QATAR UNIVERSITY

COLLEGE OF ENGINEERING

HOME ENERGY MANAGEMENT SYSTEM EMBEDDED WITH A MULTI-OBJECTIVE DEMAND RESPONSE OPTIMIZATION MODEL TO BENEFIT CUSTOMERS AND OPERATORS

BY

AYA ALI MAHMOUD ALI AMER

A Thesis Submitted to

the College of Engineering

in Partial Fulfillment of the Requirements for the Degree of

Masters of Science in Electrical Engineering

June 2021

COMMITTEE PAGE

The members of the Committee approve the Thesis of

Aya Ali Mahmoud Ali Amer defended on 15/04/2021.

Prof. Ahmed Massoud

Thesis/Dissertation Supervisor

Dr. Khaled Bashir Shaban

Thesis/Dissertation Co-Supervisor

Approved:

Dr. Khalid Kamal Naji, Dean, College of Engineering

ABSTRACT

AMER, AYA A., Masters : June : 2021, Masters of Science in Electrical Engineering

Title: Home Energy Management System Embedded with a Multi-objective Demand Response Optimization Model to Benefit Customers and Operators

Supervisor of Thesis: Ahmed M. Massoud.

This thesis aims to develop a Home Energy Management System (HEMS) that optimizes the load demand and distributed energy resources considering utility price signal, customer satisfaction, and distribution transformer condition. The electricity home demand considers Electric Vehicles (EVs), PV-based renewable energy resources, Energy Storage Systems (ESSs), and all types of fixed, shiftable, and controllable appliances. A multi-objective demand/generation response is presented to optimize the scheduling of various loads/supplies based on the pricing schemes. The customers' behavior comfort-level and a degradation cost that reflects the distribution transformer Loss-of-Life (LoL) are integrated into the multi-objective optimization problem. First, conventional optimization approaches are utilized to solve the multi-objective optimization problem. To overcome the conventional optimization limitations, a data-driven analysis, which utilizes deep reinforcement learning (DRL), is used. The results show that the DRL-based HEMS is more efficient in minimizing the energy cost while adapting to the user comfort within the desired level.

DEDICATION

To my mom and dad

ACKNOWLEDGMENTS

I would like to express my deepest gratitude and appreciation to my thesis supervisors, Pro. Ahmed Massoud and Dr. Khaled Shaban for their support, motivation, and comments to improve this thesis. Their guidance along the way was precious.

This thesis was made possible by NPRP11S-1202-170052 grant from Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

$h(H)$	Index (set) of time slots.
$n(N)$	Index (set) of demand response appliances.
ρ	Balance parameter for customer/utility benefits.
F_1	Transformer LoL mitigation objective function.
F_2	Electricity cost objective function.
F_3	Customer dissatisfaction cost objective function.
λ_h	Hourly electricity cost (¢/kWh).
$P_{n,h}^{sold}$	Power sold to the grid (kW).
C_h^{Tx}	Transformer LoL cost at time h.
$cdc_{n,h}$	Dissatisfaction cost for appliance n at time h.
$C_{n,\mathrm{h}}$	Electricity cost for appliance n at time h.
$E_{n,\mathrm{h}}$	Consumption for appliance n at period h (kWh).
$e_{n,max}$	Maximum consumption for appliance n (kWh).
$e_{n,min}$	Minimum consumption for appliance n (kWh).
$u_{n,h}$	Appliance status in household $\in \{0,1\}$.
ζ_n	Appliances dissatisfaction coefficient.
$T_{n,\mathrm{start}}$	Operation starting time.
$T_{n,int}$	Initial time of working period.
$T_{n,end}$	End time of the working period.
$T_{n,total}$	Appliance required operation time.
$P_{n,h}^{EV,sold}$	Power used by appliances fed from electric vehicles (kW).

$P_{n,h}^{EV}$	EV charging power (kW).
$E_{n,h}^{EV}$	Energy accumulated in electric vehicle battery (kWh).
$E_{n,h}^{EV,Max}$	Maximum energy in electric vehicle battery (kWh).
$E_{n,h}^{EV,Min}$	Minimum energy in electric vehicle battery (kWh).
$P_{n,h}^{EV,c}$	Electric vehicle charging power. (kW).
$P_{n,h}^{EV,d}$	Electric vehicle discharging power (kW).
$R^{EV,c}$	Electric vehicle charging rate factor.
$R^{EV,\mathrm{d}}$	Electric vehicle discharging rate factor.
$\eta^{EV,c}$	Electric vehicle charging efficiency factor.
$\eta^{EV,d}$	Electric vehicle discharging efficiency factor.
$SOE^{EV,int}$	Initial state of energy in electric vehicle battery (kWh).
$SOE_{n,h}^{EV,min}$	Minimum state of energy in electric vehicle battery (kWh).
$SOE_{n,h}^{EV,max}$	Maximum state of energy in electric vehicle battery (kWh).
$SOE_{n,h}^{EV}$	Available state of energy in electric vehicle battery (kWh).
$P_{n,h}^{ESS,sold}$	Power sold back to the grid from an energy storage system (kW).
$P_{n,h}^{ESS,c}$	Energy storage system charging power. (kW).
$P_{n,h}^{ESS,d}$	Energy storage system discharging power (kW).
$R^{ESS,c}$	Energy storage system charging rate factor.
$R^{ESS,\mathrm{d}}$	Energy storage system discharging rate factor.
$\eta^{ESS,c}$	Energy storage system charging efficiency factor.
$\eta^{ESS,d}$	Energy storage system discharging efficiency factor.

$SOE^{ESS,int}$	Initial state of the energy storage system (kWh).
$SOE_{n,h}^{ESS,min}$	Minimum state of energy storage system (kWh).
$SOE_{n,h}^{ESS,max}$	Maximum state of energy storage system (kWh).
$SOE_{n,h}^{ESS}$	Available state of energy storage system (kWh).
P_h^{PV}	PV output power (kW).
h_n^{arr}, h_n^{dep}	EV arrival / departure time.
P_h^{PV}	PV output power (kW).
$P_h^{PV,used}$	Household appliances power fed by PV (kW).
$P_h^{PV,sold}$	Power sold to the grid from PV (kW).
P_h^{G}	Flow of energy between household and grid.
l_h	Utility power limiting factor for the power drawn from the grid (kW).
θ_H	Transformer winding hot spot temperature.
θ_A	Ambient temperature in $^{\circ}$C
$\Delta\theta_H$	Winding hot-spot rise over top-oil temperature in $^{\circ}$C
θ_{TO}	Hot spot rise over top oil temperature in $^{\circ}$C
$\Delta\theta_{H,u}$	Ultimate hot-spot rise over top-oil temperature.
$\Delta\theta_{H,i}$	Initial hot-spot rise over top-oil temperature
τH	Winding hot spot time constant in hours.
$\Delta\theta_{H,R}$	Rated hot spot temperature rise over top oil temperature.
k_u	Ratio of ultimate load to rated load in per unit.
k_i	Ratio of initial load to rated load in per unit
R	Load loss ratio.

$\Delta\theta_{TO,u}$	Ultimate top oil rise temperature,
$\Delta\theta_{TO,i}$	Initial top oil rise temperature.
τTO	Oil hot spot time constant in hours.
$\Delta\theta_{TO,R}$	Rated top oil temperature rais over ambient temperature.
F_{AA}	Transformer aging acceleration factor
$F_{AA,r}$	Transformer aging acceleration factor for time interval Δt_r.

CHAPTER 1 : INTRODUCTION

In this chapter, a background pertinent to smart grid and demand-side management applications is introduced. Also, the thesis motivation, objectives, and contribution are discussed.

1.1 Background

The electric grid consists of three main phases: generation, transmission, and distribution phase, as demonstrated in Figure 1-1. The electricity is transmitted from the main substations and distributed to different customers [1]. The first power grid with the alternating current was constructed in 1886 [2]. Since then, with the advances in technology, the electric grid has a significant expansion and changes. The current US electric grid has more than 9,200 generation units, 300,000 miles of transmission lines, and its generation capacity is estimated to be more than 1 million megawatts [1].

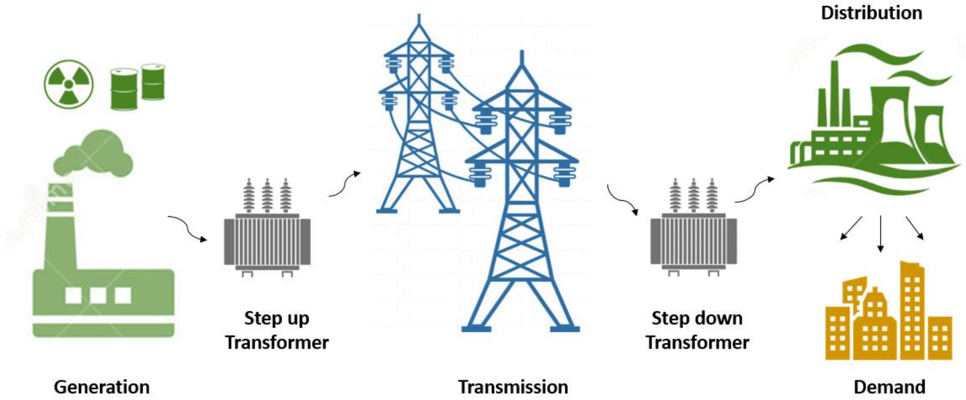

Figure 1-1 Simpled diagram of a conventional power grid

Since its invention, the electric power grid has been considered an engineering marvel. The electrical grid was implemented using centralized architecture, which consists of high voltage power plants. The power is transported to customers using high voltage transmission lines. At that time, this centralized architecture was meeting the demand requirements. Also, it provided a rapid network expansion and good quality of electrical supply to the end-user. In recent years, the electric grid has been subjected to a set of emerging challenges and problems that it is not designed and engineered to overcome. Some of these challenges are the aging of network infrastructures, growing energy demand, optimal deployment of expensive assets, new electrical uses (i.e., electric vehicles), energy management systems, pollution, and the greenhouse effect.

Two-thirds of the used fossil fuel is wasted in the existing grid and cannot be transferred into electric energy [4]. About 8% of the generation is wasted during the transmission from the power plant to the customers. Besides, 20% of the existing grid capability is only available to cover the demand-side requirements at peak time, which occurs only 5% of the time. Additionally, the existing power grid experiences failures due to the hierarchical topology of its assets [4]. Moreover, the existing grid only provides one-way communication. It does not allow communication between utilities and customers, which leads to a lot of inconveniences and economic losses.

The above-mentioned challenges cannot be tackled within the existing electric grid. Therefore, the next generation of the electric power grid, also called smart grid, is introduced to the electrical market to overcome the existing challenges and enhance electric system performance. The smart grid is an information-producing and intelligent entity rather than only an operation-based system. A smart grid utilizes new technologies and strategies to effectively integrate the grid with distributed energy generation and energy storage to balance the load (Figure 1-2) [2]. The conventional power grid only has one way of communication, while the smart grid has enhanced sensing and computing abilities to enable a two-way communication network. Different elements and components are connected via the communication network and smart sensors to enable advanced control.

The smart grid's main features are providing utilities with advanced control and complete visibility over its assets and resources. Advanced metering infrastructure (AMI) has become a popular research topic [4],[5]. AMI has a significant impact on the system performance and asset management. AMI utilizes a two-way communication, which can monitor and record system parameters (i.e., voltage and current), remotely connect and disconnect services, send alarm information from end-user to operator/operator to end-user within a near real-time operation. Additionally, unlike the conventional power grid, the smart grid is designed to be self-healing and robust to any inconsistencies. Therefore, utilities should invest in smart grid technologies to make electric grids efficient, reliable, sustainable, and resilient.

Along with these smart grid features, other new and highly penetrating components need to be considered. For example, the development in Renewable Energy Sources (RESs) such as Photovoltaics (PVs), Energy Storage Systems (ESSs), and energy harvesting technologies have led to a rapid increase in the integration of PV systems into residential and commercial premises. Moreover, Electric Vehicles (EVs) are becoming widely used. EVs differ from traditional loads, as they may consume and provide electricity.

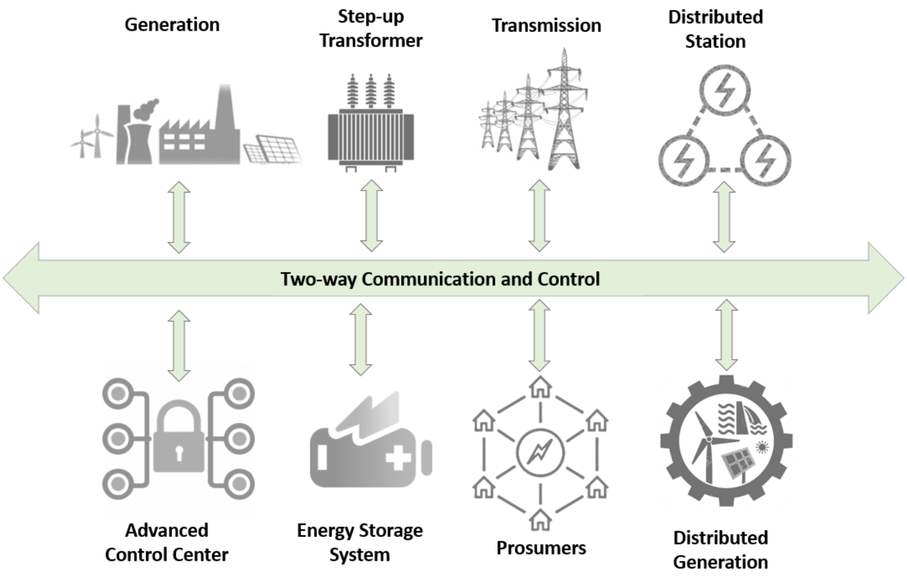

Figure 1-2 Smart grid diagram

A key challenge for smart grids is consumer participation in the system as an active element affecting the system performance. Therefore, some crucial concepts should be considered, such as Demand-Side Management (DSM), prosumers, and energy citizenship, which require the customers to adopt energy as a fundamental part of their life. The load which customers require from the electric grid always varies. Therefore, the utility should manage both generation resources and consumer consumption. The utility has managed this issue by using Demand Response (DR) strategies to compensate for short-time supply-side gaps. Although DR is effective, if done traditionally, it can affect customers' satisfaction or comfort. In the 30-plus years, utilities have performed peak reduction by controlling particular devices and systems [6]. A smart grid combines appliances and systems to develop integrated solutions, which enable utilities to look for more value suggestions across more advanced technology. This is known as DR optimization, a strategy to create technical and economic benefits by leveraging the demand in many ways.

Primarily and due to their massive impact, DR programs often target industrial and commercial customers. This is because a high cost and energy reduction can be made when DR is implemented with industrial customers. However, for residential customers, who typically have distinct needs and requirements, utilities have to consider different customer consumption profiles. One of the main challenges to implement DR in the residential sector is consumers' involvement. However, a class of end-users, who are not technology supporters, are not usually willing to effectively participate in DR programs. Therefore, to implement a DR program in a residential sector, an additional incentive with a high comfort level is required to support customers' engagement besides the electricity cost reduction achievement.

Effective demand optimization is meant to integrate the power systems with several advanced resources to demonstrate the grid's operations capabilities. Some of the required demand optimization abilities are [7]:

- Network awareness: The ability to coordinate assets on an electrical network is necessary. Also, utilities need to utilize distributed generation in the grid at different locations to expand the power system operations.
- Customer awareness: Different contractual agreements should be developed for the residential, industrial, and commercial sectors to represent and meet their needs.
- Forecasting: The meter data and the behavior-based analytics should be utilized to forecast customers' participation rates in the DR programs. This method is a strong indicator to help the utility to improve the market and the DR contractual agreements.

Integrating the above capabilities into a DR program will demonstrate more benefits beyond what traditional DR programs could anticipate, such as balancing electricity supply and demand, optimizing utility assets, minimizing energy generation cost, reducing reliance on fossil-fuel, and increasing the accommodation of RESs [8]. However, the significance of this type of DR program needs to be investigated both for the customers and the operators. Finally, educating the customers about the savings that can be achieved by their participation in DR programs is an important key to have the DR optimization working effectively. Also, operators should avoid DR's conventional model and consider demand optimization more inclusively and comprehensively. Power system operators, government, customer, service and technology companies, and other stakeholders should work together to improve the electrical grid capabilities and make it more efficient and reliable by utilizing DR strategies.

1.2 Problem Statement

DR schemes' primary aim is to match the electrical power supply with the consumption. Traditionally, utilities adjust generation rates according to changes in demand. This practice is costly as it leads to turning generation units on and off, importing power from other utilities, or applying load-shedding. With the advent of the smart grid, it is technologically enabled for operators to adjust demands. For instance, non-essential loads can be reduced, and the energy consumption is shifted from peak hours to lower-demand times. This is principally done with customers' approvals and based on time-based dynamic pricing schemes.

DR schemes also aim to maximize the utility profit by the optimal deployment of expensive and critical assets, such as transformers, as they impact the power system adequacy and reliability. They should be deployed efficiently to receive a reasonable return on investments. Any failure in the distribution transformer can results in power outages, in addition to expensive and time-consuming repairs and replacements. DR programs are vital to effectively increase the smart grid performance while considering all these components and utilizing the enabling technologies.

The residential sector's energy consumption accounts for 16–50% of the consumed power by all sectors in the US and presents about 30% worldwide [9]. Due to this reason, this thesis targets optimizing energy consumption for residential users. It is complex to deal with residential loads as they usually have different needs and relatively fine-grained requirements to ensure comfort. Home energy management systems (HEMSs) manages the residential loads and oversee the entire facility's energy and data flow [10]. These systems facilitate communications with customers through various channels to confirm their participation in a DR and inform them about an event, energy usage, pricing, etc.

This thesis proposes HEMS embedded with a multi-objective DR algorithm to achieve the appropriate balance between customers and operator benefits. The DR objective function's optimal weight can provide a realistic trade-off solution without violating the distribution network regulation rather than a cost-effective oriented solution. Therefore, this research deals with the DR problem in a large-scale context to solve residential buildings' demand optimization problems from end-user and utility perspectives. The primary objective is to use different strategies to reduce customers' electricity bills and power peaks, considering their techno-economic, environmental, and social effects. These objectives should be fulfilled without comfort losses for the customers or power grid regulations.

1.3 Thesis Motivation

The actual energy situation and future previsions are alarming because the increase in energy demand does not satisfy the sustainability objectives. Indeed, based on the International Energy Agency (IEA) report, the world demand for energy was estimated to be 12 Giga-ton equivalents of petroleum (GTEP) in 2010 with a 13% use of renewable energies. In 2035, the world energy demand is estimated to increase to 17 GTEP, with an 18% use of renewable energy [11]. Although most of the electricity generation depends on fossil fuels, the share of RESs has increased recently. In 2019 investments in RESs, mainly investments in solar PV, hydro, and wind power, were more than fossil fuel investments. Also, for the first time, renewable energy generation has increased to 10.4% [12]. The total renewable energy sources' consumption is growing strongly. In 2018, renewable energy share was increased from 4.5% to reach 5%. The European Union's total share in renewable energy is about a third of the total generation capacity [11],[12]. Figure 1-3 illustrates the primary RES share in 2018 and 2019 for the main countries.

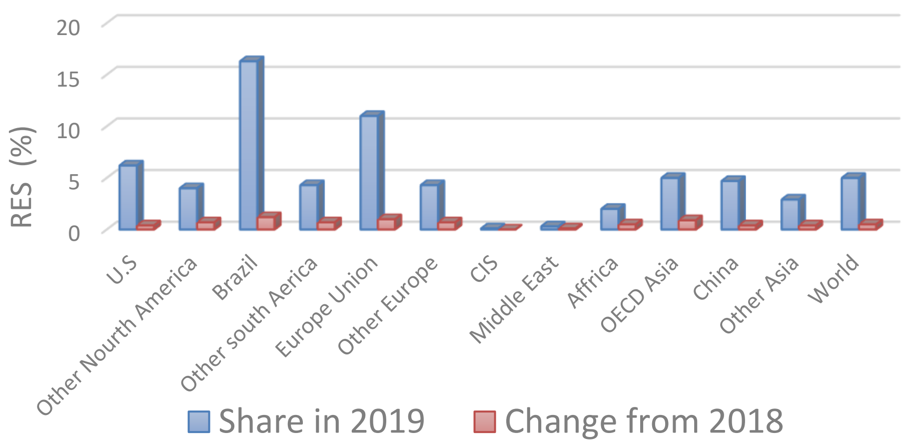

Figure 1-3 RES share in the key countries and regions [12]

Integrating RES into the power grid may disturb the system operation balance since the electricity, in this case, is produced under specific conditions, for example, when the weather is sunny or windy. Thus, variations in electricity production should be considered. Besides RESs, the interest in EVs is increasing, and it is expected to increase more as more new models are entering the market. Figure 1-4 demonstrates the growth in the EVs market in the period of 2012 to 2013 in key countries. The growth is shown as the year-on-year percent change in EVs market shares. Netherlands' relative market share has reached more than 400%, making it one of the leading countries in 2013. The EVs market share is increased from 1 % to around 6%. Germany comes in second place with a 105% market share increase [13]. It is estimated that EVs' consumption rate in Europe would increase to 4% in 2030 and 8% in 2050 from the total generated power [14].

Like RESs, EVs charging/discharging would affect the power system stability. If charging/discharging of the EVs is not coordinated, this could lead to additional demand on the power system. This increase in demand will require grid expansion, for instance, increases transformer capacity [14]. These issues can be tackled at the micro-grid level by looking at individual residential households to minimize their consumption and make it more sustainable.

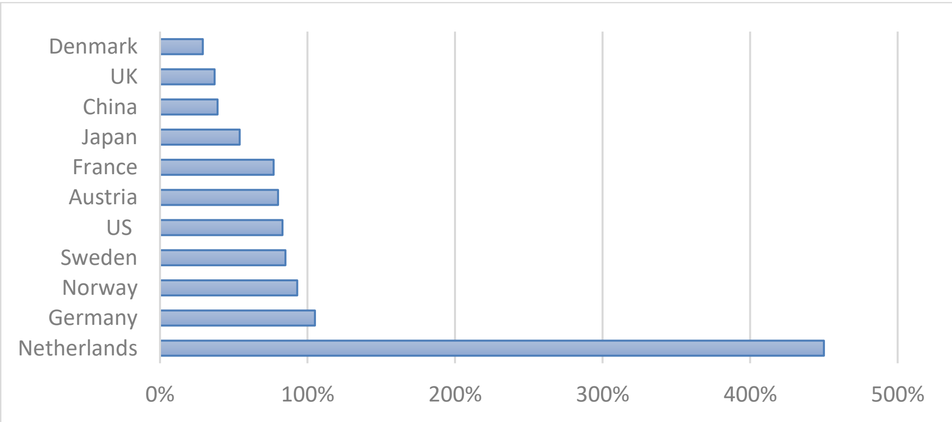

Figure 1-4 Growth rate of EV market share (2012-2013) of electric vehicles [13]

There are many concerns regarding the residential sector encouragement to enhance the power system's reliability and security in recent years. It is reported that a 5% decrease in the peak demand by participating in DR programs could lead to $35 billion savings in the US over 20 years [16]. Furthermore, penetration of EV in the residential sector and the increase in residential appliances could result in load shortages if operated simultaneously. Thus, the residential customer should be considered when investigating system-level demand reduction, and HEMS is a crucial topic in the research area to ensure a well-functioning power system.

The available information about the variation of the energy consumption in households is relatively limited. The new advances in Machine Learning (ML) technologies and AMI increase electricity usage visibility and provide more valuable information for both supply and demand sides. Finally, there is a lack of negotiation and dialogue between end-user and power system operators (PSO). The DR implementation is necessary to obtain this dialogue.

1.4 Thesis Objectives

This thesis aims to design an incentive-based residential DR algorithm that can benefit both the utility operators and the end-users. To accomplish the thesis objective, the following objectives are followed:

1. Perform literature review on the existing demand response optimization algorithms focusing on HEMS and different utilized resources.
2. Study the models of different smart household appliances, controlled and managed by the demand response algorithm.
3. Design a flexible HEMS framework to smooth the power consumption profile and optimize energy consumption without compromising user comfort.
4. Develop a transformer thermal model to calculate the LoL cost and integrate it into the DR program.

5. Evaluate the performance of the HEMS model through different simulation cases using conventional optimization.
6. Investigate the conventional optimization limitations and propose a second solution based on a model-free machine learning technique.
7. Presents the HEMS problem formulation based on the DRL algorithm and the home energy management framework.
8. Present simulation results and performance analysis of the proposed conventional and machine learning-based HEMS in achieving the objectives.

1.5 Thesis Contributions

This thesis deals with the DR problem in a large-scale context to solve residential buildings' demand optimization problems from end-user and utility perspectives. The main contributions of this study are presented in the following points:

1. Development of HEMS that schedules the household load by optimizing the trade-off between the energy consumption cost and the Customer Dissatisfaction Cost (CDC), considering the distribution transformers' asset condition.
2. The HEMS considers different types of loads (fixed, time shiftable, controllable, and EV) and facilitating RES and ESS integration with an efficient energy management system. Also, a bi-directional power flow among after-the-meter RESs and the residential appliances is considered.
3. The conventional HEMS methods are based on a system model. These methods can show good performance because of the assumption of accurate prediction and accurate input data. However, these assumptions are impracticable and unreasonable. Therefore, a Deep Reinforcement Learning (DRL) data-driven method is utilized to maximize energy efficiency in a residential household

where the variation of the real-time electricity prices and residents' activities is considered.

4. Most RL-based HEMS in literature deals with multiple agents where each household appliance represents an agent acting in the same environment, which has proven to be a challenging task to solve. The proposed scheme works with a single agent and uses a reduced number of state-action pairs, making it more effective in HEMS applications since the decentralized control is not crucial at a single household level.

1.6 Thesis Structure

The outlines of this thesis are as follows:

Chapter 1 Provides the general introduction, problem statement, motivation, objectives, and contributions of the thesis and thesis structure.

Chapter 2 Presents a literature review and explains the concept of smart homes, smart appliances, smart meters, and energy management systems and demand response in smart homes. Also, explains the different pricing schemes and their ability to reduce peak demands. The effect of demand response on customer comfort is presented.

Chapter 3 Presents the study of different residential appliances models and their implementation in a multi-objective demand response framework. Residential load models are used in the HEMS optimization, where each appliance's optimal operation schedule is determined.

Chapter 4 Presents the mathematical programming and implementation process of the HEMS algorithm. Presents the models' parameter and the input data for the model, along with the transformer parameter. Also, the considered cases in the simulation are introduced, along with calculating energy and cost.

Chapter 5 Presents the simulation results, discussion, and evaluation of the proposed algorithm using a conventional optimization technique. Different cases are examined to validate the efficiency of the proposed HEMS model. First, a single objective optimization where the effect of allowing flexibility to the household appliances on the cost is examined considering customer comfort. Then, the transformer LoL cost is incorporated into the algorithm. Also, the effect of DERs on the HEMS is present in one of the cases.

Chapter 6 Investigates reinforcement learning (RL) rules and application in the power system and provides a background and related work for the RL algorithm. The main elements that jointly drive the performance of an RL algorithm are introduced. Moreover, the details of the Markov decision process as a decision-making model for RL are discussed. The RL categories are presented along with the application of deep neural networks in RL algorithms.

Chapter 7 Addresses home energy management using deep reinforcement learning (DRL). Also, Presents the HEMS problem formulation based on the DRL algorithm and the home energy management framework. Elucidates how a DRL agent can be utilized to produce optimal solutions in a home energy management system. Finally, the implementation of the algorithm in MATLAB is presented.

Chapter 8 Provides simulation results and performance analysis of the DRL algorithm to optimize power consumption in smart households. Two scenarios are considered. First, minimizing the electricity cost and customer comfort is considered. Then, the transformer LoL cost is incorporated into the algorithm. Also, a comparison with conventional approaches is presented.

Chapter 9 Presents the conclusions based on the simulations done in chapter 5. The last section of chapter 6 finishes with some final remarks and future work of the research. Figure 1-5 shows a sequence diagram to highlight the outline with the thesis's objectives and contributions.

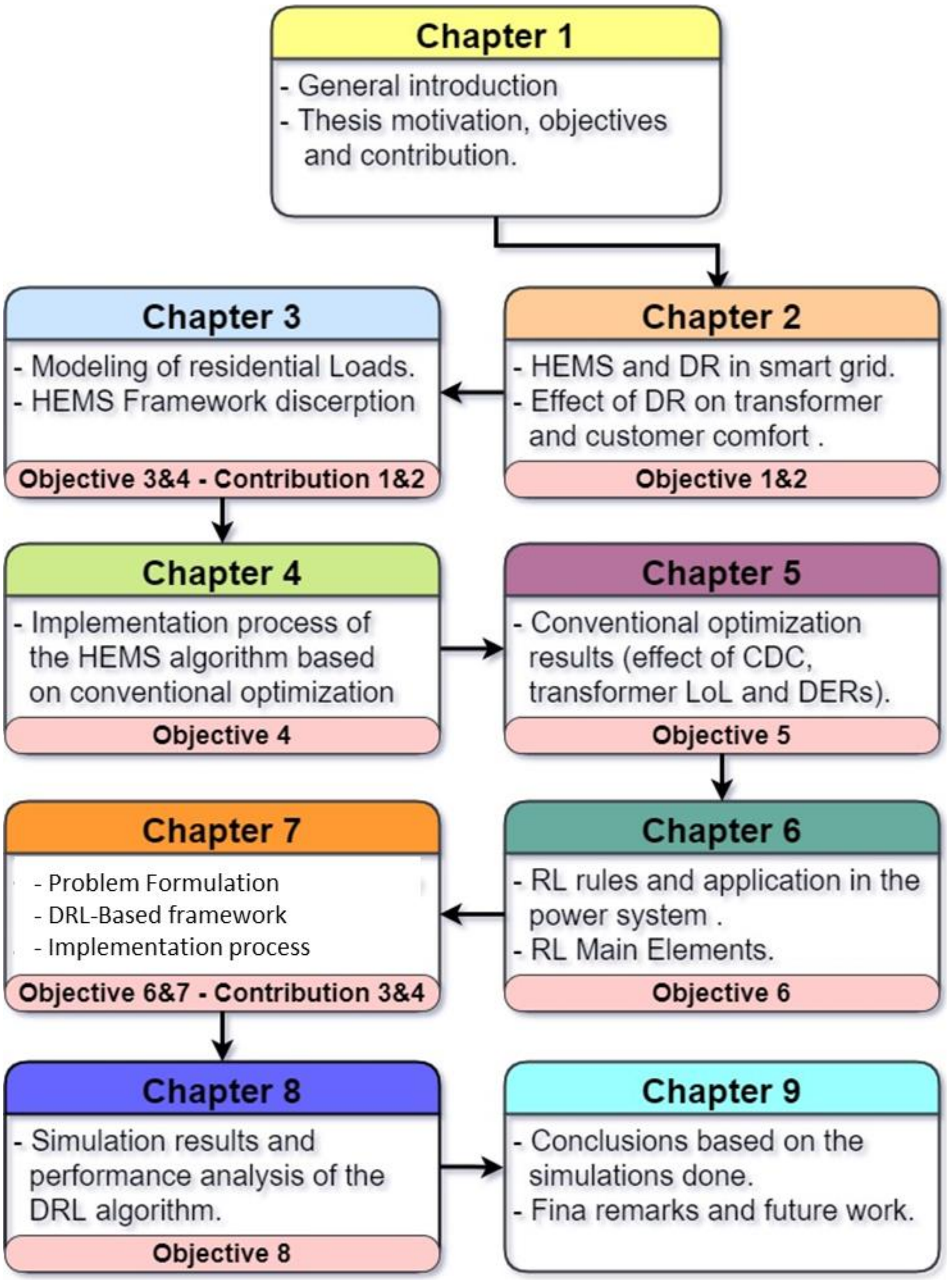

Figure 1-5 Thesis outline sequence diagram.

CHAPTER 2 : LITERATURE REVIEW

This chapter gives a detailed examination of different demand response and HEMS models and addresses developing DR and HEMS. Also, the effect of the proposed HEMS on the distribution transformer is discussed.

2.1 Smart Home

A smart home is a household equipped with different home automation systems controlled by the main controller [17]. Customer interaction with the grid is the most critical attribute of smart homes. Customer involvement is facilitated by a smart meter, which is considered a connection point between the smart home and the smart grid. Smart meters collect the consumption data for utility and end customers. With the upcoming technologies of home automation, all home appliances will have the ability to send and receive information. A smart home can provide the owner with more opportunities, such as expense-saving, comfort, and reduced carbon emission. For example, smart home appliances can be programmed to respond to specific commands or signals from utility to cut power consumption during high peak load or shift appliances operation to low-cost times. The smart home will have different levels of complexity based on the number and type of appliances, a home communication platform, and the desired automation level. A home energy management in the smart home may include a simple notification from the home user's DR program. The HEMS or a smart meter directly communicates with a specific appliance to turn it on or off [18]. As a more complex example, a smart meter communicates with HEMS to perform specific parallel actions. The HEMS can communicate with smart appliances in different ways, e.g., wired or wireless, to perform DR over the Internet using an existing broadband connection.

Moreover, the future smart homes will be equipped with renewable generation and ESS, as shown in Figure 2-1, to make the smart home working as a small connected micro-grid. The higher share of renewable energy in the smart home will minimize the grid's purchased energy, which will result in more cost minimization and peak demand reduction. The excessive renewable energy generation can be injected into the smart grid or stored in ESS. Smart homes are usually managed by HEMS, optimizing energy consumption, and giving the users feedback about their electric consumption [18]. The HEMS incorporate and manage three main tasks:

- Optimization: find the most suitable time to use electric appliances considering the electric pricing fluctuation and the peak loading. It aims at minimizing the wasted energy and electricity bills while maintaining users' comfort.
- Control and automation: a microcontroller-based system to control the HEMS interface and oversees the use of home appliances and their consumption.
- Communication: manages the wireless network part by providing dynamic information about home energy consumption based on power line communication. A ZigBee interface is a perfect example of such a communication system.

