T.C
ALTINBAS UNIVERSITY
Institute of Graduate Studies
Electrical and Computer Engineering

ENERGY MANANGEMENT OF SMART GRID BASED ON MULTI AGENT SYSTEM AND REINFORCEMENT LEARNING

AL-SAADI, Ali Abdulhasan Salman

Master of Science

Supervisor

Asst. Prof. Dr. Oğuz KARAN

Istanbul, 2021

ENERGY MANANGEMENT OF SMART GRID BASED ON MULTI AGENT SYSTEM AND REINFORCEMENT LEARNING

by

ALSAADI Ali Abdulhasan Salman

Electrical and Computer Engineering

Submitted to the Graduate School of Science and Engineering

in partial fulfillment of the requirements for the degree of

Master of Science

ALTINBAŞ UNIVERSITY

2021

The thesis titled "ENERGY MANANGEMENT OF SMART GRID BASED ON MULTI AGENT SYSTEM AND REINFORCEMENT LEARNING" prepared and presented by "Ali Abdulhasan Salman AL-SAADI" was accepted as a Master of Science Thesis in Electrical and Computer Engineering.

____

Asst. Prof. Dr. Oğuz KARAN

Supervisor

Thesis Defense Jury Members:

Asst. Prof. Dr. Oğuz KARAN	School of Engineering and Natural Sciences, Altinbas University	__________________
Asst. Prof. Dr. Abdullahi Abdu IBRAHIM	School of Engineering and Natural Sciences, Altinbas University	__________________
Asst. Prof. Dr. Zeynep ALTAN	Faculty of Engineering and Architecture, Beykent University	__________________

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Approval Date of Institute of Graduate Studies: ____/____/____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

AL-SAADI Ali Abdulhasan Salman

ACKNOWLEDGEMENTS

I might want to thank my administrators: Asst. Prof. Dr. Oğuz KARAN Please let me express my profound feeling of appreciation and gratefulness to both of you for the information: direction and unrestricted help you have given me. I want you to enjoy all that life has to offer and further achievement and accomplishments throughout your life.

ABSTRACT

ENERGY MANANGEMENT OF SMART GRID BASED ON MULTI AGENT SYSTEM AND REINFORCEMENT LEARNING

AL-SAADI, Ali Abdulhasan Salman

M.Sc., Electrical and Computer Engineering, Altinbas University,

Supervisor: Asst. Prof. Dr. Oğuz KARAN

Date: 06/2021

Pages: 55

The energy sector is undergoing fundamental shifts the depletion of fossil fuels and considerations for the environment have allowed it to use renews like solar and wind energy resources. A micro-grid constitutes a component of an intelligent grid and is ready to play an important role in the generalization of renewable resources. However, because renewable energy is intermittent in nature, it affects the dynamics and stability of the micro-grid and thus new approaches for coordination and control are required when integrated in the micro-grid. The existing systems lack run-time adaptive behavior and suffer from communication overhead. To meet these challenges and to achieve an optimal balance between generation, energy storage and load demands, we need to incorporate efficient communication and control strategies into micro-grid monitoring. In this thesis a, Reinforcement learning is implemented for optimal energy management of a micro-grid. It is extended into Multi-Agent Reinforcement Learning (MARL) for distributed optimization of micro-grids. The performances of the reinforcement learning methods are compared with conventional methods. The effectiveness of MAS and MARL are investigated in micro-grid for autonomous adaptation in the dynamic environment, and for economic and environmental optimization.

Keywords: Smart grid, Micro-grid, Virtualization Storage, energy Renewable sources, Reinforcement Learning.

TABLE OF CONTENTS

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

LIST OF ABBREVIATIONS

RESs	:	Renewable energy sources
RE	:	Renewable energy
SGAM	:	Smart Grid Architecture Model
DG	:	Distribution Generations
ESSs	:	Energy Storage Systems
PV	:	Photovoltaic

1. INTRODUCTION

A gift from science and technology to the human community is electrical power. It is a very important thing in the modern day to day life. Feel the importance of electrical power in our life during power outages. The common life of the peoples is almost stopped in power outages. One of the most complex systems in the world is power system structure, and it is a large system used to provide the electricity to the consumer. Three important modules in the power systems are power generation, transmission, and distribution. These modules are very much important for the process of distributing power generation from a remote location to the consumers.

Electrical Power is one of the primary things for the economic growth of a nation. The fossil fuel is the major source for the generation of electricity, and it leads to environmental pollution and the cost of energy generation is too high. In India total installed capacity 326.8 GW in that 59% of power generated from fossil fuel coal and 17% are contributed by the renewable resources like wind, solar, Biomass and small hydro plants. Nonrenewable resources like oil, gas, etc., are getting condensed frequently. Renewable resources like solar, wind, etc., are always available to generate power. It will produce power at a low cost and eco-friendly. Revolution in the power system is in demand to hope up to integrate non-renewable resources into the present power grid to reduce the losses and provide uninterrupted [1].

Most of the power generation stations are located so far from the consumer's end and it leads to transmission and distribution losses. In India A&T losses about to 15% from its total generation. Inefficient functions of power system equipment due to an aging of the equipment are the response to produce the distribution losses. The present power grid is a dummy and inert it does not distinguish the peak time and off-peak time of the consumers so excess powers are getting wasted and also increases the environmental pollution. The power markets also challenge to reduce the power losses and power demands, especially during peak load hours, to upgrade the power system structure. Intelligent price technology and an end to end two-way communication are providing the solution for new market challenges. Reconfiguration of the power system is in need of the future power system by modifying the centralized control center into distributed control by using an intelligent multi-agent system.

The greenhouse gas reduction is essential for the ecological development of the world. In this all the concentrate more on to reduces their greenhouse gas mission. In this power sectors contributes 28% of world greenhouse gas emission. Figure 1.1 shows the world greenhouse emission. The need of electrical power in every country are increased rapidly this is meet out by adding new thermal power plants. This produce more amount of CO2 emission and it increases the overall greenhouse emission.

The main source of electricity was fossil fuels. The combustion products of these fuels were CO2, CO, NO, NO2, SO2 and so on. The collection of a large source of pollution in the atmosphere would lead to degradation of the environment. The best choice in this case could be the use of renewable energy. Different system for the use of renewable energy sources such as wind, solar, geothermal, dynamic electricity, and magneto, have been developed with an ideal application. Wind and solar are mostly the main resource of the microgrid.

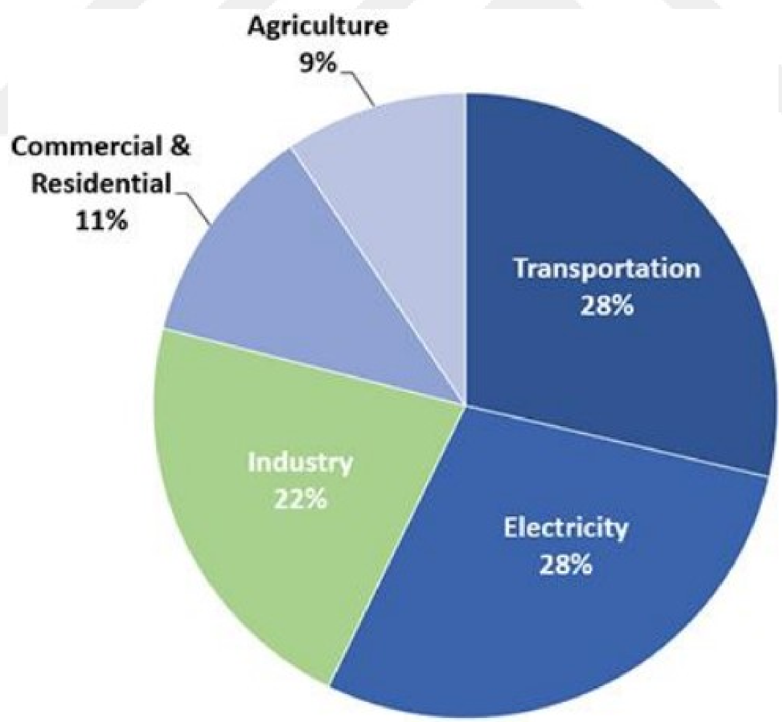

Figure 1.1: World Greenhouse Gas Emission [1]

1.1 ELECTRICAL POWER SYSTEM STRUCTURE

Most of the countries electric power delivered by the grid, it is invented when the power generation is low cost. More than 100 years still the grid was not upgraded to meet out the new challenges. Industrial and domestic loads are increased rapidly it leads to produce more greenhouse gases and grid, not able to handle future demands. Nowadays, the Power system is constructed with centralized control with all-time availability of the electricity. The system is designed with highly secured for efficient utilization of the power generation. An economic point of view large scale generating systems is connected to reduce the energy cost. In centralized control of power system, balancing of energy demand and production need to be maintaining in all the times. Consequently, generated power is distributed through transmission and distribution network properly. In this system, generation is followed by demand. The flexibility of the power demand needs to add or removing of power generation in the present system leads to a malfunction of the entire system. India faced the largest blackout history in the year 2012 due to overdrawing of power from the grid, 300 million people are affected.

Present system facing many challenges due to flexibility in power generation side is granted in the centralized control system, interconnecting renewable power generation in the system, it needs to adopt the bi-direction flow of power. Wind and solar power are connected in the system directly or indirectly. Increase in the generation leads to local challenges like power quality and grid stability. Reducing fossil fuel cost and increasing the overall efficiency of the system using renewable energy to supply the energy demand. The capacity of the traditional power system is modified to meet peak demand it requires both generation and grid capacity.

1.2 SMART GRID

The fundamental expansion is essentially needed, in today's power system structure. In addition to that, it will hope up with present technology in all aspects to drive the fundamental changes in the grid structure. A smart grid is an effective solution provider for present power system issues. The term smart grid a traditional power grid with advanced information and communication technology from the generation to the distribution. A smart grid was introduced in the year 2005. Integration of information and two-way communication technology into the electrical system is known as a smart grid [3]. The flow of electricity from generation to end user becomes a two-way conversation, saving consumers money, energy, delivering more transparency in terms of end-user use, and reducing carbon emissions.

Three main areas of the smart grid are transmission, distribution and system operations. In transmission introduces the high voltage DC and AC grids for asset management and advanced metering infrastructure, demand management is done in the distribution. Intelligent systems, self-healing, secure communication is in the system operations.

The function of the smart is more complicated than the present grid, it needs advanced sensing elements, sensors, intelligent electronic devices, switching devices and monitoring devices with high speed secured communication network. Need more care to install and operation of the smart grid any miss match of data from any devices develops the complicated issues like grid shut down. The goals of the smart grid, collecting the data from the sensors, measuring devices, and consumer utilization details, process it and providing relevant information to the control center, it maintains the grid stability, forecasting demands and resolving the harms by the structured manner. The figure 1.2 and 1.3 shows the structure of the present grid and smart grid.

1.2.1 Characters of Smart Grid

A smart grid provides intelligent technology with monitoring, control, communication, and self-healing grid, some of the characteristics of a smart grid are:

• The smart grid gives much importance to a consumer to participate in the operation and control of power generation and deliver with choice of supply. Demand management and demand-side control were done using smart meters, micro-generation, smart appliances, and storage information. This information helps to choose the consumer load pattern to reduce the peak load issues.

• It interconnects the large generation station, medium and micro generation stations, renewable generations, storage system and plugs in an electrical vehicle. It will reduce the environmental pollution of the whole electrical supply systems. It provides better adaptability to the interconnection of different range of power generation systems similar to plug-and-play.

• Efficient asset management in the smart grid is done using intelligent usages of the electrical power structure like rerouting and working autonomously. It will provide the power by asking the question like when, where, how much the power is needed depends upon the conditions.

• It has self-healing capability with highly secure and reliability of the supply in up normal condition. It isolates the faulted portion of the grid to maintain continues operation of the system with a high level of security for cyber-attack and physical issues.

• It develops the open access market for the business peoples, private power producers, power agencies and small domestic consumer to buy and sell the power with flexible pricing.

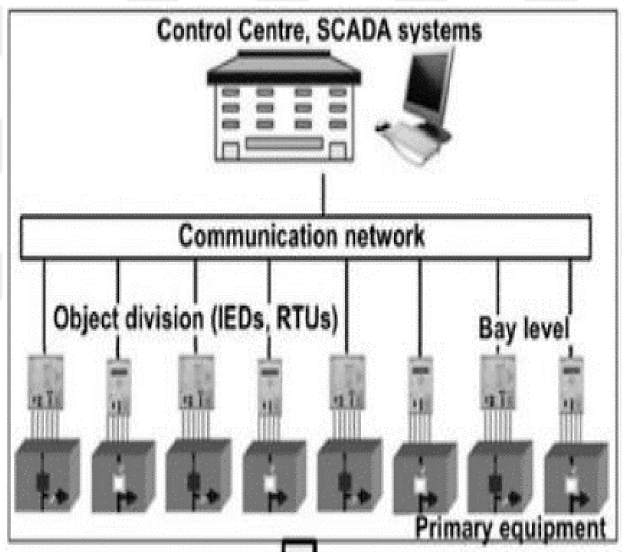

Figure 1.2: Present grid Structure [3]

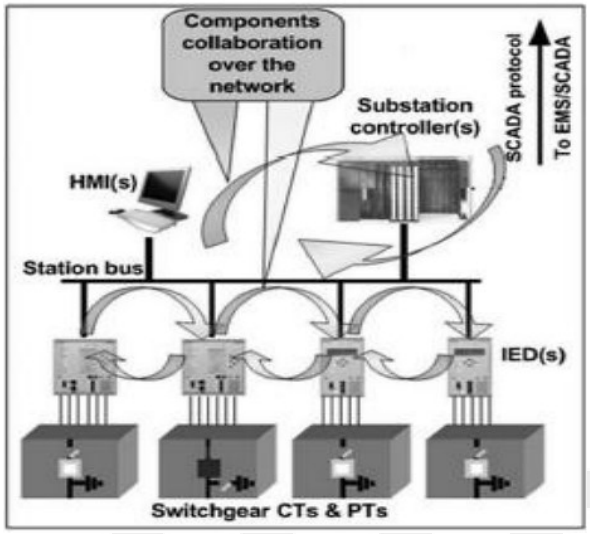

Figure 1.3: Smart Grid Structure [3]

1.2.2 Difference Between Present Grid Verses Smart Grid

Typical characteristics of the present grids are not up to the present technology. Absences of new technology led to provide fewer benefits to the customer and performance of the grid low. Below table shows the characteristics of a traditional grid with the preferred Smart Grid.

