International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8506
***Article Received:** 30 July 2023 **Revised:** 20 September 2023 **Accepted:** 10 October 2023*

BQL-DRS: A Novel Balanced Q-Learning Based Demand Response System for IoT based Smart Grids

[1]N.S. Gowri Ganesh, [2]Dr. Helina Rajini Suresh, [3]M. Bharathi, [4]Jeneetha Jebanazer J, [5]S. Deepa, [6]S Sankar
[1]Department of Artificial intelligence and Data Science, Saveetha Engineering College, Chennai,
gowriganesh@gmail.com
[2]Department of ECE, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Chennai
helinarajini@gmail.com
[3]Department of Electronics and Communication Engineering Jeppiaar Institute of Technology, Sriperumbudur,
Tamil Nadu, India
bharathimjbn@gmail.com
[4]Panimalar Engineering College, Professor, Dept. of ECE, India,
jeneethaseelan@gmail.com
[5]Panimalar Engineering College, Dept. of ECE, India,
dineshdeepas1977@gmail.com
[6]Saveetha School of Engineering, Dept. of CSE, India,
sankars.sse@saveetha.com

Abstract— The modernization of electricity networks and the integration of renewable energy resources in Internet of Things (IoT) based smart grids have led to increased variability in market prices, necessitating effective demand response (DR) strategies. To address this challenge, this paper proposes a novel Balanced Q-Learning based Demand Response System (BQL-DRS) that combines both optimistic and pessimistic targets in the Q-learning algorithm to achieve a balanced decision-making process in IoT based smart grids. It optimizes DR actions by efficiently managing consumer demand in real-time, considering IoT data from grid conditions, energy prices, and consumer preferences. The significance of the BQL-DRS lies in its ability to handle dynamic and uncertain IoT based grid environments, enabling it to make informed and cautious decisions while pursuing energy efficiency and cost-effectiveness. By effectively addressing both pessimistic and optimistic scenarios, the BQL-DRS ensures grid stability, load balancing, and substantial cost savings compared to representative models.

Keywords- Intelligent Mobile Sink, Routing Protocol, IMSARP, Cluster Head, CH, WSN, Wireless Sensor Network

1. INTRODUCTION

In recent years, the demand for electricity in home and commercial settings has increased dramatically as a direct result of the growing use of electric cars and household equipment [1]. The expansion of power producing facilities, the development of energy storage technologies, and the use of smart grid technology for effective electricity management are all strategies that may be used to satisfy this need [2]. The transition of traditional electricity grids into more intelligent and resourceful smart grids is made possible in large part by the Internet of Things (IoT) technologies. IoT-enabled smart grids minimize energy use, improve grid resilience, and make it easier to integrate renewable energy resources in a seamless manner by harnessing real-time data and sophisticated analytics.

Smart grids are grids that combine sophisticated technologies and serve as a vital solution to improve grid performance, communicate with customers, and support Demand Response (DR) programs [3]. DR programs are designed to motivate consumers to modify their power use during peak load hours, which contributes to system stability. Smart grids are a crucial answer to these problems. This strategy may be used in homes, companies, and even whole industries. In today's sophisticated smart grids, DR systems are an absolute need for effectively controlling the variable output of renewable energy sources and maintaining grid stability. They allow for the smooth incorporation of renewable energy sources, the effective management of peak loads, the reduction of costs and emissions, the enhancement of grid flexibility, and the empowerment of consumers for energy efficiency and a sustainable future.

The following is a description of a few different kinds of demand response systems. Programs known as "Time-of-Use" (ToU) pricing [4] set different prices for energy at different times of the day. The prices are often higher during times of strong demand and lower when demand is reduced, for as during off-peak hours. Consumers are urged to move their energy-intensive activities, such as operating appliances or charging electric cars, to times when rates are lower in order to take advantage of these cheaper rates and reduce the burden that is placed on the grid during peak hours.

Critical Peak Pricing (CPP) [5] programs involve higher electricity rates during critical peak periods when electricity demand is at its highest. These periods are typically limited to a few hours on specific days when the grid is under significant stress. Participants in CPP programs receive advanced notice of these critical peak events and are incentivized to reduce their energy consumption during those times to avoid higher costs. In Peak Time Rebate (PTR) [6] programs, consumers receive financial incentives or rebates for reducing electricity usage during peak demand periods. Participants receive a rebate or credit on their electricity bill based on the amount of energy they save during peak hours.