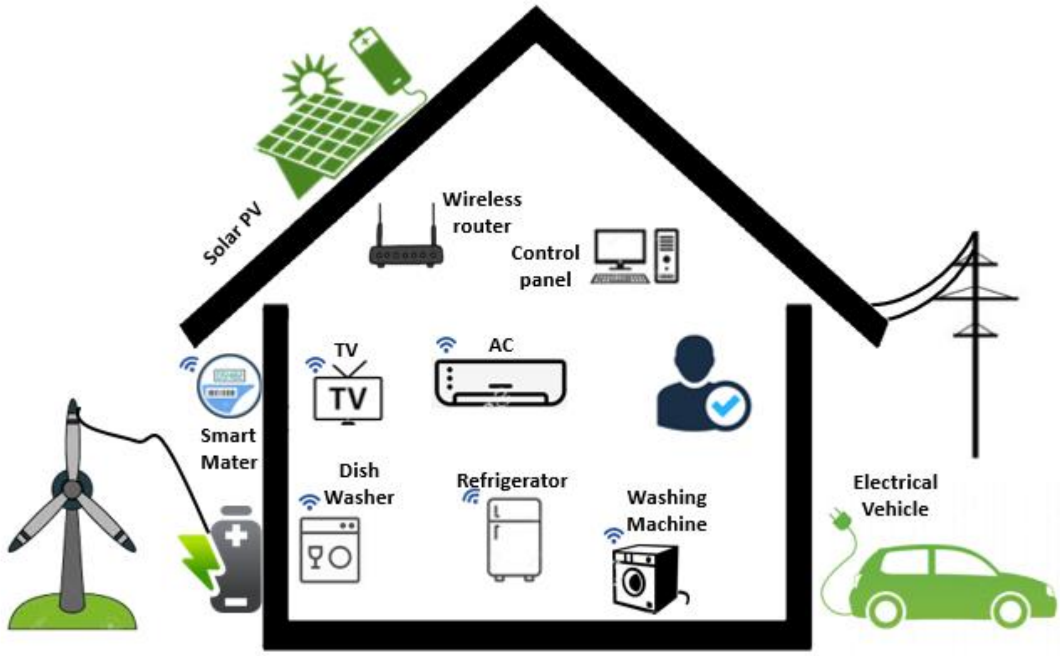

Figure 2-1 Smart home structure

2.1.1 Smart Appliances

A smart appliance refers to a home appliance that monitors, protects, and automatically adjusts its operation according to the homeowner [19]. The main characteristics of smart appliances include the following:

- Can be programmed to sets of instructions by the user.
- Send and receive alerts from and to the user.
- Provide an energy consumption based on the user guidelines.

Customers can get economic advantages and environmental benefits by adapting smart appliances in their homes. However, a small percentage of customers are aware of the environment. Most consumers adopt smart appliances for economic benefits [20]. To trigger consumers to buy smart appliances, utilities should offer attractive tariffs to customers or other incentives.

All residential smart appliances are considered receivers and controlled by a transmitter, such as remote control, keypad, or smartphone application [21]. For instance, as shown in Figure 2-2, if the user needs to switch any appliance ON or OFF, the transmitter (smartphone) will transmit a signal to the appliance in the form of a code, including the required instruction be performed. Each appliance is identified by a unit number [21].

Figure 2-2 Smart home devices accessed through smart phone

2.1.2 Smart Meter

Smart meters are considered intelligent devices which used to monitor and control energy usage in homes. Its monitoring function is based on collecting measured energy data, performing energy analysis based on the algorithm uploaded, and preparing a real-time energy usage [5]. A smart meter performs more complicated monitoring functions compared to a normal automated metering reading (AMR). Moreover, the smart meter can be combined into an AMI to supply real-time information and services to utilities and customers [5]. Table 2-1 presents a comparison between AMR and AMI.

Table 2-1 Comparison between AMI and AMR.

Automatic Meter Reading (AMR)	Advanced Metering Infrastructure (AMI)
Support one-way data flow.	Support two-way data flow.
Information flow from AMR to the utility.	Information flow among home appliances, AMI, and utility.
Interacts with Neighbourhood Area Network (NAN).	Interacts with NAN and Home Area Network (HAN) and or Business Area Network (BAN).
Consumers are unable to control their electricity usage through DR.	Consumers can implement DR.
Benefits majorly utility.	Benefits both utility and consumer.
Simple architecture.	Complex architecture.
Negligible security risk.	High-security risk.

Smart meter devices' development employs advanced technologies that impact information transfer between consumers and utility. Communication technology has been considered mostly in smart meter advancement due to its significant impacts on the social, economic, and environmental points of use. In many articles regarding the smart meter, the major concerns are market as well as social benefits. Technological concerns of the smart meter are yet widespread. Moreover, most consumers have also focused on the economic benefits of smart meters rather than their performance. In [22], a study is developed to identify and measure the smart meter's social significance to the consumer. The factors considered to reflect social benefits were service reliability, existence, workability of feedback, the presence of DR, new products, and macroeconomic impacts.

In smart meter's technological advancements, an overview of a smart meter accompanied with the investigation on the implementation functionalities are proposed by [23]. An investigation is also made to observe satisfaction in the adapted services. The power is given to the "internet of things" on which goods are available in the same marketplace. A smart meter is attached in the home gateway to merge the home appliances through a home automation network (a communication between appliances and smart meter) with the utility via data exchange. Therefore, a smart meter is taken as a home gateway to connect the domestic appliances and utility through the internet. Apart from using a smart meter for energy monitoring in houses, the work suggests that a smart meter should be a multi-utilities device, meaning that it should work in electric energy, thermal energy, and natural gas. A smart meter's main advantage to operate under multi-utilities is to assist customers in an uncontrolled energy market. Furthermore, the work proposed a model for combining the hardware provider, service providers, and smart meter for each end-user.

Advanced measurement is advantageous in a smart grid. It takes place on monitoring transformer health in the grid. The monitoring process involves measuring the power lines' temperature, moisture content and computing electrical devices' thermo-images. Moreover, it should account for load capability and insulation aging factors. Based on these measurements, proper actions can minimize transformer failure risk by 2.5 times through a properly selected maintenance strategy [24]. Ref. [25] presents a technique to solve the power flow problems using smart meters. The observations were made on the utility's practicability to trim the end-users load with this work's remote signals. The mathematical approach was a modified one from the normal Optimal Power Flow (OPF) and could consider the practicability to purchase energy from different providers and convey it to the end-user. The main outcome of the optimization is to reduce the running cost of the distribution companies.

2.2 Demand Response (DR)

DR is about providing incentives to customers to schedule their loads to reduce costs and improve the electric power system [2]. The DR has been applied for decades, but due to EV and RES, assets aging, and advances in smart grid technologies, DR has gained more attention in recent years. Traditionally, DR programs are intended to be used during emergencies. These approaches are mainly described as peak shaving or as valley filling. With RES, DR can shift loads to operate at hours with high energy production from RES.

Figure 2-3 shows the different strategies of DR. The left graph shows the traditional way of DR, and the right presents DR in a system with distributed generation sources, e.g., ESS and RESs. Compared to the traditional case (left graph), in the distributed generation case (right graph), any shift in demand may increase the system total peak. In the distributed generation case, the increased peak demand does not always affect the whole grid since the electricity is produced locally. However, the grid could be affected when there is a large centralized renewable energy generation.

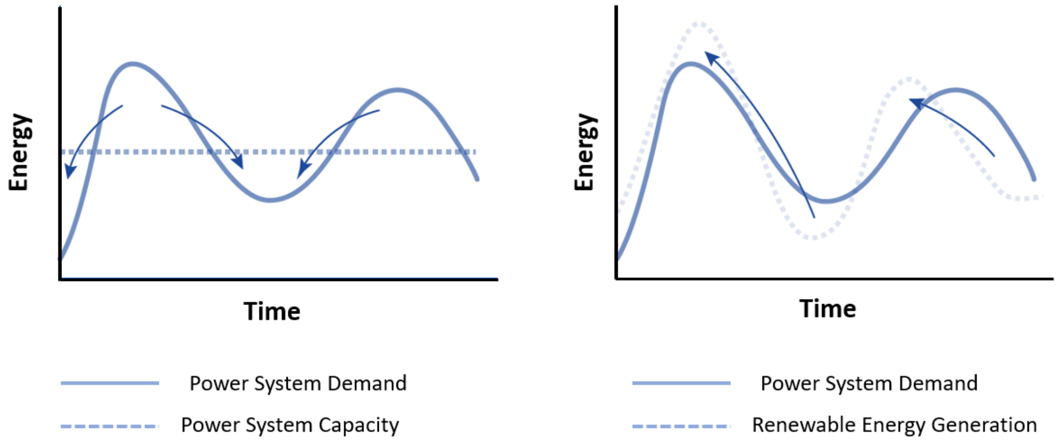

Figure 2-3 Difference between DR in traditional power system (left) and power system with considerable renewable energy share (right) [26]

There are different approaches to incentivize consumers to participate in DR. The DR programs can be arranged in two main groups depending on how the load adjustments are made. The first group is the price-based DR programs, and the second group incentive-based DR programs, as demonstrated in Figure 2-4. The incentive-based DR takes place when the utility has immediate control over the customers' loads. In this case, customers are rewarded with points or with a reduction in their electricity bill. This type of DR works when undesirable conditions exist in the network or when an emergency occurs.

On the other hand, price-based DR programs refer to user electricity usage changes to respond to electricity price change. Customers manage their consumption during electricity price periods to minimize the electricity bill [26]. Here, the customer response is entirely voluntary. More elaboration about both DR types is discussed next.

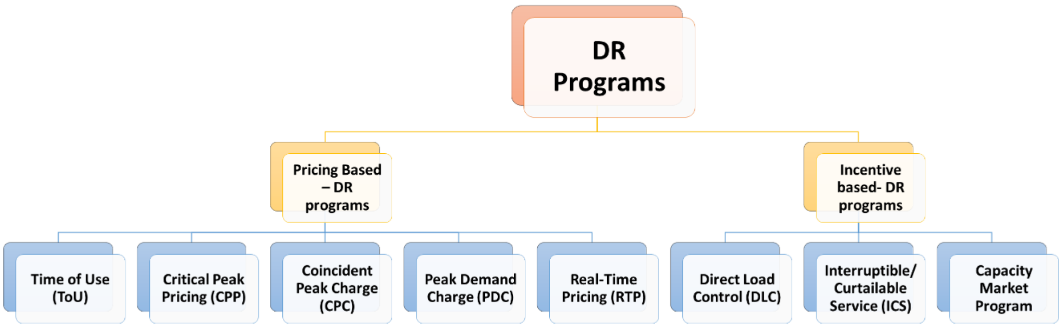

Figure 2-4 Avilable DR programs

2.2.1 Pricing Based Demand Response Programs

Pricing-based DR programs work by altering the load from high peak to off-peak times. Examples of pricing-based programs are discussed next.

Time of use Tariff (ToU) – ToU is a mutual method of DR pricing schemes is the TOU, which is used by the utilities in several countries such as Sweden, UK, the USA, Spain, and Italy [27]. This method is based on offering different electricity prices to customers. Usually, the day is divided into three main periods, where each period has a different electricity price, as shown in Figure 2-5 [28]. PSO uses this method to boost customers to change their consumption in lower peak periods. This tariff could be effective in a conventional power system where the consumption and production patterns are predictable. With the existence of DERs, this tariff structure is too fixed to describe the changes in the modern power system with DERs' share.

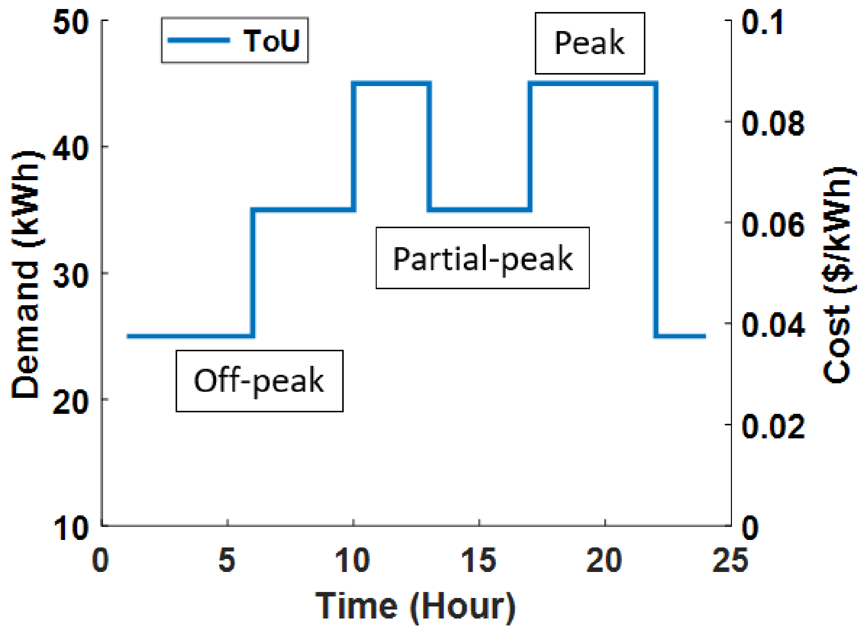

Figure 2-5 Time of Use (ToU) scheme

Critical Peak Pricing (CPP) – In this program, customers are incentivized by PSO to change their electricity usage to avoid transmission/distribution system overload. An example of a CPP situation is shown in Figure 2-6. CPP is applied during congested scenarios when the demand increases dramatically. PSO notify the users when this phenomenon occurs by sending electricity prices to them. One of the disadvantages of this approach that it is applied for a limited number during the year [29]. Therefore, it is not suitable to be used on regular operation times to enhance system performance.

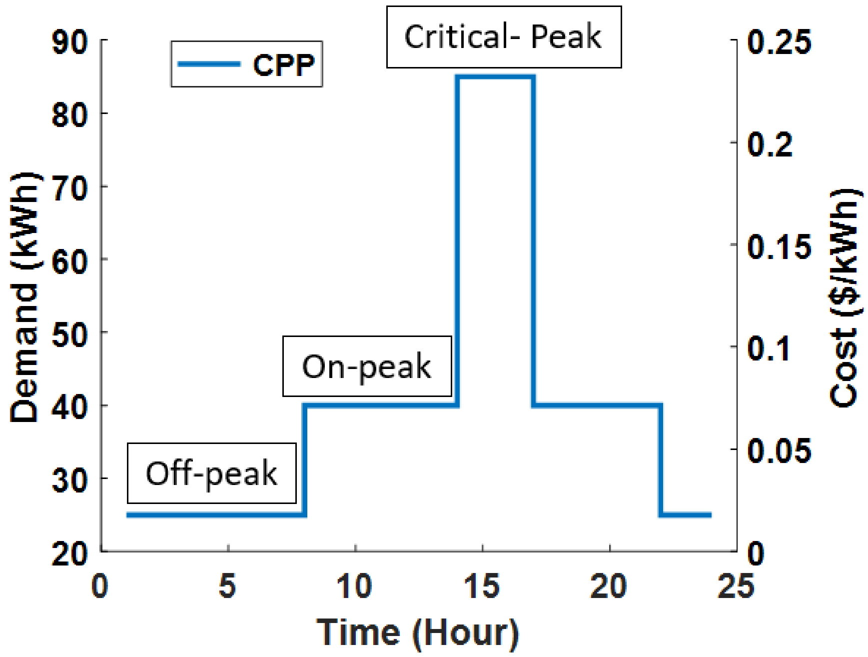

Figure 2-6 Crtical Peak Pricing (CPP) scheme

Power tariffs – The method is also called demand charges, where customers are motivated to minimize power consumption by increasing the electricity price at the time of their peak demand. This charge in this type is calculated by determining the day/month's hour when the customer's consumption is highest [30]. Generally, this scheme contributes to minimizing the overall peak system. However, it does not motivate the user to change their consumption to meet the generation. Moreover, minimizing the peak for each customer does not mean, in particular, minimizing the overall system peak since the customer’s peak may arise at a different period compared to the power system's peak.

Coincident Peak Charge (CPC) – This charging scheme is the same as the power tariff charge but manages the power system peak (the coincident peak) instead of the peak hour for a specific customer. This is done by determining the utility's peak hour at the end of each day/month. After that, all customers are charged for their electricity consumption during coincident peak time [31].

Real-Time Pricing (RTP) – The RTP program provides the participants with an electricity tariff representing the electricity market's real situation (see Figure 2-7). One of the scheme's disadvantages is the time difference between announcing the prices and the real consumption time. A long-time lag would give a price that does not precisely reflect the electricity market situation [30]. On the other hand, a shorter time lag will reflect the demand/supply condition but will make it more difficult for users to shift the energy usage. Users could overcome this difficulty by using HEMS.

Since the electricity market depends on the demand and available products, the electricity price variation may increase when electricity production from DERs increases. For instance, been observed when the generation of wind power increased, the electricity prices went negative [36]. Therefore, RTP is considered the best pricing scheme that supports the integration of RES and DR operations in the smart grid. The work done based on different DR pricing schemes will be dissuaded in the below.

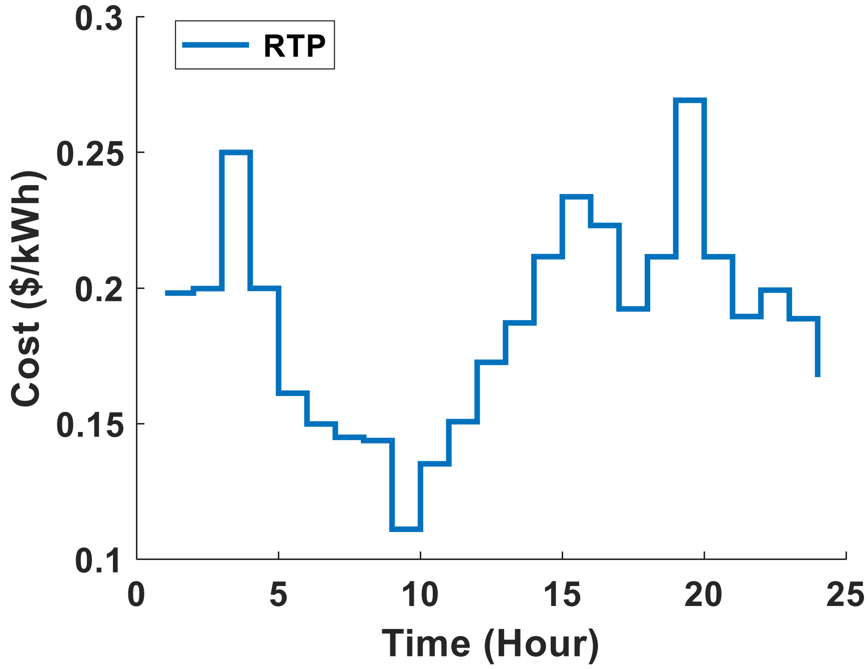

Figure 2-7 Real Time Pricing (RTP) scheam

As discussed, the electricity price is a key element in any DR program. Most of the electricity market studies integrate the quadratic functions to describe the relationship between electricity price and electric consumption. Different pricing methods are introduced to alter the consumption from high to off-peak periods. In [37], the ToU rate is effectively studied for residential users in various U.S. cities. ToU rate was implemented for the first time in the residential sector in 1975 in Vermont. Ref. [38],[39], [40], and[41] include other studies that focused on the effects of ToU rates. The studies conclude that the most critical issue to having a successful ToU rate is to have an efficient pricing method to alter consumers' consumption behavior.

To alleviate the peaks caused by EVs charging, work in [38] utilizes the ToU rate to schedule EV charging in low electricity price time. On large-scale applications, this leads to maximizing the generated load utilization and minimizing the cost of electricity generation. To better understand the effect of ToU rates on EV charging,

[39] represents the total fueling costs for EV under various ToU rates. The simulation results demonstrate that ToU pricing schemas are not effective in changing off-peak PHEV charging behavior. ToU rate models' effectiveness depends on the vehicle type, ToU scheme, and the peak period duration.

Work in [40] estimates the effect of ToU on residential customers. It is found that the load profile patterns of the participant and non-participant customers were almost the same under the ToU rates. Work in [41] presents a ToU DR program. The results show the ToU rate is profiting to the utility. These results contradict the common belief that ToU is unprofitable to utilities.

Besides ToU rates, the RTP rate is another pricing scheme, which has been intensively investigated by researchers (e.g.,[42],[43], [44], and [45]). Work in [45] points out the major barriers to fully implement RTP and utilize its benefits. First, the lack of experience between customers about how to respond to the RTP scheme. Also, the lack of building automation technologies could limit RTP potentials. Thus, they propose a residential energy management framework to overcome these problems. The trade-off between reducing electricity bills and reducing the appliance's operation waiting time is achieved using the RTP scheme. The coefficients' optimal choices for each day of the week are obtained by implementing a price prediction model to the actual RTP. The results indicate that combining the proposed DR and the price predictor model contributes to a considerable reduction in user's electricity bills.

2.2.2 Incentive Based Demand Response Programs

Besides varying the electricity tariff, incentive-based DR programs are also successful in the load schedule. Examples of incentive-based DR programs include:

Direct Load Control (DLC) – In DLC, customers get rewards when they give the PSO the control on their electrical loads during emergencies in the power system

[32]. This pricing model is simple to be implemented because the AMI is not required. One of the drawbacks of this approach is that all DLC participants are always rewarded even if they did not contribute during the contingency. Moreover, if loads are connected again, a controlled reconnection is required [33]. DLC has advantages on the power system by minimizing the cost due to smaller capacity provisioning, but overall, they do not contribute to the system performance during normal operation. Therefore, it is not commonly used nowadays.

Interruptible/Curtailable Service (ICS) – ICS method is also based on decreasing the power consumption during contingency [34][20]. ICS is usually only utilized for large customers such as industrial and commercial customers. PSO incentivizes the participants by offering a discount tariff. The participants are penalized when the required load reduction is not attained [35]. Similar to DLC, this pricing method cannot contribute to enhancing system operation during normal conditions.

Capacity Market Program – In this DR program, customers submit bids depending on the potential load reductions as an alternative for expensive generators. At the peak time, the utility selects the users based on their bids and sends a prior notice for load shifting. After that, users shit their loads in order to match the utility's prior notice. Users are penalized if they fail to minimize and change their energy usage during the given time.

Ref [46] proposes an incentive-based DR to generate flexibility in retail customers voluntarily. Their scheme's main advantages are improving the social benefits where the customers are not subjected to the price changes. However, the proposed scheme increases the communications burden. The Stackelberg game is used in [47] for determining proper incentives to increase the utility profits and minimize the customer electricity bills. The utility interactions with the end-user are modeled as a 1-

leader, N-follower Stackelberg game. The genetic algorithm is adapted at the leader's side to increase the benefits. An analytical solution is developed to minimize customers' bills at the followers' side. The results indicate that this method is useful for both the utility and end-user.

2.3 Home Energy Management System (HEMS)

Typically, HEMS serves several functions to end-users, such as monitoring energy consumption, managing the operation states of appliances, receiving information (such as tariff prices), and optimizing household appliances' power consumption based on the environment and time factors, and tariff prices. HEMS can also optimize the household appliances and manage the DERs and ESSs simultaneously. Figure 2-8 presents the basic architecture of HEMS. It consists of four main components: a monitor module, a prediction model, a demand response model, and the control unit. The monitor module monitors the household appliances' actual behavior, the storage system, and the PV installation. In practice, smart meters monitor the appliances that continuously measure power consumption over a certain period. The prediction module calculates the power production of the PV according to the monitored module's input data. These input data are solar radiation data and environmental variables, such as the PV panels' temperature, extracted through weather forecasting or sensors. The environmental variables are monitored because they affect the power output of the PV panels. The DR module contains the DSM software that computes the optimal schedule according to the specified optimization objectives, e.g., electricity cost and homeowners comfor.

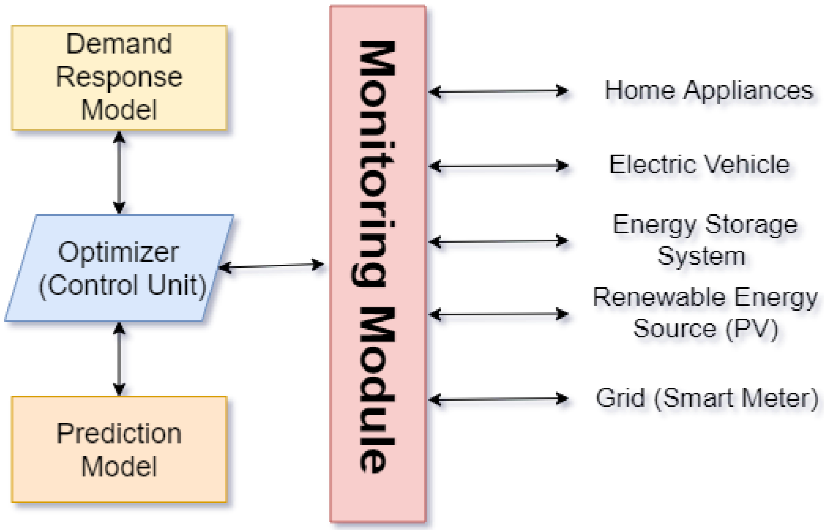

Figure 2-8 The architecture of the HEMS

Although HEMS's industrial field is still in its early stages, the HEMS market is rapidly expanding in the past few years. With the DR service, advanced pricing schemes, and different effective energy optimization techniques, HEMS will have core supporting techniques for further development. Many studies have worked on designing HEMSs using a different algorithm. Among these algorithms are the optimization-based algorithms, which consist of the objective function and a set of different constraints. Any optimization algorithm always aims to allocate the best optimum solutions for the optimization objective function that meets problem constraints. The HEMS application aims to find an optimal load schedule while meeting the home appliances model constraints and user comfort.

Several researchers have proposed HEMS with price-based DR algorithms for cutting the consumption cost. This is done by shifting different loads from high price periods to periods with low prices. MILP optimization is utilized to reduce electricity costs as in [36] and [48]. These studies considered only house appliances with thermostatically controlled loads and ignored others whose loads shifted. MILP optimization is also used in [49] to minimize energy cost by incentive-based optimal scheduling technique. Although several constraints are c, user comfort is neglected in this study. In [50], a DR model is presented to handle the energy consumption of shiftable and controllable appliances in a smart home. The customer discomfort is formulated by using the Taguchi loss function [51]. Moreover, another appliances scheduling problem is presented in [52]. The customer bill is minimized by shifting the appliances' operation at low price periods and avoiding the periods with high prices. However, this application could make new undesirable peaks in the minimum price slots, which will affect the utility side.

DERs, also called “decentralized” or “embedded generation,” refer to any power resource installed in the power grid that generates or stores energy via grid-connected devices. Example of DERs is EVs, ESSs, and RESs. The use of DERs another solution that contributes to a decrease in power system peaks. Recently, multiple works have been carried out to tackle RESs and ESSs integration in household operation with DR programs. A DR strategy for a nondeferrable load with PV and ESS capabilities has been presented in [53]. The study aims to reduce the expected energy cost by changing the discharging and charging time for ESS from renewable energy management unit considering ToU-based DR. Work in [54] proposes a DSM algorithm and for households with PV model. First, the expected daily consumption and PV generation are forecasted. Then, appliances' energy consumption is aggregated. After that, the load profile is optimized by minimizing the value of aggregated energy. The optimization attempts to use the PV output whenever possible. However, user comfort is not considered. In [55], [56], DR strategies are proposed with a bi-directional power flow possibility in a single HEMS to decrease the total electricity cost and improve the load pattern, however neglecting customer comfort. In [57], a HEMS is developed to provide optimum scheduling of the household appliances and sell the excess local PV generation to the grid.

Similarly, ref [58] presents HEMS architecture that allows consumers to interact with each other and suppliers. The architecture enables the integration of RERs with the electric grid. In [59] and [60], the effects of EVs charging on residential networks have been examined. A HEMS has been suggested in [59] to optimize the appliances' operation considering plug-in EVs.

2.4 Effect of Demand Response Programs on Distribution Transformer

Transformers are considered as one of the most significant assets in the electric power system. They are categorized into types: power transformer and distribution transformer. The difference between the two types is that the distribution transformer is designed to operate at a maximum efficiency of 60% to 70%, as usual [61]. The distribution Transformer is utilized at a low voltage level of the power system (distribution level). Any Failure in the transformer results in an interruption of power supply to customers, which leads to substantial economic losses. Thus, distribution transformer efficiency is an essential concern for the PSO.

Along with other causes, the main cause of failure of distribution transformers is prolonged overloading. The transformer's thermal management can utilize its operation by managing the transformer loss of life (LoL). The main two parameters used to approximate a transformer's lifetime are transformer winding top oil temperature (TOT) and transformer winding hot spot temperature (HST). A precise thermal model of distribution transformer can minimize LoL while utilizing the maximal loading [62].

External and internal factors can affect the transformer's thermal behavior, such as ambient temperature and the transformer's loading level. Increasing ambient temperature results in an increase in the operating temperature of transformers. Ref. [63] states that for every 10℃ increase in ambient temperature, the lifetime of distribution transformers will be reduced two times faster than the normal operation. According to [64], when the HST reaches 140℃, the relative aging rate reaches 100. This means that operating the transformer for one hour at overloading conditions equals 100 operating hours at normal operation. Any increase in the ambient temperature or transformer loading levels leads to an increase in the transformer HST. Thus, the transformer loading conditions depend on the HST and ambient temperature [64].

There are several aging models to manage transformer thermal behavior. There are transformer aging models specified in the International Electrotechnical Commission (IEC) and the Institute of Electrical and Electronics Engineers (IEEE) standards. These models are based on transformer relative aging rate. This rate presents the transformer’s equivalent aging rate when operating at a temperature different from the rated temperature. The relative aging rates for each standard are presented and discussed below.

IEC Standards 60076- This standard provides a guide for distribution transformer loading based on operating temperatures and thermal aging. This algorithm takes the transformer load, tap position, TOT, ambient temperature, and cooling operations as an input, and it gives the HST and LoL as an output [65]. The aging rates in this model can be calculated at two conditions:

a) *IEC non-thermally upgraded insulation* – The relative aging is attained at 98°C reference temperature.

$$F_{AA} = 2^{(\theta_H - 98)/6} \tag{2.1}$$

Where F_{AA} is the relative equivalent rate and θ_H is transformer's winding hot-spot temperature.

b) *IEC thermally upgraded insulation* – The relative aging is attained at 110°C reference temperature.

$$F_{AA} = 2^{\left(\frac{15000}{110+273} - \frac{15000}{\theta_H+273}\right)} \tag{2.2}$$

Where F_{AA} is the relative equivalent rate and θ_H is transformer's winding hot-spot temperature.

Utilities can utilize this model to calculate the insulation paper losses and several pre-loading and overloading scenarios. One of this model's technical challenges is not easily applied to all network distribution transformers [66]. However, many tests have been applied to the above model, and the output showed precise results for HST values.

IEEE standards C57.91 – This model is commonly used to calculate the thermal behavior of transformers. Based on the IEEE standard, the transformer's ambient temperature value is recommended to be less than 30℃ and limited to 40 ℃. Also, it recommends 110 ℃ as a maximum HST, with a maximum value of 65 ℃ for the winding temperature over the value of ambient temperature. The main principle of this model depends on the relationship between the transformer load and its operating temperature. According to IEEE C57.91-2011 standards, for 65 ℃ average winding rises, the reference HST is 110 ℃. The transformer LoL percent is defined as the equivalent aging at the reference HST. The transformer cooling type can affect transformer loading conditions. Based on the cooling type, the IEEE C57.91-2011 standard suggests different exponents to use in HST temperature calculations, as indicated in Table 2-2 [67].

Table 2-2 Transfomer's HST Exponents [67]

Types of transformer cooling	m	n
Oil Natural Air Natural (ONAN)	0.8	0.8
Oil Forced Air Forced (OFAF)	0.8	0.9
Oil Forced Water Forced (OFWF)	0.8	0.9
Oil Directed Wate Force (ODWF)	1.0	1.0
Oil Directed Air Force (OAF)	1.0	1.0

Where *m* and *n* are exponents used in calculating the change in hot spot rise over top oil temperature and top oil rise over ambient temperature with changes in load, respectively. The aging rates in this model can be analyzed at two conditions:

a) *IEEE non-thermally upgraded insulation* – The relative aging rate is attained at 95°C reference temperature.

$$F_{AA} = \exp\left(\frac{15000}{368} - \frac{15000}{\theta_H + 273}\right) \tag{2.3}$$

Where F_{AA} is the relative equivalent rate and θ_H is transformer's winding hot-spot temperature.

b) *IEEE thermally upgraded insulation* – Unity relative aging rate is attained at 110°C reference temperature.

$$F_{AA} = \exp\left(\frac{15000}{383} - \frac{15000}{\theta_H + 273}\right) \tag{2.4}$$

Where F_{AA} is the relative equivalent rate and θ_H is transformer's winding hot-spot temperature.

As can be seen, the IEC and IEEE relative aging rates for the thermally upgraded insulation are identical, and non-thermally upgraded insulation rates are different. This is due to the different reference temperatures used in non-thermally rates in the two standards.

DR can be used to mitigate the aging effect by managing the transformer's thermal behavior. In relation to previous literature, there has been work in DR literature aiming to increase transformer utilization. According to [68], utilizing DR programs could decrease the utility investment for installing new transformers by 75%. Work in [69] aims to minimize peak load by considering scheduling the appliances under RTP. In [70]-[72], DR's impact on distribution transformers aging has been examined. Different types of loads were controlled and shifted to reduce the LoL of transformers. However, in these studies decreasing the electricity price was not considered as an objective nor customer satisfaction. Ref [73] proposed a DR optimization model based on transformer hottest-spot temperature to improve transformers' utilization. In [74], the authors proposed a primary substation model to increase transformers utilization and life extension. However, EV loads were not incorporated into these studies, and again customer comfort was neglected. In [75], a price-responsive DR scheme has been used to investigate EV's impact on transformer aging. The results showed that the increase in EVs capacity might significantly increase the transformer aging.

2.5 Customer Preferences

The thesis assumes that customers take into account their energy costs and their convenience in the energy schedule. A class of customers prefers to pay more for their convenience. Authors in [76] claim that none of the presented DR methods considers the customers' point of view in literature. They develop an intelligent system for the residential level to manage energy consumption efficiently. They believe that their system will learn the consumers' behavior. Their method aims to encourage customers to expand renewable energy share and decrease the dependency on nonrenewable sources.

The authors in [77] propose a model to optimize the residential loads' consumption based on the end user's preferences. Their method makes the homeowner execute DR actions automatically. The DR is based on the customer's preferences and utility programs. They claim that their work achieves a reduction in electricity payments and a reduction in PAR. The study in [78] presents the effect of minimizing power consumption on the customer satisfaction level. In the future, this model expects to improve the effect of minimizing the power consumption on customer satisfaction further if the utility provides the customers with more consumption information for the appliances.

2.6 Barriers and Limitations

Despite their beneficial opportunities, DR strategies should meet various constraints to penetrate the market. These constraints are mainly presented in understanding the customers' energy use, economical and technical valuation, and policymaking. The authors in [79] noted some main challenges to implementing the advanced DR programs, namely: the need to understand the terminology from different perspectives, transparent pricing, marketing, awareness by policymakers, and development of the enabling technologies. Ref [80] underlines the solution for such constraints, and the unreliability of customer behavior lies in improving the program design and policies instead of developing new technology.

The unexpected behavior and awareness of the customers are some of the main challenges. The customers have used to the primarily flat, fixed rates. Participating in the DR program and all the groups of customers do not accept other energy market opportunities. Another class of customers still presents a challenge and needs additional incentives for engagement, even with the electricity cost reduction achievement. Hence, full engagement in this program is limited to the technology-supporter class of customers.

