Reinforcement Learning for Microgrid Energy Management

Usman Inayat*, Sajid Mahmood, Malik Tahir Hassan***

* Department of Computer Science, University of Management and Technology Lahore, Pakistan.
** Department of Informatics and Systems, University of Management and Technology Lahore, Pakistan.

Abstract- The consumer functions as an agent in a control strategy for solar microgrid energy management that employs optimization and learning to take optimal actions to minimize grid power usage through continuous interaction with the environment. Consumer behavior involves learning directly, allowing them to make decisions and take actions that are best for themselves in terms of scheduling. Consumers evolve through their interactions with environmental factors. This study examines a solar microgrid system that is connected to the grid and comprises a battery, a renewable generator, such as a solar photovoltaic system, and an energy consumer. A deep Q-learning-based energy management system, which is a Reinforcement Learning model, is used to maximize battery scheduling when solar power availability and dynamic load are involved. An algorithm for reinforcement learning is fed by the solar power and load. The solar microgrid operates optimally when the battery and solar power generator are both more useful and the electricity billing costs are reduced. The results of a system-dependability test using numerical data on price, cost, and load are displayed through simulations. The proposed energy management system accounts for the reduction in electricity bills from the main grid and uncertainty in the load and solar power.

Index Terms- Energy Management System (EMS), Microgrid, Reinforcement Learning (RL), Deep Q-Learning, Deep Q Network (DQN)

I. Introduction

A microgrid (MGs) is used to connect consumers with an electricity supply by using small-scale electricity network. Microgrids also defined as a one system that is controlling the group of micro-sources and loads [1]. One system microgrids used in the production of electricity, electrical vehicles and large batteries for storage, software and hardware for monitoring and distribution of energy and energy consumption by end users with the help of connected distributed energy resources like wind turbines, Fuel burning generators, solar arrays. In the case of power outage, microgrids have potential of maintaining the electricity fluency. Microgrid can be located behind the meter or in front of the meter [2]. Typically, microgrids are set up in one of two ways. The term "customer microgrid" refers to a microgrid that is entirely located at one particular site and is likely managed by a utility customer. The second type of microgrid is made up of a portion of the regulated grid that incorporates multiple technologies located at different locations.

A microgrid is a low-voltage distribution network connected to a point-of-common-coupling (PCC) that is located immediately downstream of a substation for distribution. Controllable loads, distributed energy storage (DES), and distributed generators (DGs) are some of the components that make up a microgrid. The distinct features and properties of a microgrid's constituent parts pose a special challenge for grid management and operation. In any particular microgrid, the required energy management plan may change considerably from a traditional power system based on the features and distributed energy resources (DERs) and DES nodes 's perforation [3]. According to Asmus (2010) and Lasseter (2002), a prototypical microgrid operates in two modes: interconnected mode, which is connected to the main grid via the distribution substation transformer, and islanded mode, which is autonomous when the microgrid is cut off from the main grid throughout a blackout. The microgrid continues to operate and function as a stand-alone unit when in the islanded mode. Because of hardware restrictions and safety considerations, the islanding process is not allowed in a traditional power distribution system. Modern solid-state transformers, or sophisticated power electronic devices, combine high computational capacity, digital data processing, metering, protective relaying, two-way power flow, switching operations, and two-way communication. Under a range of operational situations, a microgrid's interconnection switch can accommodate islanding and resynchronization. Based on the current conditions of microgrid components operations, a microgrid EMS is control software that can serve the load economically, ideally distribute power output of DG unit and trigger the system's resynchronization reaction automatically when it switches between connected and islanded modes of operation.

Multiple techniques and models proposed to address the energy management in microgrids. Reinforcement learning plays a major role to manage and optimize energy resources in past years. Researchers presented an agent based system for addressing energy management challenges in solar microgrids. The objective of the agent is to meet the energy demand within the solar microgrid by optimizing battery usage while ensuring the provision of satisfactory services and reduction in cost of electricity with the implementation of Q-learning. [1,4,5,7,10,11,14]. Real-time scheduling approach for of a Microgrid (MG) energy management presented in [3], taking into account the fluctuations in load demand, the production of renewable energy, and the cost of power. They put forth a learning-based approach that is independent of a clear uncertainty model. In [6], energy management system (EMS) aims to optimize the coordination of these flexible sources by prioritizing resources, implementing control signals of demand and considering prices of electricity. The asynchronous advantage actor-critic technique was modified to include replay of experiences and training of semi-deterministic phase, which led to remarkable model performance and convergence to almost optimum policies.