However, DR programs face limitations in consumer participation, technology requirements, timing constraints, and rebound effects as evident from an investigation of Lu et al [7]. Incentives may not be compelling enough, and equity concerns can arise. Data privacy and program complexity also pose challenges. To optimize DR's effectiveness, policymakers must enhance consumer engagement through education and attractive incentives. Investing in advanced technologies should be balanced with cost considerations. Addressing equity issues is vital to ensure broad participation. Robust data privacy measures must be implemented to build trust. Timely coordination and streamlined program designs can bolster DR's impact in managing peak demand, enhancing energy efficiency, and promoting grid stability for a sustainable energy future.

Deep learning based DR systems offer promising solutions to resolve the limitations of traditional DR programs [8]. By harnessing the power of advanced algorithms and IoT technology, these systems can enhance consumer engagement, optimize incentive structures, address timing constraints, mitigate rebound effects, ensure equity considerations, strengthen data privacy, and streamline program designs. Implementing deep learning based DR systems can significantly improve the effectiveness and efficiency of DR initiatives, ultimately leading to better grid stability, enhanced energy efficiency, and a more sustainable energy future.

Of late, Reinforcement Learning (RL) based DR programs play a pivotal role in reshaping energy consumption patterns, promoting grid stability, and advancing the transition towards a more sustainable and reliable energy future [9-10]. Their dynamic adaptability, cost-effectiveness, and positive impact on the grid and the environment make them a significant component of modern energy management strategies.

Balanced Q-learning (BQL) [11] is a RL algorithm that strikes a compromise between optimistic and pessimistic targets to achieve a balanced approach. Traditional Q-learning's optimistic targets assume the best-case scenario, leading to overestimation bias, while pessimistic targets can result in overly conservative policies. BQL combines both approaches as a convex combination, enabling exploration while considering risks and uncertainties. This algorithm improves performance and stability in domains like robotics, finance, and energy management, offering effective and reliable learning outcomes.

By leveraging the BQL approach, the DR system can achieve enhanced stability, improved risk management, efficient resource utilization, flexibility in action spaces, real-time responsiveness, balanced trade-offs in objectives, and increased consumer engagement. The balanced learning process can minimize overestimation bias and conservativeness, ensuring robust performance in dynamic smart grid environments. The DR system can optimally allocate resources, adapt to changing conditions, and navigate conflicting objectives. Consumers are encouraged to participate due to the harmonious balance between energy-saving measures and comfort, making BQL a promising algorithm for effective and reliable DRSs in smart grids.

This research proposes the BQL-DRS, the first of its kind to optimize demand response actions effectively and reliably, contributing to grid stability, energy efficiency, and sustainability in the smart grid ecosystem. The contributions of this paper are as below.

1. The BQL optimizes demand response actions while considering risks and uncertainties, enhancing DRS stability in dynamic grid environments.
2. The BQL improves risk management by making cautious decisions, ensuring grid stability and mitigates potential instability or safety issues, even during unforeseen events or disturbances.

The rest of this paper is organized as follows. Section II presents a literature review on DR strategies and RL in the context of smart grids. Section III outlines the proposed BQL-DRS architecture with its components. Section IV presents the experimental results on evaluation of the BQL-DRS using objective metrics, interpretation of the results and presents the research findings. Finally, section VII concludes the paper and suggests future research directions.