Table 1.1: Present Grid Vs Smart Grid

Characters	Traditional grid	Smart Grid
Types of control	Electromechanical, solid state drives	Digital/Microprocessor
Communication system	One-way and local two-way communication	Global/integrated two-way communication
Control center	Centralized generation	Accommodates distributed generation
Production	Limited protection, monitoring and control systems	WAMPAC, Adaptive protection
Monitoring System	Blind	Self-monitoring
System rebuilding	Manual restoration	Automated, self-healing
Equipment monitoring	Check equipment manually	Monitor equipment remotely
Control	Limited control system contingencies	Pervasive control system
Reliability	Estimated reliability	Predictive reliability

1.2.3 Evaluation of Smart Grid

Frequent urbanization and infrastructure development in the present power system, they are keen to the production of power from the generating station; it is located in various geographical locations, away from the consumer. The development of a smart grid is controlled by influences from political, geographical location, economic, society, and gap of education. Basic topology of the present electrical power system is unchanged. Private power producers are isolating their power generation, transmission, and distribution using advanced level of technology like automation, estimate, and revolution in each level.

Exciting power strictly follows the structured system, the bottom of the system connected with the consumer load and the top of the system represent the electrical power plants. This system is a one-way communication from the power plant to the consumer as like an open loop system it doesn't know the real-time status of the load. Even though the grid has the capability to withstand the peak demands with it is an allowable load. The demand occurrence is in unpredictable time due to environmental constraints, grid unable to manage the demand it becomes inefficient and unstable. These unexpected failures cause equipment failure and develop the blackouts. Addressing the issues and keep, the asset introduces the various levels of control and command functions are used [4]. Supervisory Control and Data Acquisition System (SCADA) are the most available power system for reliable operation.

In general, most of the losses in the power system due to transmission and distribution, Distribution losses are more comparing to transmission losses. Modification of the power system is done form the bottom to top. Inability consumers are connected in the grid. They draw more power from a grid to develop the demand it leads to increase in power cost. For demand side management and revenue generation were achieving by to introduce new topology in distribution. Advanced Metering Infrastructure (AMI) used to address the above issues by two-way communication between utilities. Utilities know their demand with real-time power cost and ability to buy and sell the power to generate the revenue and mete out the demand in the grid.

Figure 1.5 Shows the order of operations of the power grid with varies management levels. Overall goal of the power system is to provide the quality power with reasonable price for the betterment of both consumer and power producer.

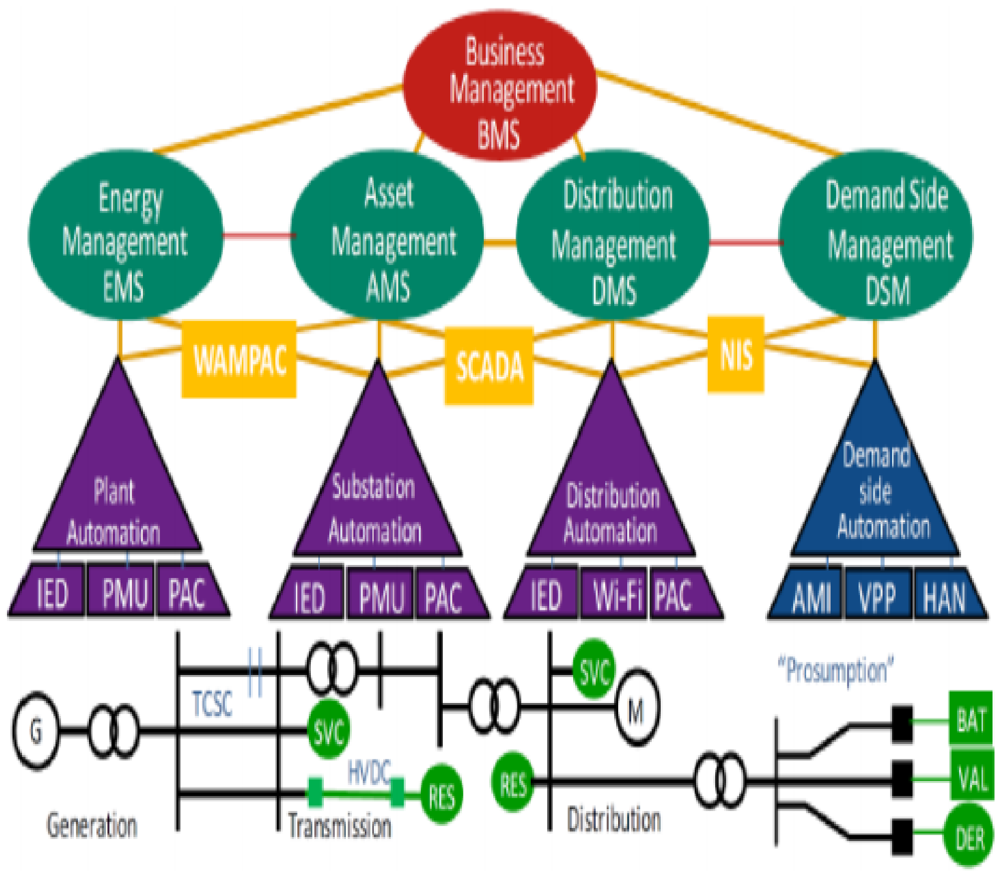

Figure 1.4: Grid operation Hierarchy [6]

1.2.4 Components of Smart Grid

1.2.4.1 Monitoring and control

The present power system network designed to deliver the power from generation units to the end consumer with defined voltage levels. In Future grid PV, power generation units are located near to the consumer, due to these voltage variation issues are encountered by a consumer. This will develop two challenges in the network are Consumer near the PV units, receives the high voltage level and voltage fluctuation in the entire systems. Intelligent transmission systems need self-monitoring, self-healing, load predictability, and adaptability to handle the issues. The future system needs to analyze the power load and forecasting the voltage level and demand to ensure the quality of supply.

1.2.4.2 Transmission Components

The backbone of the entire power system operation is transmission components; it interconnects the load centers and generation units. The system ensures the reliability of the supply at the sudden change in the system variables due to environmental issues, without interruption of the supply. In smart grid, implementation needs analysis tools and measuring components in the transmission system for reliable supply. Various techniques are used to measure the electrical parameters like phasor measurement techniques, sensors and communication devices.

1.2.4.3 Interfacing intelligent devices

Monitoring and control in generation and distribution were handled by the intelligent devices refer to data received from the real-time systems. An overall operation of the smart grid is decided by the function of the electronic device. The most important part of the smart grid structure is an intelligent component. Devices are in a user-friendly and adaptability to interconnect in various conditions. Selection of the devices improves the performances of the smart grid. A set of standards and models are needs to develop for easy implementation of the intelligent system.

1.2.4.4 Distribution System Components

Distribution is the end of the power supply chain to the consumer. Most of the power losses are due to the present metering infrastructure. The accuracy of the meter reading is less and no option to the interconnection of the local power generation in the grid. Advanced Metering Infrastructure (AMI) is capable of self-healing, voltage stabilization, fault identification, automatic billing, real-time pricing, and bidirectional metering. Communication between the load center and consumer happens at a frequent time interval, consumer data are stored in the control center for billing and reference purpose. Wireless Mesh network communication technology used to reduce network traffic, information transferred through the nearby meters to information and control center. Frequent data transfer by the meters causes' data loss and actions need to maintain data storage and retrieval.

1.2.4.5 Storage system

Storage of more electricity at less cost is a challenging task for electrical engineers. The unpredictability of the load and generation is the main cost for the storage. Power availability from renewable resources and peak demand do not occur at the same time.

Need to implement more storage solution technology like a super capacitor, battery storage, flywheel, and superconducting magnetic storage, etc. Power marketing, which interchanges the power between the private suppliers and consumer through load centers.

1.2.4.6 Demand-side management

Demand-side management (DSM) in developed to modify the customer to involve in energy management to reduce the expensive generation and addition of capacitive generation. This DSM will reduce the fuel, emission, power generation cost and provide the reliability of the supply. An industrial consumer is a piece of advice to maintain their demand and supply in predefined frequency level others wise need to pay the more cost for their energy.

With increasing the photovoltaic generation in the consumer side leads to voltage variation in the power system network. The table 1.2 shows the Types of components functions and it is capabilities. The demand management techniques used to resolve the issues by control the power generation and concentrate on battery storage and other storage methods.

Table 1.2: Types of components functions, and it is capabilities

Components	**Capabilities**
One - way communication	Automated Billing
Two-way communication	Outage detection and restoration, Demand response
Distributed control	Distributed automation, customer information system

Network management	Load management, Preventing/ self-healing, Substation automation
Intelligent Application	Distributed co-generation, Emission control
Intelligent Agents	Intelligent Appliances, Customer portals
Secure system	Cyber-attack

1.3 THE ENVIRONMENTAL IMPACT OF SMART GRID

In India, 70 % of electrical power is generated by means of coal-based power plants. Power requirement in India increasing rapidly it will produce CO2 emissions. Coal is imported and its quality is poor it leads to produce fly ash while burning, transposition cost. The fly ash pollutes the air leads to harm the people. Power generation depends on the requirement by the consumer. Effective management of the power by the consumer leads to reduce power generation and CO2 emission.

One of the main objectives of the smart grid is to reduce CO2 emission and increase more renewable power generation. Power generation by the rooftop solar power plant and micro wind power plants in the consumer side is to generate power. These types of generation are reducing T&D loss and power generation loss.

1.4 MICRO GRID

The microgrid is a part of a smart grid with minimum power capacity compared to the main grid. In the combination of the greater number of microgrids is called a smart grid. A grid with the power capacity of low and medium voltage levels are considered as the microgrid. It will operate as an islanded mode or distributed with the main grid. In this renewable resource like solar, wind and batteries are connected in the microgrid.

1.4.1 Benefits Of Micro Grid

A microgrid has the potential to

a. Increase the amount of renewable energy contributing to power needs, both locally and on the grid overall.

b. Combine renewable and other power sources.

c. Provide power in places where there is no grid.

d. Provide more reliable power in places served by a grid.

e. Increase the stability and reliability of the power grid.

f. Reduce the transmission loss its nearly to 2%

A very simple definition of a microgrid is an independent energy system providing grid backup or off-grid power.

1.5 TECHNICAL BACKROUND

In the energy field, ICT growth has greatly affected the way in which energy is generated and distributed. We are moving towards more decentralized, sustainable, smart power. A transition to intelligent two-way supply networks is the smart grid paradigm. Intelligent grids can be included in the micro grid and transformed into intelligent micro grids such as auto healing, side management requirements, plugin and playback, dynamic pricing, dynamic adaptations and distributed optimisation. An intelligent grid is an integrated intelligent micro grid system.

Single, large, single-powered systems and few experiences in the running of the same Grid are extensively experienced by the energy industry. In future it will be only a large, virtual interconnection power plant consisting of DERs, prosumers and other energy resources. A high penetration of renewable energy resources requires new coordination and control approaches. The energy network must be dynamically optimized for economic and environmental benefits. Micro-grid surveillance in a distributive environment should incorporate communication and control, to create an optimal balance between generation, storage of energy and load demand. Power and energy industrial automation is becoming increasingly complex today and requires advanced ICT concepts and methods.

In recent years, the Multi Agent System (MAS) has received a lot of attention because of its capacity to improve operating efficiency significantly in a distributive environment. MAS offers new microgrid capability through the calculation of data and computer resources and the flexible communication of relevant results to the respective entities when necessary. MAS enables flexible, robust and environmentally adaptable systems to be developed. It is difficult to dynamically adapt when system components are heterogeneous, unknown and change over time. MAS offers the ability to decentralize calculations through autonomy and intelligence integration into distributed applications.

1.6 AIM AND OBJECTIVES OF THE RESEARCH

The aim of this research work is (1) to implement a MAS for advanced energy management of micro-grid in stages, starting from a proof of concept, comprehensive verification and validation through experimental set up, and (2) to incorporate smart grid features into the micro-grid to make it smart micro-grid for economic and environmental optimization.

1.7 OBJECTIVES OF THE RESEARCH

The main objectives of this research work are:

a. Implement a Multi-Agent System (MAS) running in JADE platform for effective energy management of solar micro-grid under environmental dynamics to maintain stability, reliability and fault tolerance.

b. Incorporate heterogeneous renewable resources of solar and wind, achieving economic and environmental optimization.

c. Implement Reinforcement Learning (RL) for optimal energy management of solar micro-grid.

d. Implement Multi-Agent Reinforcement Learning (MARL) in solar micro-grid for distributed optimization.

1.8 CONTRIBUTIONS OF THE THESIS

a. The key contribution of the present work is the insight it provides into a hybrid model, linking a Multi-Agent System (MAS) in JADE with a micro-grid model in MATLAB/Simulink, for effective energy management and demand side management of a micro-grid. It is customary to simulate a MAS in Java Agent Development Environment (JADE) platform. Yet, JADE based simulation of an MAS proves only the theoretical concept and is not comprehensive and cannot be validated. Therefore, MAS implemented in JADE needs to be linked with MATLAB-Simulink for deeper understanding of the dynamic simulation models. Since, however, the MATLAB-Simulink does not support parallelism, required by the MAS, we use a middleware, Multi Agent Control using Simulink with Jade Extension (MACSimJX), for linking the MAS running under JADE with the micro-grid model in MATLAB/Simulink. This allows the solar micro-grid model designed with Simulink to be controlled by agents of the MAS operating under JADE. Simulink feeds the environmental information to the JADE agents which send back the results of the coordinated action to the Simulink for comprehensive validation of all the smart grid features in the micro-grid, turning it into a smart micro-grid. The command signals can be given to the actuators.

b. The present work is distinguished from the published work by the linking strategies for the hybrid models in MAS and MATLAB, and the control architecture in the MAS for validation and verification of smart micro-grid. We have spelt out in detail the design of the micro-grid model in MATLAB, the control strategies for sensing and actuating, the validation and verification of smart grid features. We have also presented control algorithms for advanced energy management strategies. The model can be scaled up and applied to smart grid industry, where many heterogeneous Distributed Energy Resources (DER) are present, leading to economic and environmental optimization.

c. The MAS implemented in JADE is linked with Arduino processor for practical verification of advanced energy management and demand side management strategies.

d. Reinforcement learning is implemented for optimizing energy management of solar micro-grid for economic and environmental optimization.

e. Multi-Agent Reinforcement Learning is implemented in solar micro-grid for distributed optimization.

