Deep Reinforcement Learning-Graph Neural Networks-Dynamic Clustering triplet for Adaptive Multi Energy Microgrid optimization

Tessel Kaal[1], Azarakhsh Rafiee[2]

[1]Department of Architectural Engineering and Technology, Delft University of Technology, The Netherlands- tessel-kaal@outlook.com
[2]Department of Architectural Engineering and Technology, Delft University of Technology, The Netherlands- a.rafiee@tudelft.nl

Keywords: Deep Reinforcement Learning, Graph Neural Network, Dynamic Clustering, Microgrid, Renewable Energy Sources.

Abstract

Centralized energy systems are often limited by their dependence on large, centralized power plants and extensive transmission networks, making them vulnerable to single points of failure and less resilient to disruptions. Microgrids offer resilience, enhanced energy efficiency, and improved integration of renewable resources compared to centralized energy systems, enabling localized energy management and reduced reliance on fossil fuels. Deep Reinforcement Learning (DRL) has shown its potential for microgrid energy optimization by enabling intelligent, adaptive control over energy resources and energy exchange. By learning from interactions with the environment, the DRL agent dynamically adjusts the power outputs of distributed energy resources, manages energy storage systems, and balances energy exchange between microgrid elements and with the main grid, aiming to minimize costs and ensure reliable power availability. However, incorporating spatial relationships into DRL action space significantly increases computational demands. In line with this, we have introduced a novel method that integrates DRL, Graph Neural Network (GNN) and dynamic clustering to optimize microgrid operations. GNNs are specialized deep learning models that adapt to graphs of varying sizes and structures. This adaptability enables GNN-equipped DRL agents to effectively learn from and apply knowledge to a wide range of network topologies. The agent can be used for subsets, or sub-microgrids, taking into account the scalability and efficiency of the optimization process, enabling distance and routing optimization without an aggregated model. This approach addresses the computational challenges associated with large action spaces and varying topologies in microgrid management.

1. Introduction

1.1 Need for Resilient Energy Systems

With the European Union (EU) raising its renewable target, set within the Renewable Energy Directive, to 42.5 percent by 2030 and moving away from gas, the energy landscape in the Netherlands is undergoing significant changes toward sustainability and renewable energy integration. Part of the broader ‘Fit for 55’ package, the plan aims to significantly reduce the EU’s dependence on fossil fuels and enhance its overall climate objectives (European Council, 2022).

The Netherlands has set ambitious goals, including revisions to critical energy legislation, such as the Energy Performance of Buildings Directive and the Energy Efficiency Directive, to adopt a more vigorous approach towards renewable energy. The necessity of this shift is highlighted by the benefits of renewable energy being low-cost and domestically produced, which reduces dependency on external suppliers (European Commission, 2023). The transition requires the electrification of energy systems, but the Dutch electricity grid is currently ill-prepared for a scale-up of this magnitude.

Recognizing the limitations of centralized power grids in meeting the nuanced energy demands of urban environments, this study proposes a new model to manage microgrids that perform independent and sustainable. The inherent intermittency and variability of renewable sources present considerable challenges to the existing grid infrastructure, necessitating the development of more flexible and efficient transmission systems to ensure consistent and reliable energy delivery (CBS Statline, 2023). Microgrids offer numerous benefits over conventional power grids, including enhanced reliability, reduced transmission losses, environmental benefits, and increased flexibility (Shahzad et al., 2023), (Ali et al., 2022). They allow for the integration of diverse and DERs, fostering a more resilient and sustainable energy infrastructure.

1.2 Reinforcement Learning for Multi Energy Microgrid Management

Recent advancements in microgrid optimization emphasize the superiority of Reinforcement Learning (RL) and Deep Reinforcement Learning (DRL) over traditional heuristic algorithms. (Tajjour and Singh Chandel, 2023) note that RL’s key advantage lies in its robust decision-making and stable convergence capabilities, even when faced with numerous variables. What sets RL apart is its consistent ability to achieve optimal solutions, an attribute well-supported by a variety of studies (Li et al., 2023), (Subramanya et al., 2022), (Klemm and Wiese, 2022).