2 RELATED WORKS

In [12], a comprehensive overview of employing rl algorithms for demand response (dr) applications in the smart grid is presented. The authors explore different rl techniques and their modeling methodologies, emphasizing their potential advantages and associated challenges. A most recent review in this context also advocates rl based dr for resource optimization in iot based smart grids [13]. The research on dr includes a precise classification of economic signals for managing electricity demand. A study by yan et al [14] proposes a price based dr utilizing the tou method and appliance scheduling, optimizing consumer restrictions and price changes for improved device decision-making. Due to their outstanding performance, machine learning-based dr techniques have recently attracted attention [15]. For load balancing in smart grids, game theory-based dr methods are helpful. In [16], a two-stage dr strategy employs a stochastic one-leader multiple-follower stackelberg gaming model for decision-making and a noisy inverse optimization technique for load prediction. A generative adversarial network (gan), which updates the model regularly for decision-making in the face of missing data, is used to enhance dynamic electricity price forecasts [17]. In smart grids, these methods provide effective load control and dynamic power pricing options. In recent years, rl has been widely used to address the dr problem, enabling agents to learn and adapt in unknown environments. Rl's effectiveness in understanding consumer preferences in dynamic settings makes it a state-of-the-art method for dr programs. Comprehensive analyses were conducted on articles focusing on dr programs, including those based on prices, incentives, consumer satisfaction, consumer classification, and practical case applications. In [18], an rl architecture optimizes heating, ventilation, and air conditioning (hvac) control in a building for energy savings and thermal comfort with demand response. The study achieves up to 22% weekly energy reduction compared to traditional control methods. In this line, a study in [19] proposes a method using rl to manage a multipurpose energy storage (mpes) system for demand response programs. Industrial consumers can gain added profits through market participation while optimizing electrical load management. The benefits of tou rates are explored, showing that consumers can maximize cost savings by shifting consumption to lower-priced time slots. In [20], a neural network is trained to develop discrete-time control strategies using rl. This study focuses on optimizing thermostat configurations, considering factors like thermal comfort, energy consumption, and environment. A novel objective function truncation method is introduced to improve algorithm robustness. Furthermore, the proposed rl algorithm is utilized to learn the thermostat settings during dr periods based on electricity prices and tou approach. In [21], a dr scheduling model for residential community which uses an energy management system aggregator is proposed. The model dynamically controls power demand and distributed energy resources to match renewable power generation with consumption, while reducing operational costs through electricity trading in day-ahead and real-time markets. The problem is formulated as a mixed-integer linear programming task, and a two-level model predictive control (mpc) integrating rl is used to address uncertainties in system operation. The results demonstrate that operating houses in aggregate mode yields greater benefits for the community. Extensive review of several literature shows that rl based dr systems in iot networks face limitations in dealing with high dimensionality, lengthy training times, data inefficiency, safety concerns, and adapting to non-stationary conditions. The bql approach emerges as a promising solution, offering faster convergence, enhanced data efficiency, safety, stability, and adaptability. Despite these advantages, there are currently no existing works on dr based on bql, creating a crucial research opportunity to explore its untapped potential. By integrating bql with dr strategies, this study aims to advance efficient and robust dr implementations in iot-based smart grid systems.

III PROPOSED DQS-BRL

This section presents the formal definition of the proposed DR system for smart grids and the BQL-DRS architecture.

A. Problem Definition

Let SG be a smart grid system with n sensors $(Sensor_1, Sensor_2, \ldots, Sensor_n)$ collecting data related to energy consumption, environmental conditions, and renewable energy generation. The system also includes m actuators $(Actuator_1, Actuator_2, \ldots, Actuator_m)$ responsible for implementing DR actions in the smart grid. The objective of the BQL-DR system is to optimize DR actions in SG to efficiently manage power demand and distributed energy resources while ensuring grid stability and sustainability.

Let S be the state space representing the possible combinations of sensor readings, A be the action space representing the set of demand response actions that can be taken, and $R: S \times A \rightarrow \mathbb{R}$ be the reward function that provides a scalar reward for each action taken in a particular state. The goal is to find an optimal policy $\pi: S \rightarrow A$ that maps states to actions, maximizing the expected cumulative reward over time as in (1), where $\mathbb{E}$ is the expected reward obtained when the system is in state s_t and the BQL agent takes action a_t according to the policy π. It is the expected cumulative reward over time, considering the probabilistic nature of state transitions and action selections. The expectation is taken with respect to the probability distributions of states and actions given by the sensor readings from S and the policy π, respectively.

$$\pi^* = \underset{\pi}{\operatorname{argmax}} \sum_{t=0}^{T} \; \mathbb{E}_{s_t \sim \mathcal{S}, a_t \sim \pi(s_t)}[R(s_t, a_t)] \qquad (1)$$

The BQL-DRS updates the Q-values for state-action pairs according to equation (2), where $Q(s, a)$ is the Q-value for state s and action a, α is the learning rate, controlling the weight of the new information in the Q-value update, and λ is the balance parameter, determining the trade-off between optimistic and pessimistic targets. Here, $\underset{a}{max}\{Q(s\ , a\ )\})$ represents the maximum Q-value among all possible actions a that can be taken in the current state s. It indicates the highest expected reward the agent can achieve by selecting the best action a. The agent uses this value to update the Q-value for the current state-action pair (s', a') in the Q-learning algorithm, helping it to learn and make better decisions over time.

$$Q(s', a') \leftarrow (1 - \alpha) \cdot Q(s, a) + \alpha \cdot ((1 - \lambda) \cdot R(s, a) + \lambda \cdot \underset{a'}{max}\{Q(s\ , a\ )\}) \qquad (2)$$

The state diagram illustrating the above process is shown in Figure 1.