Major contribution of the proposed system is that as the local renewable energy source (solar generator) is available, the consumer can store power in the battery to partially meet his need. Consumer can then release the stored energy as needed. Both the solar power generation and the storage are within the consumer's control. The primary goal of a solar microgrid is to maximize solar power utilization, optimize battery and solar photovoltaic (PV) system performance, and qualitatively meet load needs.

The rest of the article is structured as follows: Section II presented the related work. Section III described the proposed system for energy management by using deep network. The performance evaluation and results of the performed experiment are discussed in Section IV. The research conclusions are narrated in Section V. Finally, the future discussions of research are reported in Section VI.

II. LITERATURE REVIEW

This section focuses on recent research that examines the management of electrical energy using renewable resources. In [1], a single-agent system was implemented for solar microgrids, which included a PV source, battery bank, desalination unit, and local consumers. The objective of the system was to optimize the battery utilization while maintaining satisfactory service levels. The system employs Reinforcement Learning to identify the optimal policies and map states to actions. The reward function considers the battery state, water level, and power demand coverage. The simulation was conducted using MATLAB with a PV source as the sole energy source. Authors in [4] proposed distinct Q-learning methods for microgrid-independent agents, allowing them to function as learners and coordinate their behavior. Fuzzy logic, which is a computational tool, is used to express complex processes without complex models. The effectiveness of the fuzzy Q-learning algorithm was demonstrated in this study.

The energy management method proposed in [3] for the real-time scheduling of microgrids considers uncertainties such as power costs, renewable energy production, and load demand. To reduce daily operational costs, this method utilizes deep reinforcement learning (DRL) and a Markov decision process (MDP). At each time step, the reward function was equal to the negative value of the scaled operating cost. A microgrid model with an energy-management system (EMS) optimizes the resource coordination. In [6], the researchers presented a model that prioritized resources, implemented demand control signals, and considered electricity prices. The asynchronous advantage actor-critic technique was modified to achieve impressive performance and near-optimal policies. The microgrid has three layers: physical, information, and control. Researchers explored different deep reinforcement learning (DRL) algorithms in [8] and compared their performance. The authors defined the reward function and proposed three DRL-based methods to solve the Markov decision process (MDP) model. This study provides insights into the effectiveness of different algorithms for energy management in microgrids.

An energy management approach for home microgrid systems was presented in [9], which utilized the Shapley value and reinforcement learning (RL) based on Model Predictive Control (MPC) to establish an energy trading policy aimed at minimizing operating, environmental, and EV charging expenses. Similarly, a decentralized energy management system (EMS) was proposed for a smart microgrid in [10], using a reinforcement learning (RL) strategy. The EMS aims to maximize benefits for all entities within the microgrid, while considering the stochastic behavior of renewable energy sources and customer consumption.

A cost-based framework was developed to enhance the performance of standalone microgrids using reinforcement learning and two rule-based Energy Management Systems (EMSs). The primary objectives of this framework are to minimize electricity costs, increase photovoltaic (PV) utilization, and improve the system efficiency [11]. The EMS collects data from 24 states, each representing an hour of the day [12]. To address the challenge of sparse rewards in deep reinforcement-learning-based microgrid energy management, a multistage reward mechanism was proposed [13]. Additionally, the CDQN technique was suggested to address the increasing complexity of energy trading within and between microgrids [13]. Finally, a distributed multi-agent approach was proposed to reduce the microgrid's dependence on the main grid and to incorporate consumer and Distributed Energy Resource (DER) agent variabilities into optimal policies [14].

III. METHODOLOGY

The effective management of energy within microgrids poses a complex challenge, requiring innovative solutions to optimize resource utilization and enhance system efficiency. The aim of this study is to provide a comprehensive overview of the reinforcement learning process and proposed energy management system by using Deep Q-Agent to address the unique characteristics and constraints of microgrid systems.