One of the main concerns of utilities that the DR program is not a reliable source and can affect the network negatively in some cases. This leads utilities to create DR programs in their isolation, e.g., conventional DR programs and did not seek to understand customers' needs. To get an influential DSM, PSO should start to move away from DR’s conventional models and consider more holistic and effective DR optimization techniques. Moreover, for DR to be effective, it should reflect the market situation. For instance, emergencies or crises, e.g., California in the 2000 – 2001 crisis [81], encouraged people to participate in a DR program.

DR has the potential to provide significant benefits for utility and customers, but it has to meet several constraints. Research on this area raises different questions, and despite the work done, there is still missing work and significant limitations that consolidate the advantages and challenges for integrating a DR platform into an existing electricity market. The efficient deployment of a DR optimization algorithm for utilities and customers' benefit is challenging. This is due to the increased complexity and uncertainty in supply and power demand and intermittency of RESs, changes in users’ behavior, and the dynamic nature of electricity prices. Most of the work presented in the literature concentrated on optimizing the load and reducing consumers' bills considering their comforts. A few of these publications considered utility asset conditions in the presence of RESs, EVs, and ESSs.

This emphasizes the need for developing a HEMS with a DR algorithm that improves the economic performance locally at the end-use level considering utility assets. Thus, this research investigates the combined impact of optimizing energy consumption considering DERs, EVs, and ESSs integration, while satisfying end-users comfort/engagement and backup electric utility assets. The next chapter of this thesis presents the problem formulation and methodology of the performed research.

CHAPTER 3 : MULTI-OBJECTIVE RESIDENTIAL DEMAND RESPONSE FRAMEWORK DESCRIPTION

In this chapter, an overview of the residential DR problem is presented. The model of different residential loads is explained. These models are a fundamental description of the respective appliances, where the dynamics are neglected for simplicity. The overall system description is also provided, and the formulation of a multi-objective DR algorithm is presented along with mathematical programming and code implementation details.

3.1 Defining the Problem of Appliances' Scheduling

The cumulative load for residential appliances depends on how the appliances are operated over time. For instance, if all appliances start simultaneously, the coincident demand could be increased to exceed the PSO's power limits and sometimes adversely affect the home electrical system and utility assets. Therefore, the appliance consumption should be appropriately controlled to keep the peak demand and electricity cost to a minimum without sacrificing user comfort. HEMS is implemented to minimize the electricity price and utilize the usage of the grid's energy. A scheduled plan for the power usages of the different appliances, EV, ESS, and PV, is achieved considering the RTP-price and transformer load.

The utility sends a price signal and power limit to several HEMSs to perform demand response optimization in this context. Involved customers get financial benefits, along with minimized discomfort concerns in the proposed framework. To define the scheduling of home appliance load as to when appliances start and end (called appliance window) in a 24-hour time scale in a day, the following terms are defined:

Set of Appliances - ***A*** represents the set of electrical appliances used in the smart home. Then, A = {Lights (L), Dish Washer (DW), Clothes Dryer (DRY), Refrigerator (REFR), Cooker, Washing Machine (WM), Water Heater (WH), Air Conditioning (AC), EV, PV panel}. In this case, *n* denotes the index of the appliance shown in set A.

Duration of Operation - The day is divided into 24-time slots. Thus, each time slot represents 60 minutes. The appliances can be set to start at any time within this time frame and end their operation cycle before or on the 24th time slot.

Execution Window of each Operation - Appliances should start in a user-specified window (operating time). Therefore, users specify the desired execution window for each appliance, i.e., the interval when the appliance can run. Specifically, for each appliance, there is a minimal starting time (before that time, the operation cannot start) and a maximal ending time (by that time, the operation should be finished). The HEMS can switch the appliances on or off at any time as long as it is in the user's pre-determined starting and ending times. Similarly, the user has to specify the arrival and departure time for the EV.

Defining Problem - The appliance execution period of 24 hours is considered here, divided into $h = 1, 2, 3, \ldots, H$. N denotes the number of appliances for scheduling, $n = 1, 2, 3, \ldots, N$. If we want to define appliance use by name, that is, some load in use of appliance n during the time slot h. Therefore, $E_{n,h}$ represents power consumption assigned to an appliance n during the time slot h. The unit for the power consumption $E_{n,h}$ is kW. The unit changes to kWh (energy) by multiplying it with a factor of time (60/60=1). In addition to $E_{n,h}$, $u_{n,h}$ is a binary decision variable to indicate if a particular appliance is being processed or not. $u_{n,h}$ = 1, if an appliance n and at time slot h is operating, otherwise. $u_{n,h}$= 0. It can be, alternatively, stated that in the time interval of operation of appliances in their respective execution windows, the binary variable $u_{n,h}$= 1 and beyond the window, it is $u_{n,h}$=0.

$$u_{n,h} = \begin{cases} 1\ if\ appliance\ is\ ON \\ 0\ if\ appliance\ is\ OFF \end{cases} \tag{3.1}$$

The optimal load management of the optimization problem is linear. It consists of an objective function defined by different linear and nonlinear constraints. The optimization problem requires the solution of linear equations, describing the optimal and secure operation of the home network. The general load management problem can be presented as follows:

$$\begin{gathered} min = \{f(x)|x \in X\} \\ Subjected\ to: \\ g(x) \geq 0 \\ h(x) = 0 \end{gathered} \tag{3.2}$$

Where;

f(x) presents the optimization objective function.

g(x) presents the inequality constraint for the objective function.

h(x) presents the equality constraint for the objective function.

3.2 Residential Demand Response Framework Description

A smart community has *M* residential buildings served by a transformer, as shown in Figure 3-1. According to practical needs, the future residential building will be equipped with smart meters, RESs ESSs, HEMSs, and several electrical appliances. A smart meter is used to measure and transfer electricity consumption from utility to household and vice versa. The HEMS acts as the brain of the whole system where the proposed DR scheme is embedded in. It oversees the entire facility's energy and data flow and manages home appliances' power consumption considering a pre-specified set of constraints and requirements. Moreover, the existing EV and ESS provide two-way energy trading, enhancing flexibility and economic benefits.

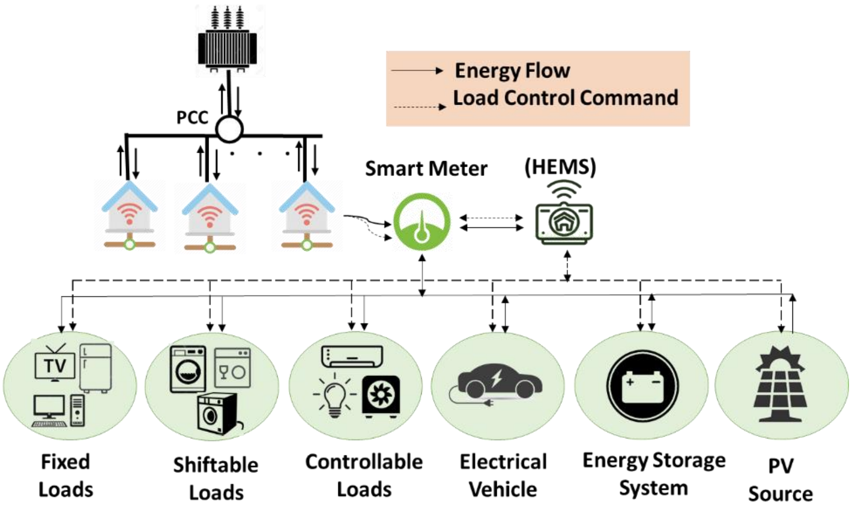

Figure 3-1 The proposed HEMS DR model and its considerations

In this context, the PSO presents different incentives to the users to optimize their consumption to alleviate the transformer LoL cost. The user cannot perform DR actions without a utility operator contract. According to the DR contract, the electricity bill is made from the following pricing charge:

1. Off-peak/usage charge: In this case, the utility announces the electricity tariff (λ_h) for the next 24 hours. This price can be fixed or time-varying, depending on the electricity market. Usually, the value of λ_h equals to several cents per kWh. This thesis assumes the RTP charge at off-peak times.
2. Peak demand charge: This charge is applied throughout the customer's peak hour. This price is used to encourage the residential user to change their power uniformly. The utility can profit from this since it reduces the capacity provision investments. Here, λ_h is much higher than in the off-peak charge. It can reach several dollars per kWh.
3. Coincident peak demand charge: This charging scheme is the same as the peak

charge scheme but manages the peak hour for the utility instead of the peak for each consumer. This is done by determining the utility's peak hour at the end of each day/month. After that, all users pay for their electricity consumption at this time. The rate of these charges can reach a price that is more than the peak demand charge rate.

This research considers a net-metering technique where the price of purchasing/selling energy from/to the grid is the same. However, the energy sold to the grid can be charged with any other pricing scheme. The customers can achieve significant cost reduction by knowing the market peak hours minimize their electricity consumption at that hour. The proposed DR program's schematic diagram is presented in Figure 3-2. Firstly, the forecasted hourly PV output and electricity price are received. The optimization problem is then formulated based on PV generation, asset degradation costs, electricity price, the initial state of energy (SoE) of EVs, and ESSs’ batteries, together with the user preferences data. Each of these elements is explained in detail in the below subsections.

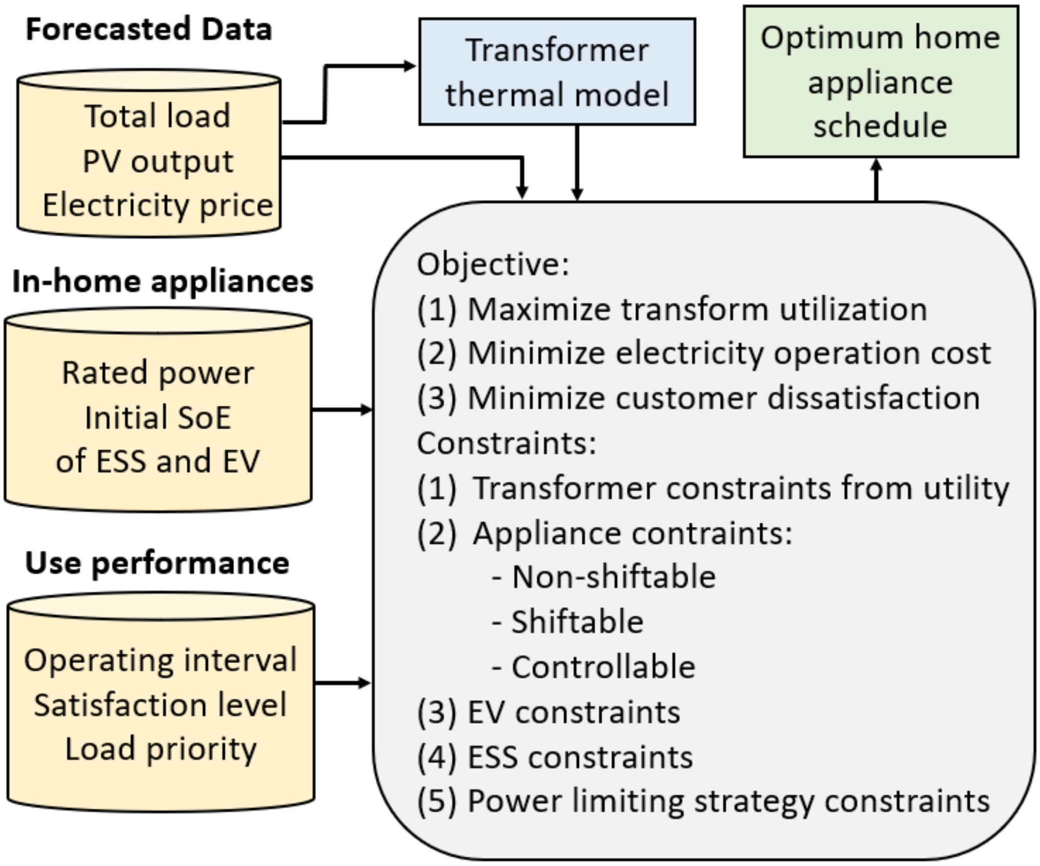

Figure 3-2 The multiobjective HEMS Demand Response structure

3.3 Thermal Model of a Distribution Transformer

Transformer thermal modeling is crucial for PSO to examine the operating condition for distribution transformers. This subsection presents the details for modeling the distribution transformer’s thermal characteristics according to the IEEE standard C57.91-2011. As mentioned in the previous chapter, the main factor for transformer aging is insulation degradation. Degradation of the insulation is sensitive to the transformer's thermal condition. The transformer thermal condition is one of the main parameters used to calculate a distribution transformer's LoL. Heat affects the thermal degradation of the power transformers. Therefore, controlling a transformer's operating temperature is one of the fundamental techniques of extending its life. The transformer thermal model can provide utilities with an overview of the transformer operating condition to overcome potential problems and extend transformer life. Therefore, it is essential to have an accurate model to present the thermal behavior of the transformer.

There are different models utilized to determine the thermal behavior of the transformers. As found in the literature, IEEE standard C57.91 is considered one of the popular models used to describe the transformer's thermal behavior. This model considers the transformer's load as the transformer's overall temperature increases as the transformer's loading current increases [67]. Also, it considers different thermal parameters such as average winding temperature, top oil rise, and bottom oil rise for evaluating the loss of life.

3.3.1 Hot Spot Temperature (HST) Calculation

The transformer winding HST model is used to calculate the %LoL of the distribution transformer, and it consists of three variables: ambient temperature (θ_A), winding hottest-spot temperature ($\Delta\theta_H$), and top oil temperature (θ_{TO}), as given by (3.3).

$$\theta_H = \theta_A + \Delta\theta_H + \theta_{TO} \tag{3.3}$$

Where $\Delta\theta_H$ calculates the increase of the winding temperature and oil temperature caused by the increase of transformer current as given by (3.4). The initial value of top oil temperature rises over ambient is given by (3.5), and the ultimate top oil temperature rise is given by (3.6).

$$\Delta\theta_H = \left(\Delta\theta_{H,u} - \Delta\theta_{H,i}\right)\left(1 - exp^{-\frac{t}{\tau H}}\right) + \Delta\theta_{H,i} \tag{3.4}$$

$$\Delta\theta_{H,u} = \Delta\theta_{H,R}[k_u]^{2m} \tag{3.5}$$

$$\Delta\theta_{H,i} = \Delta\theta_{H,R}[k_i]^{2m} \tag{3.6}$$

Equation (3.7) presents the top oil rise over ambient temperature. It can be seen from (3.7) that any increase in the transformer current will increase the overall transformer temperature. The ultimate and initial top oil rise temperatures are given by (3.8) and (3.9), respectively, where n is a factor that depends on the type of cooling of the transformer.

$$\Delta\theta_{TO} = \left(\Delta\theta_{TO,u} - \Delta\theta_{TO,i}\right)\left(1 - exp^{-\frac{t}{\tau TO}}\right) + \Delta\theta_{TO,i} \tag{3.7}$$

$$\Delta\theta_{TO,u} = \Delta\theta_{TO,R}[\frac{{k_u}^2 \cdot R + 1}{R + 1}]^n \tag{3.8}$$

$$\Delta\theta_{TO,i} = \Delta\theta_{TO,R}[\frac{{k_i}^2 \cdot R + 1}{R + 1}]^n \tag{3.9}$$

3.3.2 Transformer Aging Calculation

This model is based on the effect of the HST on the transformer aging. The relative relation of the aging acceleration factor and transformer's HST (reference temperature 110° C) is expressed in (3.10).

$$F_{AA} = \exp\left(\frac{1500}{383} - \frac{1500}{\theta_H + 273}\right) \tag{3.10}$$

Equation 3.11 calculates the equivalent aging factor at the reference temperature used in the transformer's LoL calculations.

$$F_{EQA} = \frac{\sum_{r=1}^{N} F_{AA,r}\Delta t_r}{\Delta t_r} \tag{3.11}$$

Where Δt_r is the given time interval, N is the number of the time slots in the time interval, and r is the index of the time interval. Equation (3.12) is used to calculate the percent LoL. The normal insulation life is considered 180000 hours (20.55 years) [42].

$$\%LOL = \frac{F_{EQA} \times t \times 100}{Noral\ Insulation\ life} \tag{3.12}$$

3.4 Home Appliance Model

The appliances' operation and user preferences limit their scheduling. The mathematical model for each category of appliances and their constraints is explained in the following subsections.

3.4.1 Fixed Appliances (Passive Loads)

Fixed loads have critical operating status, which is always "on" and cannot be scheduled or controlled, such as refrigerators and alarm systems. The cost of such loads is presented by (3.13).

$$C_{n,h} = \lambda_h \cdot E_{n,h} \tag{3.13}$$

3.4.2 Shiftable Appliances (Active Loads)

Shiftable loads, i.e., washing machine, dryer, and dishwasher, can be scheduled for off-peak periods where the electricity tariff is lower. The consumption cost for these appliances is represented by (3.14). Also, the limits for the shiftable appliance n during the scheduling window should meet the constraints in (3.15) and (3.16). For example, assuming $[T_{n,int}, T_{n,end}] \in H$ is the desired operating interval in which the shiftable appliance n is expected to start the operation. It implies that this appliance should start any time after $T_{n,int}$And should complete its operation before $T_{n,end}$.

$$C_{n,h} = \lambda_h \cdot u_{n,h} \cdot E_{n,h} \tag{3.14}$$

$$u_{n,h} = 0 \quad \forall h \in \left[1, T_{n,int}\right) \cup (T_{n,end}, H] \tag{3.15}$$

$$u_{n,h} \leq 1 \quad \forall h \in \left[T_{n,int}, T_{n,end}\right] \tag{3.16}$$

To reflect the customer dissatisfaction cost in the scheduling program, equation (3.17) is presented, reflecting the customer's dissatisfaction with waiting for the appliance to start the operation. Minimizing (3.17) leads to minimizing the waiting time as possible, which supports customer comfort. For example, if the DW usually pm-operate during the period [5 pm-10 pm], operation time is shifted to other time slots (i.e., 7 pm). According to this, the waiting time equals to $T_{n,start} - T_{n,int}$ (i.e., 2 h). Note that

$(T_{n,start} - T_{n,end})$ should not less than the operating duration time, $T_{n,end}$, as presented in (3.18).

$$cdc_{n,h} = \zeta_n\left(T_{n,start} - T_{n,int}\right) \tag{3.17}$$

$$T_{n,int} \leq T_{n,\text{start}} \leq (T_{n,end} - T_{n,total}) \tag{3.18}$$

3.4.3 Controlled Appliances (Active Loads)

This category includes Water Heater (WH), heating, ventilation, air conditioning (HVAC), and Light (L). This appliance has flexible power consumption, such as lights and air conditioners. Their power consumption can be regulated between the maximum and minimum in response to price changes, as presented in (3.19). Controlling these loads support minimizing the customer electricity bill. Hence, the electricity operation cost of a controllable appliance *n* is given by (3.20). Nonetheless, power reduction can cause dissatisfaction for the customer, as presented by (3.21).

$$e_{n,min} \leq E_{n,h} \leq e_{n,max} \tag{3.19}$$

$$C_{n,h} = \lambda_h \cdot E_{n,h} \tag{3.20}$$

$$cdc_{n,h} = \zeta_n(E_{n,h} - e_{n,max})^2 \tag{3.21}$$

3.5 Electrical Vehicle Constraints

The EV model supports the engagement of customers in the energy market. Charging and discharging of the EV battery is controlled based on the electricity price and the asset condition. The EV charging/discharging cost is presented by (3.22). Equation (3.23) presents the difference between the EV's maximum energy and the actual charged energy, multiplied by a dissatisfaction factor that penalizes not having a fully charged EV at departure time. Constraint (3.24) characterizes the power used from discharging the EV battery (V2H mood or V2G mood). The amount of energy available in the EV battery at time h is presented by (3.25) and (3.26). The EV battery's stored energy is limited between the minimum and maximum values to prevent deep discharging or full charging, as presented by (3.27).

$$C_{n,h} = \begin{cases} \lambda_h \cdot E_{n,h}^{EV/c} \\ -\lambda_h \cdot E_{n,h}^{EV/d} \end{cases} \tag{3.22}$$

$$cdc_{n,h} = \zeta_n (SOE_{n,h}^{EV} - SOE_n^{EV/max})^2 \; if \; h = h_n^{dep} \tag{3.23}$$

$$E_{n,h}^{EV/used} + E_{n,h}^{EV/sold} = \eta^{EV/d} \cdot E_{n,h}^{EV/d} \; \forall h \in [h_n^{arr}, h_n^{dep}] \tag{3.24}$$

$$SOE_{n,h}^{EV} = SOE_n^{EV/int} + \eta^{EV/c} \cdot E_{n,h}^{EV/c} - E_{n,h}^{EV/d} \; , if \; h = h_n^{arr} \tag{3.25}$$

$$SOE_{n,h}^{EV} = SOE_{t-1}^{EV} + \eta^{EV/c} \cdot E_{n,h}^{EV/c} - E_{n,h}^{EV/d} \; \forall h \in (h_n^{arr}, h_n^{dep}] \tag{3.26}$$

$$SOE_n^{EV/min} \leq SOE_{n,h}^{EV} \leq SOE_n^{EV/max} \tag{3.27}$$

3.6 Energy Storage System Constraints

ESS is modeled similarly to the EV, as presented by (3.28) -(3.33). However, according to the DR program, the ESS is available all day at the house to be utilized (charging/discharging).

$$C_{n,h} = \begin{cases} \lambda_h \cdot E_{n,h}^{ESS/c} \\ -\lambda_h \cdot E_{n,h}^{ESS/d} \end{cases} \tag{3.28}$$

$$cdc_{n,h} = \begin{cases} \zeta_n \left(SOE_{n,h}^{ESS} - SOE_n^{ESS/max}\right)^2 \; if \;\; SOE_{n,h} > SOE^{ESS/max} \, , \\ \zeta_n \left(SOE_{n,h}^{ESS} - SOE_n^{ESS/min}\right)^2 \;\; if \;\; SOE_{nh} < SOE^{ESS/min} , \\ 0 \qquad\qquad otheiwise, \end{cases} \tag{3.29}$$

$$P_{n,h}^{ESS/used} + E_{n,h}^{ESS/sold} = \eta^{ESS/d} E_{n,h}^{EV/d} \tag{3.30}$$

$$SOE_{n,h}^{ESS} = SOE_{t-1}^{ESS} + \eta^{ESS/c} \cdot E_{n,h}^{ESS/c} - E_{n,h}^{ES/d} \; \forall h > 1 \tag{3.31}$$

$$SOE_{n,h}^{ESS} = SOE_n^{ESS/int} , if \; h = 1 \tag{3.32}$$

$$SOE_n^{ESS/min} \leq SOE_{n,h}^{ESS} \leq SOE_n^{ESS/max} \tag{3.33}$$

3.7 PV Model Constraints

PV energy resources further support customer engagement in the energy market. Equation (3.34) implies that the generated power from PV can be used by the household appliances or injected back to the grid based on the DR contract, utility, and prosumer.

$$P_h^{PV} = P_h^{PV/used} + P_h^{PV/sold} \tag{3.34}$$

3.8 Power-Limiting Strategies

The load profile's smoothness is one of the main criteria that should be considered in DR strategies. In addition to the transformer thermal model, utilities can provide several incentive power-limiting strategies to the households according to the DR program they participate in. This will contribute to the demand peak reduction, which brings benefits for the whole power system. The incentives for the end-users are direct since they pay less by following a particular strategy. For instance, during coincident peak periods, the utility can impose limits on the users' power to further reduce the power system's peak. A power-limit method may also control the customer's daily power consumption pattern, as presented in (3.35) and (3.36). However, If the user wants to consume power above the limit, the excess energy will be charged by a higher charge, e.g., coincident peak charge, so a significant cost reduction can be realized if the user reduces its demand peak hours.

$$max_h(P_h^g - P_h^{sold}) \leq \Gamma \tag{3.35}$$

$$P_h^g \leq l_h.u_h \tag{3.36}$$

3.9 Weighted Objective Function

The proposed DR framework described in previous subsections can be formulated as an optimization problem with three main objectives and constraints. The multi-objective function is presented as a cost model that consists of electricity charge cost considering transformer LoL (F_1) and customer dissatisfaction cost (F_2), weighted and combined into a single objective function, as shown in (3.37).

$$min\ \ \rho\, F_1 + (1-\rho)\, F_2 \quad \rho\ \in [0, 1] \tag{3.37}$$

This integration allows the customers and operators to decide the combinations that better fit their economic interests and meet their technical requirements. This can be done by adjusting the customer/operator balance parameter ρ to achieve the trade-off between F_1 and F_2. These sub-objective functions are presented as:

$$F_1 = \sum_{n=1}^{N} \sum_{h=1}^{H} C_{n,h} + C_h^{Tx} \tag{3.38}$$

$$F_2 = \sum_{n=1}^{N} \sum_{h=1}^{H} cdc_{n,h} \tag{3.39}$$

Electricity charge Cost ($C_{n,h}$) - $C_{n,h}$ presents the electricity charge cost. If the value of F_2 is positive, the consumed power is from the grid. Otherwise, If the value of $C_{n,h}$ is positive, the consumed power is from the grid. Otherwise, if the value of $C_{n,h}$ is negative, the consumed power is from after-the-meter-generated power based on DERs and BESSs. $C_{n,h}$ is subjected to different appliance, EV, and ESS constraints (Eq.(3.22)-(3.34)).

$$C_{n,h} = \sum_{n=1}^{N} \sum_{h=1}^{H} C_{n,h}{}^{non} + C_{n,h}{}^{shift} + C_{n,h}{}^{Con} + C_{n,h}{}^{EV} + C_{n,h}{}^{ESS} \tag{3.40}$$

Transformer LoL Mitigation Cost (C_h^{Tx}) - C_h^{Tx} presents the transformer LoL cost, which is directly related to the asset condition. Based on the utility's received electricity price, the DR program intends to allocate as much load as possible at the low-price periods. However, high penetration of scheduled household loads during these periods raises a concern of new high-power peaks that leads to distribution transformer overloading with high LoL. Hence C_h^{Tx} is used to reflect the transformer LoL degradation cost in the DR optimization program. This degradation cost at a particular time depends on the transformer load, ambient temperature, and the transformer load at that time. The transformer capacity total cost is utilized instead of a fixed threshold constraint, and the LoL cost is estimated using (3.41) [53].

$$C_h^{Tx} = \frac{Transformer\ invesment\ cost}{Normal\ life\ expectancy} \times LOL_h\% \tag{3.41}$$

Customer Dissatisfaction Cost ($\boldsymbol{F_2}$) - F_2 is introduced to capture the degree of discomfort caused to the customer due to the DR schedules. For example, when a shiftable appliance is scheduled to operate at different times than the user's initially decided time, it inconveniences the customers. The same thing could happen when controllable appliances operate with less power. Asa result, the customers may stop participating in the DR programs. Moreover, to further enhance customer comfort, an appliance importance parameter is introduced to give the customer more control of the appliance operation.

$$F_3 = \sum_{n=1}^{N} \sum_{h=1}^{H} {cdc_{n,h}}^{non} + {cdc_{n,h}}^{shift} + {cdc_{n,h}}^{Con} + {cdc_{n,h}}^{EV} + {cdc_{n,h}}^{ESS} \tag{3.42}$$

CHAPTER 4 : HOME ENERGY MANAGEMENT BASED ON CONVENTIONAL OPTIMIZATION

This chapter explains the implementation of HEMS based on conventional optimization techniques. Section 1 explains the used optimization approaches along with their implementation on MATLAB. Section 2 demonstrates the implementation of the algorithm code in MATLAB. The Assumptions and considerations made during the simulation are provided in section 3. Section 4 demonstrates the household appliances and transformer input data along and the assumptions made. Different studied scenarios and guides for energy cost calculation for the proposed DR algorithm are presented in 5.

4.1 Optimization Approach

In this thesis, MILP and interior-point optimization (IPO) are used efficiently to solve the multi-objective function. First, the MILP is utilized to optimize load profiles for shiftable appliances. Then, the IPO is utilized to optimize the load profile of controllable appliances. To solve the energy management problem optimally, there is a need for a mathematical programming language and a solver. There are many mathematical programming languages such as AMPL, TOMLAB, and MATLAB (A Toolbox for optimization). MATLAB Optimization Toolbox solves linear programming (LP), quadratic programming (QP), nonlinear programming (NLP), MILP, and nonlinear equations [82].

The solvers such as GUROBI and IBM CPLEX are state-of-the-art solvers for LP, QP, MILP, and mixed-integer quadratic programming (MIQP). GUROBI/CPLEX can solve MILP problems with many binary variables of a reasonably large size. The main reason for using MATLAB Optimization Toolbox is its rapid algorithm development, which provides faster and accurate solutions for a wide range of objective functions and constraints.

4.1.1 MILP Optimization Algorithm

MILP is used to generate the shiftable appliance operating schedules. In our case, the time scale is divided into 24 slots. Each time slot represents a 1-hour duration. The appliances can operate at any starting and ending time scale, starting from 1 to 24. Depending on the appliance operating time's length, the number of time slots available for starting the appliance is given by (3.18). For example, if we have a shiftable appliance with the following parameter:

Table 4-1 Shiftable appliance parameters

Parameter	Value
n	1
$E_{n,h}$	1.5 kW
$T_{n,int}$	1 h
$T_{n,end}$	24 h
$T_{n,total}$	2 h

If the scheduling window starts at hour 1 and ends at hour 24, then by using equation (3.18), we have 22-time slots available for starting the appliance. Therefore, 22 binary variables are required for the appliance assignment at 22 different starting time slots. In this case, we consider that the electricity price fluctuates every hour (λ_h). Optimization scheduling requires the use of the MILP solver and the addition of shiftable scheduling constraints. The optimization function is given by a 24x1 matrix, as seen in equation

(4.1).

$$C_{n,h} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{24} \end{bmatrix} * [E_{1,1} \quad E_{1,2} \quad \cdots \quad E_{1,24}] = \begin{bmatrix} \lambda_1 \cdot E_{1,1} \\ \lambda_2 \cdot E_{1,24} \\ \vdots \\ \lambda_{24} \cdot E_{1,24} \end{bmatrix} \tag{4.1}$$

This equation is subjected to two different constraints. These countries as classified as power and time constraints. First, the preferred operating time of the appliance is presented by (3.15-3.18). Second, if there is any power limit imposed by the utility, presented in (3.36), it needs to be considered. Other constraints linked to customer preferences, safety, and power requirements can also be considered.

To implement the appliance model in MATLAB, the function "*intlinprog*" is used, a MILP solver. Figure 4-1 presents a schematic diagram for the MILP method, which uses the following basic strategy to solve the problem:

1. The problem size is reduced using Linear Program Pre-processing.
2. LP is utilized to solve the problem initial relaxed (non-integer).
3. Pre-processing is performed for Mixed-Integer Program.
4. Searching for integer-feasible solutions using Branch and Bound algorithm.

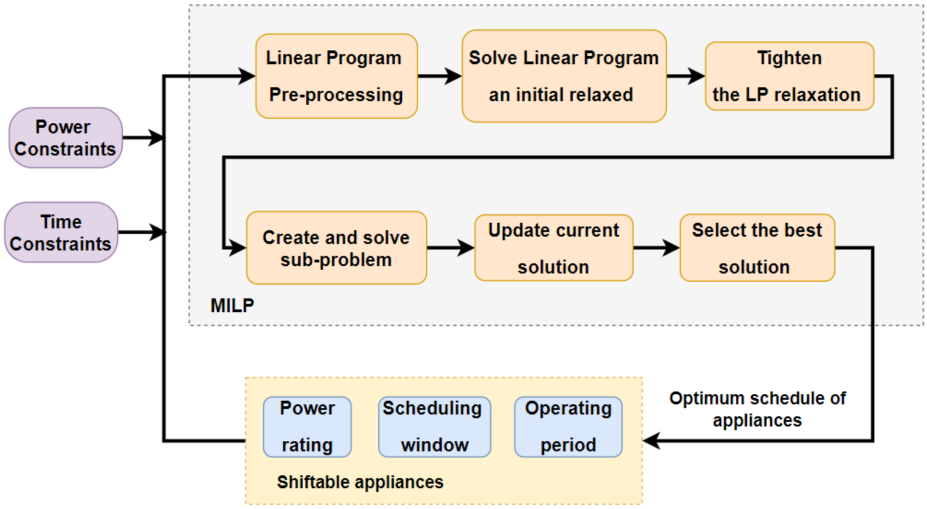

Figure 4-1 Schematic diagram of the MILP method for shiftable appliances

The syntax for "*intlinprog*" in MATLAB is given as in (4.2) and (4.3) [82].

$$\min f(x) \text{ such that } \begin{cases} x(intcon) \\ A \cdot x < b \\ Aeq \cdot x = beq \\ lb \leq x \leq ub \end{cases} \tag{4.2}$$

$$\text{X=intlinprog (f,intcon, A,b, Aeq,beq,lb,ub)} \tag{4.3}$$

Where

- ***X*** represents the vector of variables (to be determined).
- ***f*** represents the coefficient vector.
- ***intcon*** refers to the vector of integer constraints.
- ***A*** refers to the vector of linear inequality constraints.
- ***b*** refers to the vector of linear inequality constraints.
- ***Aeq*** refers to the vector of equality constraint.
- ***beq*** refers to the vector of linear equality constraint.
- ***lb*** refers to the lower bounds.
- ***ub*** refers to the upper bounds.

In this case, f is equal to the electricity cost vector; A is equal to the appliance power profile, b is equal to the required operating period. Since there are no equality constraints, Aeq and beq are set to [$\emptyset$]. Table 4-2 shows the required arguments and their values of the "*intlinprog"* function in MATLAB.

Table 4-2 Arguments of *"intlinprog"* in MATLAB

Arguments	value
f	$[\lambda_1 \quad \lambda_2 \quad \dots \quad \lambda_{24}]^T$
A	$[E_{1,1} \quad E_{1,2} \quad \dots \quad E_{1,24}]$
b	$T_{n,total}$
Aeq	$\emptyset$
beq	$\emptyset$
lb	0
ub	∞

The above constraints instruct the optimization solver to select an operating time slot out of available time slots to obtain the optimum appliance schedule. As explained above, the same methodology is applied to determine the operating time for all the shiftable appliances separately. For the PV panel and EV/ESS discharging, variable "A" will be negative as it is not load consuming appliance but a local source of power.

4.1.2 Interior Point Optimization Algorithm

The IPO is a gradient method used to solving large-scale nonlinear convex multi-variable functions. It is known for its simplicity in mathematical modeling. It starts with an initial guess and iterates based on a given scheme. The iterations stop when certain constraints are reached [83]. To avoid the violation of constraints, the objective function is augmented by a barrier term. At each iteration, there are two main types of steps, which are used to solve the optimization problem: a direct step and a conjugate gradient step. By default, the IPM takes the direct step first. If it is not applicable, it takes the conjugate gradient step. When the direct step cannot be used, the approximate problem is not locally convex near the current iterate[83]. A flowchart illustrating the IPM algorithm employed is given in Figure 4-2.