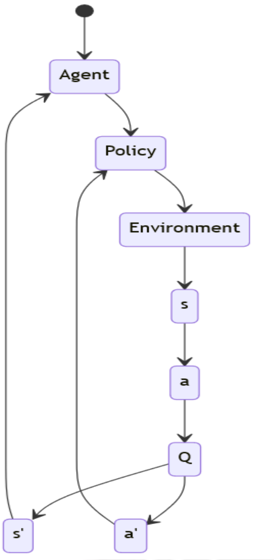

Fig.1. BQL-DRS State Diagram

B. BQL-DRS Architecture

The architecture of the BQL-DRS comprises the following interconnected components designed to optimize DR actions in IoT-based smart grid environments as in Figure 2.

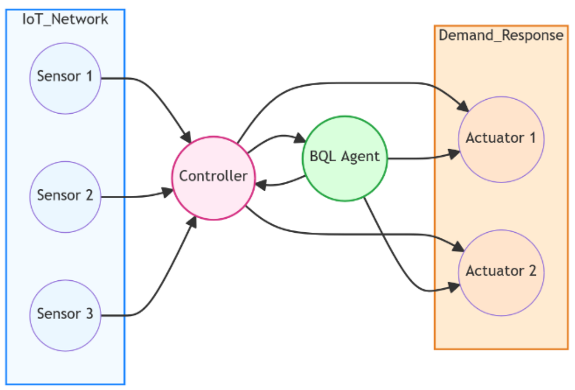

Fig.2. BQL-DRS Architecture

1. **IoT Network:** The IoT network represents the IoT sensors responsible for collecting data from the smart grid about various aspects such as energy consumption, environmental conditions, and renewable energy generation.
2. **BQL Agent:** The central component of the architecture is the BQL agent. It receives input from the IoT sensors and employs the BQL algorithm to make optimal decisions regarding DR actions. The BQL agent's main goal is to dynamically manage power demand and distributed energy resources, ensuring a balance between renewable power generation and community consumption.
3. **Controller:** The controller acts as the decision-making entity that coordinates interactions between the BQL agent, IoT sensors, and demand response actuators. It receives the output from the BQL agent and communicates with other components to execute the optimal demand response actions based on the agent's decisions.
4. **Demand Response Actuators:** These actuators are responsible for implementing the DR actions in the smart grid. The controller interacts with these actuators to execute the optimized strategies determined by the BQL agent. The BQL-DR system leverages the BQL agent as the core decision-making component, which takes input from IoT sensors, optimizes demand response actions, and communicates with the controller to execute these actions through the demand response actuators. This architecture ensures an efficient and balanced approach to demand response in IoT-based smart grid systems, promoting energy efficiency, grid stability, and sustainability.

IV EXPERIMENTAL RESULTS AND ANALYSIS

This section presents the evaluation of the BQL-DRS in the smart grid context. It showcases the metrics, empirical findings, and comparative analyses, providing insights into the system's strengths and limitations for future enhancements in DR strategies within IoT-based smart grids.

A. Experimental Setup

The simulated smart grid environment includes a photovoltaic system (100 kW) and a wind turbine (50 kW) as virtual power generation sources, along with two energy storage units (500 kWh each) and controllable loads (200 kW to 500 kW). Electric vehicles with charging demands ranging from 50 kW to 150 kW were also considered. The BQL-DRS was configured with α=0.2, γ=0.9, ϵ=0.1, and λ=0.5 for BQL. Comprehensive data collection and monitoring systems were used to track Q-values, rewards, state transitions, and action selections. The experiments encompassed varying renewable energy generation, real-time price fluctuations, and energy consumption patterns. The performance of the model was also evaluated for different values of α and λ.