A. Microgrid EMS

A microgrid is a small-scale collection of power loads and resources that often runs in tandem with the conventional centralized grid like macrogrid, however it may also disconnect and run independently when the environment whether natural or man-made demands it. A consumer with a dynamically fluctuating demand, a transformer supplying energy via outside the grid, a solar photovoltaic system in the availability of output power and having facility of storage with a level of battery charge are all components of an urban solar microgrid [15, 16]. Figure 1 depicts the proposed architecture of the microgrid energy management system. ESS represents energy storage system which stores energy from solar system (PV) and utilized in Load (Domestic and commercial). Energy Management System (EMS) is used to control energy consumptions, utilized through storage batteries and switch grid whenever needs in terms of minimizing the cost.

B. Deep Q-Network

There are several choices to take into account when creating the microgrid environment's framework. A number of articles incorporated various aspects of the microgrid, including limitations on the generation side and limitation on the demand side [17]. Reference [18] takes into account both the limitations of renewable energy sources and microgrid operation.

Numerous studies examined limitations on energy storage, such as rates of charging and discharging, caps on the cost of electricity, or emissions of carbon [19]. The three most important microgrid constraints of power flow limits, power output constraints from various distributed generation devices, and restrictions on power exchange with the external grid are covered in this study. The proposed microgrid model includes variable loads, solar PV, wind turbines, energy storage devices, and some traditional distributed generators connected to PCC's utility.

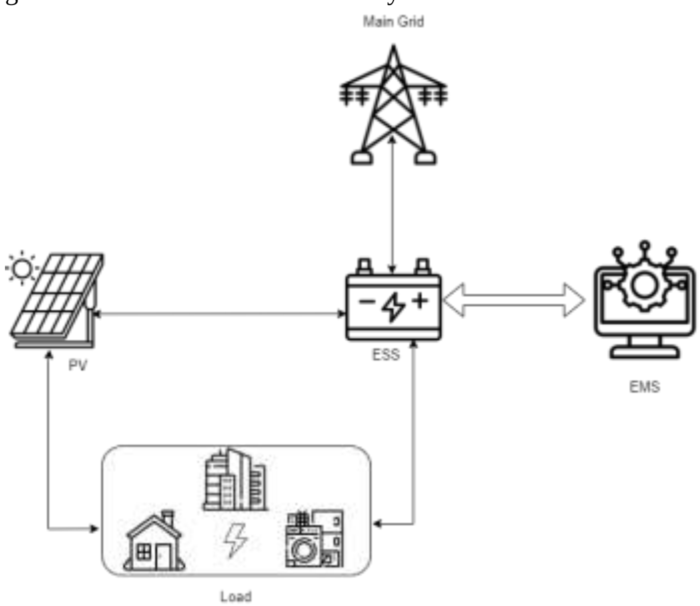

Figure 1. Microgrid Energy Management System

The following equations limit the traditional generators that are taken into account in the proposed formulation:

$$P_{max}^{DG_h} \geq P_t^{DG_h} \geq P_{min}^{DG_h} \qquad (1)$$

$$S_{max}^{{DG_h}^2} \geq Q_t^{{DG_h}^2} + P_t^{{DG_h}^2} \qquad (2)$$

In (1) and (2), $P_t^{DG_h}$, $Q_t^{DG_h}$ represents the active and reactive power output of generator h, at time t. $S_{max}^{DG_h}$ denoted the rating of generator. The following equation illustrates how a quadratic function provides the cost function of running conventional generators:

$$C_t^{DG_h} = \left[a_h\left(P_t^{DG_h}\right)^2 + b_h P_t^{DG_h} + c_h\right]\Delta t \qquad (3)$$

Constants are represents as a_h,b_h and c_h. The following formulae also impose limitations on energy storage devices:

$$P_{max}^{E} \geq P_t^{E} \geq 0 \qquad (4)$$

$$E_{max} \geq E_t \geq E_{min} \qquad (5)$$

$$E_t = E_{t-1} + \eta_{ch} u_t P_t^E \Delta t - (1 - u_t) P_t^E \Delta t / \eta_{dis} \qquad (6)$$

In above equations, charging and discharging power represented by P_t^E and state of charge given by E_t. u_t is used to show that ESS charging, u_t=1 for charging and u_t=0 for discharging. Efficiency of charging and discharging are presented by η_{ch} and η_{dis}. Time period denoted with Δt.