These advanced learning methods represent a significant move towards model-free or data-driven management approaches (Chen et al., 2022). (Nakabi and Toivanen, 2021) highlight that learning directly from a microgrid’s operational data allows these systems to fine-tune control strategies without the need for an explicit system model. This advantage is twofold: it enhances Energy Management Systems (EMS) scalability and reduces maintenance costs, marking a shift towards more autonomous and intelligent energy management systems.

The adaptability of RL and DRL is especially crucial in the dynamic environment of microgrid operations (Chen et al., 2022). Employing RL and DRL strategies provides a practical

https://doi.org/10.5194/isprs-annals-X-G-2025-427-2025

approach. These methods learn and adapt to the system's behavior iteratively, allowing them to solve the control problem efficiently without any prior knowledge of the system (Israr and Yang, 2021).

1.3 DRL and Limitations

The DRL agent autonomously determines the best actions to take within the energy management system, learning to maximize a reward function that reflects the optimal way of operating the microgrid. Through continuous learning and adaptation, the DRL model improves its decision-making process, effectively managing energy resources to ensure stability and reliability. This capability is of value in the dynamic urban environments where demand patterns and resource availability can be highly variable. However, dense urban districts, with their high resource demands, also pose complexities in mapping system operations, revealing the limitations of DRL in efficiently handling high-density scenarios. Introducing distance considerations or dynamic microgrids into the RL paradigm complicates the optimization process further. The action space, representing all possible actions the RL agent can take, increases exponentially when considering shortest paths or changing topologies. This complexity arises from the combinatorial nature of calculating the shortest distances between various nodes in the system, where the action space can potentially grow factorially with each additional node (Almasan et al., 2022). A challenge arises when combining this concept with a Deep Q-Network (DQN)-trained RL agent, as they are not designed to handle changing topologies. (Almasan et al., 2022) highlight that current advancements in DRL for networking are limited to network configurations encountered during training and struggle with unfamiliar structures. This limitation stems from the reliance on conventional neural network architectures, such as densely connected layers, which do not effectively capture the complexities of graph-based data.

1.4 DRL-GNN for Internet Networks

The recent studies on the integration of DRL and GNN (Fathinezhad et al., 2023), and (Ji et al., 2024) in Internet of Vehicles technology demonstrate the potency of this integration. (Almasan et al., 2022) propose combining DRL with GNNs for routing optimization in internet networks. This combination allows for route allocation and training an agent for unknown topologies (Almasan et al., 2022). The architecture they provide can learn and generalize over unknown topologies due to optimization over an additional dimension. However, this approach requires significant computational power, and the input network of nodes and action space is limited. In their research, the agent is trained in scenarios with a single topology of 14 nodes (Nsfnet) and then analyzed in larger topologies up to 100 nodes. Performance dropped by 15% in these larger topologies. See (Almasan et al., 2022) for the algorithm for the DRL-GNN agent and the action space representation.

1.5 The gap and DRL-GNN with Clustering as New Proposed Method

While DRL demonstrates suitability for flexibility and adaptation to the dynamic nature of microgrids, the quickly growing action space poses a challenge in managing the complexity and scale of potential actions. To address these limitations, a combination of DRL with a GNN and Clustering is proposed. This hybrid approach aims to manage the expanding action space and improve the efficiency and scalability of the optimization process. The main objective is to minimize the burden on the grid and enhance grid flexibility by increasing the energy balance while integrating Renewable DERs and meeting variable energy demand for diverse topologies. The developed model may offer significant value by facilitating exploration of various microgrid configurations, through the identification of optimal setups and of optimal management.