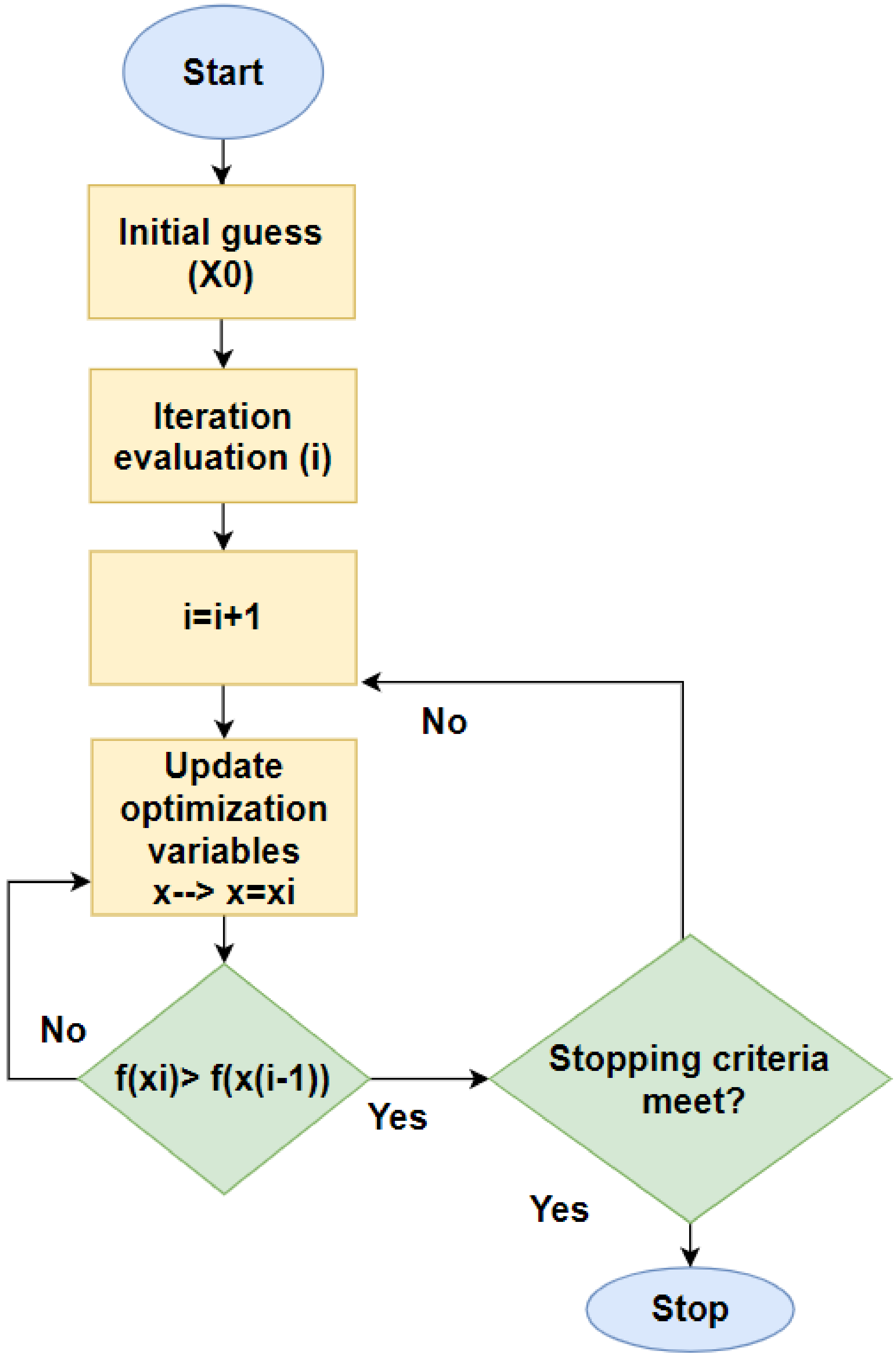

Figure 4-2 Flowchart of IPO Algorithm.

In this work, IPO is used to optimize the controllable appliance load profile. The controllable appliances' power consumption is controlled for 24 hours, starting from 1 to 24. Depending on the electricity price, appliance importance parameter, the appliance operation is regulated between maximum and minimum power level, as presented by (3.18). For example, if we have a controllable appliance with the following parameter:

Table 4-3 Controllable appliance parameters

Parameter	Value
n	2
$e_{n,min}$	0.8 kW
$e_{n,\max}$	2 h
$T_{n,int}$	1 h
$T_{n,end}$	24 h
$T_{n,total}$	24 h

First, the IPO starts with an initial guess X_0 for the optimization variable X. The algorithm attempts to find a minimum value of X described in the objective function. In this example, X is the appliance power level, and it is bounded between $e_{n,min}$ (0.8 kW) and $e_{n,max}$ (2 kW) . In order to implement the appliance model in MATLAB, the function "*fmincon*" is used. The syntax for the "*fmincon*" in MATLAB is given as in (4.4) and (4.5) [83].

$$\min f(x)\ shuch\ that \begin{cases} c(x) \leq 0 \\ Ceq(x) = 0 \\ A \cdot x \leq b \\ Aeq \cdot x = beq \\ lb \leq x \leq ub \end{cases} \tag{4.4}$$

$$X = \text{fmincon}\,(fun, X_0, \text{intcon}, A, b, Aeq, beq, lb, ub, nonlcon) \tag{4.5}$$

Where

- ***X*** represents the vector of variables (to be determined)
- ***fun*** represents the objective function.
- ***X₀*** refers to the initial guess for the objective function.
- ***Nonlcon*** refers to nonlinear inequalities and equalities constraints.
- ***A*** refers to the vector of linear inequality constraints.
- ***b*** refers to the vector of linear inequality constraints.
- ***Aeq*** refers to the vector of equality constraint.
- ***beq*** refers to the vector of linear equality constraint.
- ***lb*** refers to the lower bounds.
- ***ub*** refers to the upper bounds.

For this example, *fun* is equal to the objective function of electricity cost and customer dissatisfaction cost, presented by (4.6). Constraint (4.7) instructs the solver to select the operating power level between *lb* and *ub* to obtain the appliance's optimum operation.

$$C_{n,h} = \lambda_h \cdot E_{n,h} + \zeta_n (E_{n,h} - e_{n,max})^2 \tag{4.6}$$

$$0.8 \leq E_{n,h} \leq 2 \tag{4.7}$$

4.2 Algorithm Implementation

The flowchart for the proposed HEMS algorithm is presented in Figure 4-3, which shows the proposed optimization problem's implementation process as described perversely. For 24 h time slots, the date for electricity price, PV output, and transformer load are received. The transformer thermal model receives the predicted transformer load, and the transformer load level is considered. After that, loss of life is calculated according to its load condition. Then, optimization algorithms are utilized to perform

DR decisions for different appliances, ESS, and EV. Finally, the algorithm checks that the power limits imposed from the utility are not violated.

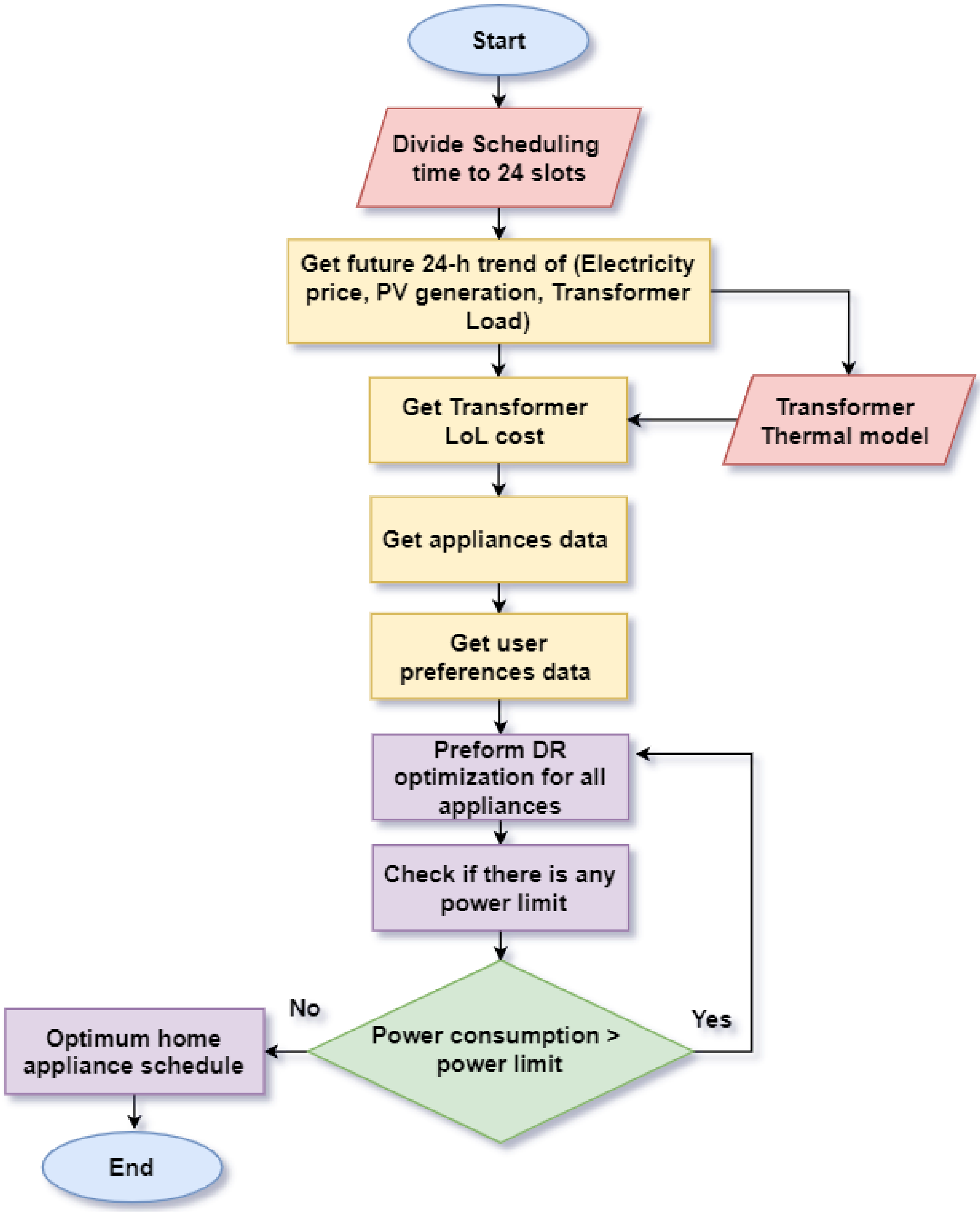

Figure 4-3 Flow chart for the proposed DR algorithm

4.3 Assumptions and Considerations

To execute the formulated DR optimization problem, several considerations and assumptions are made. These assumptions and considerations come under two main categories: (i) predictions of the model input data and (ii) executing the scheduled load proposed by the HEMS. Each of these points is examined by prior literature, as highlighted in the following paragraphs.

The user has a DR contract with the utility. This DR-contract is categorized by different periods: off-peak, demand peak, and coincident peak. This work assumes that the hourly energy cost (λ_h) is known for each end-user. This price is not the accurate price that the user will pay, but it indicates the next day's expected prices. Utility announces the actual price in real-time, a common procedure in real-life [55]. Besides, to aid customers, the utility should have the Peak Alert-DR program in which they issue alerts notifying participants of potential peak hours each day/month. This program simply provides forecasted alerts. In practice, prices and coincident peak predictions can be made using historical data that is generally available by the utilities operating demand response programs. Other parameters needed by the algorithm are considered perfectly known and can be fairly predictable in practice. For example, the algorithm needs the predicted renewable generation and transformer load as input. The input data prediction can be made in many ways, e.g., [56], [57]. In practice, the household owner may set the appliances' information according to their needs, or an intelligent system that learns the household owner’s behavior could be adapted.

Given the predictions for the coincident peak, transformer load, electricity price, renewable generation, etc., the proposed algorithms make ideal energy scheduling decisions to fulfill the optimization problem objectives since the optimization problem is convex. In a simple form, it can be solved efficiently. For simplicity, the ESS degradation cost is neglected. The remaining work is implementing and consolidating the HEMS on the household level by giving the energy schedule.

4.4 Description of Major Household Appliances Parameters

In this thesis, for load profiles of appliances, a mid-size home is considered with the main electrical appliances. The home also has a photovoltaic (PV) panel for electricity generation. Each appliance has a specific operating time to complete its cycles. Also, it has different power consumption levels according to its operation cycle. The exact power for each appliance can be determined by measuring their demand experimentally for one operational cycle. In this thesis, it is considered that appliances are working on rated power during their operation.

Appliances with major contribution in terms of energy consumptions such as dishwasher, cooker, refrigerator, plugs, washing machine, clothes dryer, water heater, lights, air conditioners, and electric vehicle, are considered in our model to study their DR and optimize their operation over a period of time to optimize the total energy cost and the load profile.

Other electrical appliances such as electric kettles, laptops, microwaves, etc. considered non-shiftable appliances (fixed). Loads of these appliances are small compared to the major load discussed in this thesis. Also, these appliances are interactive and depend on users. Thus, they have little scheduling flexibility. Therefore, it is considered as “other” loads.

Some attributes are considered for the appliances, for example, are as follows:

- ID number.
- Scheduling window.
- Importance parameter.
- Power rating.

- Operating time duration.

Many devices have multiple switching time constants, such as $T_{n,int}$, $T_{n,end}$, $T_{n,start}$ and $T_{n,total}$.

$T_{n,int}$: It is the initial time of the scheduling window.

$T_{n,end}$: It is the end time of the scheduling window.

$T_{n,start}$: It is the scheduled starting time by the DR algorithm.

$T_{n,total}$: It is the total required time for the appliance to stays in on state.

Table 4-4 presents different kinds of Fixed, shiftable and controllable appliances with different attributes.

Table 4-4 Parameters of Household Appliances

ID	Importance Parameter (ζ_n)	Power rating (kWh)	scheduling window $[T_{n,int}, T_{n,end}]$	Operating time ($T_{n,total}$)	Type
Cooker	-	1.5	-	-	Fixed
Plugs	-	1	-	-	Fixed
REFR	-	0.75	-	-	Fixed
other	-	2	-	-	Fixed
WM	0	1.5	17-22	3	shiftable
DW	0	1.2	7-12	2	shiftable
DRY	0	2	20-24	2	shiftable
WH	2	0.6-1	6-9, 20-22	-	Controllable
AC1	2	0.8-2	0-24	-	Controllable
AC2	2.5	0.8-2	0-24	-	Controllable
AC3	3	0.8-2	0-24	-	Controllable
L	0.3-2	0.2-0.8	6-12	-	Controllable

Also, Chevy Volt electric vehicle is considered. The EV batteries are charged through the home-electricity socket. Therefore, it is named a plug-in EV. The EV's maximum charging rate is limited to 3.3 kW. The EV takes 4 hours to charge at a maximum charging rate of 3.3 kW fully. EV is considered a shiftable load to charge when the electricity price is the lowest. With high electricity prices, EV is considered a source of energy to supply the household load. In this case, the load is considered negative as it is not a consuming load but a local power source. The charging efficiency is considered 90%. It.

The home is also equipped with ESS, which charges from the PV source and discharges during high price periods. The ESS capacity can be varied between 0 and 6 kWh according to the user needs. The ESS is controlled similarly to EV. However, the ESS is available all day at the house to be utilized. The EV and ESS parameters are shown in Table 4-5.

Table 4-5 ESS and EV data of each household

Type	ESS	EV
Maximum power accumulated in the battery (kWh)	3	16
Maximum energy of Charging/Discharging (kWh)	0.6	3.3
Minimum discharging Level (%)	40	30
Maximum charging Level (%)	90	90
Initial SOE (%)	90%	50%
Arrival time	-	2pm
Departure time	-	6am

For this thesis's purpose, the PV source is designed to meet about 10% of home demand for 24 h. The PV output power is utilized when the PV generation is greater than appliance consumption. When electricity produced by PV is less than the home's demand, the HEMS will consume power from the grid. The utilized power profile of the PV panel is presented in Figure 4-4.

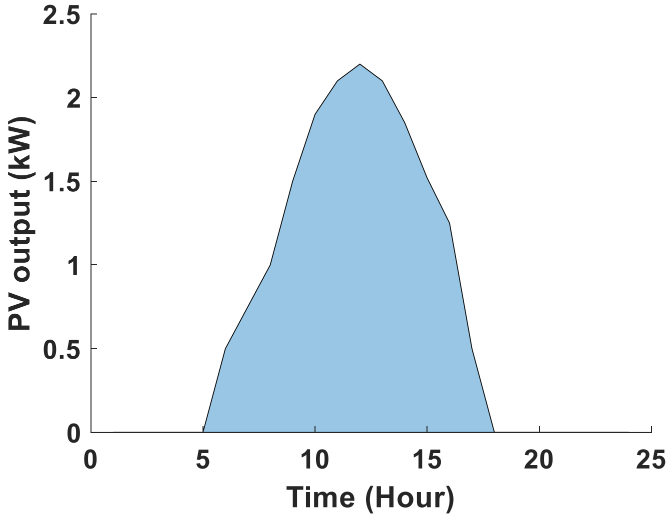

Figure 4-4 PV panel output

4.5 Transformer Parameters

A 30 KVA, 11/0.433 kV ideal distribution transformer is considered [81]. The transformer parameters are presented in Table 4-6. Parameters were obtained from a transformer manufacture specification sheet for distribution transformers [81]. Additionally, according to the IEEE standard C57.91-2011, the ideal distribution transformer's normal lifetime is 20.6 years, which equals 180000 hours approximately.

Table 4-6 Transformer parameters

Item	Value	Item	Value
System Voltage (max.)	12 kV	Average daily ambient air temperature (°C)	40
Rated Voltage HV	11 kV	Rated TOT rise over ambient temperature (°C)	35
Rated Voltage LV (v)	433-250	Rated HST rise over ambient temperature (°C)	40

Item	Value	Item	Value
Line current HV (A)	1.57	Exponent n	0.8
Line current LV (A)	40	Exponent m	0.8
Over fluxing limit	12.5%	Total loss at rated (W)	695
Max. ambient air temp (°C)	50	Ratio of load to no-load loss	8
Min. ambient air temp (°C)	-5	Top oil time constant	24

4.6 Simulated Cases

The proposed DR code is run to optimize appliances consumptions for minimizing the energy cost (lower utility bill), customer dissatisfaction cost, and transformer LoL cost based on equations (3.38), (3.39), and (3.40) for the following (see Figure 4-5):

- **Case (0)**: is considered a reference case in the study, without the DR algorithm.
- **Case (1):** This case minimizes the total daily operation cost considering the CDC (F_3) and ignoring the transformer LoL cost (F_1). Also, different appliances' important parameters are considered in the simulation.
- **Case (2):** This case shows the effect of considering transformer LoL cost (F_1).
- **Case (3):** This case emphasizes the benefits of the DR algorithm by utilizing assets such as PVs, ESSs, and EVs.
- **Case (4):** Finally, a performance evaluation is conducted to show the proposed algorithm's effectiveness in minimizing the optimization objectives.

Reducing the cost of electricity bills is one of the main objectives of the optimization problem. Therefore, the electricity operation cost will be studied and calculated for the five different scheduling scenarios shown above.

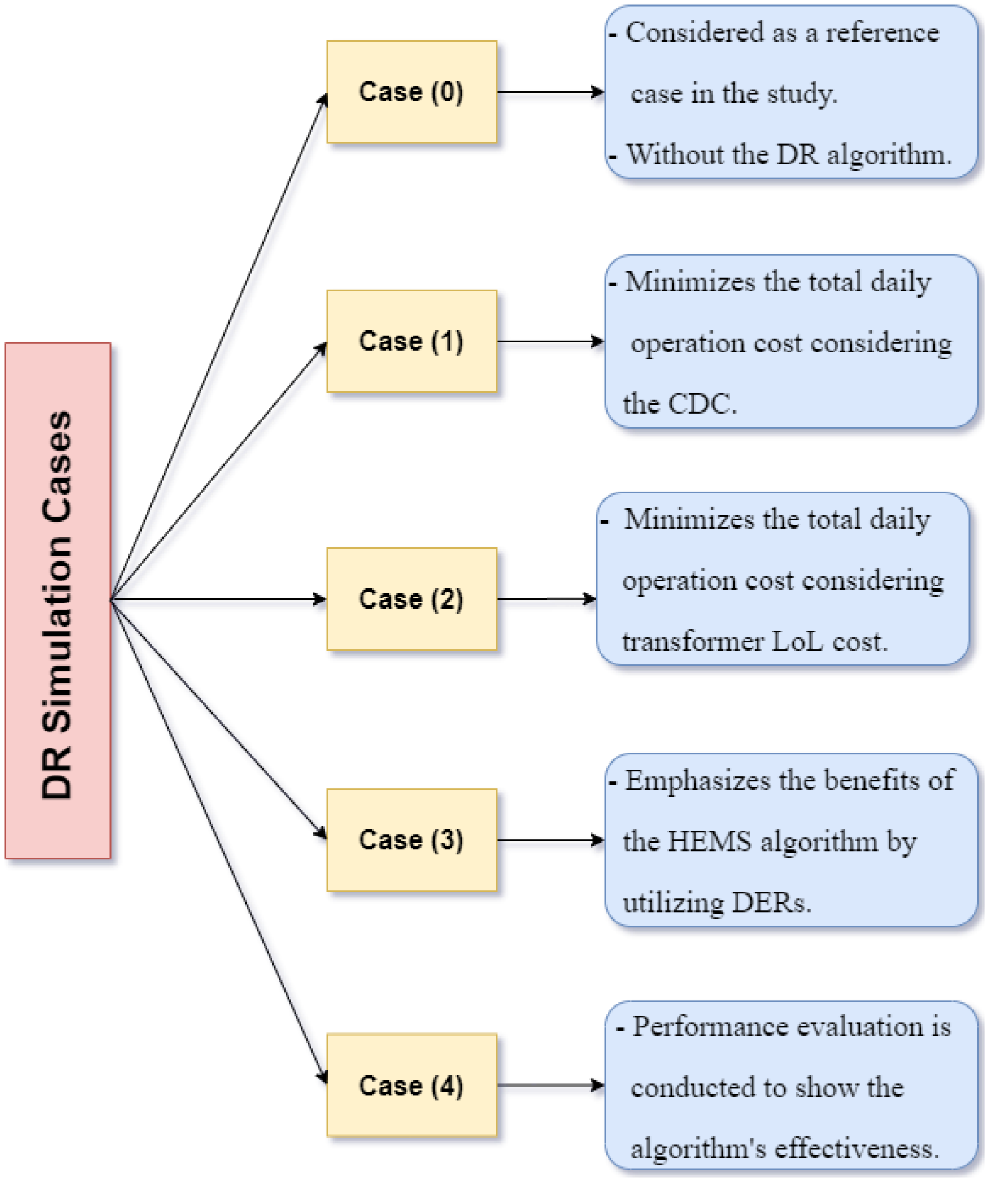

Figure 4-5 Demand response simulation cases

4.7 Calculation of Energy Consumption and Cost

Electricity consumption (kWh) of any appliance is determined by calculating the area under the curve of load (kW) profile *E* versus operation time in *h* hours. Electricity consumption of appliance n is given as $\int P_n \cdot dh$; where P_i is the load profile of appliance *n*. If the load varies with time, the electricity consumption is approximated as the area under the curve comprising summation of the area of series of rectangles formed on time slots. The smaller the time slot size, the better the approximation for calculating the area under the curve. In this thesis, the time slots are considered 1 hour each.

Reducing the electricity bill's cost under the RTP tariff is achieved through shifting and changing appliances' power consumption levels when the tariff is low. The optimized scheduling of appliances under RTP and different constraints is obtained, as explained in chapter 3. After having established the starting and ending times of the appliances associated with their respective load profile on the time scale having 24-time slots with zeros in unoccupied/unassigned slots, the operation cost of electricity is calculated as the following:

Let $P_{n,h}$ denotes the load profile vector of appliance *n* for all different starting time *h* of appliances, on a time scale of 24 slots, then the sum of energy (kWh) of N appliances is given as

$$\sum_{n=1}^{11} P_{n,h}$$

The energy (kWh) summation is a vector of order 1 × 24, and RTP (₵/kWh) is also a vector of order 1 × 24. Therefore, the following equation gives the total cost of energy used by all appliances in a day, and all energy requirements are met from the national grid.

$$Operation\ Cost = \left[\sum_{n=1}^{11} P_{n,h}\right] \cdot [RTP]^T$$

When a part of electrical energy is being met through DERs like PV or EV, the power imports from the grid will be reduced equal to the power supplied by the DERs.

CHAPTER 5 : RESULTS AND DISCUSSION: CONVENTIONAL APPROACH

This chapter demonstrates the simulation and performance evaluation results carried out by MATLAB to analyze the optimal load management in smart home integrated considering different cases. The proposed DR algorithm is implemented to compare three operating cases with a base case representing a reference without using the proposed algorithm. The utility and customer balance parameter (ρ) is set to 0.5 for all the following cases unless specified otherwise.

The ability of the HEMS to sell the extra energy back to the grid complicates the simulation study. Also, there are many policies regarding the injected energy to the grid, requiring separate analysis. In this respect, selling energy back to the grid is not considered in the analysis.

5.1 Base Case: Without the DR Program

This case is considered a reference in the study, which presents the household load profile before participating in the DR program (no reduction or shift on energy consumption). Figure 5-1 presents the accumulated energy consumption of household loads and EV and energy cost signals. Based on the cost signal, the energy consumption cost is calculated, as explained in the previous chapter. The total consumption cost for the base case is 6.1 USD.

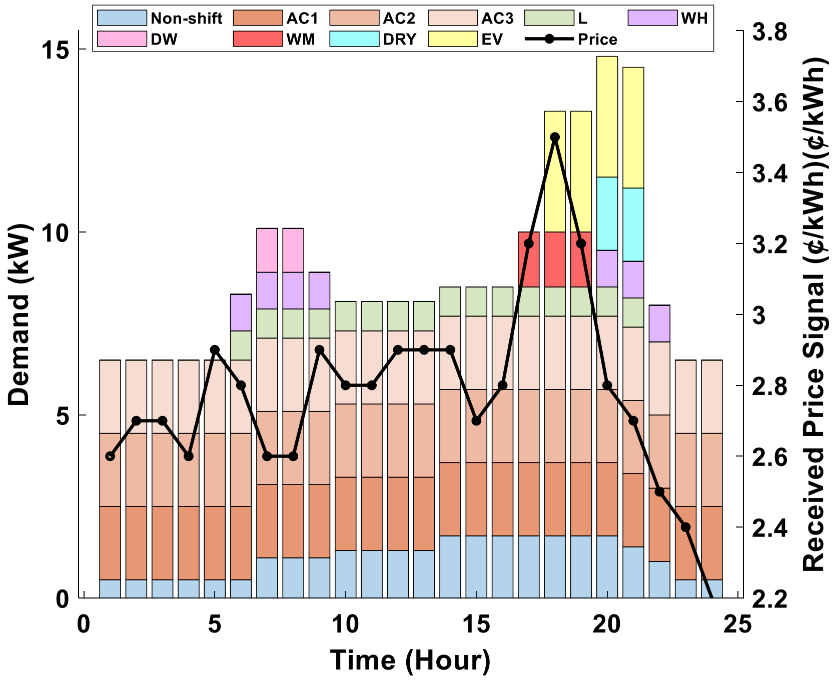

Figure 5-1 Aggregated energy consumption of all loads before DR (Base Case)

5.2 The Impact of Customer Dissatisfaction Cost

The main objective, in this case, is to reduce the daily operation cost considering CDC and ignoring transformer LoL cost. The customer comfort level is reflected in the optimization algorithm, considering ζ_n parameter, as illustrated in Table 4-4. Figure 5-2 shows the aggregated household load along with the electricity price signal after participating in DR Case 1. The proposed DR algorithm attempts to optimize energy consumption and achieve the lowest electricity charge cost. Specifically, all active loads (controllable and shiftable loads) are scheduled to operate outside electricity high-price slots. The controllable loads are scheduled to consume less energy from 15:00-20:00 times and consume more energy from 1:00-8:00 and 21:00-24:00. The shiftable loads (DRY and WM) are scheduled to operate at 23:00 and 24:00, where the tariff is low. The DW remains operating in slots 7 and 8 as it is the lowest in its working period, as specified in Table 4-4. The DR algorithm reduced the operation cost by scheduling the

EV charging in the lowest price slots (22:00-24:00 and 4:00) during the time that EV is available at home. In this case, there is no power limit on total energy consumption from the utility

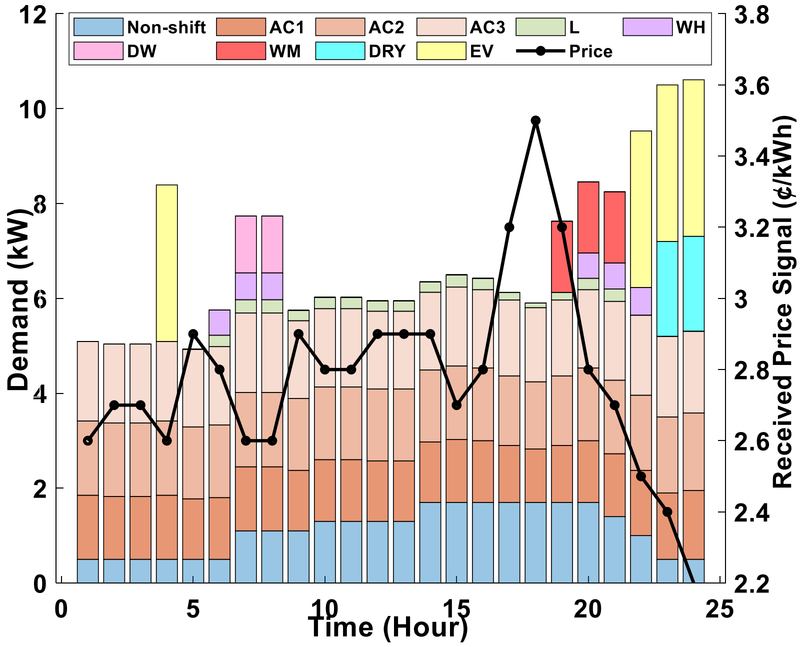

Figure 5-2 Load profile considering the impact of customer dissatisfaction

To present the HEMS algorithm's efficiency, Figure 5-3 is plotted to highlight the behavior of the considered controllable appliances (AC1, AC2, AC3, WH, L) in each time slot. It can be observed that the consumption of appliances is high during the first four-time periods. After that, the consumption is reduced due to the increase in electricity cost at time 5:00. When the price reaches its maximum amount around 18:00, each appliance's energy consumption is reduced to its specified minimum operation value Table 4-4. Finally, from time 20:00 to time slot 22:00, the appliances' consumption starts to increase since the electricity tariff decreases.

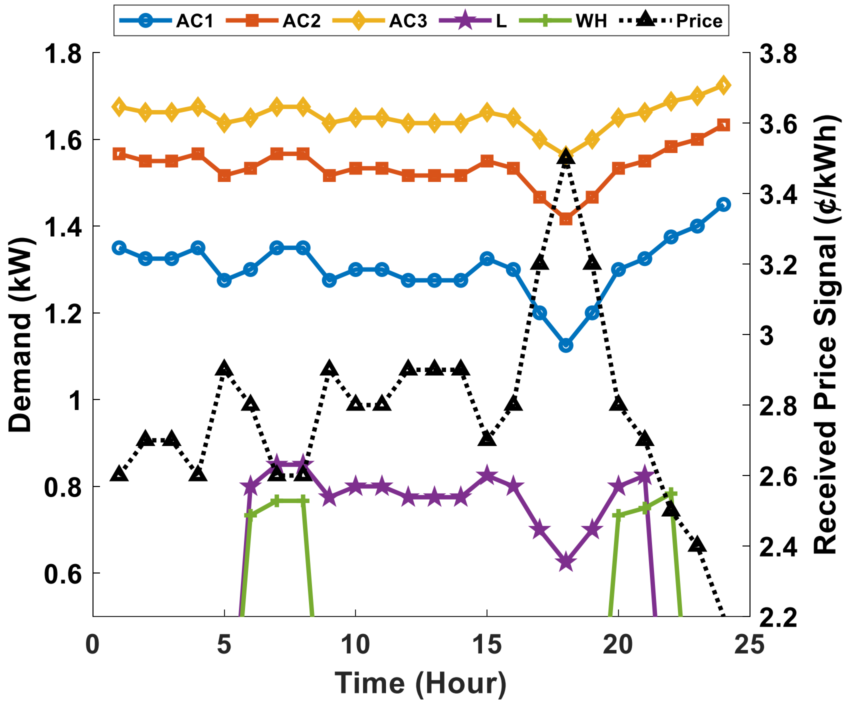

Figure 5-3 Energy consumption of controllable appliances during the day

Figure 5-4 presents AC1 energy consumption under different importance coefficients (ζ_n). As can be seen, when the value of ζ_n increases, the AC1 energy consumption increases. The appliance importance coefficient directly affects customer comfort. Similarly, Figure 5-5 presents the effect of ζ_n on shiftable appliances. in first graph, the DW operating time is scheduled according to the least price time in its scheduling interval (17:00 – 22:00). In the second graph, ζ_n set to 0.2. According to equation (3.21), the CDC increases as the waiting time of the customer for a device to start and finish the operation increases. Thus, the appliance is scheduled to start at 17:00 instead of 19:00 to minimize the CDC. Table 5-1 presents the effect of ζ_n on operation cost. It can be observed that the operation cost increase as ζ_n value increase in both shiftable and controllable appliance. Customers can set the values ζ_n for each appliances according to their needs and requirements. This is to guarantee that they suffer less dissatisfaction with the proposed HEMS algorithm.