B. Performance Evaluation

The BQL-DRS is evaluated with the metrics **Demand Variation (DV), Load Factor (ALF), Peak Load Reduction (PLR) and the Cost Savings (CS) evaluated as in equations (3)-(6).**

DV is used to measure the effectiveness of the proposed pricing schemes and DR actions. It calculates the total change in electrical energy consumption, representing the relationship between the original consumption and the consumption after implementing a DR scheme as in (3).

$$DV = \frac{E_{\text{orig}} - E_{\text{new}}}{E_{\text{orig}}} \qquad (3)$$

The LF metric indicates the ratio of the average load (AV_L) of consumers to their maximum load Max_L, representing high peak consumer demand and the effectiveness of the pricing scheme in displacing electricity demand as in (4).

$$LF = \frac{AV_L}{Max_L} \qquad (4)$$

The PLR quantifies the reduction in peak load achieved through DR strategies as in (5) where Max_L is the original peak load and Max_{DR} is the peak load after deploying the DRS in a smart grid.

$$PLR = \frac{Max_L\text{-}\ Max_{DR}}{Max_L} \qquad (6)$$

The CS metric represents the ratio of the cost savings achieved through DR actions to the total cost of electricity consumption without any DR actions in the smart grid system as in (6).

$$CS = \frac{Cost_{DR}\text{-}\ Cost_{T}\text{-}}{Cost_T} \qquad (6)$$

A low DV value indicates that the demand response actions are effectively reducing the total energy consumption and enhancing energy efficiency. On the other hand, a high LF value signifies that the pricing schemes and demand response strategies are successfully improving the load factor, leading to a more balanced and consistent utilization of energy resources.

A low PLR value indicates that the DR actions are effectively reducing the peak load during high-demand periods, contributing to better grid stability and more efficient use of energy resources. On the other hand, a high CS value signifies that the pricing schemes and DR strategies are yielding significant economic benefits by reducing the total operational costs of electricity consumption in the smart grid.

By optimizing DV, LF, PLR, and CS values, the smart grid can achieve multiple advantages, such as minimizing energy wastage, enhancing grid reliability, reducing peak demand charges, and promoting cost-effectiveness. Evaluating and monitoring these metrics enable researchers and grid operators to assess the success of DR initiatives and pricing strategies, facilitating data-driven decisions and continuous improvements in the smart grid's performance and sustainability.

The performance of the model is evaluated under different values of α and λ as in Table 1 and 2 respectively.

Table 1. Performance Metrics under different values of α

LR (α)	**DV % ↓**	**LF %↑**	**PLR % ↓**	**CS %↑**
0.1	17	80	15	24
0.01	13	87	12	33
0.001	9	94	10	35
0.0001	6	96	8	40

It is seen that as the learning rate decreases, the BQL-DRS exhibits improved performance in terms of reducing energy consumption (DV), enhancing energy efficiency (LF), and achieving higher cost savings (CS). However, there is a slight decrease in the percentage reduction of peak load during high-demand periods (PLR) as the learning rate decreases. These findings suggest that selecting an optimal learning rate is crucial, as a lower learning rate may lead to slower convergence but can result in more accurate and optimal demand response actions in the long run. The results underscore the significance of tuning the learning rate to meet the specific requirements and objectives of the smart grid system.

Table 2. Performance Metrics under different values of **λ**

Balancing Factor (λ)	**DV % ↓**	**LF % ↑**	**PLR % ↓**	**CS % ↑**
0.1	14	85	16	32
0.3	12	87	15	35
0.5	10	88	14	37
0.7	9	90	13	39
0.9	7	92	12	42

Table 2 shows that as the balancing factor λ increases, the DV decreases, indicating a more significant reduction in total energy consumption through DR actions. The LF metric increases, demonstrating improved efficiency in utilizing energy resources. The PLR also decreases, suggesting a smaller reduction in peak load during high-demand periods. Conversely, the CS metric increases, reflecting higher cost savings achieved through demand response actions.

Further, the proposed BQL-DRS is compared with the state-of-the-art models in Table 3. It is observed that the proposed model surpasses the representative models with a good performance gain with respect to all the performance metrics. It is further noted that BQL-DRS is closely succeeded by the RL model proposed in [21], while the other RL based approaches exhibit heavy performance degradations.