The following equations, which control the power exchange with the utility, are taken into consideration:

$$P_{max}^{U}, \forall t \geq P_t^U \geq -P_{max}^{U} \qquad (7)$$

$$(S_{max}^{U})^2 \geq {P_t^U}^2 + {Q_t^U}^2 \qquad (8)$$

In above equations, active reactive power exchanges with the utility represented by ${P_t^U}^2$ *and* ${Q_t^U}^2$. Maximum change of complex power is denoted by S_{max}^{U}.

The following formula, where Rt is the current price, provides the cost of acquiring power from the utility.

$$C_t^U = P_t^U . R_t . \Delta t \qquad (9)$$

Controlling the energy resources also takes power flow restrictions into account. Power flow limitations at each branch ij serve as a representation of this, as denoted below:

$$S_{max}^{{ij}^2} \geq Q_t^{{ij}^2} + P_t^{{ij}^2} \qquad (10)$$

The voltage is limited in the following ways to guarantee that voltage limitations are within specified values:

$$|V^i|_{max} \geq |{V^i}_t| \geq |V^i|_{min} \qquad (11)$$

where V^i the lowest and maximum values are provided to limit the voltage at bus i at time t.

Users have the flexibility to export labeled datasets in formats tailored to their preferred AI model architectures. As AnnoVate progresses, the trained model developed during the labeling process becomes available to users for their applications. Users are able to export their labeled dataset and the trained AI model using the export module of AnnoVate as shown in Figure 2. The methodology prioritizes ethical practices, including obtaining user consent, protecting user data, and addressing potential biases in the training data.

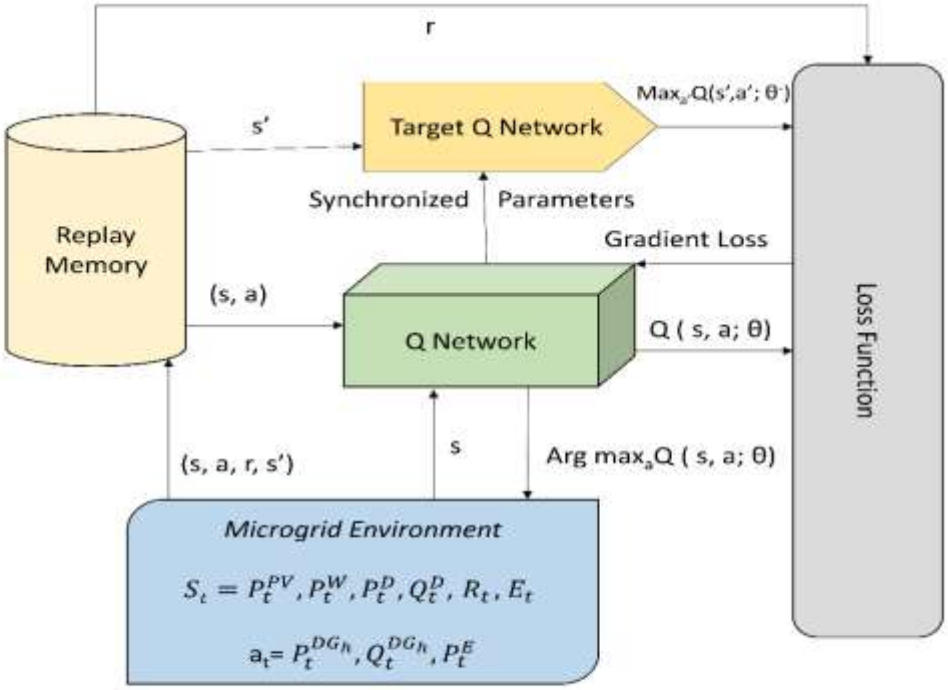

Figure 2. DQN Architecture for Microgrid Energy Management System

The following is the formulation of an MDP using the microgrid model discussed above. PPV provides the state variables at every time step t as ($P_t^{PV}, P_t^W, P_t^D, Q_t^D, R_t, E_t$), the power output from the solar PV plant is denoted by P_t^{PV}, the output power from the wind turbine is denoted by P_t^W, the active and reactive power sources are denoted by P_t^D and Q_t^D, the real-time electricity price is denoted by R_t, and the energy level of the energy storage system is denoted by E_t. One might formulate the model to incorporate solely present state variables or historical factors from the past.