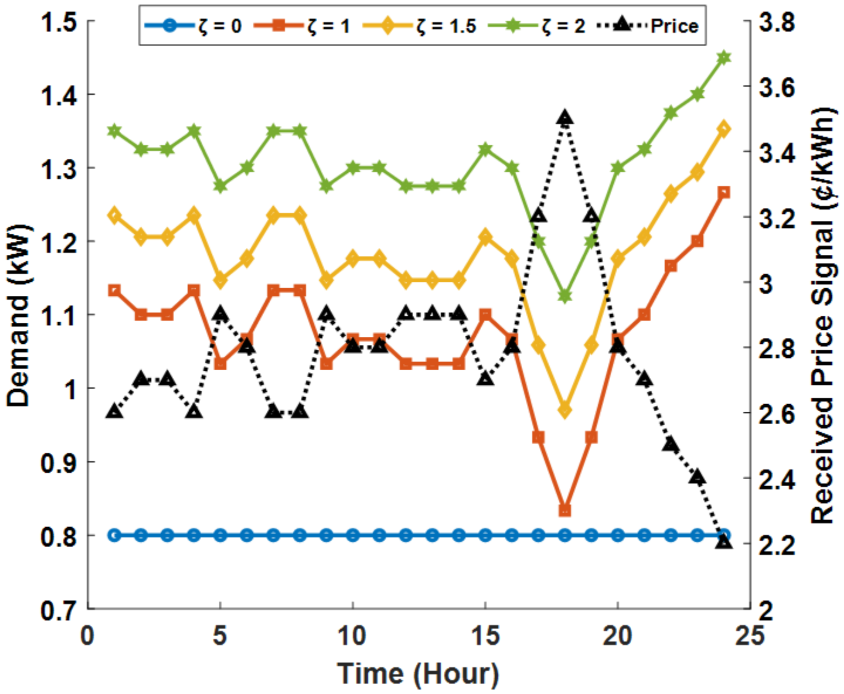

Figure 5-4 Effect of appliance importance coefficient (ζ_n) on power consumption of AC1

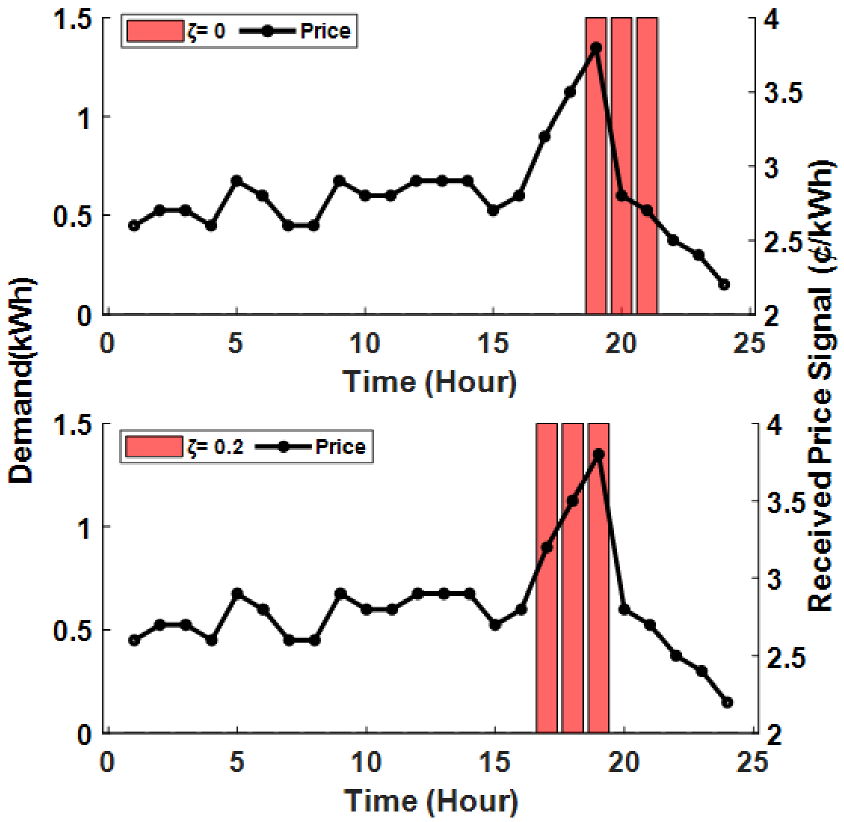

Figure 5-5 Effect of appliance importance coefficient (ζ_n) on scheduling DW

Table 5-1 Effect of appliance importance coefficient (ζ_n) on the operation cost

Appliance type	Controllable (AC)				Shiftable (DW)	
ζ_n	0	0.5	1.5	2	0	0.2
Operation cost (¢/ kWh)	53.7	71.1	78.4	86.6	8.7	9.9

The total electricity consumption cost of implementing case 1 (Figure 5-2) is 4.5USD, 25% less than the reference case. Customers' engagement in the DR program depends on the class of customers. Technology supporters have the potential to participate in such a HEMS-DR program and other energy market opportunities. However, another class of end-users still presents a challenge and needs additional incentives for engagement, even with the electricity cost reduction achievement. Hence full engagement in this program is limited to the technology-supporter class of customers. Furthermore, having individual control of the DR algorithm per individual end-user may cause the low-cost periods to operate as a sink for all customers to operate their appliances during these intervals and generate new load peaks detected by utility assets.

5.3 The Impact of Transformer LoL Cost

In this case, the objective function of transformer LoL mitigation cost (F_1) is considered in (1). The HEMS employs a method to fulfill the consumer's expectations and decreases the transformer's LoL by monitoring transformer load and electricity price. The DR program's impact on a 30-kVA distribution transformer based on the IEEE standard C57.91-2011 is used to investigate its life cost loss. The transformer is assumed to be supplying three houses where their household loads are monitored and controlled by the proposed DR algorithm. Based on case 1, the transformer has a peak demand in lower price slots as the DR algorithm operates all the smart appliances in these periods. Therefore, the load and LoL factor during these periods increased compared with the reference case's low-price periods.

The DR algorithm overcomes this issue by including the LoL as transformer deterioration cost using (5). Figure 5-6 shows case 2 optimal scheduling for the household load after adding LoL cost. It can be observed that the EV charging load is shifted to slots 1:00-3:00 of high electricity price compared to slots 22:00-24:00 used in case 1. While this load shift may increase the electricity cost, it satisfies the LoL cost and overcomes the distribution transformer's overload condition and other assets.

Case 2 results in a peak demand reduction of 18% compared to case 1, positively impacting the utility assets. The new load profile of case 2 has an electricity cost of 4.6

USD, which is higher than case 1; however, lower than the base case by 23% and provides direct benefits to both the end-user and utility operator.

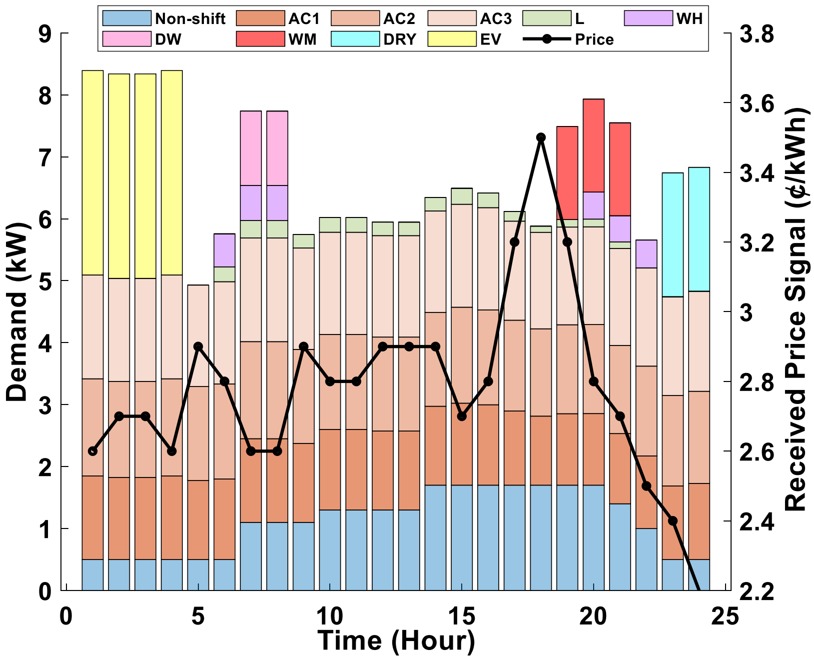

Figure 5-6 Load profile considering the impact of transformer loss of life cost

5.4 The Impact of DERs

The DERs are essential in developing active customers to engage in the local energy market and hence integrated into the HEMS-DR program. This case emphasizes the DR algorithm's benefits by utilizing assets, such as PVs, ESSs, and EVs. The energy supplied by the PV source, EV, and ESS is used to minimize the supplied power from the grid and balance the demand/generation relationship. It is assumed that PV-to-home (PV2H), ESS-to-home (ESS2H), and EV-to-home (V2H) capabilities are available. Figure 5-7 presents the results of this case, and the following are observed:

- The PV generated power is used to partially cover the demand and charge the

ESS as long as it is available.

- When prices are high, ESS's available energy is utilized to cover part of the load and reduce electricity consumption cost, as illustrated during the 17:00–19:00 time period.
- As the EV reaches the household with sufficient energy, it supplies the household needs through V2H mode during 19:00-2:00 time. It is also observed that the HEMS-DR algorithm avoids chagrining the EV in high price slots.

The electricity cost of case 3 is 4.2 USD, which is 31% lower than the base case, 9% lower than case 1, and 10% lower than case 2.

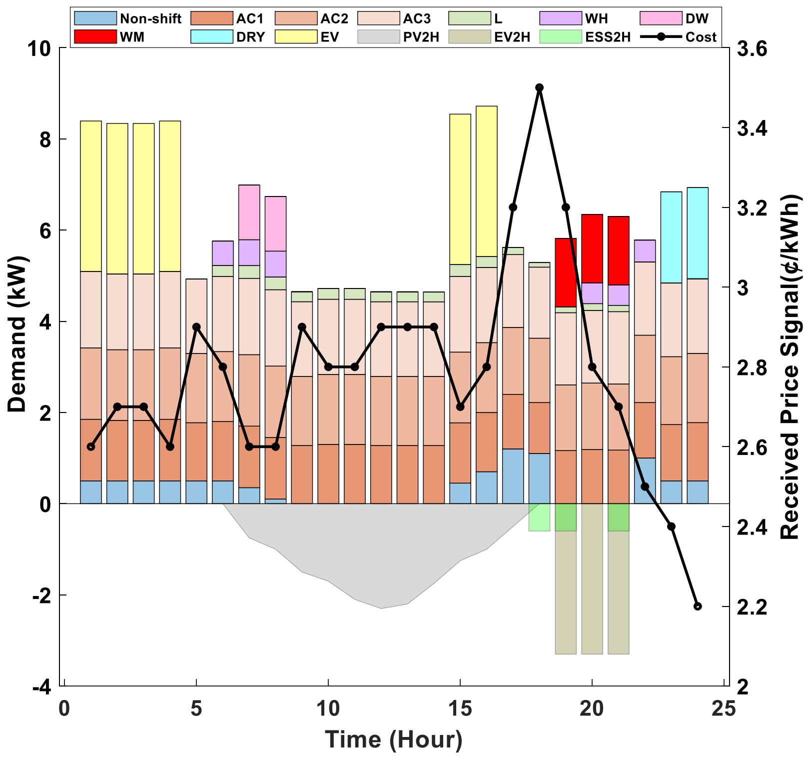

Figure 5-7 Load profile considering the impact distributed energy resources

Table 5-2 EV consumption costs using different parameters

Initial SOE (%)	Arrival / Departure time	Cost (₵)
30%	2 P.M	425.84
30%	5 P.M	428.81
30%	7 P.M	434.80
50%	2 P.M	419.60
50%	5 P.M	421.63
50%	7 P.M	427.71
70%	2 P.M	411.75
70%	5 P.M	415.56
70%	7 P.M	423.58

To complete the analysis, different initial SoE and EV arrival times are examined. This is done to study the effect of the consumer's behaviors on the operational cost. Table 5-2 presents the analysis results. The operational cost increases as the EV arrives later in the day. This is expected since the EV contributes to supplying the household at less costly periods, covered by the grid's energy during higher price periods. Also, as EV initial SoE increases, the cost decreases. Specifically, an increase in SoE by 20 % renders a 2.3% reduction in the cost. The reduction is calculated, considering that the EV arrival time is 2:00 P.M.

5.5 Performance Evaluation

5.5.1 The Impact of Power Limiting Strategy

The proposed HEMS aims to schedule the load in the low-price periods and respond to the utility's specific load-shape requirements. To evaluate this, a power limit restriction is enforced throughout the time horizon to limit the drawn energy from the grid to a maximum value of 8 kWh to reduce the customer peak demand, as presented in (3.35) and (3.36). Violating the power limit is allowed under a penalty that is equal to the demand charge tariff. The demand charged is assumed to be equal to 110% of the day's highest electricity price (3.8 cents). As shown in Figure 5-8, under this strategy, the total peak is further decreased to 8kW, and as a result, the EV charging is reduced, and the charging period is increased. However, the load stays the same as the previously discussed case (Case 3).

Also, it may be noticed that in previous cases where the power limiting strategy is not applied, increasing the capacity of ESS or EV battery will increase the maximum power (customer peak). This leads to more cost reduction for the customer. From the utility perspective, the flexibility offered to the customer by increasing the ESS capacity results in inconsistent and fluctuating load profiles. This will create challenges for the power system, such as increasing customer peaks and load balancing problems. As shown in Figure 5-8, when the power limit strategy is imposed, the EV charging pattern has changed. For instance, if the EV/ESS capacity increases, the HEMS will be forced to increase the charging time to eliminate the peaks to avoid being charged with higher electricity rates.

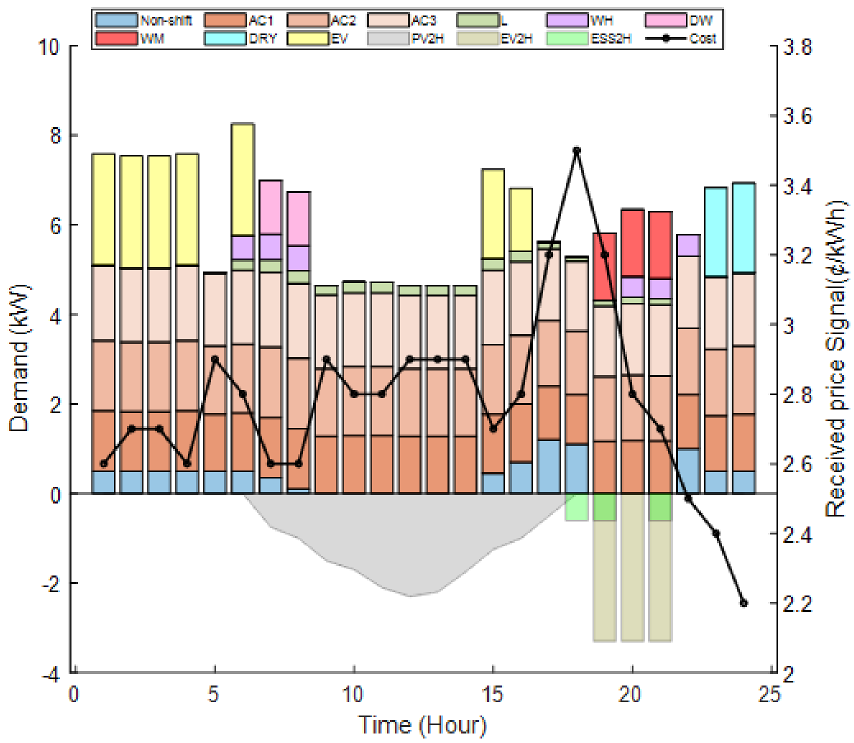

Figure 5-8 The impact of the power limiting (8kW limit)

5.5.2 The Impact of the Balance Parameter ρ

The balance parameter for customer and utility benefits directly impacts the objective function in (1). Figure 5-9 and Table 5-3 show the electricity consumption and electricity cost of the three previously mentioned cases using the proposed DR program under different values of the customer-utility balance parameters (ρ). A large value of ρ magnifies transformer LoL mitigation cost (F_1) and the energy operation cost (F_2) in the objective function. On the other hand, using a small ρ magnifies customer dissatisfaction cost (F_3) in the objective function, where the end-user has more impact on the cost function compared with the utility. Hence selecting the value of ρ is an essential element in the optimization algorithm and should be defined based on a mutual agreement that reflects the customer and utility operator benefits.

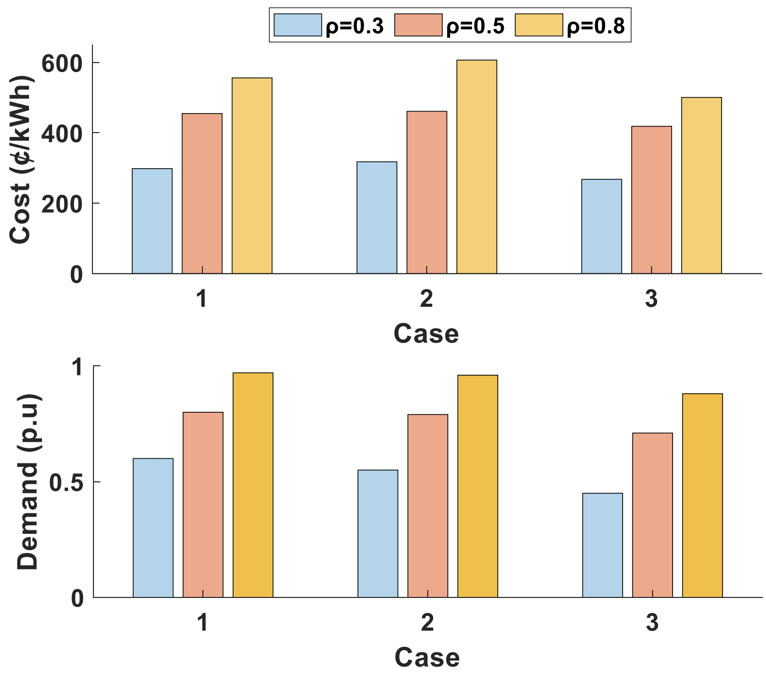

Figure 5-9. The impact of the balance parameter on electricity consumption

Table 5-3 The impact of the balance parameter on electricity consumption

Case	Total Load (p.u)			Total Cost (¢/ kWh)		
	0.3	0.5	0.8	0.3	0.5	0.8
Case 1	0.60	0.80	0.97	298	318	268
Case 2	0.55	0.79	0.96	318	461	418
Case 3	0.45	0.70	0.88	461	607	501

5.5.3 The Impact of DR Program on Distribution Transformer

The impact of the DR program's three cases on a distribution transformer (DT) loading condition is studied. The loads of three houses are used, and the DERs are randomly varied. All the households are considered to be equipped with DR-controlled appliances. So, all households have the DR capabilities to shift or reduce the consumption of the appliance properly. The aggregated load of the three houses supplied by a 30-kVA distribution transformer is estimated, and the results of the DR program of the previous cases are illustrated in Figure 5-10. The price pattern is also included in the figure to help in result analysis. To evaluate the proposed DR algorithm's efficacy, the base case is considered with a severe loading where the DT loading condition shows 4-hours continuous overloading above 100% rating during the 18:00-22:00 time interval and another 10-hours above 80% rating. Case 1 of the DR program supports customers' benefit in energy cost reduction. However, the results present an alarming indication of a possible overloading condition that accelerates transformer degradation.

Furthermore, part of the overloading conditions time intervals has a low-price signal, which may work as a sink for other customers to operate their appliances and magnify the DT’s load peak condition. Implementing case 2 of the DR program shows a reduction in the DT load beyond the 80% rating and an increase in the energy cost due to load shifting to high-price electricity cost intervals. This is shown in the figure during the time intervals 1:00-4:00 and 20:00-24:00. Case 1 and case 2 of the DR program show similar responses during the time interval 4:00-20:00. Including DERs in the DR program (case 3) shows the optimal operation scenario as expected. The DR program generates a DT loading below all the cases during the time interval from 7:00-19:00 and merges with the lowest case (case 2) during the time intervals 4:00-7:00 and 19:00: 24:00.

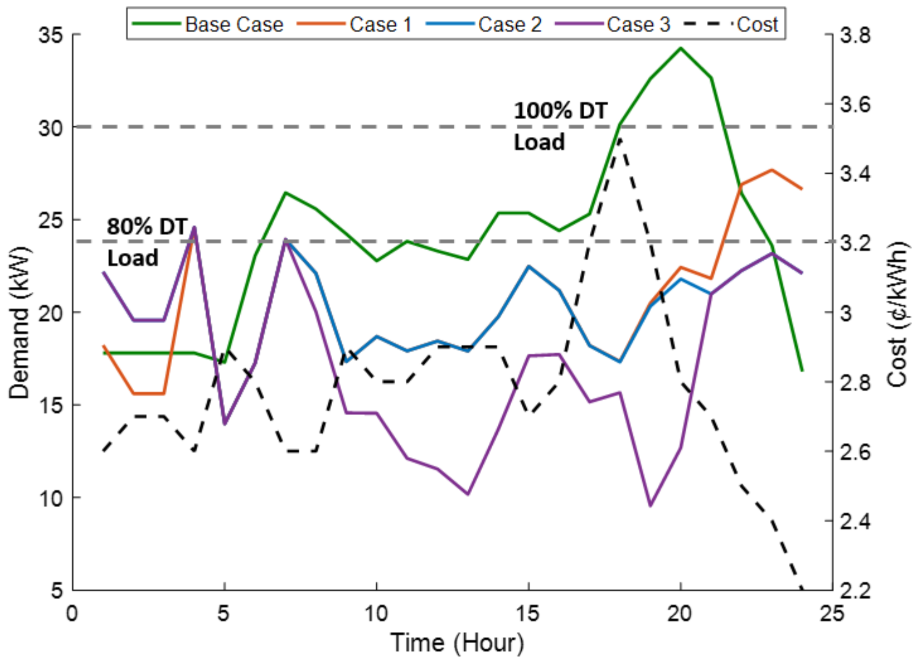

Figure 5-10 The impact of the DR on the distribution transformer load profile

According to Eq. 5 to Eq. 14, the transformer's TOT and HST will increase during overloads. Minimizing the transformer load reduces HST based on the transformer's thermal modeling. Figure 5-11 shows the thermal HST of the considered transformer. The results show the effect of the DR program on the transformer. As shown, the HST reached around 73°C without DR, and with the DR, the average and maximum HST for transformers have been reduced in all the cases. DR case 3 delivers the maximum drop in the transformer load, showing a 17% reduction in the winding temperature. The drop in the transformer temperature explains the thermal limits for transformer loading and protects the transformer insulation. The HST curve has the same shape of the transformer load in all the cases because of the minimal value of the exponential function, almost 0, in (3.4) and (3.7).

Since the HST has a considerable impact on transformers' aging, minimizing HST will minimize transformer LoL%. As shown in Figure 5-12, the transformer LoL can be minimized up to 25% by the proposed DR algorithm. A comparison of the total transformer load, HST, and LoL% is shown in Table 4-5. The benefits of utilizing DR case 3 are more noticeable where the maximum load is reduced to 0.83.

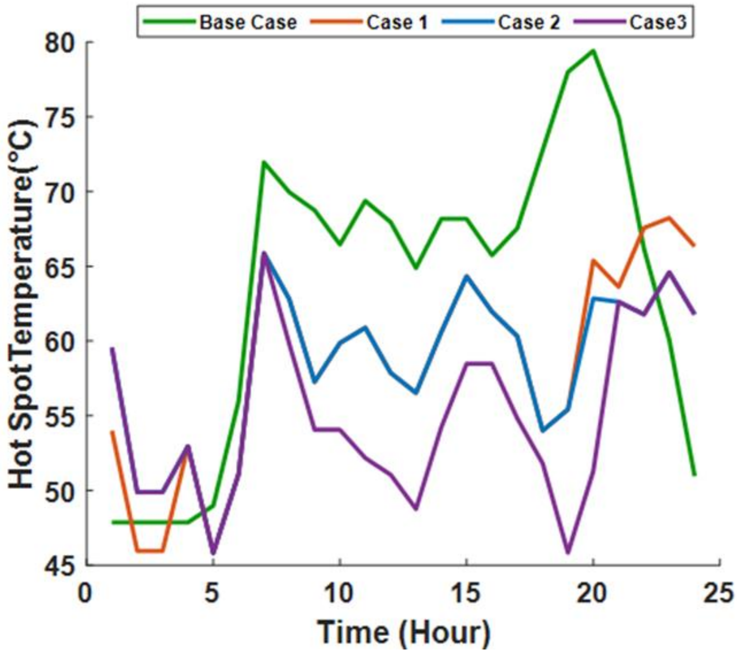

Figure 5-11 The impact of the DR on distribution transformer HST

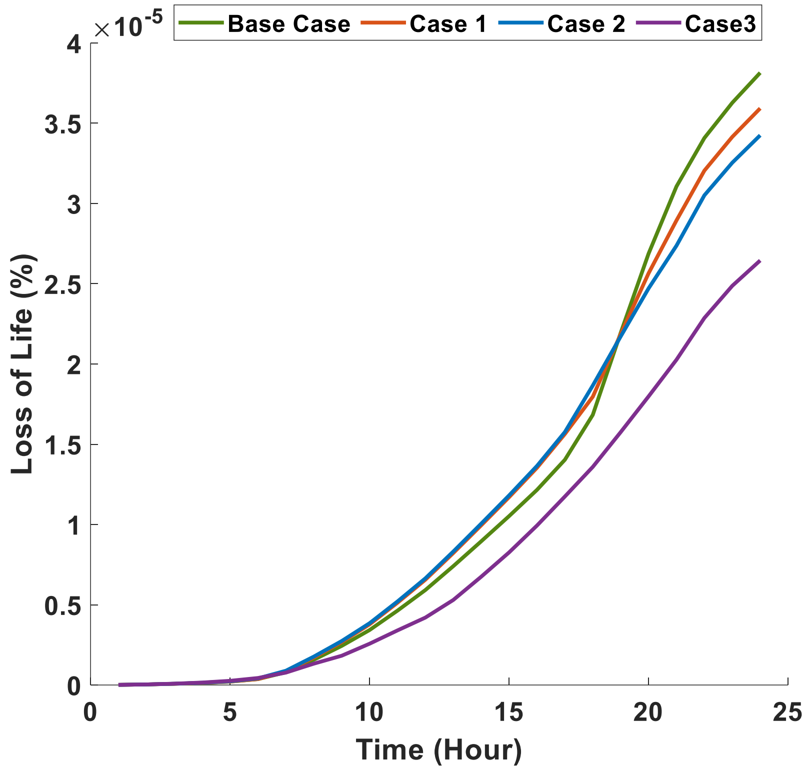

Figure 5-12 The impact of the DR on distribution transformer LoL%

Table 5-4 Comparison with and without DR

	Max. Load (p.u)	Maximum HST (°C)	Max. LoL (%)
Without DR	1.13	79.40	0.000038
With DR (Case 1)	0.90	68.20	0.000035
With DR (Case 2)	0.83	65.50	0.000033
With DR (Case 3)	0.83	65.40	0.000027

5.6 Challenges and Opportunities

One of the chief benefits of the proposed model is its flexibility. It can be automatically adjusted and adapted to consider more appliances and constraints and generate the best possible scheduling solutions. The equations can be solved in more than one way. The result was obtained using a laptop (i7 at 2.6 GHz, 16 GB of RAM, 64-bit Windows) with a computation time, on average, of 5 s. With a more powerful processor and high-capacity RAM, the computation time will be further reduced. However, implementing such an algorithm for larger-scale applications may increase the computational burden. Since the scheduling problem's computational complexity is non-polynomial (NP)-hard, as the number of constraints and variables increases, the developed model's computational time exponentially increases. However, a detailed analysis of the computational complexity would be necessary to evaluate its potential for future large-scale applications fully.

A variety of control and optimization strategies have been developed to solve the load scheduling problems, such as mathematical method [85], programming method [48], [49], or heuristic methods like genetic algorithm [86]. However, there is several challenges to implement these optimization approaches. Therefore, it is necessary to have an effective methodology that produces robust DR decisions and minimizes real-time uncertainty errors. The proposed method requires the system model to be known in order to manage. When it comes to practical systems, the exact models may not be obtainable. It is required in the power system control sector to develop a control algorithm, which is implementable in practice and real environments.

The above-mentioned methods use an explicit mathematical equation to model the system. In this case, these models' accuracy cannot be guaranteed since the household appliances' efficiency and different variables keep changing over time. In the simulation, these optimization methods can show good performance due to the assumption of accurate input data prediction. On the downside of these methods, the model dynamics have to be modeled with great precision knowing all the environment information that limits its application on large-scale systems, where the system parameters could be partially or completely unknown. For instance, renewable energy is considered most effective if applied on a large scale. If the model parameters have to be adjusted from one case to another, conventional optimization techniques will limit large-scale implementation [87]. These techniques may also suffer from high complexity and computational cost in real-time applications due to the significant number of variables involved.

Moreover, some of these methods ignore a few unquantifiable parameters or adapt ungeneralizable and inaccurate model formulas such as the users' comfort and satisfaction. Therefore, conventional optimization techniques are case-specific and need to be adjusted when the system environment changes during abnormal situations. For example, in 2020, with the current COVID-19 pandemic, the most current society's daily life is affected. Around 30% of the global population has been put in lockdown with different levels of nation-wide quarantines [88]. This ongoing situation is causing a social readjustment of most societies' daily routines, practices, behaviors, and expectations. The lifestyle is changed globally as people are mostly staying and working from home. This leads to a significant increase in residential load demand. This leads to significant impacts on the power system (from production to consumption) [89].

Therefore, it is necessary to develop a robust DR methodology, enabling a better integration between customer and utility. This will improve the performance of the home energy management and power system operation under such conditions in the future. The model-free ML approaches appear to be a good solution to beat the traditional optimization method to optimize power consumption effectively and manage such big data in real-time effectively [90]. The model-free methods are entirely based on data and do not need precise modeling of the system, making it more generalizable and can adapt to environmental changes.

CHAPTER 6 : REINFORCEMENT LEARNING

To overcome traditional optimization method limitations discussed in chapter 5, a Reinforcement Learning (RL) algorithm is introduced to optimize power consumption effectively and manage such big data. This chapter explains RL rules and their applications in the power system. Section 1 provides background and related work for the RL algorithm. Section 2 discusses the details of the Markov decision process as a decision-making model for RL. Section 3 explains the main elements that jointly drive the performance of an RL algorithm. RL categories are explained in section 4. Section 5 discusses the integration of deep neural networks in RL algorithms.

6.1 Background and Related Work

To address the upcoming complex challenges in power systems such as high generation from Renewable Energy Sources (RESs) and the increasing number of price-responsive-demand participants, recent studies look into computational intelligence and Machine Learning (ML) techniques as potential problem-solvers. Among ML techniques, Reinforcement Learning (RL) is a technique that learns from an interactive environment through trial and error. Over the last years, RL has become one of the valuable research directions of ML. Algorithms in ML are often divided into supervised, unsupervised or reinforcement learning. Supervised learning is an algorithm using input data and labeled output data (target). The supervised algorithm attempts to find a relation between the input data and output data in a manner that generalizes well to unseen input data. Examples of supervised learning are regression and classification algorithms. Algorithms using unsupervised learning attempt to find structure in unlabeled data. Examples of unsupervised learning are clustering and anomaly detection. The terms supervised and unsupervised do not describe well the mechanisms of RL algorithms. An RL agent learns from communicating with the environment and receiving rewards based on the taken action. The agent's goal is not to use labeled data in some sense or explicitly finding general structures in the data. As a result, RL is considered a category of its own [91]. The relation between supervised, unsupervised, and RL is shown in Figure 6-1.

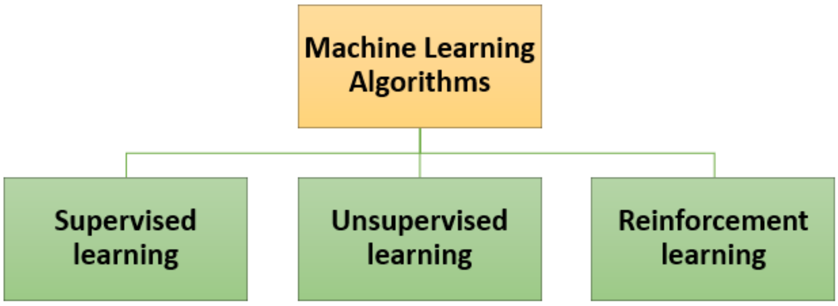

Figure 6-1 The three main categories of machine learning

RL's motivation is that RL is a method to solve problems that follow Markov decision processes (MDPs). One of the main advantages of the MDP framework is its generality. It can handle different reward functions, such as nonlinear, stochastic dynamics, and non-quadratic functions [91]. This makes the RL have the ability to handle a wide range of problems such as planning, management, optimization, and control problems. Besides its generality, RL is a model-free algorithm where the agent does not need prior knowledge about selecting an action. Hence, RL is an effective technique to find near optimum solutions for different nonlinear systems when the system parameters are unknown, mostly in the power systems. There are considerable amounts of historical interaction data that demonstrate the informative behaviors of power consumption in the residential sector, which can be a rich source of information. RL algorithms can utilize such data sets to scale to real-world problems and give solutions that generalize substantially better. Using these data for RL will enable the pre-train and test model to learn in the real world.

Recently, RL has been utilized to solve many optimization and control problems. In the optimization and control engineering context, RL bridges the gap between conventional optimization and adaptive control algorithms [1]. Although both RL and conventional optimization approaches have a common goal (optimal decision making), their fundamental working principles are different. RL is a data-driven approach, where the optimization process is achieved by an agent that performs an action and receives feedback. The feedback is an indication of how great the new state of the environment is. For example, the agent will get a positive numeric reward when it takes optimal actions (acts in a desired behavior) and negative numeric rewards for action the agent should stray away from (acts in an undesired behavior). When the agent gets negative rewards as a feedback of its action, it will be less likely to choose that action later. Similarly, when it gets a positive reward, it will more likely choose a similar action given the same observed state. By allowing the agent see many states and explore different actions, it can eventually learn behavior that yields a lot of positive rewards. The RL agent's main objective is to learn a policy to pick actions based on the current state, leading to good states on average [91], [92]. RL's framework in decision-making and control is illustrated in Figure 6-2. There are three phases of the framework: training, testing, and execution. The training phase oversees learning the policy, and the testing phase is the assessment of the quality of the learned policy and how much reward the agent obtains if it follows that policy. After evaluating the algorithm, the learned knowledge can be deployed to make optimized decisions in a real physical environment.

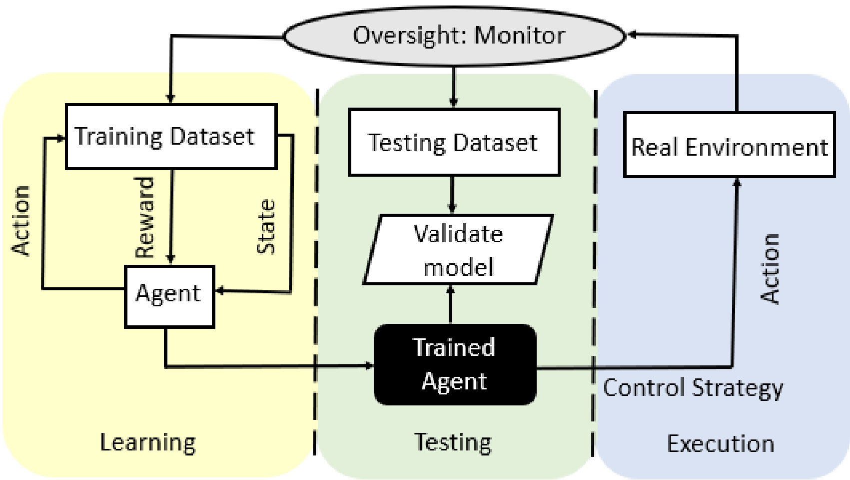

Figure 6-2 RL framework

To achieve the best policy, the RL agent takes time in the learning or training process as it has to explore the whole system, making it improper and inapplicable to problems with large state and action spaces. As a result, most RL applications are inadequate in real-world problems. Recently, deep learning [93] is considered a new revolution technique that can beat the RL limitations. This has opened a new era for RL improvement, named Deep Reinforcement Learning (DRL), where RL is combined with a deep neural network (DNNs). DRL can handle more complex and challenges problems with high-dimensional state and action spaces and enable continuous state and action spaces [94]. Recently the DRL field has attracted the researchers' attention due to its impressive success in games [95],[96], robotics [97], finance, and business management [98], [99]. In power system applications, DRL has already shown its usefulness in problems like demand response [100], energy management [101], operational control [102], cybersecurity [103], economic dispatch [104], and power system optimization [105].