This exceptional performance of the BQL-DRS is attributed to the ability of the model to handle pessimistic and optimistic targets for a balanced decision-making process. This resilience in dynamic and uncertain smart grid environments enables the system to handle various scenarios, exploring risks while making cautious decisions. Handling both pessimistic and optimistic scenarios enhances the BQL-DRS's robustness and reliability in managing DR actions. This adaptability ensures energy efficiency, cost-effectiveness, and better outcomes compared to other representative models, leading to improved energy savings, load balancing, and cost reductions.

Table 3. Comparative Analysis

Model	**DV % ↓**	**LF % ↑**	**PLR % ↓**	**CS % ↑**
RL for HVAC [18] Azuatalam et al (2020)	14	78	16	29
RL for MPES [19] Oh et al (2023)	13	81	15	33
RL for Discrete Time Control [20] Li et al (2022)	12	87	14	35
RL for Residential Community [21] Ojand & Dagdougui (2021)	**10**	**91**	**13**	**37**
BQL-DRS (Proposed)	**6**	**96**	**8**	**40**

Further, the convergence of the model is studied for different learning rates to understand its impact on the performance of the BQL-DRS. The learning rate plays a crucial role in determining how fast the BQL agent updates its Q-values based on the observed rewards. A higher learning rate may result in faster convergence, but it can also lead to overshooting and instability. On the other hand, a lower learning rate might require more iterations for convergence, but it can offer better stability and accuracy in decision-making.

By analyzing the convergence behavior for various learning rates, researchers can identify the optimal learning rate that strikes a balance between convergence speed and performance. This investigation helps in fine-tuning the BQL-DRS system to achieve the best results in different smart grid scenarios. Additionally, it provides insights into how the BQL-DRS performs under different learning rate settings and helps in understanding the trade-offs involved in choosing the most suitable learning rate for the specific application. Figure 3 depicts the convergence behaviour of the model with respect to five different values of α. The observed convergence at α = 0.1 underscores the significance of choosing the right learning rate to strike a balance between rapid convergence and system stability. This finding has substantial implications for the successful implementation of the BQL-DRS in diverse real-world scenarios within the smart grid domain. By utilizing the optimal learning rate, the BQL-DRS can effectively manage demand response actions, optimize energy consumption, and enhance grid stability in the face of dynamic and unpredictable conditions.

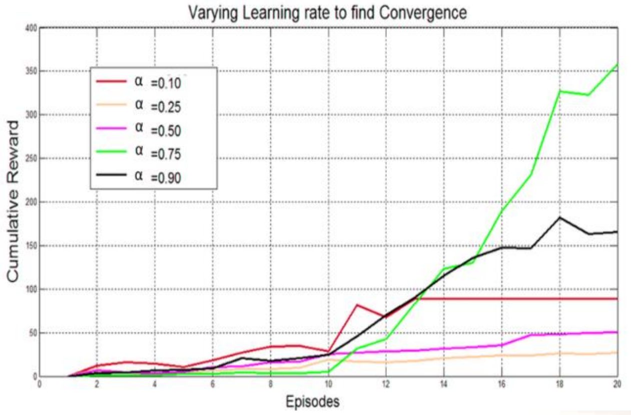

Fig.3. BQL-DRS Convergence with Learning rate

V CONCLUSION

The novel BQL-DRS proposed in this paper demonstrates exceptional performance in optimizing dr actions in smart grids. By effectively handling both pessimistic and optimistic scenarios, it achieves robust and balanced decision-making. The bql-drs outperforms representative models in energy efficiency, grid stability, cost savings, and successful dr actions exhibiting 96% lf. The analysis of convergence behavior with different learning rates emphasizes the importance of selecting the optimal learning rate for system efficiency. The bql-drs offers a promising solution for enhancing demand response strategies and contributes to a more sustainable and resilient energy future in smart grid applications.

REFERENCES

[1.] Li, W. T., Yuen, C., Hassan, N. U., Tushar, W., Wen, C. K., Wood, K. L., ... & Liu, X. (2015). Demand response management for residential smart grid: From theory to practice. IEEE Access, 3, 2431-2440.

[2.] Pawar, P. (2019). Design and development of advanced smart energy management system integrated with IoT framework in smart grid environment. Journal of Energy Storage, 25, 100846.

[3.] Hassanniakheibari, M., Hosseini, S. H., & Soleymani, S. (2020). Demand response programs maximum participation aiming to reduce negative effects on distribution networks. International Transactions on Electrical Energy Systems, 30(8), e12444.