By creating sub-microgrids, or clusters, with a net output per cluster per timestep the required computational can be regulated. As the agent has become indifferent for the topology, clusters can be determined dynamically in line with efficient energy balancing strategies that take into account factors like distance and complementary consumption profiles.

The DRL-GNN method, proposed by (Almasan et al., 2022), combined with topology-size constraints by means of clustering, facilitates a new way to optimize a microgrid with DRL. It allows for the creation of subsystems that reduce and bound the action space, thereby minimizing its influence on the training process.

2. Underlying Components

2.1 DRL and Markov Decision Process

Markov Decision Processes (MDPs) form the mathematical foundation of reinforcement learning, enabling sequential decision-making under uncertainty. MDPs consider both immediate outcomes of current decisions and the impact of future actions, represented by a dynamic state and value function in the Bellman equation (Sutton & Barto, 2018), as shown in:

$$Q^*(s,a) = \mathrm{E}[r + \gamma \max_{\acute{a}} Q^*(\acute{s},\acute{a})] \qquad (1)$$

Where $Q^*(s,a)$ represents the optimal state-action value function, which gives the maximum expected return starting from state s, taking action a, and thereafter following the optimal policy. E denotes the expected value, indicating that the returns are averaged over all possibilities, weighted by their probability of occurrence. r is the immediate reward received after taking action a in state s.

γ is the discount factor, which represents the difference in importance between future rewards and immediate rewards, and makes it possible to take in uncertainty. The factor is a value between 0 and 1, that determines the weight given to future rewards in the decision-making process. A higher discount factor places more importance on future rewards, reflecting greater confidence in the predictability of future outcomes, while a lower discount factor emphasizes immediate rewards, indicating higher uncertainty about future events.

Selecting the appropriate discount factor for a microgrid depends on the specific goals and characteristics of the system. A high discount factor can promote long-term sustainability and strategic planning, but it should be balanced with the need to address immediate operational challenges.

$\max_{\acute{a}}$ denotes the maximum value over all possible actions a' from the new state s' and $Q^*(\acute{s},\acute{a})$ represents the optimal state-action value function for the next state s' and any action a' and s' is the new state after action a is taken in state s.

2.2 Graph Neural Networks

A GNN is a type of neural network specifically designed to process data structured as graphs, capturing both the relationships (edges) and individual entities (nodes) in the data. Through iterative message passing or neighborhood aggregation, each node in a GNN updates its feature representation by aggregating information from its neighbors, allowing it to learn contextually enriched embeddings. This aggregation process enables GNNs to learn hierarchical, complex patterns that depend on the graph structure. A message passing neural networks layer l can be expressed as:

$$h_u^{(l+1)} = \phi(h_u^{(l)}, \bigoplus_{v \epsilon N_u} \psi(h_u^{(l)}, h_v^{(l)}, e_{uv})) \qquad (2)$$

where $h_u^{(l)}$ and $h_v^{(l)}$ are the feature vectors of nodes u u and v v at layer l and e_{uv} is the edge feature between nodes u and v and $\bigoplus$ is the aggregate function which combines messages from neighbours, ϕ and ψ are differentiable functions. ψ is a function that may transform or weight neighbour features before aggregation and ϕ is an activation function. This is illustrated in Figure 1.

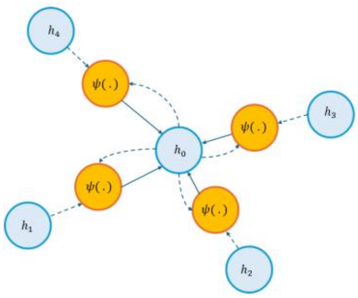

Figure 1. Graph Neural Networks

2.3 Clusters

Clustering methods have been observed to offer significant advantages for urban microgrid scenarios (Bandeiras et al., 2020; Philipo et al., 2021). According to (Yu et al. 2022), microgrid clustering can enhance the utilization and local consumption of renewable energy sources (RESs). Additionally, clustering has the potential to reduce maintenance costs and extend the overall lifespan of the network (Yu et al., 2022).