1.9 ORGANIZATION OF THE THESIS

Chapter 1 reviews the need and motivation for applying Multi-Agent Systems (MAS) and Reinforcement Learning (RL) in micro-grid environment. Chapter 2 A detailed survey of the literature is given in this chapter. This chapter also states the research objectives, contributions, and the organization of the thesis.

Chapter 3 contains two parts the first part deals with imparting learning to agents for autonomous optimization. Q-learning, a model-free reinforcement learning method, is implemented for optimal energy management of a solar micro-grid, and compared with conventional method. The consumer is represented as an agent equipped with reinforcement learning capabilities for long-term energy management. The consumer agent, by learning over a period of time, chooses the best possible action for optimal battery scheduling to achieve the long-term objective of optimizing three parameters — increasing the utility of solar power, the utility of the battery, and reducing the power consumption from the grid. The performance of Q-learning is compared with conventional method to prove the advantages. When many consumer agents are present in the environment, Multi agent reinforcement learning can be used for distributed optimization, which is discussed in the next chapter. The second part investigates the distributed optimization of the micro-grid with an algorithm for multi agent enhancement training called coordinated Q-learning (CQL). Collective decisions are taken by agents about how user demands are to be met, dynamic environmental changes and unplanned contingencies. In two different locations two solar micro-grids are considered. In each micro-grid, the consumer agent uses a Q-learning system to optimize the battery scheduling to increase the use of solar power and batteries to achieve its long-term objective of reducing grid power consumption. Then, the two agents communicate and co-operate through CQ-learning for distributed optimization of the solar micro-grid with intermittent nature of the solar power supply and randomness of load. CQ-learning is implemented for a coalition of agents and the performance is compared with individual Q-learning to show the advantage of distributed optimization. Chapter 4 presents the conclusion and suggestions for future work.

2. BACKGROUND

2.1 AGENT ORIENTED PROGRAMMING

In the evolution of programming languages, initially monolithic programming was used for simple programming and when programs became more complex, modular programming approach was used. Then, object-oriented era came which considered the world consisting of large number of entities, called objects, and the problems were solved by collaborating objects. Objects belong to some classes which define their functionality. The programmer has to identify and implement class hierarchy and write the main program for collaboration of the objects for solving the problems. In object-oriented modeling, when methods are invoked by external objects, they have no option but to execute them. Although objects encapsulate state and behavior realization, they do not encapsulate behavior activation. Objects, classes and modules provide essential yet insufficient abstraction (James Odell 2015) [1]. The relationship is defined by static inheritance hierarchies which is insufficient for flexible and dynamic operations.

An agent is a flexible and autonomous computer program capable of operating in a dynamic environment (Russel & Norvig 2010)[2]. The most recent paradigm in the computer programming language is agent-driven programming. The software agents do have their own control thread, locating not only the code and state but also their call. Agents with their own rules and goals, and with initiatives (Jennings NR 2000)[3] they appear to be active objects. An agent is a software entity encapsulated with its own state, behaviour, control and communication capabilities with other entities. They have higher level of abstraction than objects and through peer to peer communication they mediate, collaborate and cooperate to achieve goals. They are autonomous, interactive, and proactive. Agents decide when and how to act. Agents can say no or go. Since agents are autonomous they can initiate interaction and respond to a message in any way they choose. The application developer simply identify the agents and the agents organize themselves to perform the functionality. The choreography of the task execution is left to agents. Agents decides what to do to runtime changes which are unforeseen in the design time. Agents choose the best possible way when there are many alternative ways to achieve a goal, whose merit depends on the circumstances. Agent communication model is asynchronous. This means that there is no predefined flow of control from one agent to another.

An agent may initiate internal or external conduct independently at all times, not only when a message is sent. In agent-based messaging systems, asynchronous messages and event notification are provided, and agent languages have to support parallel handling. The agent model links objects (data and functionality) explicitly to parallelism (execution autonomy, thread per agent, etc.). The object-oriented modelling has its origins in system theory, whereas agent-oriented modeling reflects symbolic interactionist concepts. Since agent based software engineering is the latest software paradigm, it lacks mature tools and methodologies whereas the object oriented software engineering offers sophisticated development tools.

Autonomy and interaction distinguish the agent-based approach from traditional approaches. Autonomy means that without a central controller or commander, they can take decisions. They are driven by a number of trends in order to achieve this. For example, for a battery system, "charging the battery at low energy price" might be a tendency. Interaction means that the environment and other entities can communicate. The most basic form of interaction may be seen in object messages (method invocation). Include those agents which can react to observable events in the environment would be more complex to interact. There are reactive or deliberative agents. Reactive agent has simple architecture, easy to implement and responds quickly to the environmental stimulation. The output is produced by situation–action association. Reactive agents require less processing, less communication overhead, predefined reaction strategies and hence it can act much faster (Azevedo et al. 2017) [4]. For example, if a relay has to be activated within a short time to avoid the effect of lightning, reactive agents are used. But reactive agents are not flexible and dynamic like deliberative agents and so they are not able to behave proactively. The reactive agents are used for low level operations, with limited intelligence, for reacting extremely fast to disturbances and unforeseen relevant changes in the environment.

Deliberative agents are goal directed, and their behavior and architecture are relatively sophisticated, and the internal processing and computation are comparatively complex, which results in slow reaction. In this approach, agents are characterized by cognitive modules which are responsible for a goal-oriented rational behavior of the agent. Their behavior is comparatively smart and flexible and they can learn as well. High level management and control tasks in complex dynamic environment are done through deliberative agents. They are used in Belief, Desire and Intentions (BDI) Architecture. Belief is agent’s abstract knowledge about the environment, Desire represents the goals of the agent and Intension represent the current action plans. For higher level cognitive operations, which deal with long-term objectives, deliberative agents are used. Lower level requires horizontal communications for quicker response and higher level lean more towards hierarchical control structure. Reactive agents provides robustness at the lower level while the deliberative agents provides intelligent decision making at higher level. If an unforeseen situation occurs, the agents first try to match with its pattern. If not solved here, the problem is forwarded to next higher level. Thus, it moves from reactive architecture to deliberative architecture. In the case of local failures, reactive agents are used. For high level failure, which requires more intelligence, deliberative agent is used. The reactive actions are fine-tuned by the deliberative capabilities. The combined characteristics of both deliberative and reactive agents are used to achieve a coherent behavior as a whole (Dou et al. 2016) [5].

2.2 MULTI AGENT SYSTEM

The current power system uses the supervisory control and data acquisition system (SCADA) as a centralized system to effectively communicate and control things. No autonomous device functions were allowed by SCADA. The SCADA only sends command and control signals which lead to a failure to perform in real time. The multi-agent system overcomes the existing system's disadvantage. It offers distributed system control. As a checking agent are used intelligent electronic devices or software. This will have the right to decide on environmental grounds.

Each officer has defined objectives to fulfill in a given period of time. Agents are classified into various operation-based types, such as control agents, agents distributed, monitoring agents, centralized control agents, database agents, etc.

2.3 STRUCTURE OF MULTI AGENT SYSTEM (MAS)

The multi-agent system was used to integrate hardware and software. An agent can depend on the operating condition for hardware or software. Each agent has its own set of objectives. A number of agents at various levels with their own objectives are used in a complex power system. The agents collectively aim at implementing intelligent grid technology in the power system. The agent may be hardware or software interconnected to and capable of operating independently. Agent features include reactivity, proactivity and social skill. The basic MAS structure is shown in figure 3.1.

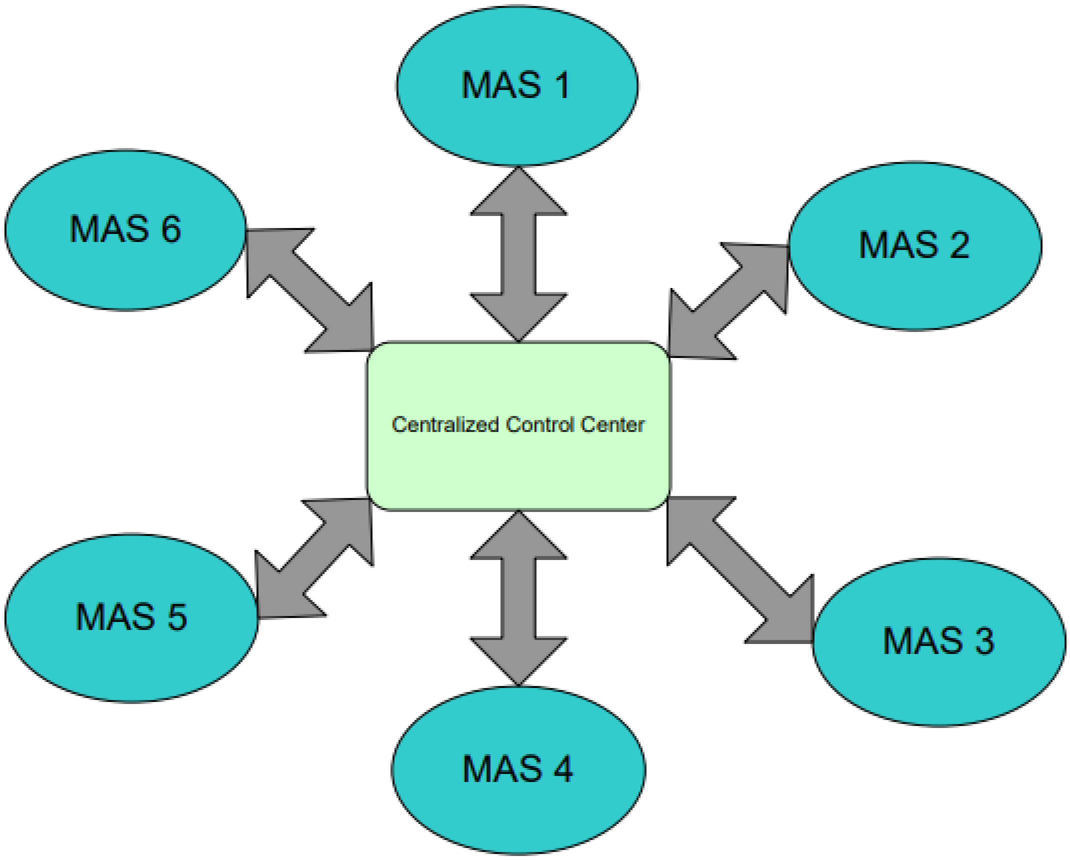

Figure 2.1: Structure of MAS

The complex problems with the power system are torn to small issues. Single MAS handles these small problems. This reduces the system's complexity easily identified and insulated. MAS has the right to decide on a separate basis, not waiting to receive the signal from the control centre. The officer can communicate the completion status of the task with the nearby agent and provide information about the status of some close agents. MAS coordination and collaboration is important to achieve the goal of all collaborative company cooperation, for the smart operation of the system example. Coordination between the MAS is necessary to ensure an effective information exchange in real time in order to achieve the overall system goal. This agency' s cooperation has the right to reject the signal from the other actors, to accept and to defend it. Cooperation is extremely critical for the MAS function in this respect.

Agencies can report on the status and achievement of goals. This helps other actors achieve the goal as quickly as possible. Common control language is used for communication throughout the operation.

MAS is flexible and so it can easily react to unforeseen sudden interferences in the runtime. It can adapt to changing real world situations and requirements. MAS is adaptable and reconfigurable to efficiently manage the evolutionary nature of execution plans. Due to the fact that they are loose, intelligent and autonomous (Paulo Leitao & Stamatis Karnouskos 2015)[6], they offer a high ruggedness and fault tolerance. MAS has the auto-properties to address changes in the system, such as self-healing, self-configuration, self-organisation and self-optimisation. For instance, self-healing allows a software system to discover, diagnose and correct failures and allows a software system to monitor the usage of resources and adapt them to ensure that their operation is optimal compared to defined requirements. Agent behaviour, which uses their resources, skills and services in order to achieve certain goals (Gerhard Weiss 2013)[7]. The manner in which the agent uses its resources, expertise and services define its behavior and the behavior of each agent is determined by its objectives. One thread per model agent is used in a parallel running agent with cooperative planning for a number of behaviours. Agents are autonomous, social, proactive and reactive in four behavioral attributes. The objective of MAS is to control a very complex system with a minimum level of data exchange and computer requirements. MAS focuses on coordination and communication between agents, environmental interaction and the organization of simulation experiments. The most widely used agent platform is the Java Agent Development Environment (JADE). JADE is the middleware for developing Java-based multi-agent distributed applications based on the architecture of peer to peer communication. With its inherent characteristics, MAS enhances the modularity, adaptability, scalability, configurability and response. In general, MAS offers a useful way of addressing modern software systems' growing complexity.

2.4 MULTI AGENT SYSTEMS IN MICRO-GRID

The power sector has undergone profound changes, as renewables such as solar and wind consumption have been affected by the depletion of fossil fuels and environmental considerations. A micro grid forms part of a smart grid and is ready to play a major part in generalizing renewable resources. However, because renewable energy is intermittent in nature, it affects the dynamics and stability of the micro-grid and thus new approaches for coordination and control are required when integrated in the micro-grid. The existing systems lack run-time adaptive behavior and suffer from communication overhead. To meet these challenges and to achieve an optimal balance between generation, energy storage, and load demands, we need to incorporate efficient communication and control strategies into micro-grid monitoring (McArthur et al. 2007a). Multi agent system can handle these challenges better than controllers, optimizers and learning systems(Gregor Rohbogner et al. 2013) [8] due to its inherent characteristics.