More spherically, work in [106] and [107] proposes DR RL-based model to control the residential thermostatically controlled loads. The model schedules the operation of a heat-pump thermostat for 24 hours ahead. Also, RL and neural network are utilized to model the HVAC system's thermal dynamics in a building. The HVAC loads are controlled to lower the electricity bill while considering the user satisfaction level in terms of the indoor temperature [108], indoor temperature, and air quality [109]. The results showed that the proposed DRL algorithms had achieved more cost savings than the rule-based control approach. Some work aimed to optimize the charging/discharging schedule for the EV using RL [110]-[112]. Work in [110] aimed to reduce the EV charging cost. The charging problem is modeled as MDP. In [111], the authors propose a decentralized charging control to schedule a plug-in EV fleet's charging. Work in [112] presented a multiagent RL architecture that aims to reduce energy generation costs and avoid transformer overloads by coordinating the EVs charging time.

Work in [113] proposed a multi-agent hour-ahead DR algorithm to schedule the household appliances. Each agent represents different types of home appliances. The energy consumption is optimized using the Q-learning algorithm. Also, the ANN model is used for real-time price prediction. Complementary with this work, authors in [114] introduced a more comprehensive model considering EV. The results showed that RL has significantly reduced the price compared to the GA and MILP algorithms. Also, few researchers have utilized RL in home energy management with DERs and ESSs. Work in [100] presented an RL model to schedule the residential loads PV source. Also, [115] proposed a HEMS with ESS and rooftop PV panels. ESS is utilized to achieve energy and cost savings. Moreover, RL is utilized in [116] to enable energy trading between the utility and its customers to balance the supply and demand and enhance grid reliability, neglecting customer comfort.

6.2 Markov Decision Process

In RL, the problem is described as a Markov Decision Process (MDP). An MDP is a mathematical framework describing sequential decision making and interaction with an environment, where the outcome can be stochastic [94]. The environment starts at $t = 0$ and is described by an initial state $s_0 \in \boldsymbol{S}$. The agent executes the action $a_0 \in \boldsymbol{A}$ and receives a reward $r_1 \in \boldsymbol{R} \subseteq \mathbb{R}$ based on how good that action is. The action a_0 interacts with the environment and gives a new state 1. This starts the sequence of states, actions, and rewards.

$$s_0, a_0, r_1, s_1, a_1, r_2, \cdots \tag{6.1}$$

The interaction lasts until the environment reaches a terminal state, for instance, when the self-driving car reaches its destination or if it crashes. The transitions from start state s_0 to terminal state s_T constitutes an episode in the RL algorithm. An example of an MDP with 2 states and 1 action is shown in Figure 6-3. Formally, a finite MDP $\boldsymbol{\mathcal{M}}$ is a tuple consists of 4 elements, as presented in 6.2.

$$\boldsymbol{\mathcal{M}} = \langle \boldsymbol{S}, \boldsymbol{A}, \boldsymbol{P}, \boldsymbol{R} \rangle \tag{6.2}$$

where $\boldsymbol{S}$ and $\boldsymbol{A}$ respectively are finite sets of states and actions, $\boldsymbol{P}$ is the matrix of state transition probabilities, and $\boldsymbol{R}$ is a reward function [94]. The probability of transitioning to the next state s_{t+1} and receiving r_{t+1} only depends on the previous state s_t and action a_t in a MDP [94]. In Figure 6-3, the numbers on each line are the state transition probabilities. Formally, a state s_t is MDP if and only if

$$\mathbb{P}[s_{t+1} | a_t, s_t] = \mathbb{P}[s_{t+1} | a_t, s_t, a_{t-1}, s_{t-1}, \cdots, a_0, s_0] \tag{6.3}$$

where $\mathbb{P}$ is a symbol for probability transition. This is called the Markov property of the state [94]. In other words, the history of states and actions leading up to the current state is not relevant for the probability of transitioning to state s_{t+1}. Let the transition function $p : \boldsymbol{S} \times \boldsymbol{R} \times \boldsymbol{A} \rightarrow [0, 1]$ be the probability of transitioning from state s to s'

and receiving reward r given the action a.

$$p(s', r, s, a) = \mathbb{P}[s_{t+1} = s', r_{t+1} = r | s_t = s, a_t = a] \tag{6.4}$$

If the transition function p in (6.4) is known, it can be used for planning actions in an RL algorithm.

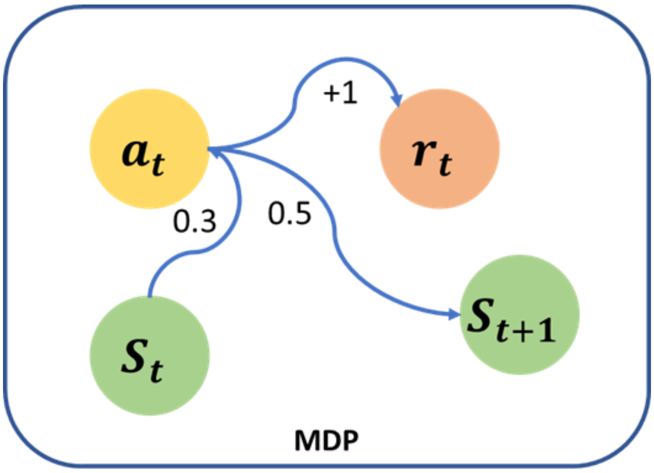

Figure 6-3 Example of MDP with 2 states and 1 action

6.3 Reinforcement Learning Elements

The agent and the environment are a fundamental part of any RL algorithm. The RL agent and environment communicate in a series of episodes. These episodes are divided into a sequence of timesteps. Each timestep, the RL agent receives information that presents the environment, and based on that information, the agent selects an action. As a result of its action, the agent collects the rewards and move to a new environment state. The interface between the agent and the enviroument goes on until a terminal state is achieved. In this state, the agent cannot take any further action. The communication between the agent and environment is visualized in Figure 6-4. More specifically, all the elements in the RL model oare explained below.

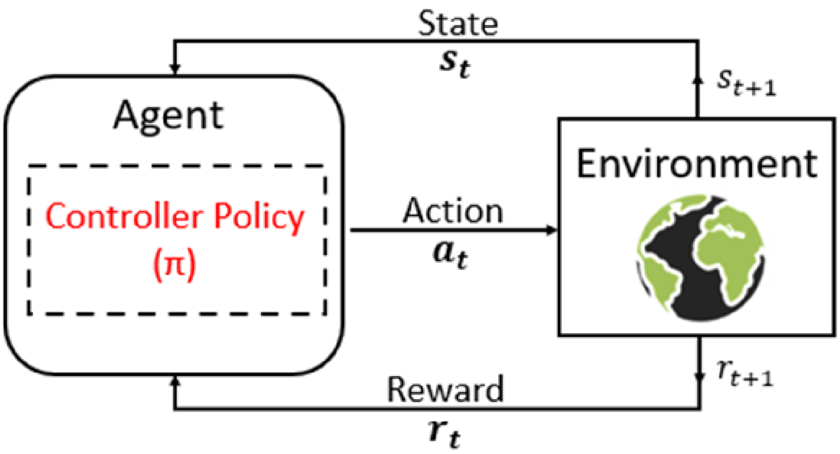

Figure 6-4 Relation between the agent and the environment in RL

6.3.1 Agent

The agent is the decisionmaker (Learner), which learns from interacting with a dynamic environment to choose actions in order to maximize future rewards. There are different types of RL agents, such as Q Learning, Deep Deterministic Policy Gradient (DDPG), Deep Q-Network Agents (DQN), and Actor-Critic (AC). It is essential to select an agent that is compatible with the action and state spaces.

6.3.2 Environment

The environment in RL is the thing that the agent interacts with. It is also called a system model. RL can be used to learn directly by interacting with the real system or a system model. The system model can be a simulation or mathematical model.

6.3.3 State Space

State-space is a set of possible states occupied by the agent at different instants of timesteps. At any time, the agent will be at one of the states from the entire state space. The state-space can be discrete or continuous. The state of the agent at time t can be represented as s_t. The entire state space is then taken as $\mathbf{S}$, where at any timestep t $s_t \in \mathbf{S}$. The state-space should contain useful information that the agent needs to make the right decision. The states should be easily fed into the DNN and catch as much information of the environment as possible.

6.3.4 Action Space

The action space is all the possible decisions that the agent can take. The agents usually select from a list of possible actions. The action space can be discrete or continuous. During the training, the agent has to make a series of actions or state transitions, $a_0, a_1 \ldots \ldots \ldots a_{N-1}$, where $a_{\mathrm{t}} \in \mathbf{A}$. Initially, the agent does not have any prior knowledge about the effects of its actions on the environment. Gradually, by training, the agent learns the desired action at each state to maximize the reward function. In some application, for istance, games the action space is simple. In other application the actions might need modifications, such as discretization or trimming.

6.3.5 Reward Function

Designing a reward function is a crucial point in any DRL problem. Reward function should guide the agent to advance in the right direction to achieve the objectives. For simple problems, the reward can be a function of the states, $r(s_t)$, or with specifying more detail, e.g., $r(a_t, s_t,)$ or $r(a_t, s_t, s_{t+1})$ for challenging problems. A well-specified reward function will help the agent to learn better and converge faster. There are no absolute restrictions on designing a reward function. In some cases, it can be straightforward; in some other cases, it is not. The reward function can be a discrete or continuous function depending on the application. Also, it can be a positive or negative value. Positive rewards motivate the agent to keep going to accumulate rewards. Negative rewards encourage the agent to avoid some actions at certain states or to incentivize the agent to reach the goal state as fast as possible. A key to making the training work can shape the reward functions in reasonable ways. For example, using sparse rewards, the agent will not get rewarded very often. This may lead the training not to converge, or the agent may be stuck in local minima. Instead, shaping the reward function so that the obtained gradual feedback will help the agent to learn faster.

Different strategies to design the reward functions to enhance agent learning and training are proposed. The authors in [117], [118] have proposed a heuristic reward for DRL algorithms to be deployed in problems with extremely large state space. Others have proposed a design for reward functions to accelerate learning, which utilizes implicit domain knowledge [119]. Work in [120] proposes a reward function design, adapted according to the degree of uncertainty in predicted data. However, it is still challenging to determine the ideal reward function for a specific environment. The reward representation is case-specific and may vary depending on the complexity of an environment.

6.3.6 Policy

A policy (π) is the agent's approach in choosing the next action based on the current state. It maps states to actions with the highest reward. The objective is to find an optimal policy π^* where the expected total cost is lower compared when following any other policy $\pi \in \Pi$. The policy can both be deterministic or stochastic. A deterministic policy maps a given state to the same action every time, while a stochastic policy maps the state to a probability distribution over the action space.

6.3.7 Value Function

The value function, $V(s)$, is a function of state-action pairs that estimate the cumulative future reward of being in a given state. Value function determines the goodness of a policy. The agent has to follow a good policy starting from the initial state to reach the goal at the minimum cost. The minimum cost obtained is also defined as $V^*(s)$, and it is called the optimal value function. It estimates the total expected future discounted reward that could be gained while following a particular policy over N time steps. The future reward is discounted with a discount factor (γ). The reason for discounting the future reward is that the real goodness of an action may not be reflected by its immediate reward. The gamma term is a hyper-parameter that can be tuned, and it determines how relevant future rewards are. If γ = 0, then the agent only considers the immediate reward as relevant. If γ = 1, then the agent will evaluate each of its actions based on the sum of all of its future rewards. For values between 0 and 1, the importance of a reward decreases exponentially with every time step. For instance, if γ = 0:5 the rewards for the next steps are weighted $0.5;\ 0.25;\ 0.125, \cdots$. Having γ smaller than 1 is also a mathematical convenience that ensures that the discounted return is finite in a continuous task, as long as the rewards are bounded. The problem environment decides the value of the discount factor.

6.3.8 Action-Value Function

The action-value function $Q\ (a, s)$, also called the Q-function, quantifies the expected discounted return given that the action a_t in state s_t and that the policy π is followed. In other words, it can evaluate a specific action in a given state, in contrast to the value function $V(s)$ that only evaluates the state. The optimal value $Q^*(s, a)$ is used to represent the maximum accumulative reward, which can be obtained. There is a crucial recursive relation between the action-value function in two successive states s_t and s_{t+1}, known as the Bellman equation.

$$Q^*(s, a) = r_{t+1} + \gamma \underbrace{max}_{a_{t+1}} Q(s_{t+1}, a_{t+1}) \tag{6.5}$$

In other words, the action-value for state s_t and action a_t is the expected sum of the immediate reward r_t and the action-value in the next state. The Bellman equation is used in several reinforcement algorithms to guide the Q-values' estimates closer to the true values. At the end of each iteration, the estimate of Q-value Q_t is updated by

$$Q_{t+1}(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[\, r_t + \gamma \underbrace{max}_{a_{t+1}} Q^*(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \quad (6.6)$$

The main concept of this update is to find the difference between the predicted Q-value, i.e., $r_t + \gamma \underbrace{max}_{a_{t+1}} Q^*(s_{t+1}, a_{t+1})$ and its current value, i.e., $Q(s_t, a_t)$. This is known as Temporal Difference (TD) learning which is a model-free type of RL. In (6.6), $\alpha \in$ [0, 1] is a learning rate that indicates the degree of overriding the old Q-values. If the value of α is 0, this is means that the agent considers only prior estimates. If the value of α is 1 this is means that the agent considers only the current knowledge to explore opportunities. The action-value function $Q(a, s)$ and state-value function $V(s)$ are similar to each other and can be used to measure the advantage of an action a. The advantage $A^{\pi}(s, a)$ of action a in state s under policy π is defined as

$$A^{\pi}(s, a) = Q^{\pi}(s, a) - V^{\pi}(s) \quad (6.7)$$

If the advantage is positive, it is better to take action a than following the action chosen by the policy in state s. Similarly, a negative advantage means that action a is worse than following the action chosen by the policy.

Q-learning is one of RL's common algorithms that follow the TD method used to learn the Q-function. When Q-learning is performed, a q-table is designed, and the q-values are initialized to zero. The q-table is presented by the shape of $[s_t, a_t]$. Then the q-values are updated and stored at each episode. After the number of iteration, the q-table becomes good enough to be considered a reference for the agent to select the optimum action-based. The details of the Q-learning process are presented in Figure 6-5.

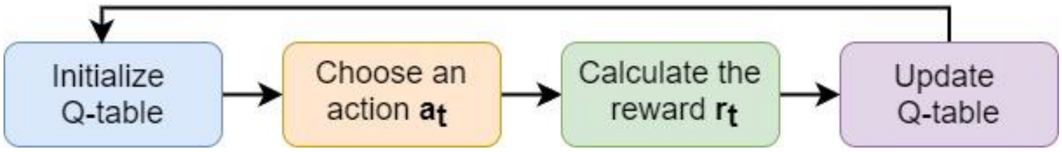

Figure 6-5 Q-learning process

6.3.9 The Exploration - Exploitation Dilemma

A problem that arises when constructing an RL algorithm is how to both be able to exploit a good policy and at the same time explore new policies. If an agent always follows its policy and picks the action it believes is the best, it will never explore new and perhaps better approaches to solve a problem. At the same time, the agent cannot merely explore new behavior all the time since its goal is to maximize future rewards. Therefore, it should also exploit the behavior that works. This is called the exploration-exploitation dilemma [91].

There are different approaches to solve this dilemma, one of which is to have two different policies. One policy is called the target policy, while the other is called the behavior policy. The target policy is used to find the optimal solution, and the behavior policy is used to explore new behaviors [91]. An RL algorithm using a behavioral and target policy is said to be learning off-policy because it can learn from the experiences made from another agent [91]. On the other hand, an RL algorithm that only learns from its own experiences is said to be learning on-policy.

6.4 Reinforcement Learning Categories

There are different classifications of reinforcement learning algorithms. The two main classifications are model-based and model-free algorithms. A model-based algorithm uses the dynamics of the system to plan actions. For instance, the transition function p in equation (6.4) gives a probability distribution over the next state and reward, which can be used for planning in dynamic programming [121]. The RL explicitly uses a model of the environment to choose actions. In this situation, transition function p should be known. When transition function is unknown, it becomes costly to use a model-based RL algorithm.

The second category is called model-free reinforcement learning. As the name suggests, it requires no model or information about the dynamics in the environment. This is useful in situations where no transition function describes the dynamics in the environment, but experiences can be sampled. Model-free algorithms can be divided into two subcategories: Value-based and policy-based. The categories are visualized in Figure 6-6.

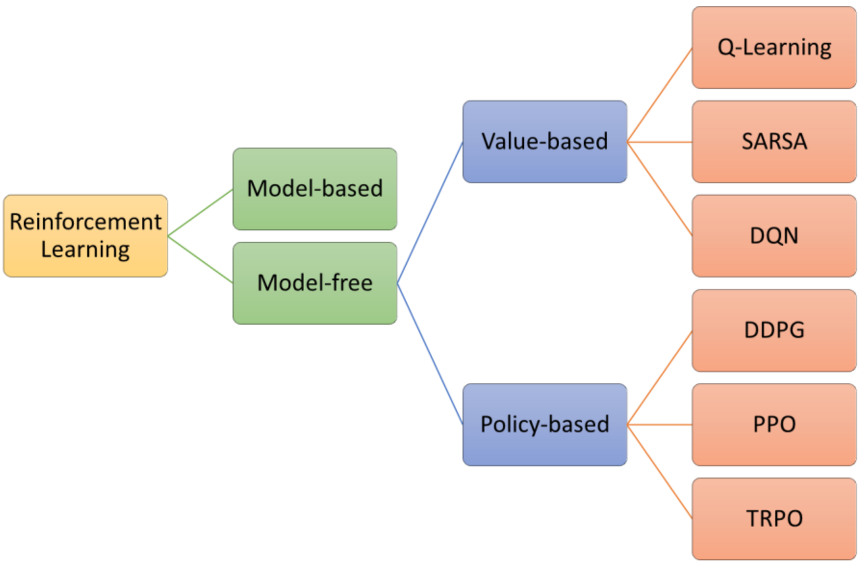

Figure 6-6 RL Categories

The first subcategory of the model-free algorithm is called value-based methods, where the approach is to approximate the action-value function Q and use that to take action. Examples of value-based algorithms are Q-learning, Deep Q-Network (DQN), and State Action Reward State Action (SARSA) [121]. An advantage of value-based methods is that they can learn off-policy, for instance, by learning from experts' behavior. Value-based methods are simple as they do not need to store any explicit policy but can learn from the action-value function Q (pick the action with the best Q-value).

The disadvantage is that value-based methods are not well suited for function approximation, such as neural networks, as they tend to be unstable [121].

Policy-based methods (also called policy gradient) directly parametrize the policy function π without involving the action-value function Q in the decision-making. Examples of policy-based algorithms are Deep Deterministic Policy Gradients (DDPG), Proximal Policy Optimizer (PPO), and Trust Region Policy Optimizer (TRPO) [121]. In contrast to value-based methods, they are stable when using function approximation but inefficient. In other words, the weakness of value-based methods is the advantage of the policy-based methods and vice versa. A natural idea is then to combine the two methods into a more robust method. This is called an actor-critic model, a mix of policy-based and value-based reinforcement learning, as illustrated in Figure 6-7. The policy π is called the actor, because it chooses the action to take. The action-value function Q is named the critic because it evaluates the action picked by the actor.

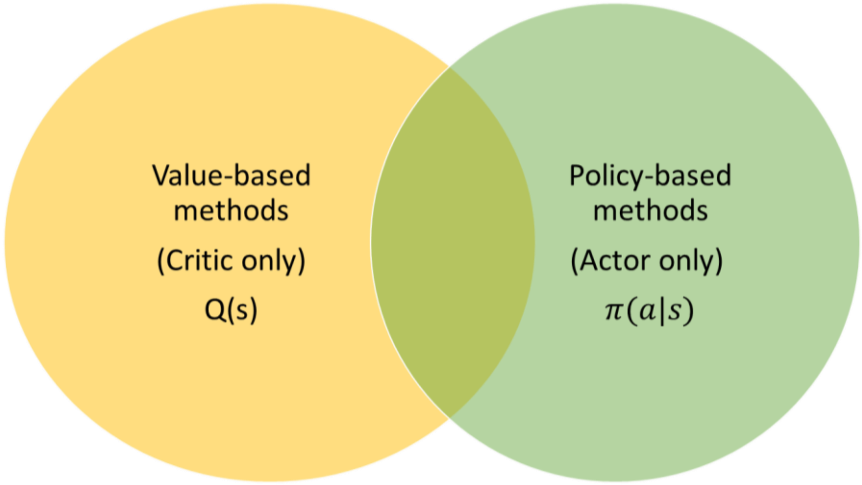

Figure 6-7 Actor-critic in relation to value-based and policy-based methods [122]

6.5 Reinforcement Learning and Deep Neural Networks.

RL has advanced to DRL, where RL is combined with a deep neural network (DNNs). DRL shares the same basic framework with RL. DRL combines artificial neural networks with an RL concept where agents learn the best actions possible in dynamic environments to attain their objectives. The possible output values from the state vector form a vast state space. In the high dimension, the agent is too slow to learn the value of each state individually. The conventional RL algorithms, i.e., Q-learning, become unrealistic when the state and action spaces are in high dimension. DQN is proposed to overcome this problem. The only difference between Q-learning and DQN is the agent's brain. The agent's brain in Q-learning is the Q-table, but in DQN, the agent's brain is a DNN, as shown in Figure 6-8.

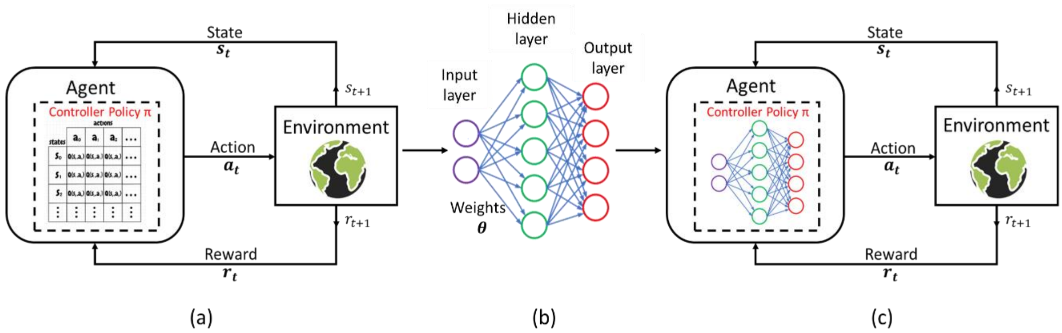

Figure 6-8 Difference between Q- learning and Deep Q-learning

The key role of neural networks in RL is that they are used as function approximators for the policy function and action-value function Q. Formally, the neural network function maps an n-dimensional input space to an m-dimensional output space. The input space can, for instance, be the pixel values of a picture or any other numerical representation of the environment state. The output space of a neural network approximating the policy function is the action space. For instance, in a car driving environment, the input space could be numerical information about the speed, lane position, distance to the closest car, etc., and the output space would have one component each for the acceleration, brake, and angle of the steering wheel.

DNN is organized in different layers, where each layer consists of several nodes or neurons, as visualized in Figure 6-8. The network has a 1-hidden layer architecture, with 5 neurons each. The input features are sent to the network's input layer, consisting of neurons for each input value. All the input neurons are connected with the neurons in the next layer by a scalar weight, represented by arrows in Figure 6-8. The first step for determining a neuron's value in the first hidden layers is by computing a linear combination of all the input features connected to it, and then adding a bias factor. It should be emphasized that the weights and bias are randomly initialized, and the whole point of training a neural network is to find appropriate values for them.

Once the linear combination is computed, it is sent through a non-linear activation function, whose output value will determine the neuron's value. Several activation functions are used, such as the hyperbolic tangent (tanh), the sigmoid function, and rectified linear unit (ReLU). Once all the neurons in a layer are found, the next layer can be calculated with the same process. When DNN is used as an approximator of the value function Q, the training process is all about updating the weights and biases such that the output layer gives the true Q-value for different states and actions. An advantage of using a DNN to approximate the Q-value is that the algorithm can evaluate an action in a new and unseen state. This is useful when the state representation is substantial.

Combing the RL with the black-box nature of DNNs as function approximators is not always valid, and that they do come at a price. Therefore, it is imperative to be aware of the challenges of combing the RL with DNNs to determine whether using DNNs as a function approximator is a safe choice. Theoretically, the combination of RL with DNNs offers a vague algorithm. It provides vast possible functions, making it challenging to find the optimum parameters for the optimization problem. Also, while it has been stated that DNNs can solve and approximate any continuously differentiable function, there is no guarantee that a particular network can learn a particular function approximation [123]. The network may end up stuck in a local minimum, and it may never increase its accuracy over a certain threshold. This makes the DNNs sensitive to their weight's initial randomized values, which leads to a significant limitation of DRL [124]. Table 6-1 summarizes a comparison between RL and DRL approaches.

Table 6-1 Comparison Among RL and DRL

Algorithm	Advantages	Disadvantages	Applications
RL	• Simple and stable • converges to the optimum action-values with probability one.	• The state and action spaces are very limited.	• Applicable for MDPs with a small number of action states.
DRL	• Advanced model of RL, which utilizes DNNs as a universal function approximation method. • Can handle even a vast state space.	• Design and implementation complexity. • Sensitive to their weight's values.	• Applicable to almost all MDPs.

CHAPTER 7 : DEEP REINFORCEMENT LEARNING BASED HOME ENERGY MANAGEMENT SYSTEM

This chapter explains home energy management using deep reinforcement learning (DRL). Section 1 presents the HEMS problem formulation based on the DRL algorithm. Section 2 outlines the home energy management framework and elucidates how a DRL agent can produce optimal solutions in the home energy management system. Finley, Section 3 presents the algorithm implemented process.

7.1 Problem Formulation for DRL-Based HEMS

The energy management in a residential house can be presented as an optimization problem with various devices with different characteristics and environmental changes. Optimal control of these devices is a key element in maximizing their energy efficiency and DR proficiencies. In this section, a single household's load profile is optimized using HEMS DRL-Agent. The appliances' operation is scheduled under a real-time pricing tariff. The objective is to minimize the electricity cost, considering the user comfort level and transformer conditions. An inconvenience price is determined by the user to be considered by the agent to schedule the loads according to user preferences and priorities.

As DRL alternative terminology, a sequential decision-making approach at a timestep of 1 hour is considered. Then a complete episode is defined as one complete day (T=24). In each time step, the HEMS determines the optimum action for appliances, EV, and ESS. For example, a time slot t, the HEMS agent examines the appliances state s_t and chooses the action a_t. Then the agent calculates the reward $r_t(a_t, s_t)$ for taking this action at that state. The details about the formulation of each DRL element are discussed below.

7.1.1 Environment (System Model)

The environment refers to the dynamic energy consumption of household appliances and equipment. It is assumed that the household is equipped with PV source, EV, ESS, and different appliances, divided into three categories: shiftable appliances, controllable appliances, and fixed appliances. Each category has its operational state $s_{n,t}$ which is presented by

$$s_{n,t} = (u_{n,t}, p_{n,t}, l_{n,t}) \tag{7.1}$$

where $u_{n,t} \in \{0, 1\}$ presents the operation status. It takes 1 if the appliance is working or 0 if otherwise; $p_{n,t}$ evaluates the appliances' operational progress, and $l_{n,t}$ measures the time or power constraints. Next, state $s_{n,t}$ is formulated for each type of appliances.

Time-Shiftable Appliances - Their operation time can be shifted, and it has two operational states that use "ON-1 or OFF-0". Assume a shiftable appliance needs a duration d_n of time slots to complete one cycle. Defining the time constraints of the n time-shiftable appliance by $t_{a,n}$ and $t_{b,n}$ where ($t_{a,n}$ > $t_{b,n}$+d_n). Accordingly, the state $s_{n,t}$ of the appliance is defined as

$$(u_{n,t}, p_{n,t}, l_{n,t}) = \begin{cases} (1, k_{t,n}/d_{n,}, t_{b,n} - d_n), t \in [t_{a,n}, t_{b,n}] \\ (0,0,0), \qquad\qquad\qquad otherwise \end{cases} \tag{7.2}$$

where $k_{t,n}$ determines whether to start the operation ($k_{t,n}$= 1) or not ($k_{t,n}$= 0); the operational constraints $p_{n,t}$ measures the required operating duration, and $l_{n,t}$ determines the end of the appliance scheduling window. By the end of the scheduling window $k_{t,n}/d_{t,n}$ should equal to 1 to satisfy the appliance operational constraints.

Controllable Appliances - For this type, the power consumption is adjustable. The state $s_{n,t}$ of these appliances is defined as

$$(u_{n,t}, p_{n,t}, l_{n,t}) = (1, E_{n,t} - E_n^{max}, E_n^{max}), \forall t \tag{7.3}$$

$$E_n^{min} < E_{n,t} < E_n^{max} \tag{7.4}$$

Their power consumption ($E_{n,t}$) can be regulated between the maximum (E_n^{max}) and minimum (E_n^{min}) in response to price changes, as presented in (7.4).

Fixed Appliances - The load of these appliances cannot be reduced or shifted. It can be regarded as a fixed demand for electricity usage. Assuming a fixed appliance n operates in the interval $[t_a, t_b]$ with the rated power, its state $s_{n,t}$ is defined by

$$\left(u_{n,t}, p_{n,t}, l_{n,t}\right) = \begin{cases} \left(1, E_{n,t}, t\right), & t \in [t_a, t_b] \\ (0,0,0), & otherwise \end{cases} \tag{7.5}$$

Electric Vehicle - The EV charging and discharging power can be controlled by satisfying certain constraints. consider the EV arrives at $t_{a,n}$ and departs at $t_{b,n}$, the state $s_{n,t}$ can be defined as

$$\left(u_{n,t}, p_{n,t}, l_{n,t}\right) = \begin{cases} (1, SOE_t, SOE^{max}), t \in [t_{a,n}, t_{b,n}] \\ (0,0,0), \qquad\qquad otherwise \end{cases} \tag{7.6}$$

The dynamics of the EV battery is modeled by

$$SOE_{n,\mathrm{t}+1}^{EV} = SOE_t + \eta^{EV} \cdot E_{n,\mathrm{t}}^{EV}, \qquad t \in [t_a, t_b] \tag{7.7}$$

$$SOE^{max} \leq SOE_t \leq SOE^{max} \tag{7.8}$$

The EV is charging if η^{EV} is positive and discharging otherwise.

Energy Storage System - The ESS state is modeled in a similar way to EV, as presented by (7.6)–(7.8). However, according to the DR program, the ESS is available all day at the house to be utilized (charging/discharging).

7.1.2 State Space Representation

In this example, the state s_t at time slot t include the required information to help the agent schedule the load to meet the problem objectives. For example, the state space may include the appliances states such as appliances load and State of Energy (SoE) of ESS and EV. The considered data is updated each 1 hour. This means that a 1-hour time resolution is provided to the agent. The 1-hour time frame is adequate to describe in detail the use of several household appliances and electricity cost variation. While there are other state representations provided in the literature which are more information-dense. By doing this, the problem's dimensionality is increased, which may increase the learning burden. For example, if the price and appliances data are updated every 10 minutes, we will have 144 steps (T=144). This means that the agent will take more time to learn the policy since it has more states to visit every episode. Therefore, it is crucial to select a time resolution that can describe the system states without losing any important information or increasing the agent's learning burden. Moreover, some studies fund that the gain from more information-dense state representations is marginal, meaning that agents can optimize the policy with reasonable state inputs [125].

The state is a description of the current situation in the MDP, and it is correlated with a set of attributes specific to the appliances, EV, and ESS. The state s_t at time step t is presented as

$$s_t = \left(s_{1,1}, \cdots, s_{N,T}, \lambda_1, \cdots, \lambda_T, E_1^{PV}, \cdots, E_T^{PV}, P_1^{tx}, \cdots, P_T^{tx}\right) \tag{7.9}$$

The attributes considered for the state encapsulates the appliances states ($s_{n,t}$), the electricity price (λ_t), the PV output (P_t^{PV}) and the transformer load (P_t^{tx}). To simplify the model and minimize the computation time, the transformer load is categorized into three levels: low, average, and high, as presented in (7.10). The transformer load in per unit will be used to calculate the hourly LoL%.

$$P_t^{tx} = \begin{cases} P_t^{tx,low}, & if\ P_t^{tx} \leq 0.8\ p.u \\ P_t^{tx,average}, & if\ 0.8\ \leq P_t^{tx} \leq 1\ p.u \\ P_t^{tx,high}, & if\ P_t^{tx} \geq 1\ p.u \end{cases} \tag{7.10}$$

7.1.3 Action Space Representation

Action Space is a combination of all the actions that the agent can decide for each appliance. The action for each appliance depends on the environment state defined in the previous section. The agent should perform the binary action {1, 0 } to turns on or off the time-shiftable appliances, which consume constant energy. The action set for controllable appliances is discretized into 5 levels of energy consumption. Similar to the controllable appliances' actions, the EV and ESS discrete actions are specified with 2 charging levels and 2 discharging levels. The actions are subject to power balance, the physical constraint of devices, demand satisfaction, user preferences constraints, and transformer loading condition. The state a_t is defined as

$$a_t = \left(k_{t,n}, E_{t,n}, E_{t,n}^{EV}, E_{t,n}^{ESS}\right), \forall t \tag{7.11}$$

Where $k_{t,n}$ determines whether to start the shiftable appliance's operation ($k_{t,n}$= 1) or not ($k_{t,n}$= 0); $E_{t,n}$ is the controllable consumption level; $E_{t,n}^{EV}$ and $E_{t,n}^{ESS}$ are the charging/discharging power for the EV and ESS, respectively.

7.1.4 Reward Representation

Unlike the traditional method of setting the dissatisfied function, the proposed method encapsulates customers' satisfaction into the rewards. The agent gradually learns the resident's electricity consumption habits through a continuous interface with the environment. Therefore, it tends to meet user needs when scheduling the appliances. Unlike the fixed discomfort function, the proposed method is more adaptable to the dynamic environment. The reward is formulated considering the electric cost, customer discomfort cost, and transformer LoL cost. The comprehensive reward function for the HEMS is the inverse of (12).

$$r_t = C_t^{elec} + C_t^{cdc} + C_t^{LoL}, \ \ \forall t \tag{7.12}$$

Where C_t^{elec} is electricity cost measured in \$, C_t^{rdc} is an index of discomfort caused to the customers by the DR scheduling measured in \$ and C_t^{LoL} presents the transformer degradation cost measured in \$.