The ($P_t^{DG_h}, Q_t^{DG_h}, P_t^E$) gives the action space, where P_t^E is the charging/discharging power of the energy management of microgrid system employing deep Q-network RL based energy storage system, and $P_t^{DG_h}$ and $Q_t^{DG_h}$ are the active/reactive power from conventional generators. To make the action space compatible with the DQN algorithm, it is built to be discrete.

The primary goal of the reward function is to maximize operating costs within the constraints of the system. As a result, the incentive function is designed to be correlated with the operational costs of the microgrid. These consist of the expenses associated with running traditional generators and buying electricity from utilities. A straightforward formulation of the reward function at every time step is given by equation (12).

$$r_t = -(\sum C_t^{DG_h} + C_t^U) \tag{12}$$

Furthermore, the reward function takes into account any infractions of the power flow limits and terminates the training episode in addition to offering a negative reward. Every time an action goes against the limits on energy storage, it also has a negative reward, discouraging the violation of such constraints.

Figure 2 shows how the agent interacts with its surroundings and the general training process. The q-network receives the microgrid's state variables and uses a greedy policy to maximize the q-value when making decisions. To train the target network and the online network, the interactions between the agents and the environment are gathered and stored in a replay memory. This procedure keeps going until the episodes conclude or the termination requirements are satisfied.

Figure 3 displays a high-level flowchart of the many phases taken to address the problem, summarizes these formulation steps.

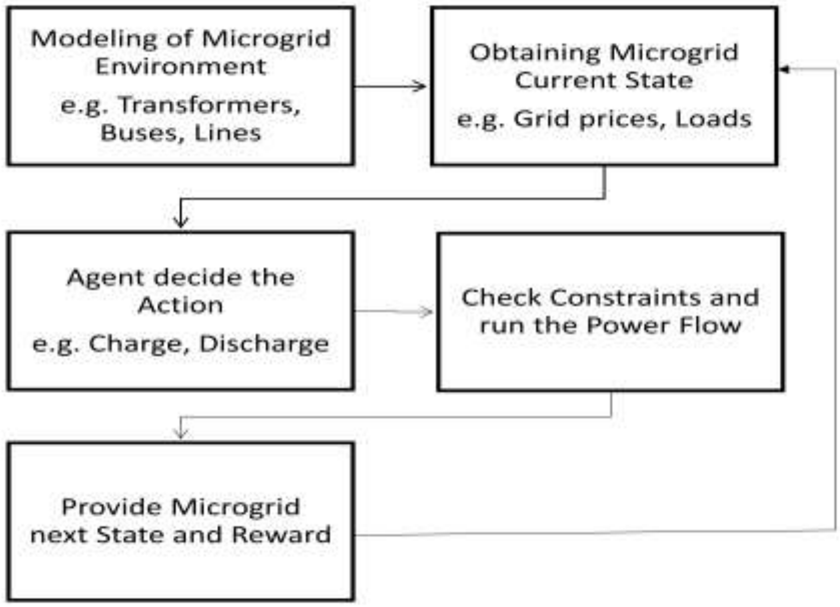

Figure 3. Flow chart of Proposed Model

IV. EVALUATION AND RESULTS

In this section, the implementation and training of the proposed model and the obtained results were discussed. The experiments were based on deep q learning model.

A. Applied Results of Deep Q Network

Training of proposed deep q agent performed through the source files of python consisting environment and loops of configuration setup. Training completed by giving the input of 1000 episodes and having hyper parameters of learning rate, discount factor, batch size, memory size, frequency of training steps, frequency of target updating, steps of memory warm up, gradient steps, hidden layers, normalization state, epsilon start and end with epsilon decay steps. Rewards, Cost and power imbalance plot with the help of panadas library and series accordingly.