2.5 REINFORCEMENT LEARNING IN MICRO-GRID

Autonomy entails the agent's capability to survive in a changing environment. In MAS, agents coordinates with one another to manage the dynamic changes in the environment. Multi agent system works in the principle that intelligence is deeply coupled with interaction. The agents can also be smarter if they can learn and take optimal decision on their own. Reinforcement learning is learning through trial and error as human beings do. Reinforcement learning is falling between supervised and unsupervised learning, where the agent learns through experience. The agent learns the optimal action by continuously interacting with the environment. The agent monitors and acts on the environment. The environment receives a reward or punishment. The agent takes the next step in optimizing long-term rewards. The agent finds the perfect policy to achieve the long-term goal after many interactions. Instead of a single agent, many agents can use reinforcement learning to optimize individually and also coordinate for global optimization. Multi-agent reinforcement learning can be used when several such reinforcement learning based agents interact for distributed optimization. Reinforcement is used for optimal energy management of a micro-grid. Multi-agent reinforcement learning is used for distributed optimization of smart-grid where many micro-grids are interconnected to form a distributed environment.

2.6 LITERATURE REVIEW

2.6.1 Multi Agent System in Micro-grid

Given the intermittent nature of energy produced from renewable resources, the dynamic and stability of the microgrid has an impact and thus new approaches to coordination and monitoring are required in their integration into the micro-grid [9]. To meet the challenges of energy generation, energy storage, and load demands and to achieve an excellent balance, we need to incorporate efficient communication and control strategies into micro-grid monitoring. Classical algorithms for centralized energy management of micro-grid are discussed in (Colson et al. 2010). Centralized control, in the presence of intermittent power resources, raises issues such as scalability, computational complexity, and communication overhead. The new metaphor of computing as social activity, as interaction between independent entities, adapting and co-evolving with one another, lies with agent technologies.

Multi-Agent System (MAS) is a viable approach for decentralized control and dynamic adaptation. MAS controlling the main operations of micro-grid is discussed in (Dimeas & Hatziargyriou 2005) [10]. (Eddy & Gooi 2011) [11] describes how MAS is used in a micro-grid with intermittent renewable energy resources to distribute decisions and to improve performance and stability. (Pipattanasomporn et al. 2009) [12] describes the design and implementation of an MAS in a hierarchically controlled micro-grid. In Logenthiran et al. 2012, [13] MAS is implemented on JADE platform and RTDS interfaces through TCP/IP protocol but is not designed to address all dynamic changes and control strategies, and it is used as a real-time digital simulator for micro grid micro time operations. (Nunna & Doolla 2013)[14] simulates, but does not practically realize, MAS simulated for market operations of integrated micro-grid. (Zhang et al. 2003) [15] discusses distributed online optimal energy management for smart grid using MAS. (Gomez-Sanz et al. 2014) [16] gives a complete review of micro-grids from MAS perspectives.

A recent microgrid control MAS survey (Abhilash Kantamneni et al. 2015) [17] is available. While MAS is widely recognized as the method of choice for the implementation of time critical smart grid applications, only theory proof of concept was simulated. The process for developing, checking, validating and deploying MAS in microgrid is not addressed properly until now. (Perkonigg et al. 2015) [18]. This is the main motivation for the work reported here. Since MAS simulations in JADE are not comprehensive, the present work attempts to link MAS in JADE with MATLAB for validation through real-time simulations, which is tedious due to lack of support for parallel operations in MATLAB. MACSimjX is a middleware which can be used to connect MATLAB to MAS to enable an agent operation in external programs, in JADE, to control the system designed with Simulink. [19] [18]. Robinson et al. 2010]. The integration of JADE with

Simulink for the energy management of micro grid distributed energy sources (DER) is discussed in MACSimJX (Eddy et al. 2015) [20], which, however, does not adequately address agent concepts, strategies for linking MAS with MATLAB/Simulink model, control strategies for the dynamic variations of the environment, and time synchronization in co-simulation between MAS and MATLAB. Moreover, the implementation of MAS does not comprehensively consider the options available in a micro-grid to enable smart grid features in a dynamic and distributed environment.

Against this background of the need to examine the co-simulation in JADE with MATLAB Simulink of MAS, current work is aiming at using MACSimJX to manage the solar micro-grid autonomically, distributed and dynamically by using solar power intermittently, the randomness in load, the dynamic grid pricing and the Non-Critical Load (NCL) control for the management of demand side conditions. The automated application is a control application for the agent, which represents each component of the micro grid by a representative and which has to process the environmental information received from Simulink and send the command signals it produces to the Simulink for validation. MAS is also linked in JADE with the Arduino processor and the environmental dynamics agent operations are verified practically. A smart test bed with the interaction of three micro-networks is implemented and the features inherent to MAS are utilized for a dynamic adaptation and distributed optimisation, so that consumers throughout the network are always able to achieve the best possible service and thus economic and environmental optimisation.

2.7 CONCLUSION

This chapter provides the detailed study of smart grid implementation concepts. The use of multi agent system in the power system and various implementation techniques are understood and also development of the concepts using logic controller was learned.

3. PROPOSED SYSTEM

The MAS is linked to Arduino in JADE with an intelligent micro-grid test bed, and the environment dynamics agent operations have been verified practically for the actual field deployment. The micro grid incorporates intelligent grid features including self-healing, demand side management, plug-and-play, dynamic rates, dynamic adapting and distributed optimisation, making it a smart microgrid. Reinforcement training is used to optimize micro-grid energy management. It is extended into Multi-Agent Reinforcement Learning (MARL) for distributed optimization of micro-grids. The performances of the reinforcement learning methods are compared with conventional methods. The effectiveness of MAS and MARL are investigated in micro-grid for autonomous adaptation in the dynamic environment, and for economic and environmental optimization.

3.1 REINFORCEMENT LEARNING BASED OPTIMAL ENERGY MANAGEMENT OF SOLAR MICRO-GRID

Reinforcement learning is a computational approach for automated goal-directed learning and decision-making. Reinforcement learning is falling between supervised and unsupervised learning. Here, the agents learn through experience. The agent learns the optimal action by continuously interacting with the environment. As an agent, the consumer as an agent interacts with the environment continuously and is taught, in an optimization-based control approach, to take optimum measures automatically in order to reduce energy consumption from the grid. Learning is integrated into the behavior of the consumer directly so that he can decide and work for optimal scheduling in his own interests. By interacting with the environmental influencing variables, the consumer evolves. In the control framework, the uncertainties in solar power and load are considered. We see a grid solar micro-grid system containing a local consumer, a renewable (solar photo-voltaic) energy generator and a store (battery). The battery planning in the dynamic charging and available solar power environment can be optimized with a model free refinement training algorithm, i.e., Q-learning. Solar energy and charging feed the reinforcement learning algorithm. Strengthening the battery use and the use of solar power reduces grid dependency. For a system reliability test, real numerical data simulation results are presented.

3.1.1 Modeling of The Solar Micro-Grid for Reinforcement Learning

3.1.1.1 Model of the Solar Micro-grid

A micro grid is a group of electricity sources and cargoes locally that normally works in synchronization with and connected with the traditional centralized grid but which can autonomously disconnect and operate under physical and/or economic conditions. Figure 3.1 shows the architecture of the micro-grid. The urban solar micro-grid includes an electricity transformer, a Dt, an external grid transformer, a Psp solar generator with the energy output, and a storage system with the battery charge level Rt. In part, the consumer can satisfy his demand by generating electricity from the local solar generator. When the solar source is available, electricity can be stored in the battery and stored in the discharge when necessary. Storage and solar generator are controlled by the consumer. The main objective of the solar micro grid is to meet the needs of the load and maximize solar power utilization.

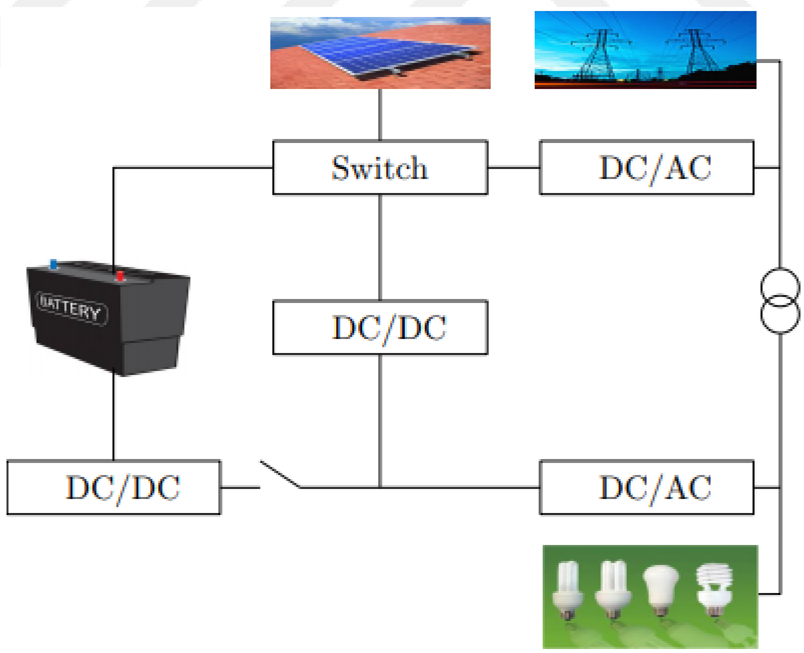

Figure 3.1: Solar micro-grid

3.1.1.2 Model for the Battery Storage

A simplified battery storage dynamic model is a discreet time system for the dynamic flow in time intervals.

$$R_t = R_{t-1} + R_t^{store.charge} + R_t^{store.discharge} \tag{1}$$

Battery storage models Equation 1 for discreet time systems (Kuznetsova et al. 2013), where Rt and Rt-1 store the battery's stored energy at times t and t−1 and R. Load t, R shop. Discharge t is the power flow between the solar generator and the battery and the consumer at time interval.

3.1.1.3 Modeling of the Consumer Agent

The external environment are considered to be the dynamic variation in charge, solar power and battery. In the consumer's decision-making, action and movement towards the goal, strengthening learning is modeled as an individual agent. In order to address the issues of feedback, enhanced learning is about sequential decision making. Markov (MDP) has become the standard formalism of sequential decision-making learning. The system status can be controlled as a set of states and actions by the environment in an MDP. Only this condition and not the history of the past is the effects of actions in a state. The goal is to manage the system so that certain performance criteria are maximized. The consumer agent uses a learning algorithm to interact, adapt and decide for its purpose, as defined in the form of MDP reward features, which include the Psp solar power output, Dt load and Rt battery load level.

3.1.1.4 Markov Decision Process

In sequential decision-making, the Markov Decision Process (MDP) is a way to model uncertainty. We formalize an MDP, taking account of discrete states and measures. The initial condition is s0 and each state is rewarded. The T(s0 |as) transition function shows the possibility of transition from state s to 0 for action a. For future rewards, a discount factor β in the range 0...1 is used. This shows that an existing prize is worth more than a future prize. If γ is near zero, future rewards are almost ignored; a near one γ places great value on future reward. The prize of a policy is the sum of each state visited by that policy's discounted expected usefulness. The best policy is to maximize the expected total discounted price.

3.1.2 Reinforcement Learning

The algorithm of reinforcement learning is used to adapt consumers to a vibrant environment by carrying out actions in the MDP environment. Strengthened learning between controlled and uncontrolled learning is falling. A simple enhancement learning system is shown in Figure 3.2. The agents watch and act on the environment. The environment receives a reward or punishment. The agent takes the next step in optimizing long-term rewards. The agent finds the optimum policy after a number of interactions for achieving long-term goals. An agent's goal is to develop an optimal policy based on interactive environmental learning. A certain number of its values, collectively referred to as its state and referred to as S(t) at t are characteristics for the environment. Each state has an intrinsic value that depends on an immediate compensation amount or cost, which is specified in t as R(t) when the state is entered. An A(t), which affects the next condition of the system: S(t+1), and thus the next reward/cost, may take a different time according to certain transition probabilities. Given the current situation, the choice of action of the agent is altered by its experience, i.e. by its previous experience and the associated cost to upgrade the decision-making process for future action the agent uses.

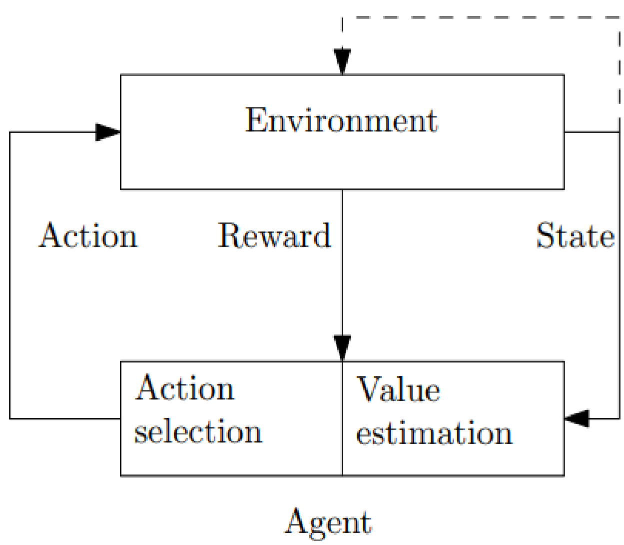

Figure 3.2: Reinforcement learning

3.1.2.1 Q Learning

Q Learning is a models-free learning, where the officer explores the environment and finds the next award plus the best the agent can accomplish. The agent does not need an ambient model in the Q study. It only needs to know which states exist and which actions are possible in each country. We assign an estimated value to each state, known as a Q value. We receive a reward when we visit a state and act. We use this prize to update our long-term assessment of the value of this action. We visit the countries frequently and continuously update the action values (Q values) until they are convergent. The Q learning algorithm(Sutton & Barto 1998) is outlined below. In the algorithm, γ is the discount factor and α, learning rate.