Electricity cost- The electricity operational cost is calculated by

$$C_t^{elec} = \lambda_t \,(E_t^g - E_t^{PV}),\ \ \forall t \tag{7.13}$$

Where E_t^g is the total energy consumed from the grid by the household and λ_t is the electricity market price and charged with the RTP tariff. First, the HEMS will share the PV output sequentially to the home appliances. The fixed appliances are served first since they always have constant load and ensure the residents' convenience. After that, the surplus energy E_t^{PV} is delivered based on the importance coefficients (ζ_n) of the remaining appliances. Finally, If the home demand is higher than the solar energy generation, the HEMS will purchase energy from the grid with electricity market price λ_t .

Customer Discomfort Cost- The resident discomfort cost (RDC) index measures i) undesired operation cost for the time-shiftable appliances, ii) the thermal discomfort of the controllable appliances, and iii) the energy utilization of the EV and ESS.

Equation (7.14) reflects the customer discomfort cost in the scheduling program. It presents the resident's discomfort due to waiting for the appliance to start the operation. The importance factors ζ_1 maps the discomfort level into money, and it is measured in \$/kWh.

$$C_t^{rdc} = \zeta_1\left(t_{start,n} - \ t_{a,n}\right) \tag{7.14}$$

Nonetheless, power reduction can cause thermal discomfort for the resident. Thermal discomfort is determined based on the deviation $\left|E_{n,t} - E_n^{max}\right|$, as presented in (7.15). When the deviation $\left|E_{n,t} - E_n^{max}\right|$ decrease, the thermal discomfort value decreases. If the deviation becomes large, the resident thermal discomfort increases. ζ_2 is the appliance's importance parameter and introduced to map the discomfort terms into money.

$$C_t^{rdc} = \zeta_2 \left| E_{n,t} - E_n^{max} \right| \quad (7.15)$$

The EV model supports the engagement of customers in the energy market. Charging and discharging the EV battery is controlled based on the electricity price and battery health condition, and customer comfort. When EV available in the household, the charging action is rewarded when the electricity price is low. Conversely, the discharging action is rewarded when the electricity price is high. Besides, the model penalizes not having a fully charged EV at departure time, affecting customer comfort. In (7.16), the squared term measures the EV anxiety range for the uncharged battery energy in $/kWh^2.

$$C_t^{rdc} = \zeta_3 (SOE_t - SOE^{max})^{\,2} \quad (7.16)$$

In this case, ESS energy underutilization is considered to prevent undercharging or overcharging of the battery, as presented in (7.17). For example, the discomfort cost will increase when the accumulated energy in the battery (SOE_t) more than the maximum value (SOE^{max}) or less than the minimum value (SOE^{min}).

$$C_t^{rdc} = \begin{cases} \zeta_4 (SOE_t - SOE^{max})^{\,2} if \;\; SOE_t > SOE^{max} \\ \zeta_4 (SOE_t - SOE^{max})^{\,2} if \;\; SOE_t < SOE^{min} \end{cases} \quad (7.17)$$

Transformer LoL cost- Ambient temperature and transformer load are the main cause of distribution transformer failure, as they affect the aging of the transformer insulation, and consequently, the transformer lifetime. In this part, the transformer thermal model is considered to calculate the percent LoL, as presented in chapter 3 equations (3.3)-(3.12). After that, the LoL of transformers is multiplied by the capital cost of transformers to calculate LoL degradation cost (C_t^{LoL}), as presented in equation (3.41). In this strategy, C_t^{LoL} is considered in the reward function as a penalty cost measured in $ to maximize the utility profit. Based on the price signal and transformer loading levels received from the utility, the agent intends to schedule the load during the lowest price period without overloading the distribution transformer during these periods. This can improve the transformer utilization during abnormal and emergencies and prolong the transformer life by minimizing its LoL.

7.2 DRL-Based Framework for HEMS

In this section, the DQN algorithm is utilized to solve the previously formulated home energy management problem. The agent's objective is to maximize the expected average or cumulative reward over an episode. It should be noted that the objective of minimizing electricity cost contradicts that of maintaining the desired comfort level, and the reward function attempts to balance the two objectives, considering the distribution transformer condition. During the operation of HEMS for i time steps, we want to maximize the accumulative reward $R = \sum_{t=0}^{t=i-1} \gamma \cdot r_t$. The optimal value $Q^*(s, a)$ is used to represent the maximum accumulative reward which can be obtained by taking action a_t in state s_t. $Q^*(s, a)$ can be calculated iteratively as presented in Equation (6.6). As shown in Figure 7-1, the input of DNN is the environment states s_t including electricity price, transformer load, and appliances states defined previously along with PV generation. The output is Q-value for each action a_t.

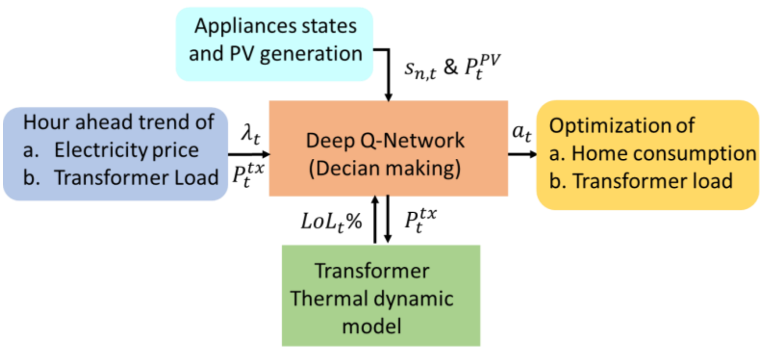

Figure 7-1 The proposed DRL-based HEMS structure

After the model is completely trained using the off-line database, it can be deployed to make optimized decisions in a real physical environment. If there are emergencies, the agents will interact with the new environment. By adjusting actions, the agent gradually increases the obtained reward and restores the optimization effect. Figure 7-2 shows the proposed DRL-based framework for HEMS. The details of appliances state transition are defined in the previous section. The details of the DRL learning and implementation process are presented in the following subsection.

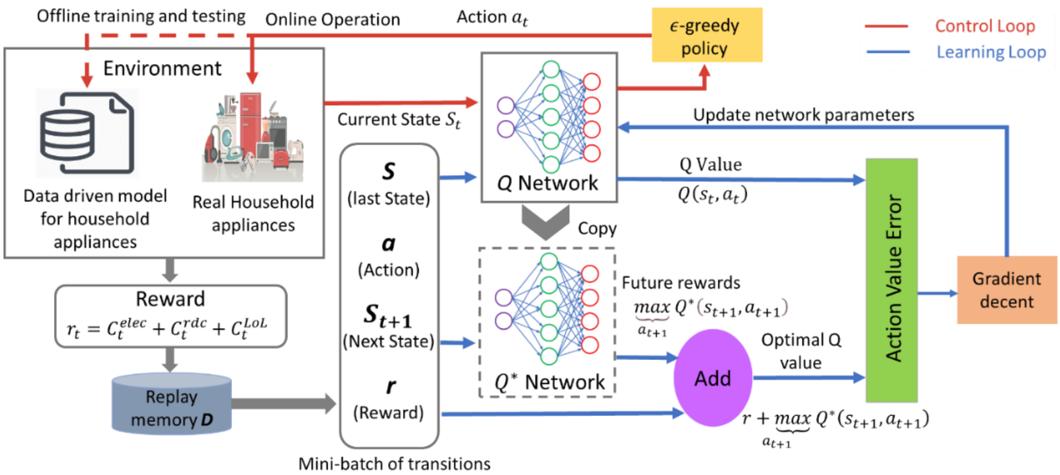

Figure 7-2 The proposed DRL-Based HEMS based framework

7.3 DRL-Based HEMS Implementation Process

The implementation process of the DRL-based HEMS is presented in Algorithm 1. The outer *for loop* determines the episodes number during the training, while the inner *for loop* performs load scheduling and power consumption management at each time step for one episode. First, the reply memory $\boldsymbol{D}$ and neural network weights $\boldsymbol{\theta}$ are initialized. As presented in (6.6), updating the Q value network requires the optimal value. From (6.5), it can be observed that the optimal value depends on the Q value. To break this dependency loop, in line 4, a copy of the neural network weights, $\boldsymbol{\theta}^{-}$**,** is designed to calculate the optimal Q value, $Q^{*}(s, a)$.

At each training episode, line 5, the states are observed, and the algorithm performs training and observes the new environment states (line 6 to 15). Specially, in the initial time slot, i.e., $t = 1$, the agent takes action, obtains a reward, and observes the next state (lines 7 to 9). The reward is calculated by (7.12). Next, in line11, the state transitions are stored in replay memory. After that, a random mini-batch of the state transition is selected from the memory. Then the neural network Q is updated. Next, the Q network is employed to determine the next action. This procedure is repeated until the agent reaches a terminal state. Finally, the agent will learn the optimal actions for hour-ahead, i.e., $t = 1, 2,3, \ldots, 24$.

The ϵ-greedy policy is used as an action selection strategy to select the optimum action. As the training starts, the agent can either discover the action space by randomly choosing an action with a probability of ϵ or choose the action which has the maximum Q-value, with probability 1 – ε. After each iteration, the exploration rate ϵ will be decreased by a decay rate until it reaches its minimum value ϵ_{min}. In this way, the agent has more probability of selecting different actions in the first couple of training episodes. As the training process advances, the agent will have a higher probability to apply the learned policy.

Algorithm 1 DQN learning algorithm

Inputs: appliances states, electricity price and transformer load

Initialize replay memory $\boldsymbol{D}$

Initialize action-value function $\mathbf{Q}$ with random weight $\boldsymbol{\theta}$

Initialize target action-value function $\boldsymbol{Q}^*$ with weights $\boldsymbol{\theta}^- = \boldsymbol{\theta}$

For episode =1 to $\boldsymbol{M}$ **do**

  Receive initial state $\boldsymbol{s_t}$

  For $\boldsymbol{t = 1}$ to $\boldsymbol{T}$ **do**

  Following an $\boldsymbol{\epsilon}$-greedy policy, selects $\boldsymbol{a_t} = \begin{cases} \text{a random action} & \text{with probability } \epsilon \\ \arg\max_a Q(\emptyset(s_t), a; \theta) & \text{otherwise} \end{cases}$

  Execute action $\mathbf{a_t}$, Obtain reward $\boldsymbol{r_t}$ and observe the next state $\boldsymbol{s_{t+1}}$

  Set $\boldsymbol{s_{t+1}}$=$\boldsymbol{s_t}, \boldsymbol{a_t}, \boldsymbol{x_{t+1}}$ and preprocess $\emptyset_{\boldsymbol{t+1}} = \emptyset(\boldsymbol{s_{t+1}})$

  Store transition <$\emptyset_t$, a_t, r_t, $\emptyset_{t+1}$> in $\boldsymbol{D}$

  Sample ransom mini-batch of transitions <$\emptyset_{\boldsymbol{j}}$, $\boldsymbol{a_j}$, $\boldsymbol{r_j}$, $\emptyset_{\boldsymbol{j+1}}$> from $\boldsymbol{D}$

  Set $\boldsymbol{y_j} = \begin{cases} \boldsymbol{r_j} & \textbf{if episode terminates at } \mathbf{j+1} \\ \boldsymbol{r_j} + \boldsymbol{\gamma} \underbrace{\boldsymbol{max}}_{\boldsymbol{a_{t+1}}} \boldsymbol{Q}^*\left(\emptyset_{\boldsymbol{j+1}}, \boldsymbol{a_{t+1}}; \boldsymbol{\theta}^-\right) & \textbf{otherwise} \end{cases}$

  Perform a gradient descent step on $(\boldsymbol{y_j} - \boldsymbol{Q}(\emptyset_{\boldsymbol{j}}, \boldsymbol{a_t}; \boldsymbol{\theta}))^2$ **w**.r.t. the network parameter θ

  end for

end for

Output: Q action value function (from which we obtain policy and select action)

CHAPTER 8 : RESULTS AND DISCUSSION: DEEP REINFORCEMENT LEARNING APPROACH

This Chapter provides simulation results and performance analysis of the proposed DQN algorithm for optimizing power consumption in the smart household.

8.1 Case Study Setup

In this paper, appliances with major contributions in terms of energy consumptions are considered in the model to study their DR and optimize their operation over a period of time to minimize the cost and optimize the transformer load curve, as presented in Table 8-1. We consider three time-shiftable appliances: a washing machine (WM), a clothes dryer (CD), and a dishwasher (DW); three controllable appliances: air conditioners (AC1, AC2, and AC3). Other appliances are in-home use, such as electric kettles, laptops, microwaves, etc. These appliances are interactive and depend on users (Fixed loads) and their load compared to the major loads is insignificant. Thus, they have little scheduling flexibility. Therefore, it is considered as “other” loads.

Also, apart from these loads, EV load is considered. The EV maximum charging power rate is 3.3 kW, and the battery rating is 16 kWh. The EV takes 4 hours to charge fully at a maximum charging rate of 3.3 kW and immediately tapers off to zero. The home is also equipped with ESS, which charges from the PV source and discharges during high price periods. The ESS capacity can be varied between 0 and 6 kWh according to the user needs. The EV and ESS parameters are shown in Table 8-2. For this work purpose, the PV source is designed to meet about 10% of home demand for 24 h. The PV output power is utilized when the PV generation is greater than appliance consumption. When the PV output is less than the home's demand, the HEMS purchases power from the grid.

During the training phase of the agent, different factors and parameters are considered to make the agent more adaptive to any changes that can be made by the user. First, a user's random choice for operating appliances is considered. For example, while training, the starting time for shiftable appliances is randomly selected between 6:00 AM – 10:00 AM and between 6:00 PM- 12:00 AM and EV departure time between 5:00A- 8:00 AM and arrival time 2:00 – 7:00 PM. Also, these appliances' operating duration and their power consumption are randomized to depict users' diversity. This will allow users to change their preferences on appliances from time to time as their interests, needs, and external conditions also change. Besides, a 25-kVA distribution transformer is considered to investigate the transformer's impact on the HEMS operation, based on the IEEE standard C57.91-2011.

A training dataset is used to train the agent, and a different testing dataset is used to evaluate and test the learned knowledge by the agent. In the training phase, electricity price data [126] from January 1st, 2017 to November 30, 2017, are used. In the test phase, the electricity price data from December 1st, 2017 to December 31st, 2017, is used. During the test stage, we randomly select one day (24 h) from the testing dataset, and a different set of appliances data are generated. We start using a network consisting of two parallel input layers (states and actions), 3 hidden layers (36 neurons, 36 neurons, and 36 neurons), and one output layer. The DQN agent's hyper-parameters are presented in Table 8-3. The training process is performed on a computer with Intel Core i9-9980XE CPU @ 3.00 GHz using MATLAB R2020b. The MATLAB Reinforcement Learning Toolbox is used to implement the RL code [127].

Table 8-1 Parameters of household appliances (DRL case)

ID	ζ_n	Power rating (kWh)	$[T_{n,a}, T_{n,b}]$	$T_{n,total}$
DW	0.2	1.5	7-12	2
WM	0.2	2	7-12	2
DRY	0.2	1.2	9-14	2
AC1	2	0.7-2	0-24	-
AC2	2.5	0.7-2	0-24	-
AC3	3	0.7-2	0-24	-
Other	-	-	-	-

Table 8-2 EV and ESS parameters (DRL case)

Type	ESS	EV
Charging Level	[0.3, 0.6]	[1.5, 3]
Maximum energy of Charging/Discharging (kWh)	3.3	0.6
Minimum discharging Level (%)	40	30
Maximum charging Level (%)	90	90
Initial SOE (%)	90%	50%
ζ_n	2.5	2.5

Table 8-3 Network hyper-parameters

Description		Value
Maximum iterations (Episodes)		4000
Time steps in each iteration		24
Learning rate		0.01
Discount factor		0.955
Epsilon	Maximum	1
	Decay	0.01
	Minimum	0.1

8.2 Performance of DQN Algorithm

Due to the algorithm's characteristics, the agent learns to adapt gradually to the environment and gradually obtain more rewards. The episode and average rewards during the training process are presented in Figure 8-1. Initially, there are many random choices; after many iterations, the agent learns to choose the converging trend and possibilities close to the optimization objective as can be noted that the total rewards rise gradually and converge after 1500 iterations. As can be seen, after convergence, there is still a small variation in the rewards. The reason is that the proposed method adopts days as the training episodes and hours as time step and the different days have significant differentiation in electricity cost. The user can obtain economic benefits by controlling the three air conditioners, washing machine, clothes dryer, dishwasher, and the charging and discharging of the EV and ESS batteries.

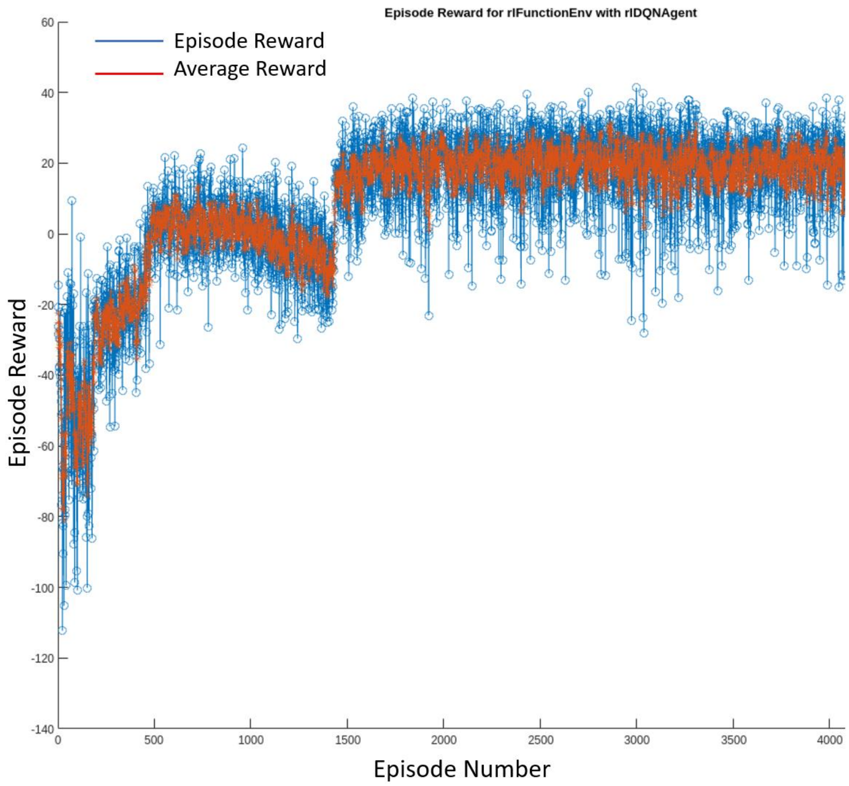

Figure 8-1 Episode and average rewards during the training process

8.3 Algorithmic Robustness

After training using the applied algorithm, the agent can adapt to the dynamic environment and complete the optimization problem. The following subsection demonstrates the trained agent's testing and evaluation to analyze the smart home's optimal load management considering different scenarios.

8.3.1 Scenario-1: Minimizing Electricity Cost and Resident Discomfort Cost

The HEMS agent is evluated for different objectives. The First objective is to reduce the electricity cost ($C_t^{Elec.}$) and resident discomfort cost (C_t^{rdc}) are considered as presented in (7.11). In this scenario, the transformer LoL cost (C_t^{LoL}) is set to 0 in the reward function. Figure 8-2 shows the aggregated household load for with and without DR along with the RTP signal for one test day. As can be seen, the overall load peak is reduced compared to the conventional consumption, where no optimization is applied.

The agent's dynamic behavior can be observed from Figure 8-3. The HEMS tends to purchase more energy when the prices are low in the time 05:00-12:00 and 21:00-24:00. When the PV generation rises, the purchased energy is reduced since the PV energy is utilized. Also, the HEMS tends to reduce the price by cutting the purchased energy by effectively schedule the charging and discharging times of the EV. For example, the EV supplies the household needs through Vehicle-to-home (V2H) mode during 15:00 and 19:00 time periods, where the electricity cost is relatively high and charging during the time 21:00-24:00, where the electricity price is low. Moreover, the ESS captures the environmental changes and starts charging when the electricity price is low, and the PV energy generation rises and discharges when the electricity price is high.

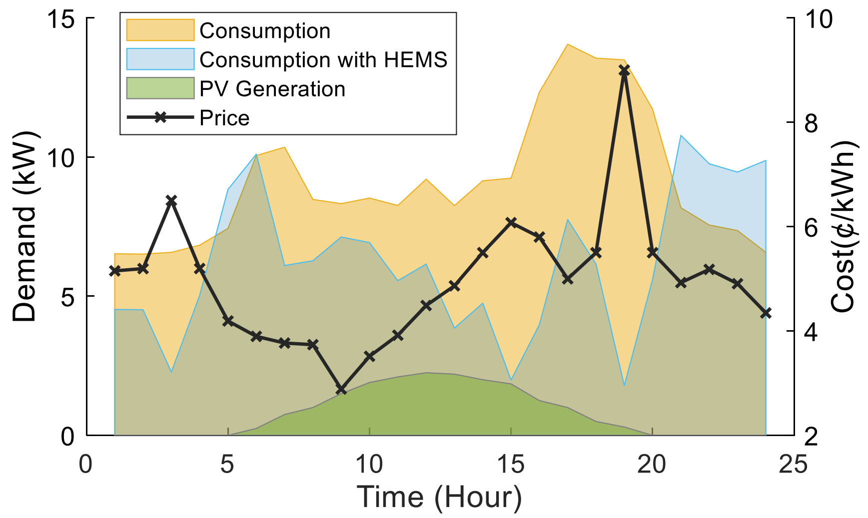

Figure 8-2 Optimization results using DQN (one test day)

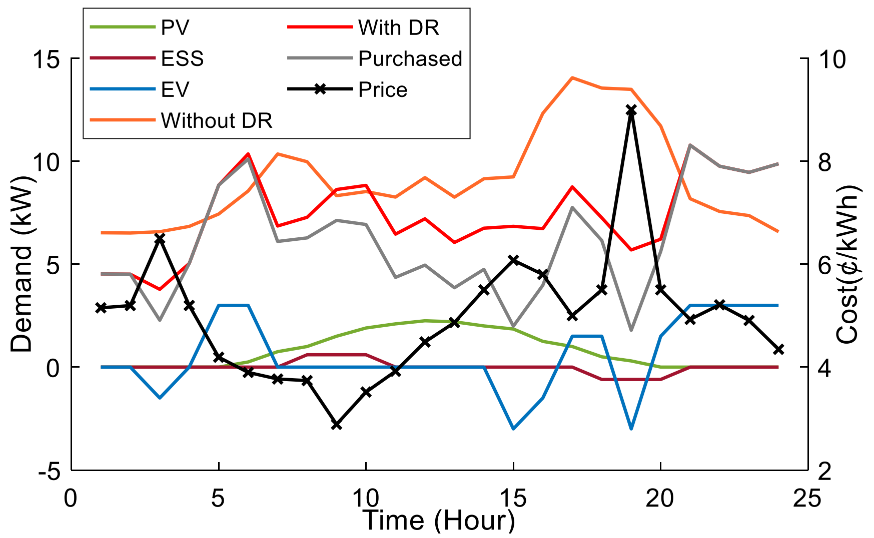

Figure 8-3 The optimization learned by HEMS

The schedule of the time-shiftable appliances is presented in Figure 8-4. The figure shows that all the time-shiftable appliances are scheduled to work during their preferred scheduling interval when the prices are low. Moreover, the behavior of the considered controllable appliances (AC1, AC2, and AC3) in each time slot for one test day is highlighted in Figure 8-5. It can be observed that the consumption of appliances is high during the time 05:00-12:00. After that, the consumption is reduced due to the increase in electricity cost at time 15:00. As the electricity price reaches its maximum at 19:00, each appliance's energy consumption is reduced to its specified minimum operation value, as presented in Table 8-1. Finally, from time 21:00 to time slot 24:00, the energy consumption starts to increase since the electricity price decrease.

The resident comfort level is reflected in the reward function, considering the ζ_n parameter, as a penalty cost depending on the appliances' importance. The controllable appliances are taken as an example to demonstrate the effectiveness of the ζ_n. Different values of ζ_n is given for AC1, AC2, and AC3, as presented in Table 8-1. Also, it can be observed that AC1 is always consuming more energy than AC2 and AC2 is consuming more energy than AC3. This is due to the different values of ζ_n where the appliances with a high value of ζ_n consumes more energy to minimize the penalty (discomfort cost).

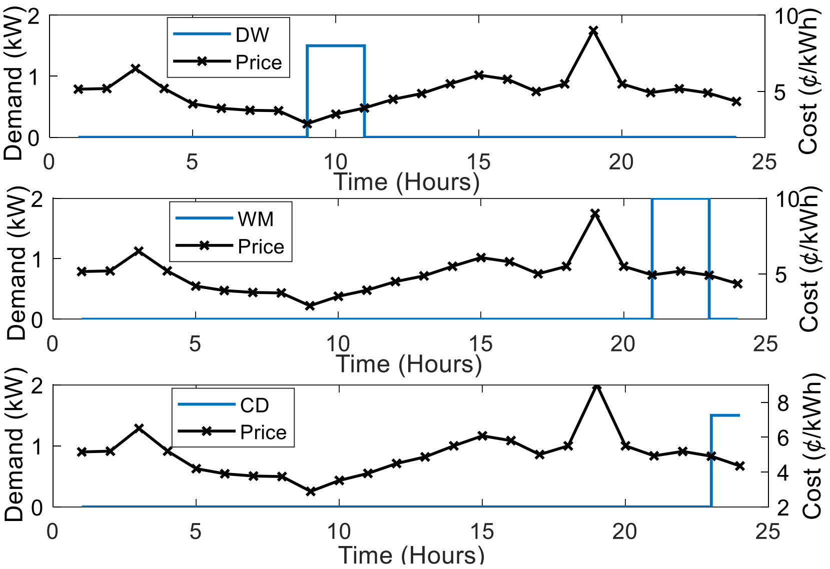

Figure 8-4 Scheduling of time-shiftable appliances during the test day

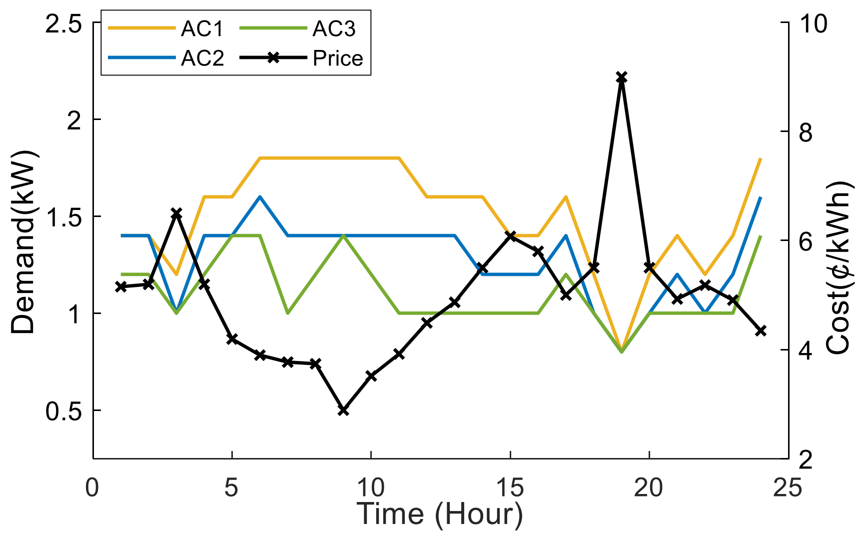

Figure 8-5 Energy consumption controllable appliances during the test day

8.3.2 Scenario -2:Minimizing Transformer LoL Cost

Having individual control of the DR algorithm per individual end-user may cause the low-cost periods to operate as a sink for all customers to operate their appliances during these intervals and generate new load peaks detected by utility assets. This may increase the load and LoL factor during these periods compared with the reference case's low-price periods. Thus, the transformer LoL cost is considered in the reward along with operation cost and resident discomfort cost, as presented (7.12). The HEMS learned an approach to fulfill the resident's expectations and decrease the transformer's LoL by monitoring transformer load, electricity price, and appliances states.

The transformer load profile for one test day is presented in Figure 8-6. The residential load curves for different customers are aggregated to determine the load on the transformer. To evaluate the proposed DRL algorithm's efficacy, the load profile is considered with a severe loading where the transformer loading condition shows 4-h continuous overloading above 100% rating during the 19:00–22:00 time interval and another 10 h above 80% rating. Figure 8-7 shows the optimization learned by the HEMS agent considering LoL cost in the reward function. It can be noted that the EV charging load is shifted to slots 4:00–6:00 of high electricity price compared to slots 21:00–24:00 used in scenario 1. While this load shift may increase the electricity cost, it satisfies the LoL cost and overcomes the distribution transformer's overload condition and other assets. This results in a peak demand reduction of 24% compared to scenario 1, positively impacting the utility assets. The new load profile has a total electricity cost of USD 5.32, which is higher than the first scenario with 5.28; however, it is lower than when no DR is applied and provides direct benefits to both the end-user and utility operator. Table 8-4 presents a comparison of electricity costs in these scenarios.

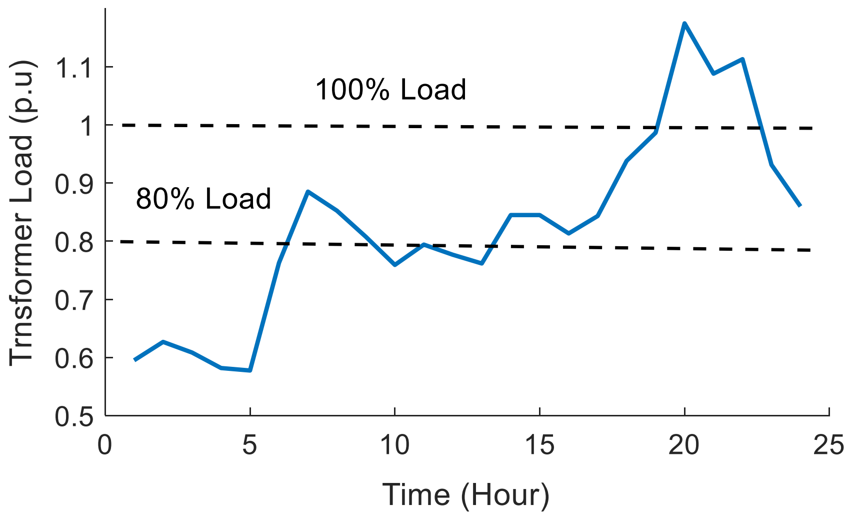

Figure 8-6 Transformer load profile

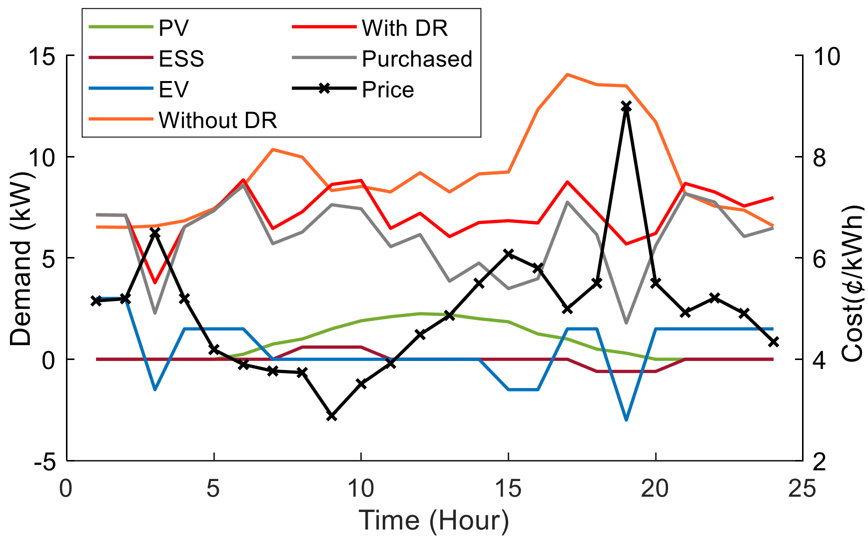

Figure 8-7 The optimization learned by HEMS considering transformer LoL cost in a typical test day

Table 8-4 Comparison of electricity cost for different scenarios

ID	Electricity Cost ($)		
	Without DR	DR Scenario-1	DR Scenario-2
DW	0.22	0.20	0.20
WM	0.17	0.139	0.139
CD	0.11	0.09	0.09
AC1	2.38	1.72	1.69
AC2	2.38	1.48	1.42
AC3	2.38	1.23	1.19
EV	0.92	0.43	0.59
Total	8.560	5.289	5.319

8.3.3 Comparison with Traditional Algorithms

To evaluate the proposed algorithm, a conventional optimization is considered a benchmark, in which the formulated problem is solved by MATLAB optimization Toolbox using MILP. It is assumed that all the required information is known for the MILP algorithm to minimize the electricity and discomfort costs considering transformer LoL as defined in Equations (7.12) -(7.16) with short-sighted actions. On the contrary, the DRL algorithm utilizes the learning knowledge to choose different actions to maximize the reward. Figure 8-8 demonstrates the total costs (the electricity and discomfort costs) under these two methods. As can be observed, although the MILP achieves less electricity cost, it leads to high discomfort cost. The reduction of the electricity cost ratio of the DQN model is only 2% less than MILP. Nevertheless, DQN achieves better cost reduction than MILP for thermal discomfort and EV anxiety ranges by 43% and 75%, respectively. This is because the DRL agent has learning capabilities and accounts for both the current reward and the future rewards, while the MILP has no learning capability.

Theoretically, it should be highlighted that the DRL has more advantages and more robust in the actual dynamical environments. However, the DRL algorithm may lose its robustness if the environment is subjected to major changes. For example, if the residents change their habits or the electricity tariff is shifted significantly away from the training regime. The HEMS performance will drop off as the agent finds itself facing new states outside of its prior knowledge. The agents will take time to restore the optimization effect. This can be avoided by effectively train the agent on different datasets.