Library of matplotlib is used to visualize the results in a graphical way. Cost, Power Imbalance and reward results demonstrated in Figure 4. Reward value maximize from -1 to 1, cost and power imbalance also minimized with the deep q agent.

B. Performance Evaluation of Deep Q Network

The performance evaluation of the deep q agent model for microgrid energy management reviews training across 1000 episodes delineates a comprehensive understanding of its proficiency in better energy utilization and cost minimization. The training phase illustrates a discernible decrement in sales price of energy from main grid, emblematic of the model's adeptness in learning intricate patterns within the training data of load, price and energy. As the episode progress, the generated action selected by the agent calculated load and consumptions. Consumer use the energy from the solar and utilize the extra usage from the grid to minimize the electricity cost. Energy storage batteries used to save the energy generated from Solar systems and utilized in future whenever the needs required.

Figure 5 represents the charge and discharge of battery, power imbalance, power generation and load details. Notably, the graph on time, mirroring the load minimization from grid, cost minimization in electricity bills and saved and better utilization of power imbalance. This nuanced performance analysis, offers a robust appraisal of the deep Q Agent model's efficacy in energy management systems, contributing substantively to the advancement of reinforcement learning methodologies in the context of microgrid energy management.

V. CONCLUSION

The energy management is performed with Q learning, a model of reinforcement learning based algorithm. A simulation model was created to depict The constantly evolving interactions between the consumer agent and its surroundings for autonomous battery schedule optimization that increases battery utility and solar power, ultimately lowering grid power usage. The solar power generator's uncertainties resulting from the random variations in temperature and irradiation are taken into consideration. The suggested framework enables the astute consumer to investigate and comprehend the stochastic environment, then apply this knowledge to choose the best energy-management strategies to lessen reliance on the grid.