Algorithm 1: Q-learning

Input: s, r
Output: *action*

Set γ and rewards matrix R.;
Initialize $Q(s,a)$ arbitrarily.;
foreach *episode* **do**
  Initialize s arbitrarily.;
  repeat

  until *(*;
  *[h]for each step of episode)s is terminal Select a in s
  using policy derived from Q;
  Take action a, observe reward r, and next state s';
  $Q(s,a) \leftarrow Q(s,a) + \alpha[\, r + \gamma \max_a Q(s',a) - Q(s,a)\,]$;
  $s \leftarrow s'$;
end

3.1.2.2 Reward Function

Q-learning optimizes the battery planning of the solar microgrid. This is an action-reward process driven by quantitative indicators of performance that evaluate the actions or sequence of selected actions and feedback on the value of a future planning decision-making process. By selecting actions a0 and a1 from the battery schedule, optimization of the number rewards is achieved. Since the battery cannot be loaded and unloaded simultaneously, only one of the actions can always be selected and carried out at step t. By selecting an optimal sequence of actions the consumer seeks to increase his performance. The reward function is the response we get from the environment for the actions taken. If it is charging (a1) then the reward function is minimum of Psp and $B_{Diffrence}$ and if it is discharging (a0) then the reward function is minimum of Dt and B_{level}. Here, Bdi f f erence is the difference between maximum possible charge and the current battery level (B_{level}.).

3.1.3 Performance Evaluation of Micro-Grid with Q Learning

For a total of 100kW solar photovoltaics, the predicted solar energy value for 2014 is derived from the National Renewable Energy Laboratory (NREL) in our 2014 campus. Various residential, commercial and industrial load patterns are investigated, and the electrical engineering department load pattern is considered on our campus. The department's electrical load Dt is calculated hourly for all electrical appliances. As shown in Figure 3.3, the available solar energy is derived hourly for a 100kW unit per year. Figure 3.4 shows the solar power produced and load pattern for a day. Q learning algorithm provides the hourly readings from Psp and Dt data for the best battery planning to achieve consumer long-term goals.

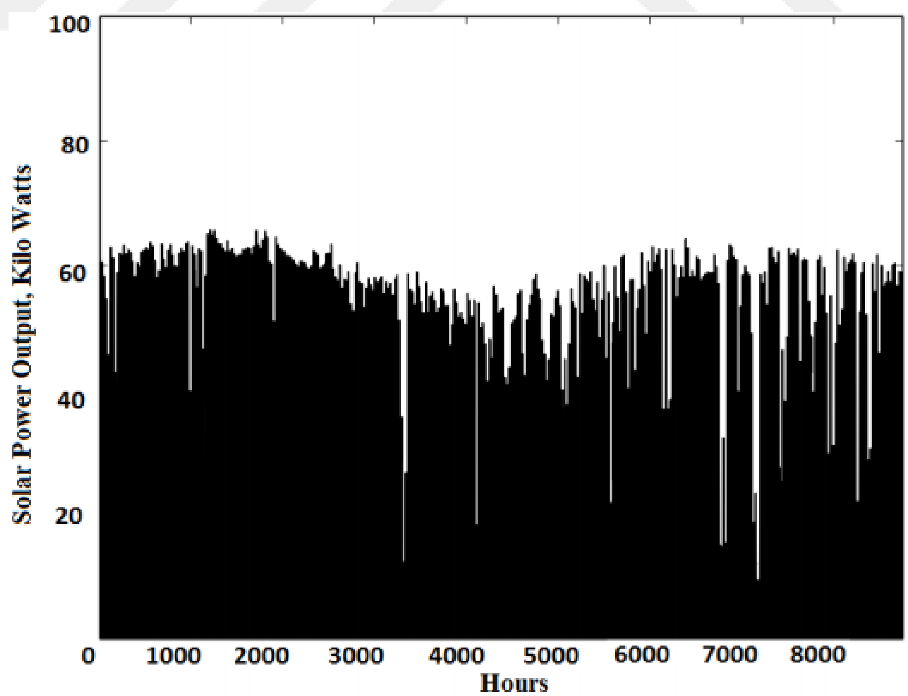

Figure 3.3: Solar power of 100 kW in a year

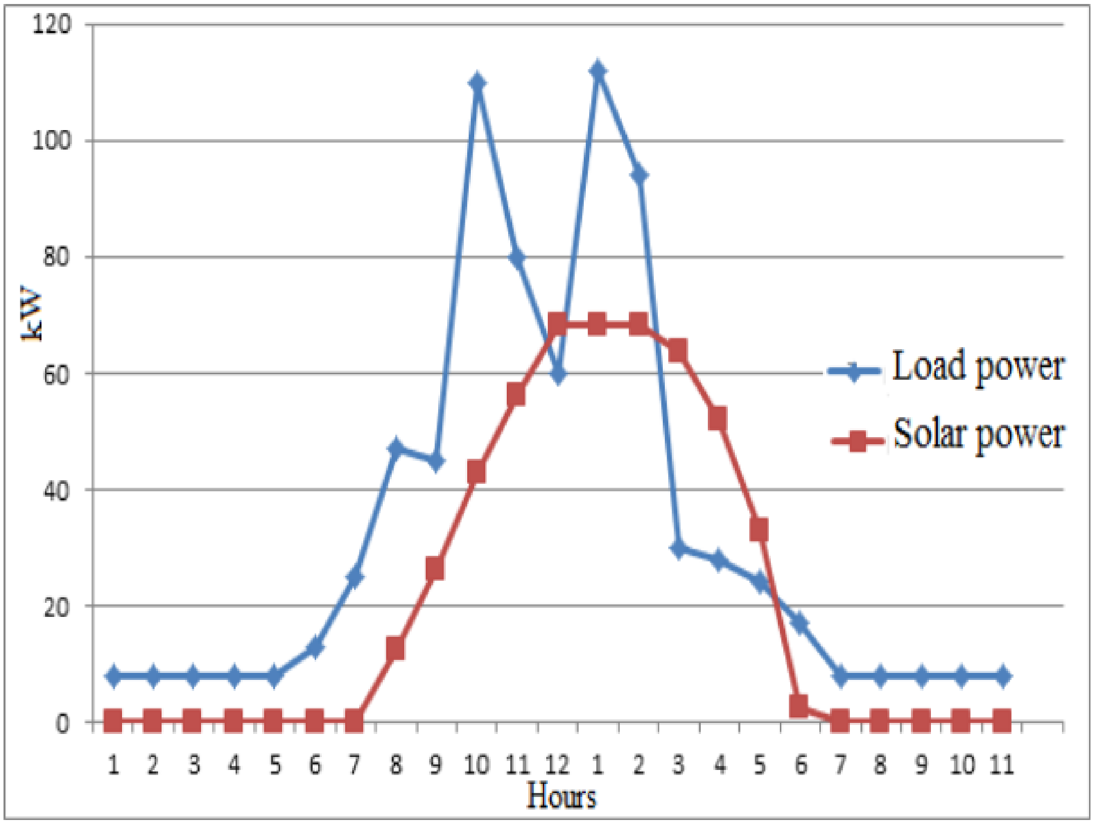

Figure 3.4: Solar power with load pattern in a day

The Q-learning programming language is implemented using Python. Two B0, S0 and Pg indicators show improvements in microgrid performance through strengthening learning. B0 estimates the increase in electricity use from the battery, which is defined as the cumulative power ratio between the batteries and the annual cumulative charge.

$$B_0 = \frac{\sum battery\ to\ load}{\sum load(D_t)} \tag{2}$$

Increased solar power utilization rate evaluated by S0, defined as the annual combined energy ratio between the solar photovoltaic system and annual cumulative power generation available.

$$S_0 = \frac{\sum solar\ to\ battery}{\sum solar\ power\ (P_{sp})} \tag{3}$$

Finally, parameter Pg indicates the cumulative annual power received from the external grid.

$$P_g = \sum Grid = \sum load(D_t) - \sum battery\ to\ load \tag{4}$$

For all these three parameters fifty simulations are performed each year and the average is calculated. Six years are simulated and the graphs are drawn for the average value of parameter per year.

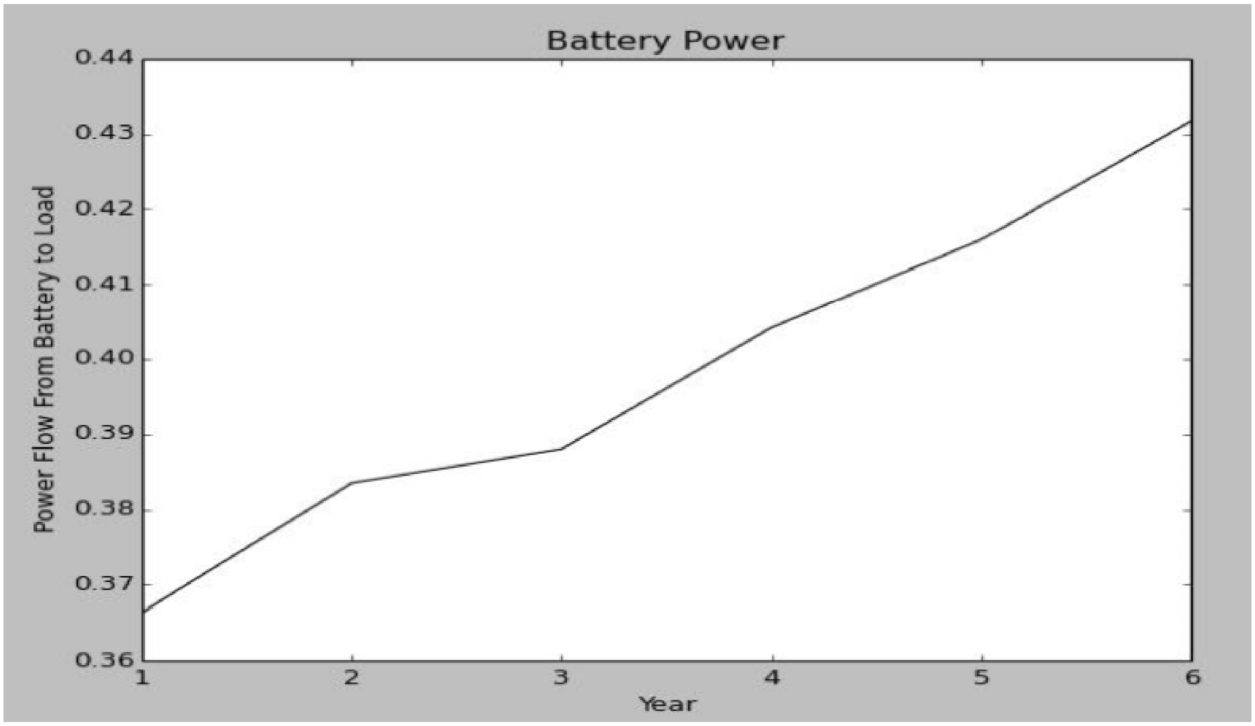

Figure 3.5: Average utilization of battery (B0)

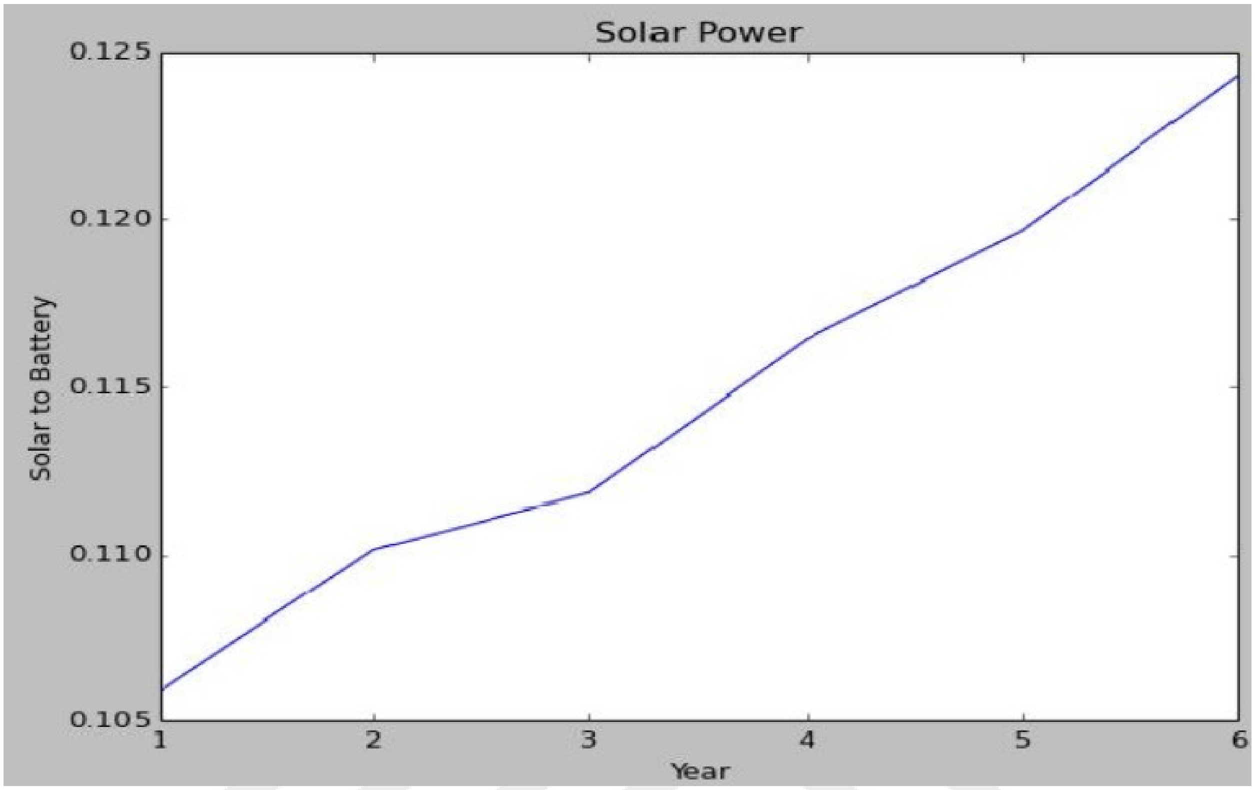

Figure 3.6: Average utilization of solar power (S0)

