8	1.558	1.772	1.835	0.699	0.773	0.643	0.691	0.525	0.9	0.931	1.138	0.399
9	1.612	1.092	1.855	0.766	0.781	0.958	0.827	0.821	0.871	1.118	1.263	0.448
10	1.602	1.036	1.899	0371	0.801	0.44	0.732	0.591	1.474	0.688	1.23	0.417
11	1.519	0.629	1.902	0.692	0.858	0.44	0.724	0.539	1.726	1.386	1.158	0.45
12	1.535	0.701	1.947	0.97	0.876	0.455	0.69	0.839	1.049	1.34	1.445	0.457
13	1.566	0.616	2.008	0354	1.015	0.562	0313	0.598	0.916	0.951	1.194	0.871
14	1.541	0.626	1.973	0.867	1.285	2.044	2.249	0.669	0.901	1.089	1.23	1.5
15	1.629	1.914	1.962	1.078	1.66	3.626	4.902	0.66	4346	0.92	1.101	1.503
16	1.729	1.404	1.963	1.195	2.964	3.803	3.626	1.693	3.031	0.937	1.963	4.635
17	1.564	2.518	1.962	1.738	3.049	2.022	1.406	1.111	1.676	1.025	5.155	4.706
18	1.828	2.804	1.966	0.467	2.883	1.951	0.656	0.784	1.611	1.224	4.666	3.522
19	1.675	0.69	1.96	0.586	1.285	1.971	0394	1.024	1.547	0.949	3.82	5.218
20	1.809	0.665	1.933	0.802	0.97	1.882	1.429	1.192	1.75	0.834	3.758	1.429
21	1.688	0.704	1.932	0.68	2.175	2.44	2.181	1.068	0.766	0.911	3.311	1.336
22	6.853	0.703	1.944	0.659	2.886	1.893	4.63	1.307	1.117	0.636	2.654	1.458
23	9.006	1.097	1.942	0.532	1.959	2.899	4.158	2.01	1.877	1.303	2.55	2.417
24	9.331	0.683	1.942	0.821	5.364	5.877	3.102	2.517	3.334	1.107	3.624	2.697

TABLE XXI. Energy load (kWh) profile for twelve consumers (different behavior)

Consumer	CI	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
1	1.88	5.09	0.435	2.016	1.435	1.002	1.548	1.826	2.232	3.929	0.656	1.824
2	1.965	4.597	2.015	0.832	5.419	0.922	1.736	1.991	2.614	8.399	1.983	1.097
3	1.044	4.429	5.719	1.479	3.479	0.788	1357	4.079	2.506	8.438	1.775	1.073
4	1.374	5.037	6.026	0.723	3.997	0.824	1.707	1.37	3.463	3.814	1.752	1.041
5	1.665	4.914	3.476	0.836	0.791	1.046	1.382	1.487	5.002	1.413	1.391	1.041
6	1.479	3.378	2.055	0.691	0.649	0.733	1.107	1.312	2.759	1.781	0.657	1.428
7	1.026	3.078	2.215	0.662	0.56	0.751	1.035	1.609	2.55	1.145	0.572	1.065
8	0.955	2.964	1.806	0.632	0.533	0.759	1.034	0.855	2.832	1.166	0.387	1.098
9	0.931	3.315	0.456	0.646	0.496	0.781	1.033	0.875	2.811	1.156	0.349	0.991
10	0.693	3.142	0.458	0.664	0.509	0.814	1.042	1.254	1.788	1.633	0.218	0.96
1 I	0.658	2.808	0.45	1.075	0.583	0.783	1.034	0.867	1.463	1.157	0.274	0.841
12	0.653	3.414	0.451	3.841	0.754	1.266	1.025	2.422	1.31	1.82	0.747	0.747
13	1.049	3.49	3.035	4.009	0.891	4.205	1.038	5.316	1.886	5.038	1.136	0.824
14	0.598	4.838	1.864	1.338	0.666	2.236	1.125	6.17	1.372	3.746	1.715	0.992
15	0.618	7.268	1.691	0.607	0.501	0.717	1.1	3.282	1.492	1.052	0.844	0.888
16	4.093	4.763	2.273	0.581	0.471	0.93	0.994	0.901	2.66	1.61	0.226	0.807
17	2.703	6.02	2.091	0.547	0 536	0.785	1.055	1.023	5.184	0.961	0.504	0.972
18	0.574	5.969	2.468	0.59	0.577	0.61	1.399	0.801	2.148	1.096	0.223	1.605
19	0.612	4.099	2.141	0.55	0.801	1.113	1.617	0.717	1.291	1.096	0.823	1.419
20	1.034	3.887	2.855	1.186	0.996	0.91	3.378	1.278	3.596	1.149	0.293	0.718
21	0.875	3.765	8.29	0.571	1.075	0.605	4.152	0.648	4.791	1.564	0.362	0.731
22	0.726	4.236	7.541	0.577	3.511	0.612	1.112	0.635	1.889	1.388	0.856	0.855
23	0.715	4.338	2.648	1.13	5.488	0.674	1.022	2.766	2.017	1.552	3.716	1.036
24	1.595	4.601	4.215	1.144	4.279	0.806	1.131	5.806	3.445	2.648	1.313	2.047