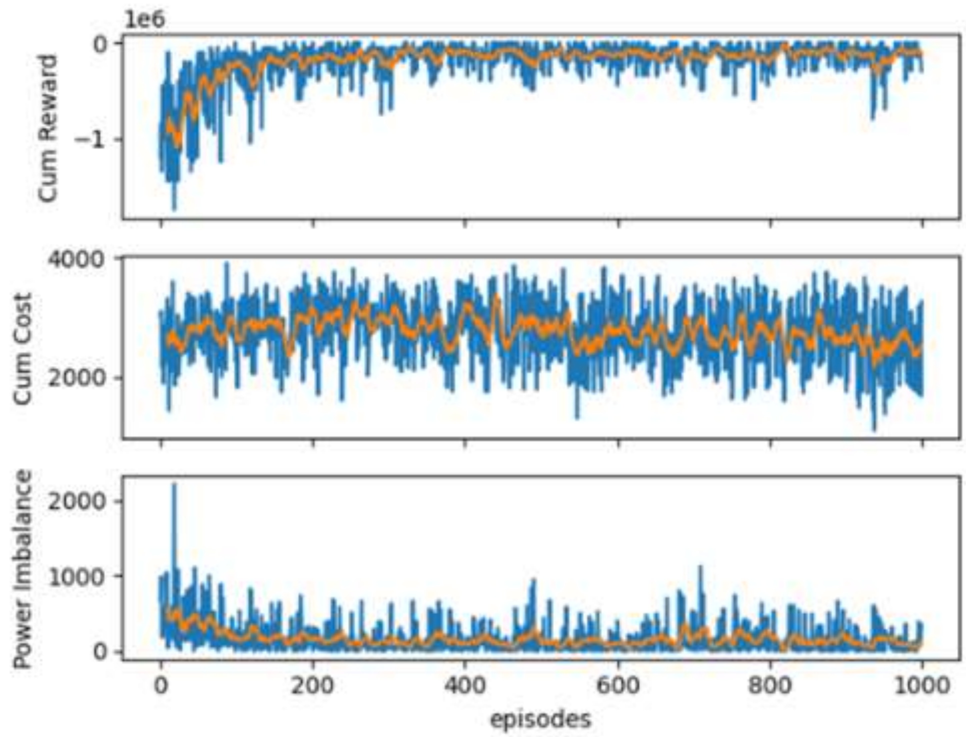

Figure 4. Deep Q-Agent Results

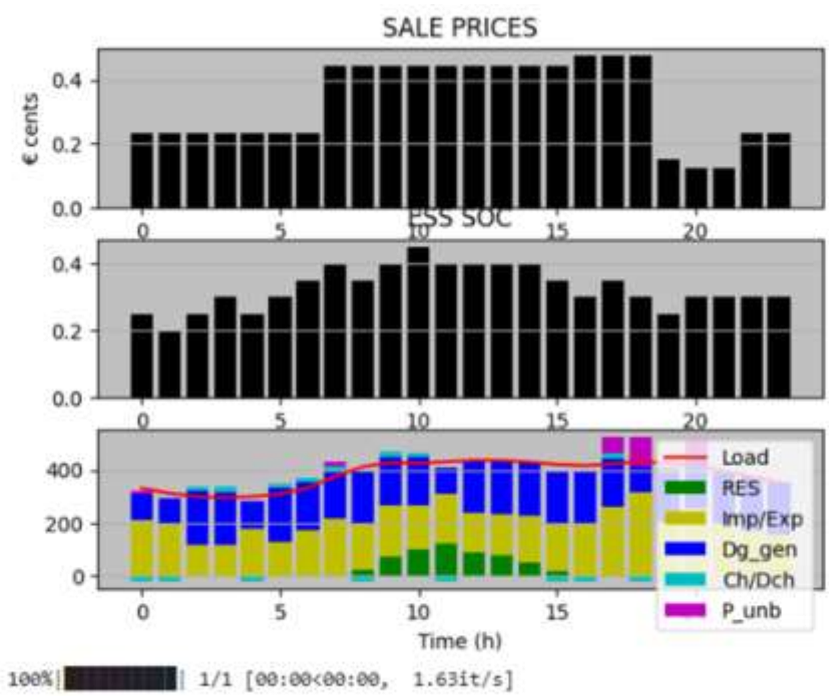

Figure 5. Performance Evaluation of Deep Q-Agent

REFERENCES

[1] P. Kofinas, G. Vouros, and A. I. Dounis, "Energy management in solar microgrid via reinforcement learning," in ACM International Conference Proceeding Series, Association for Computing Machinery, May 2016. doi: 10.1145/2903220.2903257.

[2] H. Li, Z. Wang, L. Li, and H. He, "Online Microgrid Energy Management Based on Safe Deep Reinforcement Learning."

[3] Y. Ji, J. Wang, J. Xu, X. Fang, and H. Zhang, "Real-time energy management of a microgrid using deep reinforcement learning," Energies (Basel), vol. 12, no. 12, 2019, doi: 10.3390/en12122291.

[4] P. Kofinas, A. I. Dounis, and G. A. Vouros, "Fuzzy Q-Learning for multi-agent decentralized energy management in microgrids," Appl Energy, vol. 219, pp. 53–67, Jun. 2018, doi: 10.1016/j.apenergy.2018.03.017.

[5] E. Samadi, A. Badri, and R. Ebrahimpour, "Decentralized multi-agent based energy management of microgrid using reinforcement learning," International Journal of Electrical Power and Energy Systems, vol. 122, Nov. 2020, doi: 10.1016/j.ijepes.2020.106211.

[6] T. A. Nakabi and P. Toivanen, "Deep reinforcement learning for energy management in a microgrid with flexible demand," Sustainable Energy, Grids and Networks, vol. 25, Mar. 2021, doi: 10.1016/j.segan.2020.100413.

[7] X. Fang, Q. Zhao, J. Wang, Y. Han, and Y. Li, "Multi-agent Deep Reinforcement Learning for Distributed Energy Management and Strategy Optimization of Microgrid Market," Sustain Cities Soc, vol. 74, Nov. 2021, doi: 10.1016/j.scs.2021.103163.

[8] X. Jin, F. Lin, and Y. Wang, "Research on Energy Management of Microgrid in Power Supply System Using Deep Reinforcement Learning," in IOP Conference Series: Earth and Environmental Science, IOP Publishing Ltd, Jul. 2021. doi: 10.1088/1755-1315/804/3/032042.