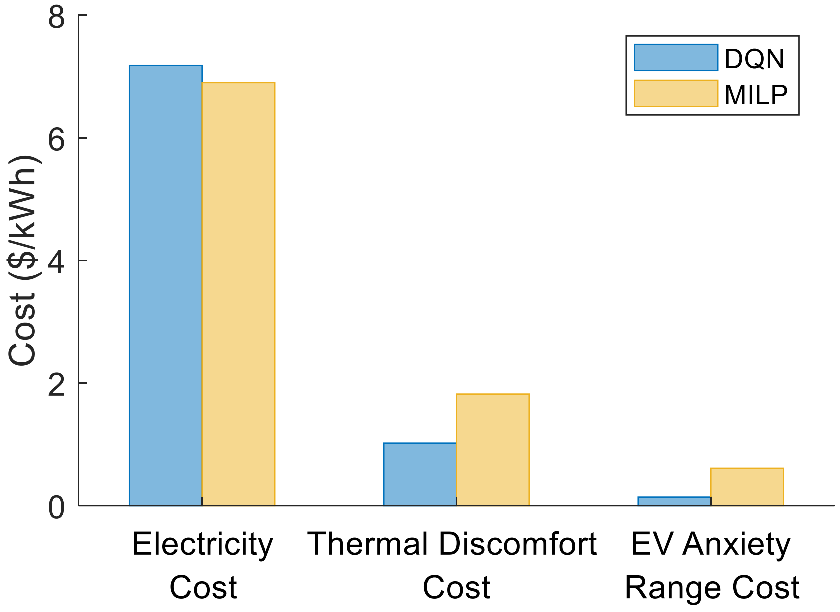

Figure 8-8 Total costs of DQN algorithm compared to MILP solver

CHAPTER 9 : CONCLUSION AND FUTURE RESEARCH

9.1 Conclusion

The thesis proposes a multi-objective HEMS model that coordinates the benefits of households and utility operators. The HEMS model seeks to minimize the individual electricity consumption cost while considering the customer’s comfort and lifestyle. The utility distribution transformer's load profile is integrated into the optimization model by incorporating the asset loss of life (LoL) cost in the multi-objective function. The proposed model's flexibility is supported by considering various demand/generation components that consume/produce electricity.

First, the problem for time shiftable loads is formulated as a mixed-integer linear programming (MILP) with the decision variable for power ON and OFF of an appliance is binary. The controllable power loads are formulated as interior-point optimization (IPO) for different power levels. The household appliances' number can be increased by considering more appliances in the network. This application is presented by three operation scenarios, compared with a reference case, and illustrated customer's and operator's benefits in terms of consumption cost, customer dissatisfaction cost, and transformer asset load leveling. The multi-objective DR model results revealed a 38% reduction in the electricity usage cost and an 18% reduction in the distribution transformer's aggregated peak demand.

Although this approach has shown a reduction in electricity cost and the aggregated peak demand of the distribution transformer, some limitations exist, as presented in chapter 5. First, the optimization problem developed in this approach is based on a static model and mathematical equations. It would be more applicable to have a dynamic presentation of the optimization problem. This approach may also suffer from high complexity and computational burden in real-time applications due to the significant number of involved variables. Additionally, this approach is case-specific and needs to be adjusted when the system environment changes during abnormal situations.

Therefore, a model-free DRL algorithm is proposed to solve the optimization problem. Different elements that drive DRL models' performance and implementation details, are identified and examined. A DQN agent is then utilized to schedule the household appliances considering customer comfort, hour-ahead electricity price, and transformer condition. The proposed algorithm's performance is validated through simulations. The results show that, compared with a conventional optimization approach, DRL is more efficient in minimizing the energy cost while adapting to the desired user’s comfort level. The proposed algorithm's application is presented by two operation scenarios, compared with a reference case, and illustrated customer's and operator's benefits in terms of consumption cost, customer discomfort cost, and transformer asset load leveling. The results show that DRL is an adequate method to address the HEMS problem since it successfully minimized both the electricity and dissatisfaction costs for a single household user.

In conclusion, optimizing residential loads provide significant financial benefits to the end-user and the utility. Encouraging consumers to participate in DR programs by providing incentives and ensuring consumers’ comfort plays a vital role. Therefore, it is required to have a proper incentive to motivate the consumers without affecting the utility required minimum savings. Although this thesis has achieved its objectives, there were some unavoidable limitations. For example, some of the model parameters were based on assumptions for theoretical research. In real-world applications, it is required to select the parameters based on proper validation processes. Also, this thesis has not performed an extensive tuning of the model, but instead used default values for hyperparameters and focused on the results that emerge from them. Therefore, it is natural to assume that improvements can be found, for instance, by increasing the training time or pretraining the DRL algorithm.

9.2 Future work

In future work, peer-to-peer energy trading and two-way energy trading with the grid will be considered an extension of the current work, further improving the HEMS's economic advantage. The Formulated HEMS framework is based on 1-hour time resolution. Therefore, it is not able to tackle problems in real-time. A finer time frame, for instance, every 10 minutes, could be implemented. Thus, the demand response program operates close to real-time.

The adaptiveness of the DRL model can provide significant benefits to the utility during abnormal conditions. However, further research is still needed and the algorithm and implementation strategy should be discussed and analyzed. Also, special consideration should be given to comparing model-driven methods and data-driven optimization methods because they have advantages in different scenarios.

Moreover, there are many hyperparameters that can be tuned in DRL, as in most machine learning algorithms. This thesis has not performed an optimal hyperparameter search. In future work an extensive tuning for the model can be done. For example, a key parameter that should be tested more thoroughly is the training time, as neural network networks are good at continuously learning from new data, and the action space is large. The presented results for DRL are only related to the DQN algorithm implemented and described in this thesis and do not necessarily represent other types of architectures or RL agents. The formulated optimization problem can consequently be applied and tested to different DRL algorithms.

REFERENCES

[1] Z. Fan *et al.*, "Smart Grid Communications: Overview of Research Challenges, Solutions, and Standardization Activities," in *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 21-38, First Quarter 2013, doi: 10.1109/SURV.2011.122211.00021.

[2] P. Palensky and D. Dietrich, "Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads," in *IEEE Transactions on Industrial Informatics*, vol. 7, no. 3, pp. 381-388, Aug. 2011, doi: 10.1109/TII.2011.2158841.

[3] H. Farhangi, "The path of the smart grid," in *IEEE Power and Energy Magazine*, vol. 8, no. 1, pp. 18-28, January-February 2010, doi: 10.1109/MPE.2009.934876.

[4] J. Corbett, K. Wardle and C. Chen, "Toward a sustainable modern electricity grid: The effects of smart metering and program investments on demand-side management performance in the US electricity sector 2009-2012," in *IEEE Transactions on Engineering Management*, vol. 65, no. 2, pp. 252-263, May 2018, doi: 10.1109/TEM.2017.2785315.

[5] A. Ghosal and M. Conti, "Key Management Systems for Smart Grid Advanced Metering Infrastructure: A Survey," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2831-2848, third quarter 2019, doi: 10.1109/COMST.2019.2907650.

[6] W. Li *et al.*, "Demand Response Management for Residential Smart Grid: From Theory to Practice," in *IEEE Access*, vol. 3, pp. 2431-2440, 2015, doi: 10.1109/ACCESS.2015.2503379.

[7] J. Runyon "Demand Response vs. Demand Optimization-How to Optimize Demand for a More Efficient Grid," *POWERGrid International*, 03-Sep-2019.

[Online]. Available: https://www.power-grid.com/2013/12/05/demand-response-vs-demand-optimization-how-to-optimize-demand-for-a-more-efficient-grid/. [Accessed: 04-Oct-2020].

[8] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A Survey on Demand Response Programs in Smart Grids: Pricing Methods and Optimization Algorithms," *in IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 152-178, First quarter, 2015.

[9] BP, "BP statistical review of world energy," BP, Tech. Rep., October 2014. [Online]. Available: http://www.bp.com/statisticalreview

[10] H. Shareef, M. S. Ahmed, A. Mohamed, and E. Al Hassan, "Review on Home Energy Management System Considering Demand Responses, Smart Technologies, and Intelligent Controllers," *in IEEE Access*, vol. 6, pp. 24498-24509, 2018

[11] International Energy Agency, "World Energy Outlook", 2016, [Online]. Available: www.ieo.org.

[12] REN21, "Statistical Review of World Energy 2020 | 69th edition" REN21 Secretariat, Tech. Rep. ISBN 978-3-9815934-2-6, 2014. [Online]. Available: https://www.bp.com.

[13] P. Mock and Z. Yang, "Driving electrification a global comparison of fiscal incentive policy for electric vehicles," *The International Council on Clean Transportation (THEICCT)*, Tech. Rep., 2014. [Online]. Available: http://www.theicct.org/driving-electrification-globalcomparison-fiscal-policy-electric-vehicles.

[14] "Energy, Transport and GHG Emissions Trends to 2050: Reference Scenario 2013." [Online]. Available:

https://ec.europa.eu/transport/sites/transport/files/media/publications/doc/trends-to-2050-update-2013.pdf. [Accessed: 04-Oct-2020].

[15] K. Clement-Nyns, E. Haesen, and J. Driesen, “The impact of charging plugin hybrid electric vehicles on a residential distribution grid,” *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 371 –380, Feb. 2010.

[16] R. Hledik and A. Faruqui, “Valuing Demand Response: International Best Practices, Case Studies, and Applications,” 2015, [Accessed 20-July-2018]. [Online]. Available: http://files.brattle.com/files/5766 valuing demand response-international best practices case studies and applications.pdf.

[17] Q. I. Sarhan, "Systematic Survey on Smart Home Safety and Security Systems Using the Arduino Platform," in *IEEE Access*, vol. 8, pp. 128362-128384, 2020, doi: 10.1109/ACCESS.2020.3008610.

[18] U. Zafar, S. Bayhan and A. Sanfilippo, "Home Energy Management System Concepts, Configurations, and Technologies for the Smart Grid," in *IEEE Access*, vol. 8, pp.

[19] “US Home Appliance Industry Issues Principles & Requirements for Achieving a Widely Accepted Smart Grid - News,” *SmartGrid*, 30-Nov-2009. [Online]. Available: https://www.smartgrid.gov. [Accessed: 04-Oct-2020].

[20] “Smart Domestic Appliances in Sustainable Energy Systems - Intelligent Energy Europe - European Commission,” *Intelligent Energy Europe*, 2008. [Online]. Available: https://ec.europa.eu/energy/intelligent/projects/en/projects/smart-0. [Accessed: 04-Oct-2020].

[21] J. Kim, E. Jung, Y. Lee, and W. Ryu, "Home appliance control framework based on smart TV set-top box," in *IEEE Transactions on Consumer Electronics*, vol. 61, no. 3, pp. 279-285, Aug. 2015, doi: 10.1109/TCE.2015.7298086.

[22] Neenan, B., (2008), “Societal benefits of smart metering investments,” *The Electricity Journal*, 21, (pp. 32-45).

[23] I. S. Jha, S. Sen, and V. Agarwal, “Eighteenth National Power Systems Conference,” in *Advanced Metering Infrastructure Analytics -A Case Study (NPSC)*, 2014.

[24] Q. Sun *et al*., "A Comprehensive Review of Smart Energy Meters in Intelligent Energy Networks," in *IEEE Internet of Things Journal*, vol. 3, no. 4, pp. 464-479, Aug. 2016, doi: 10.1109/JIOT.2015.2512325.

[25] S. Bruno, S. Lamonaca, M. L. Scala, G. Rotondo and U. Stecchi, "Load control through smart-metering on distribution networks," *2009 IEEE Bucharest PowerTech*, Bucharest, Romania, 2009, pp. 1-8.

[26] E. Nyholm and D. Steen, “Can demand response mitigate the impact of intermittent supply?” [Online]. Available: http://publications.lib.chalmers.se/records/fulltext/206861/local_206861.pdf. [Accessed: 02-Mar-2021].

[27] J. Torriti, M. G. Hassan, and M. Leach, “Demand response experience in europe: Policies, programmes and implementation,” *Energy*, vol. 35, no. 4, pp. 1575 – 1583, 2010.

[28] H. Yang, L. Wang, Y. Zhang, H. Tai, Y. Ma and M. Zhou, "Reliability Evaluation of Power System Considering Time of Use Electricity Pricing," *in IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 1991-2002, May 2019, doi: 10.1109/TPWRS.2018.2879953.

[29] J. Aghaei and M. Alizadeh, "Critical peak pricing with load control demand response program in unit commitment problem," *in IET Generation, Transmission & Distribution*, vol. 7, no. 7, pp. 681-690, July 2013, doi:

10.1049/iet-gtd.2012.0739.

[30] M. Muratori and G. Rizzoni, "Residential Demand Response: Dynamic Energy Management and Time-Varying Electricity Pricing," *in IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1108-1117, March 2016, doi: 10.1109/TPWRS.2015.2414880.

[31] Z. Liu, A. Wierman, Y. Chen, B. Razon, and N. Chen, “Data center demand response: Avoiding the coincident peak via workload shifting and local generation,” *Performance Evaluation*, vol. 70, no. 10, pp. 770–791, 2013.

[32] L. Zhu, Z. Yan, W. Lee, X. Yang, Y. Fu and W. Cao, "Direct Load Control in Microgrids to Enhance the Performance of Integrated Resources Planning," *in IEEE Transactions on Industry Applications*, vol. 51, no. 5, pp. 3553-3560, Sept.-Oct. 2015, doi: 10.1109/TIA.2015.2413960.

[33] N. Saker, M. Petit, and J. L. Coullon, "Demand side management of electrical water heaters and evaluation of the Cold Load Pick-Up characteristics (CLPU)," *2011 IEEE Trondheim PowerTech*, Trondheim, 2011, pp. 1-8, doi: 10.1109/PTC.2011.6019312.

[34] M. H. Imani, K. Yousefpour, M. T. Andani, and M. Jabbari Ghadi, "Effect of Changes in Incentives and Penalties on Interruptible/Curtailable Demand Response Program in Microgrid Operation," *2019 IEEE Texas Power and Energy Conference (TPEC)*, College Station, TX, USA, 2019, pp. 1-6, doi: 10.1109/TPEC.2019.8662184.

[35] F. Genoese, M. Genoese, and M. Wietschel, “Occurrence of negative prices on the german spot market for electricity and their influence on balancing power markets,” *in Energy Market (EEM)*, 2010 7th International Conference on the European, June 2010, pp. 1–6.

[36] M. H. Amini, J. Frye, M. D. Ilic´, and O. Karabasoglu, “Smart residential energy scheduling utilizing two stage mixed integer linear programming,” *in Proc. North Amer. Power Symp. (NAPS)*, pp. 1-6, 2015.

[37] D. Aigner, “The Residential Electricity Time-of-Use Pricing Experiments: What Have We Learned?,” *Cambridge, MA: National Bureau of Economic Research*, Jan. 1985.

[38] M. M. Collins and G. H. Mader, "The timing of EV recharging and its effect on utilities," in *IEEE Transactions on Vehicular Technology*, vol. 32, no. 1, pp. 90-97, Feb. 1983, doi: 10.1109/T-VT.1983.23948.

[39] B. M. Davis and T. H. Bradley, "The Efficacy of Electric Vehicle Time-of-Use Rates in Guiding Plug-in Hybrid Electric Vehicle Charging Behavior," in *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1679-1686, Dec. 2012, doi: 10.1109/TSG.2012.2205951.

[40] S. M. Baladi, J. A. Herriges, and T. J. Sweeney, “Residential response to voluntary time-of-use electricity rates,” *Resource and Energy Economics*, vol. 20, no. 3, pp. 225–244, 1998.

[41] R. Hartway, S. Price, and C. Woo, “Smart meter, customer choice and profitable time-of-use rate option,” *Energy*, vol. 24, no. 10, pp. 895–903, 1999.

[42] R. Yu, W. Yang and S. Rahardja, "A Statistical Demand-Price Model With Its Application in Optimal Real-Time Price," in *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1734-1742, Dec. 2012, doi: 10.1109/TSG.2012.2217400.

[43] M. Roozbehani, M. A. Dahleh and S. K. Mitter, "Volatility of Power Grids Under Real-Time Pricing," in *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1926-1940, Nov. 2012, doi: 10.1109/TPWRS.2012.2195037.

[44] T. Namerikawa, N. Okubo, R. Sato, Y. Okawa and M. Ono, "Real-Time Pricing

Mechanism for Electricity Market With Built-In Incentive for Participation," in *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 2714-2724, Nov. 2015, doi: 10.1109/TSG.2015.2447154.

[45] A. Mohsenian-Rad and A. Leon-Garcia, "Optimal Residential Load Control With Price Prediction in Real-Time Electricity Pricing Environments," in *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 120-133, Sept. 2010, doi: 10.1109/TSG.2010.2055903.

[46] H. Zhong, L. Xie and Q. Xia, "Coupon Incentive-Based Demand Response: Theory and Case Study," in *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1266-1276, May 2013, doi: 10.1109/TPWRS.2012.2218665.

[47] J. Koonamparampath, M. Sawant, K. Atharva and A. Sheikh, "A Stackelberg Game Theoretic Approach for Optimal Electricity Pricing Dynamics employing Time-of-Use Algorithm," *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, Paris, France, 2019, pp 1628-1633.

[48] N. Bhati and S. Kakran, "Optimal Household Appliances Scheduling Considering Time-Based Pricing Scheme," *2018 International Conference on Power Energy, Environment and Intelligent Control (PEEIC)*, Greater Noida, India, 2018, pp. 717-721

[49] D. Setlhaolo and X. Xia, "Optimal Scheduling of Household Appliances Incorporating Appliance Coordination," *Energy procedia*, vol. 61, pp. 198-202,2014.

[50] N. Venkatesan, J. Solanki, and S. K. Solanki, “Residential demand response model and impact on voltage profile and losses of an electric distribution network,” *Applied Energy*, vol. 96, pp. 84–91, 2012.

[51] A. Roscoe and G. Ault, “Supporting high penetrations of renewable generation via implementation of real-time electricity pricing and demand response,” *IET Renewable Power Generation*, vol. 4, no. 4, pp. 369–382, 2010

[52] H. Falsafi, A. Zakariazadeh, and S. Jadid, “The role of demand response in single and multi-objective wind-thermal generation scheduling: A stochastic programming,” *Energy*, vol. 64, pp. 853–867, 2014

[53] J. Leithon, S. Sun and T. J. Lim, "Demand Response and Renewable Energy Management Using Continuous-Time Optimization," *in IEEE Transactions on Sustainable Energy*, vol. 9, no. 2, pp. 991-1000, April 201

[54] C. O. Adika and L. Wang, "Autonomous Appliance Scheduling for Household Energy Management," *IEEE Transactions on Smart Grid*, vol. 5, (2), pp. 673-682, 2014.

[55] N. G. Paterakis, O. Erdinç, A. G. Bakirtzis and J. P. S. Catalão, "Optimal Household Appliances Scheduling Under Day-Ahead Pricing and Load-Shaping Demand Response Strategies," *in IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1509-1519, Dec. 2015.

[56] O. Erdinc, N. G. Paterakis, T. D. P. Mendes, A. G. Bakirtzis, and J. P. S. Catalão, ‘‘Smart household operation considering bi-directional EV and ESS utilization by real-time pricing-based DR,’’ *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1281–1291, May 2015.

[57] H. Akkermans, and F. Ygge, "Power load management as a computational market" *in 2nd International Conference on Multi-Agent Systems*, Kyoto, pp 393-400, 1996.

[58] G. Deconinck and B. Decroix, "Smart metering tariff schemes combined with distributed energy resources" *in 4th International Conference on Critical*

Infrastructures, Linkoping, pp. 1-8, 2009.

[59] M. Muratori and G. Rizzoni, “Residential demand response: Dynamic energy management and time-varying electricity pricing,” *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1108–1117, Mar. 2016.

[60] Z. Tan, P. Yang and A. Nehorai, "An Optimal and Distributed Demand Response Strategy With Electric Vehicles in the Smart Grid," in *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 861-869, March 2014, doi: 10.1109/TSG.2013.2291330.

[61] O. Sevastyanova, B. Pasalskiy, and B. Zhmud, “Copper Release Kinetics and Ageing of Insulation Paper in Oil-Immersed Transformers,” *Engineering*, vol. 07, no. 08, pp. 514–529, 2015.

[62] A. Elmoudi, "Thermal modeling and simulation of transformers," *2009 IEEE Power & Energy Society General Meeting*, Calgary, AB, Canada, 2009, pp. 1-4,

[63] J. W. Stahlhut, G. T. Heydt, and N. J. Selover, "A preliminary assessment of the impact of ambient temperature rise on distribution transformer loss of life," *IEEE Transactions on Power Delivery*, vol. 23, pp. 2000-2007, 2008.

[64] J. McCarthy, “Analysis of Transformer Ratings in a Wind Farm Environment,” dissertation, 2010.

[65] Standard, "60076-7 Power Transformer-Part 7:‘Loading Guide for Oil-Immersed Power Transformers’," ed: IEC.

[66] J. K. Pylvanainen, K. Nousiainen and P. Verho, "Studies to Utilize Loading Guides and ANN for Oil-Immersed Distribution Transformer Condition Monitoring," in *IEEE Transactions on Power Delivery*, vol. 22, no. 1, pp. 201-207, Jan. 2007, doi: 10.1109/TPWRD.2006.877075.

[67] "IEEE Guide for Loading Mineral-Oil-Immersed Transformers and Step-Voltage

Regulators," IEEE Std C57.91-2011 (Revision of IEEE Std C57.91-1995), pp. 1-123, 2012.

[68] J. Jargstorf, K. Vanthournout, T. De Rybel, and D. Van Hertem, "Effect of demand response on transformer lifetime expectation," *in Innovative Smart Grid Technologies (ISGT Europe)*, 2012 3rd IEEE PES International Conference and Exhibition on, 2012, pp. 1-8.

[69] Y. Liu, L. Xiao, G. Yao, and S. Bu, "Pricing-Based Demand Response for a Smart Home With Various Types of Household Appliances Considering Customer Satisfaction," *in IEEE Access*, vol. 7, pp. 86463-86472, 2019, doi: 10.1109/ACCESS.2019.2924110.

[70] C. S. Teja and P. K. Yemula, "Reducing the Ageing of Transformer using Demand Responsive HVAC," *2018 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, Singapore, 2018, pp. 569-574.

[71] J. Jargstorf, K. Vanthournout, T. De Rybel, and D. Van Hertem, "Effect of Demand Response on transformer lifetime expectation," *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, Berlin, 2012, pp. 1-8.

[72] J. Tian, H. Ren, L. Hu, F. Wang, and H. Feng, "Research on the Influence of Demand Response on the Life of Distribution Network Transformer," *2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, Chengdu, China, 2019, pp. 899-904

[73] M. Humayun, M. Z. Degefa, A. Safdarian, and M. Lehtonen, "Utilization Improvement of Transformers Using Demand Response," *in IEEE Transactions on Power Delivery*, vol. 30, no. 1, pp. 202-210, Feb. 2015

[74] M. Humayun, A. Safdarian, M. Z. Degefa, and M. Lehtonen, "Demand Response

for Operational Life Extension and Efficient Capacity Utilization of Power Transformers During Contingencies," *in IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 2160-2169, July 2015.

[75] S. Shao, M. Pipattanasomporn, and S. Rahman, "Demand Response as a Load Shaping Tool in an Intelligent Grid With Electric Vehicles," i*n IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 624-631, Dec. 2011.

[76] O. A. Sianaki, O. Hussain, T. Dillon, and A. R. Tabesh, "Intelligent Decision Support System for Including Consumers' Preferences in Residential Energy Consumption in Smart Grid," *2010 Second International Conference on Computational Intelligence, Modelling and Simulation*, 2010.

[77] Amer, M., Naaman, A., M'Sirdi, N. K., and El-Zonkoly, A. M. (2014). A hardware algorithm for PAR reduction in smart home. *International Conference in Applied and Theoretical Electricity (ICATE)*. 1-6. IEEE.

[78] Yoo, Y. S., and Lee, I. W. Heuristic Analysis of Energy Information and Preference for Smart Grid Infrastructure. ICTC2014: *The 5th International Conference on ICT Convergence*. 569-571. IEEE.

[79] *Assessment of Demand Response and Advanced Metering* (Tech.). (2008). United States: Federal Energy Regulatory Commission.

[80] Electric Power Research Institute, (2002), "New Principles for Demand Response Planning, EPRI: 1006015", Palo Alto, CA.

[81] Specification for outdoors type three phase oil immersed (including completely self-protected) distribution transformers. (2008, August). Retrieved October 13, 2020, from http://cea.nic.in/reports/others/god/dpd/guidelines_spec_1n3phase_tfs.pdf.

[82] MatlabWorks. (2020). Retrieved October 17, 2020, from

https://www.mathworks.com/help/optim/ug/intlinprog.html?s_tid=srchtitle.

[83] Uppsala University, "Parallel Global Optimization of ABB's metal process models using Matlab", 2014.

[84] MatlabWorks. (2020). Retrieved October 17, 2020, from https://www.mathworks.com/help/optim/ug/fmincon.html?s_tid=srchtitle.

[85] W. B. Shi, N. Li, C. C. Chu, and R. Gadh, "Real-time energy management in microgrids, " IEEE Transactions on Smart Grid, vol. 8, no. 1, pp. 228–238, Jan. 2017.

[86] P. Megantoro, F. Danang Wijaya and E. Firmansyah, "Analyze and optimization of genetic algorithm implemented on maximum power point tracking technique for PV system," *2017 International Seminar on Application for Technology of Information and Communication (iSemantic)*, Semarang, Indonesia, 2017, pp. 79-84.

[87] Seung-Jun Kim and G. B. Giannakis, "Efficient and scalable demand response for the smart power grid," *2011 4th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, San Juan, 2011, pp. 109-112.

[88] WHO, “Considerations for quarantine of individuals in the context of containment for coronavirus disease (COVID-19),” Who, no. February, pp. 3–5, 2020.

[89] R. Elavarasan, G. Shafiullah, K.Raju, V. Mudgal, M. Arif, T. Jamal, S. Subramanian, V. Balaguru, K. Reddy, U. Subramaniam,"COVID-19: Impact analysis and recommendations for power sector operation," in *Applied Energy*, vol .279, 2020.

[90] Z. Zhang, D. Zhang and R. C. Qiu, "Deep reinforcement learning for power system applications: An overview," in *CSEE Journal of Power and Energy*

Systems, vol. 6, no. 1, pp. 213-225, March 2020

[91] O. A. Alimi, K. Ouahada and A. M. Abu-Mahfouz, "A Review of Machine Learning Approaches to Power System Security and Stability," in *IEEE Access*, vol. 8, pp. 113512-113531, 2020.

[92] D. Zhang, X. Han and C. Deng, "Review on the research and practice of deep learning and reinforcement learning in smart grids," in CSEE Journal of Power and Energy Systems, vol. 4, no. 3, pp. 362-370, September 2018, doi: 10.17775/CSEEJPES.2018.00520.

[93] M. Khodayar, G. Liu, J. Wang, and M. E. Khodayar, "Deep learning in power systems research: A review," in CSEE Journal of Power and Energy Systems, doi: 10.17775/CSEEJPES.2020.02700.

[94] Puterman, L. M. (2005). Markov decision processes: Discrete stochastic dynamic programming. New York etc.: John Wiley & Sons.

[95] "Guest Editorial Special Issue on Deep/Reinforcement Learning and Games," in *IEEE Transactions on Games*, vol. 10, no. 4, pp. 333-335, Dec. 2018, doi: 10.1109/TG.2018.2882899.

[96] I. Oh, S. Rho, S. Moon, S. Son, H. Lee and J. Chung, "Creating Pro-Level AI for a Real-Time Fighting Game Using Deep Reinforcement Learning," in *IEEE Transactions on Games*, doi: 10.1109/TG.2021.3049539.

[97] F. Niroui, K. Zhang, Z. Kashino and G. Nejat, "Deep Reinforcement Learning Robot for Search and Rescue Applications: Exploration in Unknown Cluttered Environments," in *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 610-617, April 2019, doi: 10.1109/LRA.2019.2891991 Gg

[98] Y. Zhang, J. Yao and H. Guan, "Intelligent Cloud Resource Management with Deep Reinforcement Learning," in *IEEE Cloud Computing*, vol. 4, no. 6, pp. 60-

69, November/December 2017.

[99] Y. Deng, F. Bao, Y. Kong, Z. Ren and Q. Dai, "Deep Direct Reinforcement Learning for Financial Signal Representation and Trading," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 3, pp. 653-664, March 2017, doi: 10.1109/TNNLS.2016.2522401

[100] T. Remani, E. A. Jasmin, and T. P. ImthiasAhamed, “Residential load scheduling with renewable generation in the smart grid: a reinforcement learning approach, ” IEEE Systems Journal, vol. 13, no. 3, pp. 3283–3294, Sep. 2019.

[101] E. Mocanu, D. C. Mocanu, P. H. Nguyen, A. Liotta, M. E. Webber, M. Gibescu, and J. G. Slootweg, “On-line building energy optimization using deep reinforcement learning,” IEEE Transactions on Smart Grid, vol. 10, no. 4, pp. 3698–3708, Jul. 2019.

[102] Z. M. Yan and Y. Xu, “Data-driven load frequency control for stochastic power systems: a deep reinforcement learning method with continuous action search,” IEEE Transactions on Power Systems, vol. 34, no. 2, pp. 1653–1656, Mar. 2018.

[103] Z. Ni and S. Paul, “A multistage game in smart grid security: a reinforcement learning solution,” IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 9, pp. 2684–2695, Sep. (2019).

[104] F W. R. Liu, P. Zhuang, H. Liang, J. Peng, and Z. W. Huang, “Distributed economic dispatch in Microgrids based on cooperative reinforcement learning,” IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 6, pp. 2192–2203, Jun. 2018.

[105] F T. Hirata, D. B. Malla, K. Sakamoto, K. Yamaguchi, Y. Okada, and T. Sogabe, “Smart grid optimization by deep reinforcement learning over discrete and continuous action space,” Bulletin of Networking, Computing, Systems, and

Software, vol. 8, no. 1, pp. 19–22, Jan. 2019.

[106] F. Ruelens, B. J. Claessens, S. Vandael, B. De Schutter, R. Babuška and R. Belmans, "Residential Demand Response of Thermostatically Controlled Loads Using Batch Reinforcement Learning," in *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2149-2159, Sept. 2017.

[107] T. Peirelinck, F. Ruelens, and G. Decnoninck, "Using reinforcement learning for optimizing heat pump control in a building model in Modelica," *2018 IEEE International Energy Conference (ENERGYCON)*, Limassol, 2018, pp. 1-6.

[108] T. Wei, Yanzhi Wang, and Q. Zhu, "Deep reinforcement learning for building HVAC control," *2017 54th ACM/EDAC/IEEE Design Automation Conference (DAC)*, Austin, TX, 2017, pp. 1-6.

[109] L. Eller, L. C. Siafara and T. Sauter, "Adaptive control for building energy management using reinforcement learning," *2018 IEEE International Conference on Industrial Technology (ICIT)*, Lyon, 2018, pp. 1562-1567.

[110] A. Chi¸s, J. Lundén, and V. Koivunen, “Reinforcement learning-based plug-in electric vehicle charging with forecasted price,” *IEEE Trans.Veh. Technol.*, vol. 66, no. 5, pp. 3674–3684, May 2017.

[111] S. Vandael, B. Claessens, D. Ernst, T. Holvoet, and G. Deconinck, “Reinforcement learning of heuristic EV fleet charging in a day ahead electricity market,” *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1795–1805, Jul. 2015.

[112] F. L. D. Silva, C. E. H. Nishida, D. M. Roijers and A. H. R. Costa, "Coordination of Electric Vehicle Charging Through Multiagent Reinforcement Learning," in *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2347-2356, May 2020.

[113] R. Lu, S. H. Hong and M. Yu, "Demand Response for Home Energy Management Using Reinforcement Learning and Artificial Neural Network," in *IEEE*

Transactions on Smart Grid, vol. 10, no. 6, pp. 6629-6639, Nov. 2019.

[114] X. Xu, Y. Jia, Y. Xu, Z. Xu, S. Chai and C. S. Lai, "A Multi-Agent Reinforcement Learning-Based Data-Driven Method for Home Energy Management," in *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3201-3211, July 2020.

[115] S. Lee and D. H. Choi, ``Reinforcement learning-based energy management of smart home with rooftop solar photovoltaic system Energy Storage Syst. Home Appl., *Sensors*, vol. 19, no. 18, p. 3937,2019.

[116] S. Zhou, Z. Hu, W. Gu, M. Jiang, and X. Zhang, "Artificial intelligence based smart energy community management: A reinforcement learning approach," in *CSEE Journal of Power and Energy Systems*, vol. 5, no. 1, pp. 1-10, March 2019.

[117] Y. Wei, M. Zhao, F. Zhang, and Y. Hu, "Research of a heuristic reward function for reinforcement learning algorithms," in Proceedings of the Intelligent Control and Automation, pp. 2676–2680, Hangzhou, China, June 2004.

[118] Qicui Yan, Quan Liu and Daojing Hu, "A hierarchical reinforcement learning algorithm based on heuristic reward function," *2010 2nd International Conference on Advanced Computer Control*, Shenyang, China, 2010, pp. 371-376.

[119] J. M. Maja, "Reward functions for accelerated learning," in Machine Learning Proceedings 1994, pp. 181–189, Elsevier, Amsterdam, Netherlands, 1994.

[120] B. Mesot, E. Sanchez, C. A. Pena, and A. Perez-Uribe, SOS++finding smart behaviors using learning and evolution, MIT Press, Cambridge, UK, 2002.

[121] "Policy-based vs value-based - policy-based methods." [Online]. Available: https://www.coursera.org/lecture/practical-rl/policy-based-vs-value-based-INklC. [Accessed: 28-Feb-2021].

[122] "11 Policy-gradient and actor-critic methods · Grokking Deep Reinforcement Learning epub," · Grokking Deep Reinforcement Learning epub. [Online]. Available: https://livebook.manning.com/book/grokking-deep-reinforcement-learning/chapter-11/v-11/. [Accessed: 28-Feb-2021].

[123] M. R. G. Meireles, P. E. M. Almeida and M. G. Simoes, "A comprehensive review for industrial applicability of artificial neural networks," in *IEEE Transactions on Industrial Electronics*, vol. 50, no. 3, pp. 585-601, June 2003, doi: 10.1109/TIE.2003.812470.

[124] A. Rajeswaran, K. Lowrey, E. Todorov, and S. Kakade, "Towards Generalization and Simplicity in Continuous Control," *Advances in Neural Information Processing Systems*, pp. 6553–6564, Mar. 2018.

[125] W. Genders and S. Razavi, "Evaluating reinforcement learning state representations for adaptive traffic signal control," Procedia Computer Science, vol. 130, pp. 26–33, 2018.

[126] "Historical Market Data," *Nord Pool*. [Online]. Available: http://www.nordpoolgroup.com/historical-market-data/. [Accessed: 02-Mar-2021].

[127] "Reinforcement Learning Toolbox," MATLAB. [Online]. Available: https://www.mathworks.com/products/reinforcement-learning.html. [Accessed: 02-Mar-2021].

APPENDIX

Appendix A: Thesis Publications

1. Home Energy Management System Embedded with a Multi-Objective Demand Response Optimization Model to Benefit Customers and Operators (**Published, *Energies* 2021, *14*, 257. https://doi.org/10.3390/en14020257**).
2. Deep Reinforcement Learning Demand Response for Home Energy Management Systems: Customers and Operators Perspectives (**to be submitted to IEEE Smart Grid**).
3. Applicability and Challenges of Deep Reinforcement Learning for Load Scheduling in Smart Grids: Matlab/RL toolbox Perspective (**In preparation phase**).
4. Step by Step implementation guidance for Deep Reinforcement Learning in Power System application (**In preparation phase).**