[4.] Sediqi, M. M., Nakadomari, A., Mikhaylov, A., Krishnan, N., Lotfy, M. E., Yona, A., & Senjyu, T. (2022). Impact of time-of-use demand response program Energies, 15(1), 296.

[5.] Yusuf, J., Hasan, A. J., & Ula, S. (2021, February). Impacts analysis & field implementation of plug-in electric vehicles participation in demand response and critical peak pricing for commercial buildings. In 2021 IEEE Texas Power and Energy Conference (TPEC) (pp. 1-6). IEEE.

[6.] Wang, X., & Tang, W. (2019, October). Designing multistep peak time rebate programs for curtailment service providers. In 2019 North American Power Symposium (NAPS) (pp. 1-6). IEEE.

[7.] Mansoor, C.M.M., Vishnupriya, G., Anand, A., ...Kumaran, G., Samuthira Pandi, V, “A Novel Framework on QoS in IoT Applications for Improvising Adaptability and Distributiveness”, International Conference on Computer Communication and Informatics, ICCCI 2023, 2023.

[8.] Samuthira Pandi, V., Singh, M., Grover, A., Malhotra, J., Singh, S, “Performance analysis of 400 Gbit/s hybrid space division multiplexing-polarization division multiplexing-coherent detection-orthogonal frequency division multiplexing-based free-space optics transmission system”, International Journal of Communication Systems, 2022, 35(16), e5310.

[9.] Wen, L., Zhou, K., Li, J., & Wang, S. (2020). Modified deep learning and reinforcement learning for an incentive-based demand response model. Energy, 205, 118019.

[10.] R, G. et al. (2022). “Optimization of Solar Hybrid Power Generation Using Conductance-Fuzzy Dual-Mode Control Method” International Journal of Photoenergy, Volume 2022, Article ID 7756261, 10 Pages, 2022 https://doi.org/10.1155/2022/7756261.

[11.] Karimpanal, T. G., Le, H., Abdolshah, M., Rana, S., Gupta, S., Tran, T., & Venkatesh, S. (2021). Balanced q-learning: Combining the influence of optimistic and pessimistic targets. arXiv preprint arXiv:2111.02787.

[12.] Vázquez-Canteli, J. R., & Nagy, Z. (2019). Reinforcement learning for demand response: A review of algorithms and modeling techniques. Applied energy, 235, 1072-1089.

[13.] R, G. et al. (2022). “A Novel Approach in Hybrid Energy Storage System for Maximizing Solar PV Energy Penetration in Microgrid”, International Journal of Photoenergy, Volume 2022, Article ID 3559837, 7 pages, https://doi.org/10.1155/2022/3559837.

[14.] Yan, X., Ozturk, Y., Hu, Z., & Song, Y. (2018). A review on price-driven residential demand response. Renewable and Sustainable Energy Reviews, 96, 411-419.

[15.] Pallonetto, F., De Rosa, M., Milano, F., & Finn, D. P. (2019). Demand response algorithms for smart-grid ready residential buildings using machine learning models. Applied energy, 239, 1265-1282.

[16.] Lu, T., Wang, Z., Wang, J., Ai, Q., & Wang, C. (2018). A data-driven Stackelberg market strategy for demand response-enabled distribution systems. IEEE Transactions on Smart Grid, 10(3), 2345-2357.

[17.] Ali, S. S., & Choi, B. J. (2020). State-of-the-art artificial intelligence techniques for distributed smart grids: A review. Electronics, 9(6), 1030.

[18.] Senthilkumar, S., Samuthira Pandi, V., Sripriya, T., Pragadish, N (2023), “Design of recustomize finite impulse response filter using truncation based scalable rounding approximate multiplier and error reduced carry prediction approximate adder for image processing application”, Concurrency and Computation: Practice and Experience, 35(8), e7629.

[19.] Oh, S., Kong, J., Yang, Y., Jung, J., & Lee, C. H. (2023). A multi-use framework of energy storage systems using reinforcement learning for both price-based and incentive-based demand response programs. International Journal of Electrical Power & Energy Systems, 144, 108519.

[20.] Li, Z., Sun, Z., Meng, Q., Wang, Y., & Li, Y. (2022). Reinforcement learning of room temperature set-point of thermal storage air-conditioning system with demand response. Energy and Buildings, 259, 111903.