[9] W. Cai, A. B. Kordabad, and S. Gros, "Energy management in residential microgrid using model predictive control-based reinforcement learning and Shapley value," Eng Appl Artif Intell, vol. 119, Mar. 2023, doi: 10.1016/j.engappai.2022.105793.

[10] R. Darshi, S. Shamaghdari, A. Jalali, and H. Arasteh, "Decentralized energy management system for smart microgrids using reinforcement learning," IET Generation, Transmission and Distribution, May 2023, doi: 10.1049/gtd2.12796.

[11] Y. Khawaja, I. Qiqieh, J. Alzubi, O. Alzubi, A. Allahham, and D. Giaouris, "Design of cost-based sizing and energy management framework for standalone microgrid using reinforcement learning," Solar Energy, vol. 251, pp. 249–260, Feb. 2023, doi: 10.1016/j.solener.2023.01.027.

[12] H. H. Goh et al., "An Assessment of Multistage Reward Function Design for Deep Reinforcement Learning-Based Microgrid Energy Management," in IEEE Transactions on Smart Grid, vol. 13, no. 6, pp. 4300-4311, Nov. 2022, doi: 10.1109/TSG.2022.3179567.

[13] H. Zhou and M. Erol-Kantarci, "Correlated Deep Q-learning based Microgrid Energy Management," 2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), Pisa, Italy, 2020, pp. 1-6, doi: 10.1109/CAMAD50429.2020.9209254.

[14] E. Foruzan, L. -K. Soh and S. Asgarpoor, "Reinforcement Learning Approach for Optimal Distributed Energy Management in a Microgrid," in IEEE Transactions on Power Systems, vol. 33, no. 5, pp. 5749-5758, Sept. 2018, doi: 10.1109/TPWRS.2018.2823641./

[15] M. F. Zia, M. Nasir, E. Elbouchikhi, M. Benbouzid, J. C. Vasquez, and J. M. Guerrero, "Energy management system for an islanded renewables-based DC microgrid," in 2nd IEEE International Conference on Smart Power & Internet Energy Systems (SPIES), pp. 543-547, 2020, doi: 10.1109/SPIES48661.2020.9242919 .

[16] M. Nasir, and M. F. Zia, "Global maximum power point tracking algorithm for photovoltaic systems under partial shading conditions," in 16th IEEE International Power Electronics and Motion Control Conference and Exposition, pp. 667-672, 2014, doi: 10.1109/EPEPEMC.2014.6980572.

[17] U. T. khan and M. F. Zia, "Smart city technologies, key components, and its aspects," in IEEE International Conference on Innovative Computing (ICIC), pp. 1-10, 2021, doi: 10.1109/ICIC53490.2021.9692989

[18] H. Zhou, A. Aral, I. Brandić and M. Erol-Kantarci, "Multiagent Bayesian Deep Reinforcement Learning for Microgrid Energy Management Under Communication Failures," in IEEE Internet of Things Journal, vol. 9, no. 14, pp. 11685-11698, 15 July15, 2022, doi: 10.1109/JIOT.2021.3131719.

[19] M. S. Hossain and C. Enyioha, "Multi-Agent Energy Management Strategy for Multi-Microgrids Using Reinforcement Learning," 2023 IEEE Texas Power and Energy Conference (TPEC), College Station, TX, USA, 2023, pp. 1-6, doi: 10.1109/TPEC56611.2023.10078538.

[20] J. Qi, L. Lei, K. Zheng, S. X. Yang and X. Shen, "Optimal Scheduling in IoT-Driven Smart Isolated Microgrids Based on Deep Reinforcement Learning," in IEEE Internet of Things Journal, vol. 10, no. 18, pp. 16284-16299, 15 Sept.15, 2023, doi: 10.1109/JIOT.2023.3267625

AUTHORS

First Author – Usman Inayat, PhD Computer Science, Department of Computer Science, University of Management and Technology Lahore, Pakistan.
Second Author – Sajid Mahmood, Department of Informatics and Systems, University of Management and Technology Lahore, Pakistan.
Third Author – Malik Tahir Hassan, , PhD Computer Science, Department of Computer Science, University of Management and Technology Lahore, Pakistan.

Correspondence Author – Usman Inayat, PhD Computer Science, Department of Computer Science, University of Management and Technology Lahore, Pakistan.