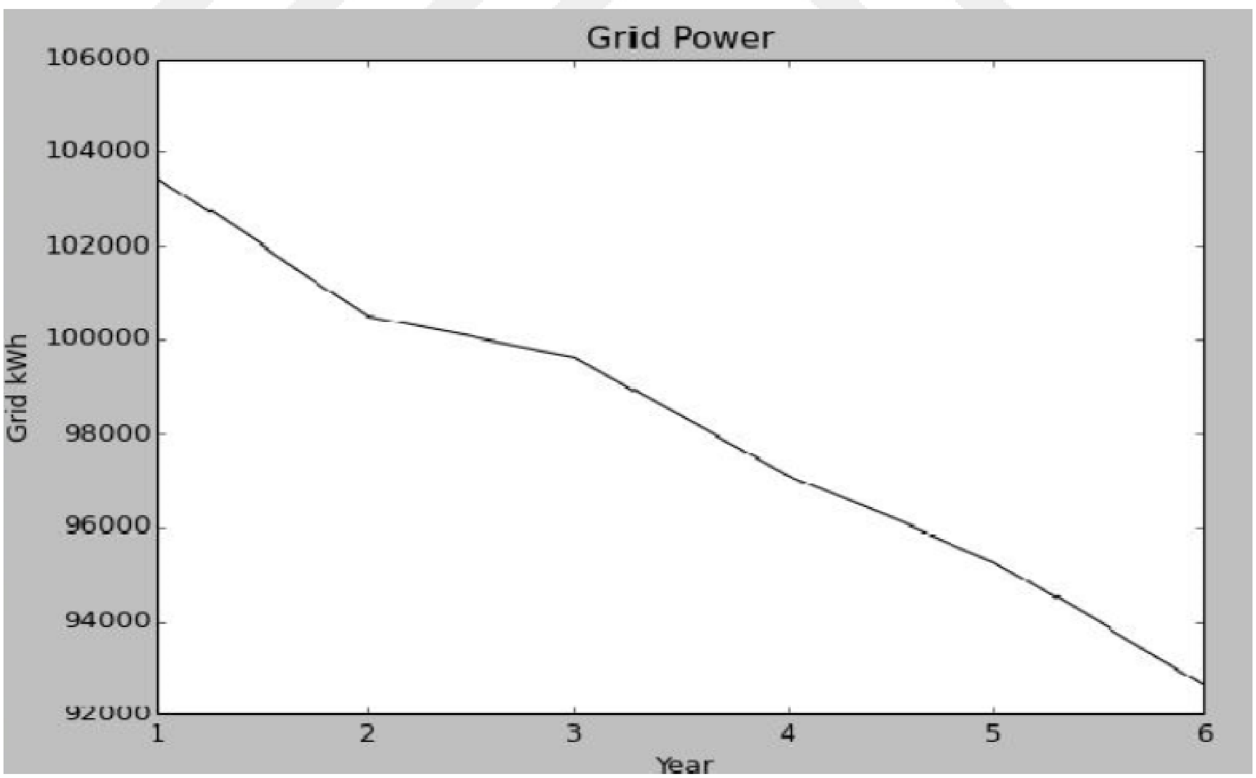

Figure 3.7: Average power from grid (Pg)

The average increase in battery and solar power utilities is shown in Figures 3.5 and 3.6; the power consumption from grid over time is shown in Figure 3.7. The power consumption from the grid is shown to decline gradually over a period due to Q-learning.

3.1.4 Comparison of Q Learning with Conventional Method

Figure 3.4 gives an estimate of 630 kWh (day) + 170 kWh (night) > 800 kWh for the total energy consumption for the department during the day and night. The total power generated in a day with the 100kW solar photovoltaic system is 490 kWh. Network energy requirement is (800 - 490) kWh = 310 = 365 days = 113150 kWh by conventional method for each year in the department. The grid energy requirement is 103400 kWh after Q learning is implemented. Therefore, in the first year when the consumer agent is enabled to learn enhanced energy, the energy required is saved at 9750 kWh and over time the energy needs of the grid are further reduced.

3.1.5 Discussion

The optimal battery planning is implemented with Q learn, a model-free learning algorithm. A simulation model has been developed to represent the dynamic interaction between consumers and the environment, thereby reducing long-term electrical consumption in the grid, in order to optimize self-employed battery planning to increase battery power and solar power consumption. The proposed Framework allows smart consumers to explore and understand the stochastic environment and use this experience to select optimal actions to reduce grid dependency. The performance of Q learning is compared with conventional method and it is observed that the grid dependency is relatively reduced over the period of time. This chapter discusses about single agent reinforcement learning for a single solar micro-grid. When two or more solar micro-grids are present in a in a distributed environment, multi agent reinforcement learning can be used for distributed optimization, which is discussed in the next section.

3.2 MULTI AGENT REINFORCEMENT LEARNING FOR DISTRIBUTED OPTIMIZATION OF SOLAR MICRO-GRIDS

A multi-agent system (MAS) is a loosely linked network of software agencies which work together to solve problems beyond each problem resolver's individual capabilities or knowledge. MAS designs and develops a decentralized optimization system. Reinforcement learning can be imparted to agents to reason and take decisions autonomously. Instead of a single agent, many agents can use reinforcement learning to optimize individually and also coordinate for global optimization. When many of these agents interact in the same environment, static environments (from a single agency's point of view) become non-stationary, because each agent has a particular influence on the environment. The main idea is that each controllable element – the power source, storage and loading – takes over an autonomous control process. MAS consists of a coordination algorithm, communication among the agents and organization of the whole system. The need for multi-agent adaptive systems combined with the complexities of working with interacting students have led to the development of a field of strengthening learning that builds on two fundamental concepts: enhancement learning and theory of games. The agents decide collectively how to answer user requests in unplanned situations (Dimeas & Hatziargyriou 2010). Coordinated Q Learning (CQ Learning) is used for coalition of agents which are using Q learning for optimization. Two agents cooperate for distributed optimization in dynamic environment created due to intermittent nature of the solar power supply and randomness in load.

3.2.1 Implementation of Distributed Optimization Using Coordinated Q Learning (CQ Learning)

The solar micro grid system comprises a two-solar photovoltaic loading and battery systems, each with a local photovoltaic (pv) solar panel system for consumer use. As an agent the consumer continually interacts with the environment and learns to act optimally using a model-free learning algorithm, Q-learning . The agent is intended to optimize the battery scheduling so that battery and solar energy are used more efficiently, in order to reduce energy consumption by grid. Multiple agents detect environmental conditions and make collective choices on how to respond with a multiagent strengthening algorithm, namely Coordinated Q Learning (CQL), to intermittent solar power and load randomness (Wiering & van Otterlo 2012). Each agent is individually optimized and helps optimize globally. Netz power is compared to distributed operation of two solar micro sets with CQ learning when two solar micro grids are operated individually by Q learning.

When acting alone in the environment, an agent learns optimum single agent policy. Each agent has a model of their expected incentives for all pairs of states using this prior understanding. The algorithm uses a test to detect if the rewards of the selected status pair change. The algorithm considers this to be the joint state where a collision occurs, where it is a risky state where the action is conditioned by other agents, if it detects a change in the immediate income received. State action pairs without collision are marked as safe states, i.e. the action of the agent in this country is independent from the status of other agents. Thus the algorithm tries first to find changes in the rewards received by an agent, based solely on its own condition before trying to detect the conditions in which these changes occur for the other agent. When an officer reaches a marked state, he will check whether he or she has to consider the other agent in one of the common states. If so, an action with this joint state is selected, and for convergence of Q values the following update rule is used.

$$Q_K^J(js,a) = Q_k^j(js,a_k) + a_t[r(js,a_k) + \gamma max_a Q(s'_k, a'_k) - Q(js,a_k)] \qquad (5)$$

In the Equation (5), Qk stands for the Q-table containing the independent states, and Q_k^j contains the joint states (js). The second Q-table is initially empty. The Q-values of the single states of an agent are used to bootstrap the Q-values of the states that were augmented to joint states (De Hauwere et al. 2010). When an agent arrives at a state which has not been marked by the agent, no updates are executed. The reward functions are the response we get from the environment for the actions taken. If it is charging (a1) then the reward function is minimum of Psp and B difference and if it is discharging (a0) then the reward function is minimum of Dt and B_{level}. Here, $B_{Diffrence}$ is the difference between maximum possible charge and the current battery level (B_{level}). Both the solar units powers and both the load units are considered for reward calculation during the joint state. CQ-learning is implemented using Python programming language.

Two solar micro-grids, one in electrical and another in the hostel building, have been taken into account in this case. Each agent optimizes the battery schedule by Q-learning so that solar power and battery utilities can be increased and the long-term goal of reducing grid energy is achieved.

The two agents then coordinate the battery planning by means of a CQ learning. The coordination improves battery programming optimization and reduces the grid power demand with CQ learning.

3.2.2 Comparison of Q Learning and CQ Learning in a Distributed Environment

In the building and the hostel buildings, we consider 100 kW and 200 kW Solar PV systems. The distribution of solar energy for the entire year on an hourly basis is shown in Figure 3.8 and Figure 3.8 of the National Renewable Energy Laboratory (NREL). All electric devices are considered and the electricity service and hostel loads are calculated on an hourly basis.

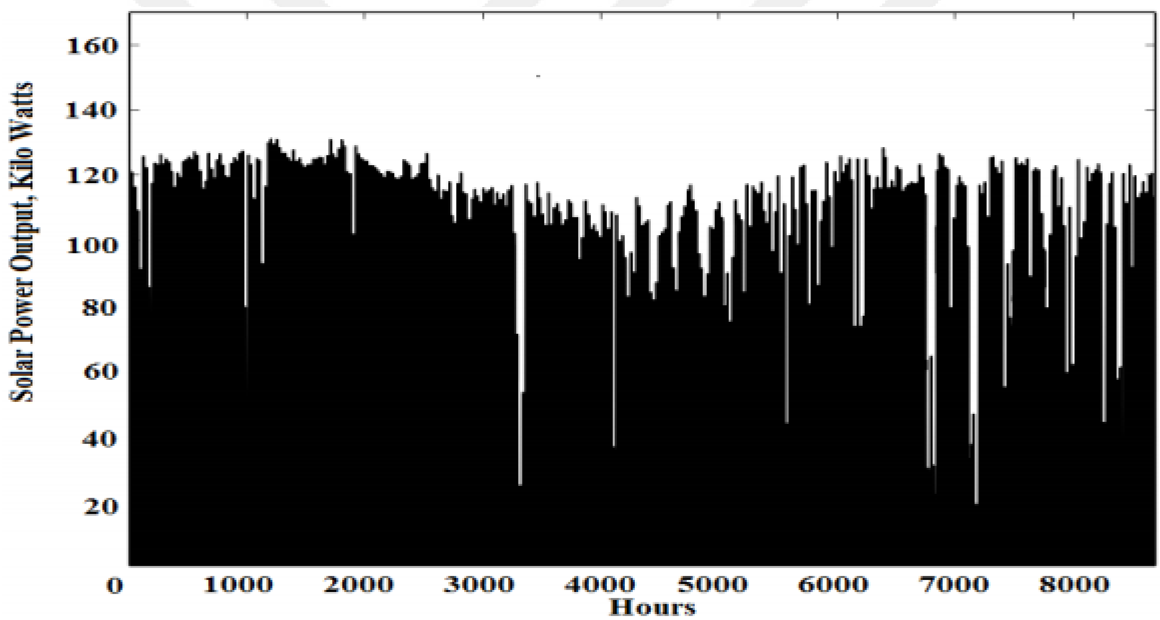

Figure 3.8: Solar power for 200 kW unit in a year

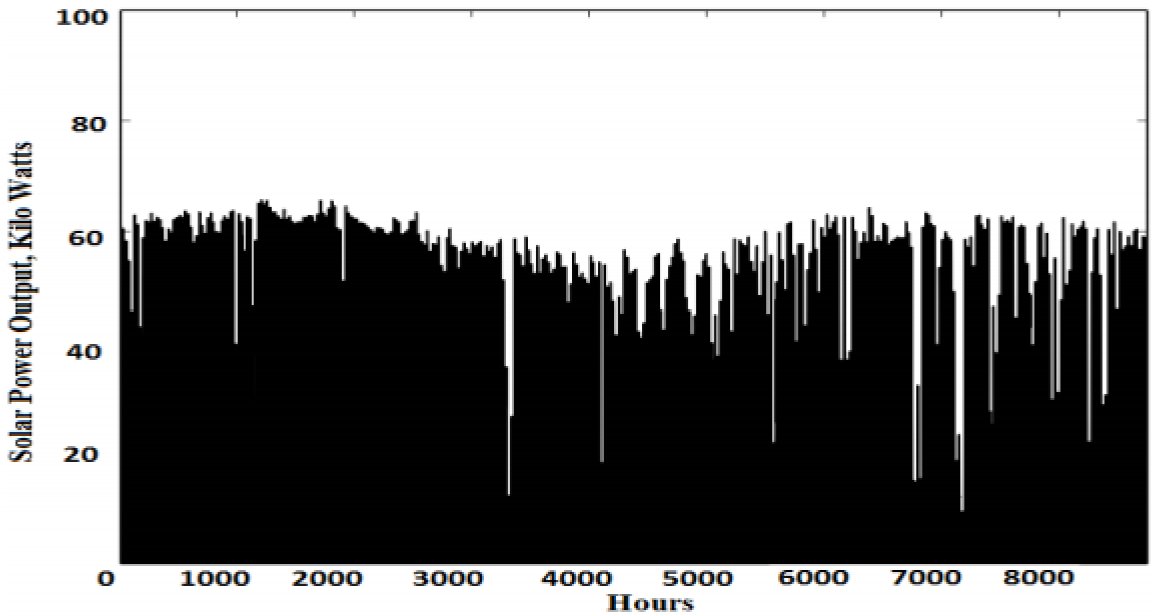

Figure 3.9: Solar power for 100 kW unit in a year

For the department and hostel, the hourly basis charges and the power generated in a day are shown in figures 3.10 and 3.11. High loads occur in the department in the morning and peak loadings occur in the hostel in the evening. Dt supplies the Q-Learning algorithm in both micro-networks with the hourly solar power, space and load (department and hostel).