In this approach, clustering methods enable the creation of subsets according to desired sizes and net output. Depending on the specific goals and characteristics of the system, factors for clustering can be selected. Examples of factors to cluster by include distance and user profile per building. By clustering in this manner, the same effect is observed as mentioned by (Yu et al. 2022), where local usage is preferred. The preference is for every sub-microgrid to become as independent as possible and for the distance of energy exchange to be limited.

An important consideration in clustering is whether dynamic clustering per time step significantly increases performance or if a longer-term approach, for example annual clustering, would suffice to determine optimal clusters while still achieving distance minimization. Hierarchical clustering, DBSCAN, and K-means clustering are all techniques that facilitate the creation of subsets. By incorporating attributes such as net output, geographical distance, consumption profiles, resource availability, and other relevant factors, these clustering methods can create more effective and meaningful clusters that align with the system's goals. Research has shown that hierarchical clustering algorithms are better suited for microgrid planning because of their flexibility with any dataset, ability to explore the entire solution space to ensure globally optimal networks, and their relative computational efficiency (Cheong et al., 2017). This method builds a hierarchy of clusters through either an agglomerative (bottom-up) approach or a divisive (top-down) approach. The algorithm's inherent simplicity and computational efficiency are significant benefits, especially when dealing with extensive datasets (Cheong et al., 2017). Its versatility is a key advantage, particularly in scenarios involving sparse and heterogeneous data typical of remote or isolated locations. Hierarchical clustering can form dense, isolated clusters optimal for microgrid configurations in these areas (Cheong et al., 2017).

2.4 Microgrid Components

Key system elements, critical for modelling and optimizing an energy grid system, include energy loads, representing total energy demand, which can be balanced across buildings to reduce reliance on the main grid. The utility grid acts as a backup, enabling energy flow during supply-demand imbalances and incurring higher costs during peak hours. Distributed Energy Resources (DERs), such as solar and geothermal energy, and components like electrolyzers, fuel cells, and batteries provide flexibility and sustainability. Electric vehicles (EVs) and hydrogen-based systems like Combined Heat and Power (CHP) units also support the system, enhancing resilience and self-sufficiency. These elements collectively shape an efficient and environmentally-friendly energy network within the urban landscape. Figure 2 presents a schematic illustration of microgrid components and their potential inter-connection at a specific time.

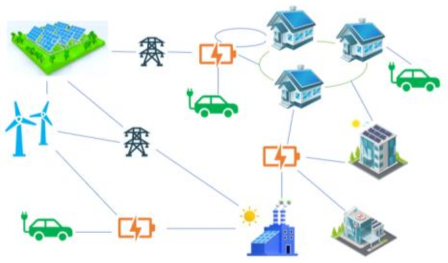

Figure 2. Schematic illustration of microgrid components and their inter-relations at a specific time.

3. Methodology and Discussions

This section includes the description and discussions of DRL-GNN-dynamic clustering integration and its adaptation to microgrid adaptive optimization.

3.1 DRL Design for Microgrid Management

The system to be implemented includes the following essential components:

Environment: This encompasses the RL environment, representing the grid setup with all DERs, loads, generators, and connections to the main grid.

States: These define the current status of the grid, providing detailed information on the output levels of each DER and generator, load demands, the level of power exchange with the main grid, battery storage levels, and external environmental factors such as weather conditions that impact renewable energy output.

Actions: The agent’s actions include adjusting the outputs of controllable DERs, managing Energy Storage Systems (ESSs), and determining the level of power exchange between microgrid elements and with the main grid.

Reward Function: This function represents the system's goals, incorporating penalties for drawing power from the main grid and rewards for utilizing renewable energy sources (RES). It is

https://doi.org/10.5194/isprs-annals-X-G-2025-427-2025

designed to promote self-sufficiency and reduce operational costs. Additionally, the reward function should prioritize local load balancing to minimize transmission losses.