TABLE XXII. Solar irradiance (W/m2) profile

Hour \ Solar	Sunny	Cloudy	Snowy
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0.107	0.048	0
9	0.549	0.432	0.264
I0	0.72	0.601	0.435
11	0.806	0.665	0.524
12	0.846	0.682	0.572
13	0.829	0.634	0.572
14	0.812	0.603	0.525
15	0.749	0.534	0.419
16	0.572	0.351	0.225
17	0.14	0.036	0
18	0	0	0
19	0	0	0
20	0	0	0
21	0	0	0
22	0	0	0
23	0	0	0
24	0	0	0

TABLE XXIII. Electricity price ($/kWh) and natural gas price ($/l)

Hour \ Price	Basic	TOU	Natural gas
1	0.0782	0.0727	0.00867
2	0.0782	0.0727	0.00867
3	0.0782	0.0727	0.00867
4	0.0782	0.0727	0.00867
5	0.0782	0.0727	0.00867
6	0.0782	0.0727	0.00867
7	0.0782	0.0727	0.00867
8	0.0782	0.0727	0.00867
9	0.0782	0.0727	0.00867
10	0.0782	0.0727	0.00867
11	0.0782	0.0727	0.00867
12	0.0782	0.0727	0.00867
13	0.0782	0.0727	0.00867
14	0.0782	0.0727	0.00867
15	0.0782	0.0727	0.00867
16	0.0782	0.0727	0.00867
17	0.0782	0.0727	0.00867
18	0.0782	0.0727	0.00867
19	0.0782	0.0727	0.00867
20	0.0782	0.0727	0.00867
21	0.0782	0.0727	0.00867
22	0.0782	0.0727	0.00867
23	0.0782	0.0727	0.00867
24	0.0782	0.0727	0.00867

APPENDIX B (COPYRIGHT STATEMENT)

Elsevier Copyright

Alqahtani, M., & Hu, M. (2020). Integrated energy scheduling and routing for a network of mobile prosumers. Energy, 200, 117451.

Author rights:

https://www.elsevier.com/about/policies/copyright

ELSEVIER

Copyright

Describes the rights related to the publication and distribution of research. It governs how authors (as well as their employers or funders), publishers and the wider general public can use, publish and distribute articles or books.

Journal author rights | Government employees | Elsevier's rights | Protecting author rights | Open access

Journal author rights

In order for Elsevier to publish and disseminate research articles, we need publishing rights. This is determined by a publishing agreement between the author and Elsevier. This agreement deals with the transfer or license of the copyright to Elsevier and authors retain significant rights to use and share their own published articles. Elsevier supports the need for authors to share, disseminate and maximize the impact of their research and these rights, in Elsevier proprietary journals* are defined below:

For subscription articles
Authors transfer copyright to the publisher as part of a journal publishing agreement, but have the right to: • Share their article for Personal Use, Internal Institutional Use and Scholarly Sharing purposes, with a DOI link to the version of record on ScienceDirect (and with the Creative Commons CC BY-NC-ND license for author manuscript versions) • Retain patent, trademark and other intellectual property rights (including research data). • Proper attribution and credit for the published work.

Quick definitions

Personal use

Authors can use their articles, in full or in part, for a wide range of scholarly, non-commercial purposes as outlined below:

- Use by an author in the author's classroom teaching (including distribution of copies, paper or electronic)
- Distribution of copies (including through e-mail) to known research colleagues for their personal use (but not for Commercial Use)
- Inclusion in a thesis or dissertation (provided that this is not to be published commercially)
- Use in a subsequent compilation of the author's works
- Extending the Article to book-length form
- Preparation of other derivative works (but not for Commercial Use)
- Otherwise using or re-using portions or excerpts in other works

These rights apply for all Elsevier authors who publish their article as either a subscription article or an open access article. In all cases we require that all Elsevier authors always include a full acknowledgement and, if appropriate, a link to the final published version hosted on Science Direct.