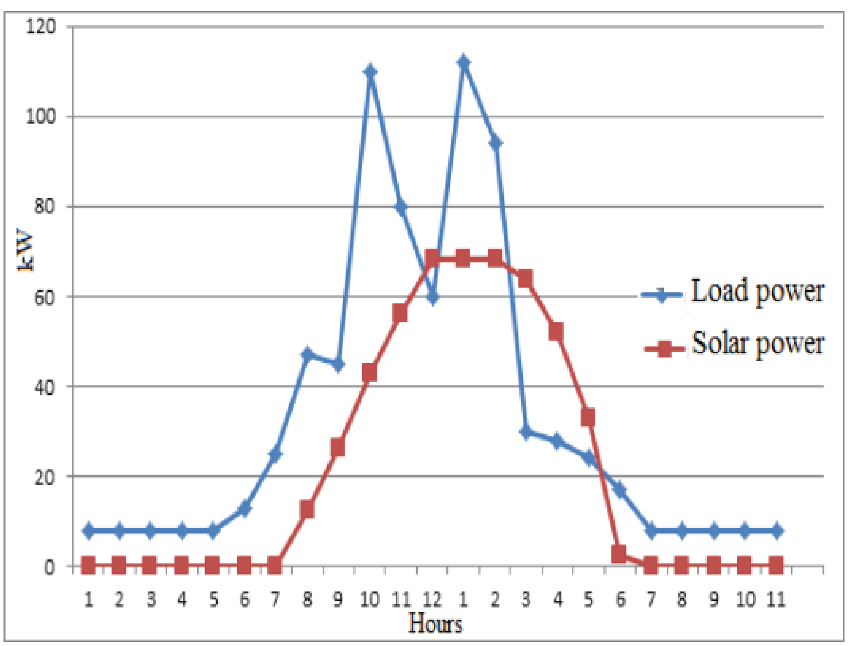

Figure 3.10: Solar power and load pattern in department

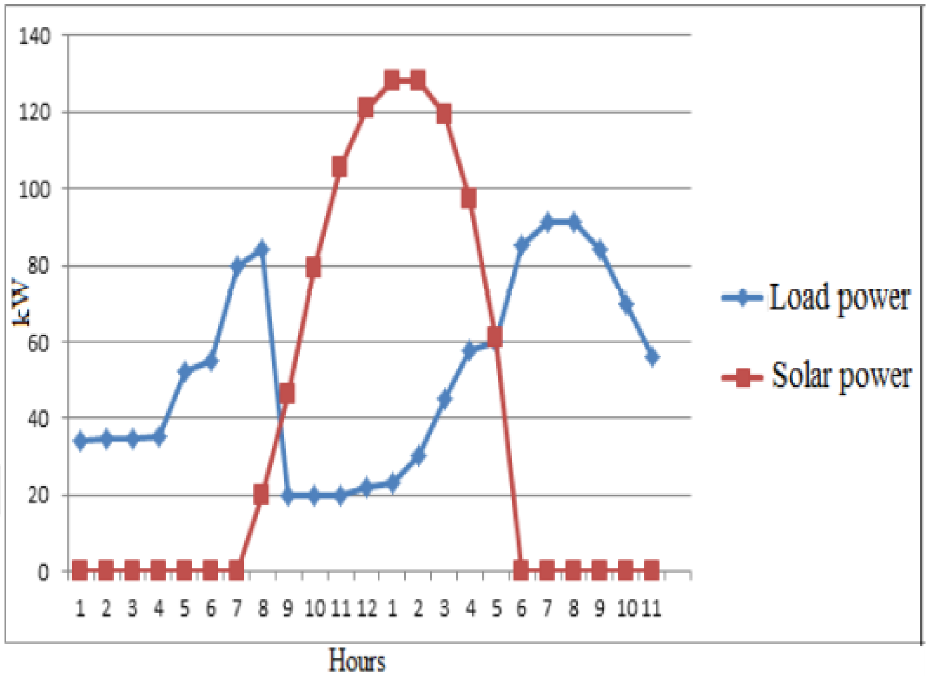

Figure 3.11: Solar power and load pattern in hostel

3 indicators 3 The Q Learning Algorithm shows B0, S0 and Pg improvements in micro-grid performance. The problem of explorers is solved with the Monte Carlo method by stochastically selecting and averaging action over a period of fifty times. The average energy supply, battery and electricity consumption values from the department's grid can be seen as shown below. The increase in solar panels assessed by S0 is defined as the annual cumulative energy ratio used by the solar panels to the annual cumulative solar energy system.

$$S_0 = \frac{\sum Solar\ to\ battery}{\sum solar\ power\ (P_{sp})} \tag{6}$$

B0 estimates the increase in electricity use from the battery, which is defined as the cumulative power ratio between the batteries and the annual cumulative charge.

$$B_0 = \frac{\sum battery\ to\ load}{\sum load(D_t)} \tag{7}$$

Finally, parameter Pg indicates the cumulative annual power received from the external grid.

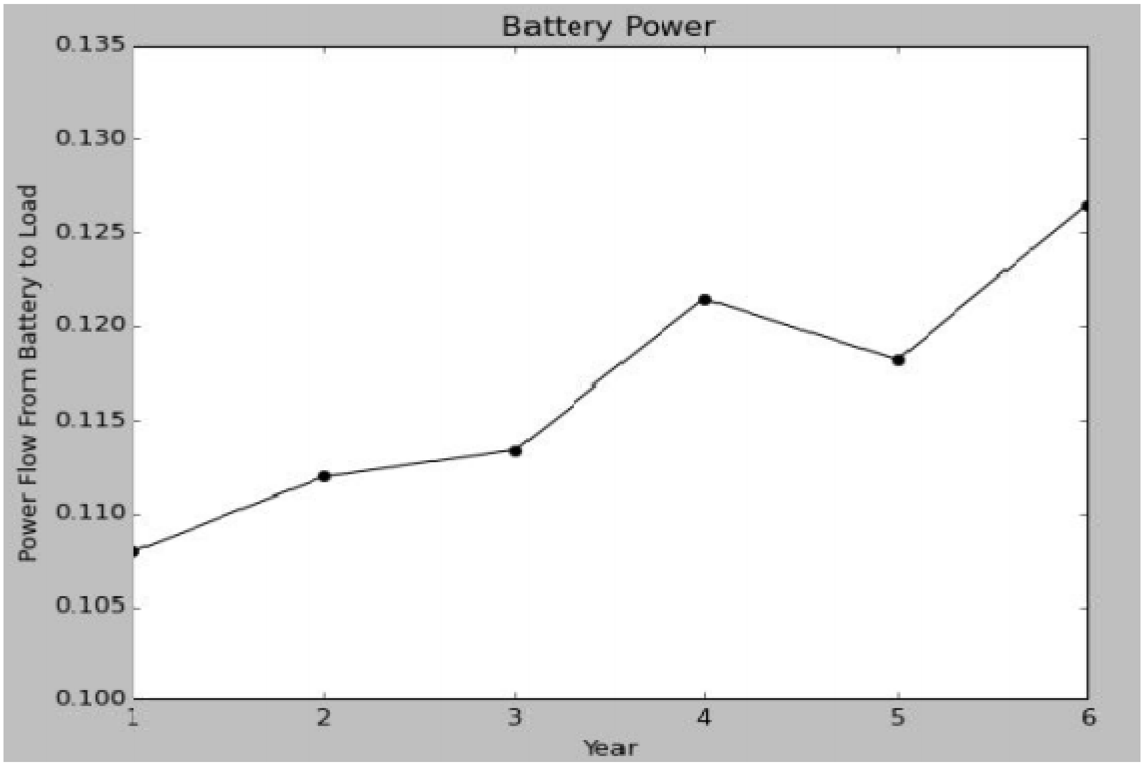

Figure 3.12: Utility of battery in department and hostel

$$P_g = \sum Grid = \sum load(D_t) - \sum battery\ to\ load \tag{8}$$

As shown in Figure 3.12., Figure 3.13 and Figure 3.14 the average solar energy utility, battery and power consumption values from the network of the hostel and the department can be found below and simulated for six years. The hostel and the department use Q-learning in order to optimize battery planning. Collective results are added to the performance indicator values of each microgrid.

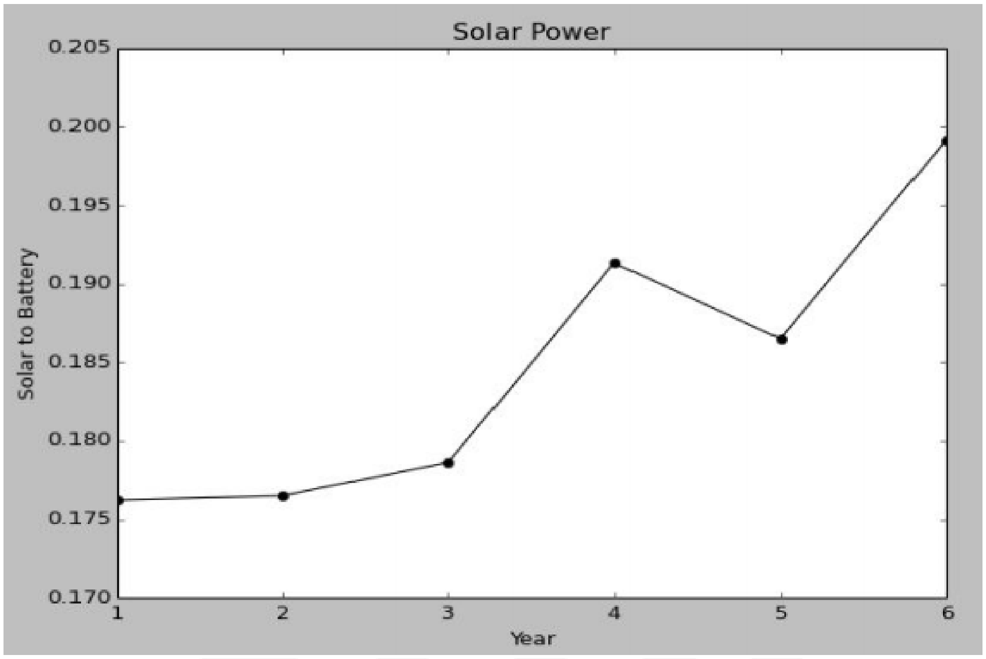

Figure 3.13: Utility of solar power in department and hostel

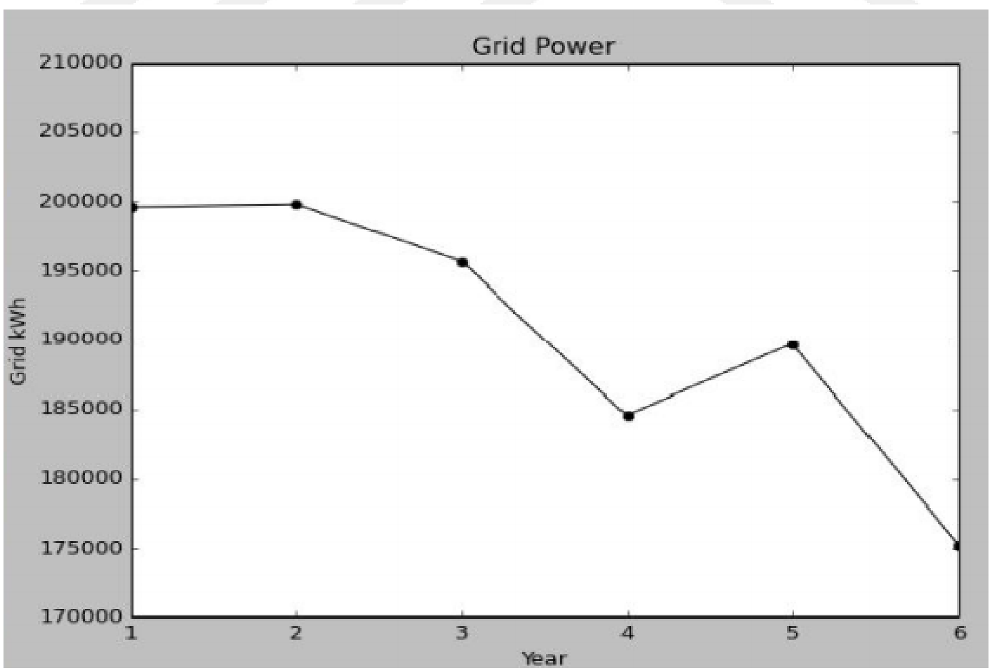

Figure 3.14: Grid power with QL in department and hostel

Figure 3.14 shows a reduction in the grid power consumption due to the Q Learning processes in both the devices (department and hostel). This is compared to grid energy consumption in a distributed environment with Coordinated Q Learning. The battery and solar power utilities are shown in Figure 3.15 and Figure 3.16 with CQ-Learning in joint state operations.

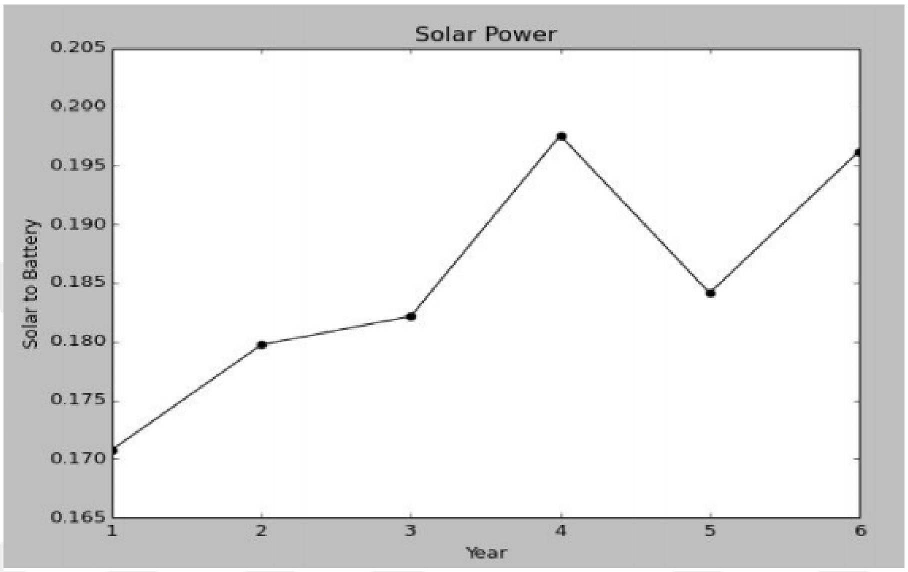

Figure 3.15: Utility of battery with CQL

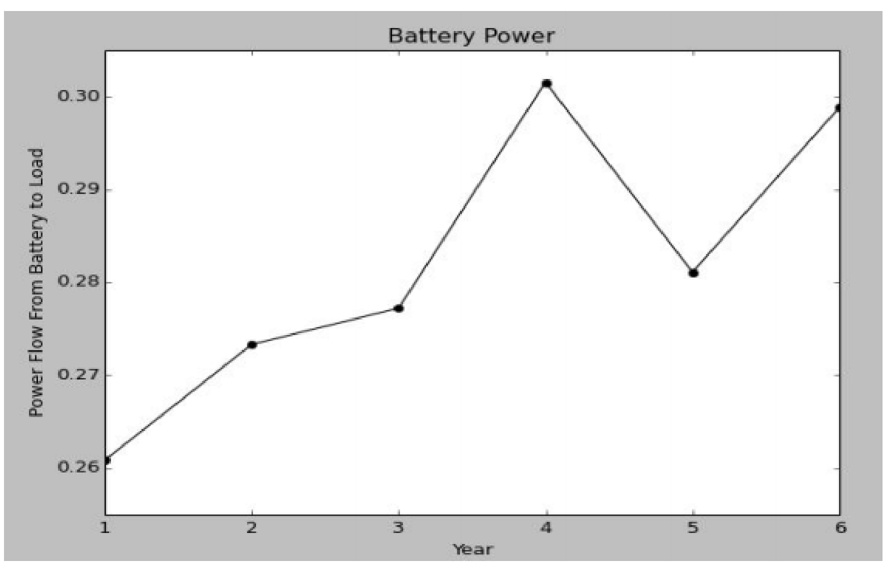

Figure 3.16: Utility of solar power with CQL

It is observed from the Figure 3.17 that the power consumed from grid is relatively reduced due to distributed optimization through Coordinated Q Learning.

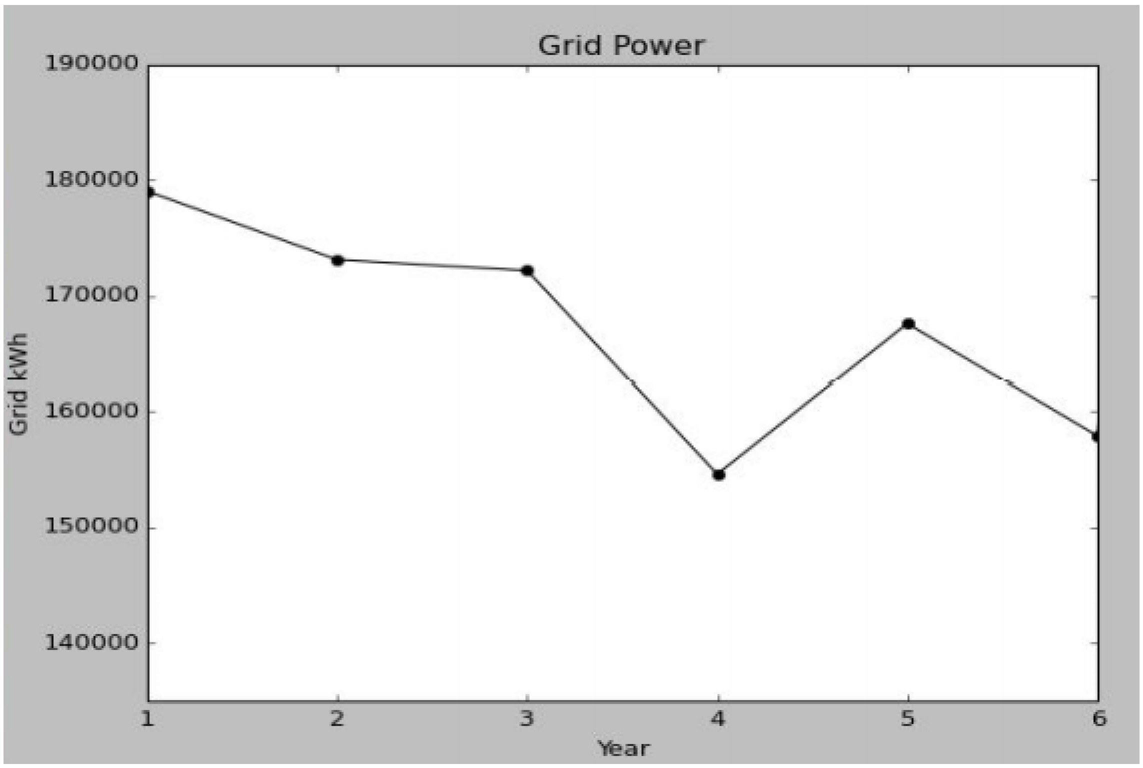

Figure 3.17: Grid power with CQ Learning

3.2.3 Comparison of Q Learning and CQ Learning with Conventional Method

The grid power saved due to CQ learning is compared with conventional methods. The department has a total energy requirement of 630 kWh (day) + 170kWh (night) = 800kWh for day and night. The total energy demand for the hostel is calculated as 322 kWh + 888 kWh = 1210 kWh, at day and at night. Department and hostel's total load are 800 kWh + 1210 kWh = 2010 kWh.

The department (100kW unit) generates total energy by solar photovoltaics at a rate of 490kWh. Calculated as 910 kWh of total energy produced from the hostel solar PV system (200 kWh unit). Total solar energy generated in department and hostel is 1400 kWh. The total grid energy requirement for hostel and department by conventional method, in a day, is $2010-1400 = 610$kWh, and in a year is $610 \times 365 = 222650$kWh.

In the first year, the grid energy required for both department and hostel in the solar micro-grid with individual Q learning is 200000 kWh and so 22650kWh energy is saved with Q learning in both department and hostel. The grid energy required using CQ Learning is 179000 kWh and so 43650kWh energy is saved due to CQ learning in the hostel and department. Therefore, when compared to individual Q learning applied department and hostel, 21000 kWh energy is saved in the first year due to CQ learning, applied collectively to department and hostel units, and the grid energy requirement is further reduced in the following years due to distributed optimization.

3.3 CONCLUSION

This chapter discussed about how reinforcement agent learning is imparted to agents in distributed environments. A multi-agent strengthening learning algorithm, namely Coordinated Q learning (CQ Learning), compared to 2 individual single-agent Q micro-grids and a conventional distributed optimization method, implements distributed optimization of two solar micro-grids. The dynamic interaction between the consumer agent and its environment enables the CQ learning to autonomously optimize battery scheduling so as to improve the battery utility, the use of solar power and reduce grid electricity consumption. A Python programming simulation model has been developed, and CQ learning involves relatively lower energy generation costs under the intermittent nature of Solar PV system and randomness of load.

4. CONCLUSION AND FUTURE WORKS

4.1 CONCLUSION

In this thesis, an MAS for advanced energy management of micro-grids is implemented in stages, starting from a proof of concept to an actual deployment in the field. Q-learning, a model free reinforcement learning algorithm, is implemented in solar micro-grid to improve the utility of solar power, thereby reducing the power consumption from grid, and the performance is compared against conventional method. It is proved that the grid dependency is reduced due to optimal energy management using reinforcement learning. In the micro-grids, distributed optimization is also implemented through the MARL algorithm, namely Coordinated Q-Learning (CQL), and compares its performance with Single Agent Reinforcement Learning and traditional methods.

There is still an infancy in science of a complex adaptive system. Agent technology provides a way to conceptualize micro grids as autonomous entities interacting and each acting and evolving individually in response to local interactions. The basis for a realistic computer simulation of the operation and behavior of these systems is a conceptualization. A systematic approach towards practical realization of MAS in smart grid is implemented in this research work. MAS linking with MATLAB for a comprehensive simulation, novel control strategies for implementing and validating smart grid features, setting up a smart micro-grid testbed for actual verification of all the smart-grid features, developing advanced negotiation algorithm and distributed optimization of micro-grids through MARL are the salient features of this work. MAS and smart grid technologies are not matured enough to measure and compare with standards and so there is hardly any way to compare the solutions of MAS.

4.2 FUTURE WORKS

MAS is considered as a promising engineering approach for developing adaptive software systems, able to handle increasing complexity, due to their open, heterogeneous, and continuous evolving nature. Future work will focus on scalability, Integration with several intelligent consumers with conflicting demands of renewable generators (solar or wind). For optimal energy management in a microsystem, enhancement learning algorithms are developed. In future, many other machine learning algorithms can be used to impart learnings to the agents. The MAS is envisioned as the future of an automated power monitor and power control system such as the Supervisory Control and Data Acquisition (SCADA). Computational and communication efficiency can be further improved by integrating MAS with Software Defined Network (SDN). In MAS, as of now, the focus is on stable, established and well proven approaches rather than experimental features. The operational context and life cycle of the solution design, cost and technology, standardization of MAS is yet to be explored for acceptance of the industry. A true test of MAS applications for micro-grid control can only come from rigorous field tests of hardware prototypes from multiple developers and vendors. Future work will focus on specific cases for specific industry customized to their need. Integration of hardware and their functionalities have to be explored. There is hardly any standard for comparing MAS-based solutions and hence there is a need for developing standards for practical applications.

REFERENCES

[1] James Odell 2015, ‘Objects and Agents Compared’, Journal of Object Technology, vol. 1, no. 1, pp. 41–53

[2] Russell, S & Norvig, P 2010, Artificial Intelligence: A modern Approach, Pearson Education Inc.

[3] Jennings, NR 2000, ‘On agent-based software engineering’, Journal of Artificial Intelligence, vol. 117, no. 2, pp. 277–296

[4] Azevedo, RD, Cintuglu, MH, Ma, T & Mohammed, OA 2017, ‘Multiagent-Based Optimal Microgrid Control Using Fully Distributed Diffusion Strategy’, IEEE Transactions on Smart Grid, vol. 8, no. 4, pp. 1997–2008.

[5] Dou, C, Yue, D, Li, X & Xue, Y 2016, ‘MAS-Based Management and Control Strategies for Integrated Hybrid Energy System’, IEEE Transactions on Industrial Informatics, vol. 12, no. 4, pp. 1332–1349.

[6] Paulo Leitao & Stamatis Karnouskos 2015, Industrial Agents: Emerging Applications of Software Agents in Industry, Elsevier Science Publishers, Amsterdam, The Netherlands.

[7] Gerhard Weiss 2013, MultiAgent Systems, MIT Press, Cambridge, London

[8] Rohbogner, G, Hahnel, UJJ, Benoit, P & Fey, S 2013, ‘Multi-agent systems’ asset for smart grid applications’, Computer science and information systems, vol. 10, no. 4, pp. 1799–1822.

[9] El-hawary, ME 2014, ‘The smart grid—state-of-the-art and future trends’, Electric Power Components and Systems, vol. 42, no. 3, pp. 239–250.

[10] Dimeas, AL & Hatziargyriou, ND 2005, ‘Operation of a multiagent system for microgrid control’, IEEE Transactions on Power Systems, vol. 20, no. 3, pp. 1447–1455.

[11] Eddy, YSF & Gooi, HB 2011, ‘Multi-agent system for optimization of microgrids’, Proceedings of the eighth IEEE international conference on power electronics, pp. 2374–2381

[12] Pipattanasomporn, M, Feroze, H & Rahman, S 2009, ‘Multi-agent systems in a distributed smart grid: Design and implementation’, Proceedings of the IEEE power engineering society power systems conference, pp. 1–8

[13] Logenthiran, T, Srinivasan, D, Khambadkone, AM & Aung, HN 2012, ‘Multiagent system for real-time operation of a microgrid in real-time digital simulator’, IEEE Transactions on Smart Grid, vol. 3, no. 2, pp. 925–933.

[14] Nunna, HSVSK & Doolla, S 2013, 'Multiagent-based distributed-energy-resource management for intelligent microgrids', IEEE Transactions on Industrial Electronics, vol. 60, no. 4, pp. 1678–1687.

[15] Zhang, Z, McCalley, J, Vishwanathan, V & Honavar, V 2003, 'Multiagent system solutions for distributed computing, communications, and data integration needs in the power industry', Proceedings of the IEEE power engineering society general meeting, pp. 44–47.

[16] Gomez-Sanz, J, Garcia-Rodriguez, S, Cuartero-Soler, N & Hernandez-Callejo, L. 2014, 'Reviewing microgrids from a multi-agent systems perspective', Energies, vol. 7, no. 5, pp. 3355–3382.

[17] Abhilash Kantamneni, Laura, EB, Gordon, P & Wayne, WW 2015, 'Survey of multi-agent systems for microgrid control', Engineering Applications of Artificial Intelligence, vol. 45, pp. 192–203.

[18] Perkonigg, F, Brujic, D & Ristic, M 2015, 'Platform for multiagent application development incorporating accurate communications modeling', IEEE Transactions on Industrial Informatics, vol. 11, no. 3, pp. 728–736.

[19] Robinson, CR, Mendham, P & Clarke, T 2010, 'Macsimjx: A tool for enabling agent modelling with simulink using jade', Journal of Physical Agents, vol. 4, no. 3, pp. 1–7.

[20] Eddy, YSF, Gooi, HB & Chen, S 2015, 'Multi-agent system for distributed management of microgrids', IEEE Transactions on Power Systems, vol. 30, no. 1, pp. 24–34