3.1.1 Microgrid Action Space

To represent the actions in a DRL framework for a microgrid, each action can be mapped to a feature vector where each element corresponds to a control variable for specific components in the microgrid. In this study, we have categorized the microgrid controllable elements into four categories of:

1. Controllable DERs Output Adjustments

$$a_{DER} = \left[a_{Solar}, a_{CHP}, \dots, a_{DER_k}\right]^T \quad (3)$$

Each a_{DER_k} represents the output of a specific DER (e.g., solar, CHP, etc.).

2. ESS Management

$$a_{ESS} = \left[a_{battery\_charge}, a_{battery\_discharge}\right]^T \quad (4)$$

These elements represent charging and discharging levels for the battery or other ESS

3. Power Exchange with the Main Grid

$$a_{Grid} = a_{Grid\_power\_exchange} \quad (5)$$

a single variable for the level of power exchanged with the main grid.

4. Energy sharing between elements

Let S be an $n \times n$ matrix where each element $s_{i,j}$ represents the amount of energy shared from element i to element j (for instance building i with building j). In this setup:

$s_{i,j} \geq 0$ represents the amount of energy sent from element i to element j.

$s_{i,i} = 0$, since a building cannot share energy with itself.

To integrate this matrix S into the action feature vector, we will flatten it into a vector. For n elements, there are $n \times (n-1)$ possible directed energy flows (since $s_{i,i} = 0$ for all i). The flattened vector $a_{Sharing}$ for the energy sharing component would then be:

$$a_{Sharing} = \left[s_{1,2}, s_{1,3}, \dots, s_{1,n}, \ s_{2,1}, s_{2,3}, \dots, s_{n,n-1}\right]^T \quad (6)$$

The overall action feature is as follows:

$$a = \begin{bmatrix} a_{Solar}, a_{CHP}, \dots, a_{DER_k}, a_{battery_{charge}}, a_{battery_{discharge}}, \\ a_{Grid_{power\ exchange}}, s_{1,2}, s_{1,3}, \dots, s_{1,n}, \ s_{2,1}, s_{2,3}, \dots, s_{n,n-1} \end{bmatrix}^T$$

(7)

3.1.2 Reward Shaping Indicators and Objectives

The reward function is essential for training the agent to achieve desired behaviors, such as minimizing grid connections, boosting microgrid independence, and cutting operational costs. Represented in euros, the reward is negative and includes costs for DERs, imports, and exports (with positive costs for exports).

A custom reward shaping mechanism is integrated into the microgrid environment to discourage grid reliance by dynamically imposing penalties for grid usage at each step. This approach encourages the agent to favor local energy production, promoting self-sufficiency and cost-effectiveness.

$$R = -\left(\sum_i c_{DER,i} . E_{DER,i} + c_{import} . E_{import} - c_{export} . E_{export} + p_{grid} . \frac{E_{import}}{E_{total}}\right)$$

(8)

The components of this reward function are:

1. Cost of DERs C_{DER}:

$$C_{DER} = \sum_i c_{DER,i} . E_{DER,i} \quad (9)$$

Where i: Index for each DER (e.g., solar, wind, geothermal). $c_{DER,i}$: Marginal cost of operating the i-th DER (in €/kWh). $E_{DER,i}$: Energy generated by the i-th DER (in kWh).

2. Cost of Importing Energy from the Main Grid, C_{import}:

$$C_{import} = c_{import} . E_{import} \quad (10)$$

Where c_{import}: Cost per unit of imported energy (in €/kWh), which may vary depending on peak/off-peak hours. E_{import}: Total energy imported from the main grid (in kWh).

3. Revenue from Exporting Excess Energy to the Grid, C_{export}:

$$C_{export} = c_{export} . E_{export} \quad (11)$$

Where c_{export}: Price per unit of energy exported to the main grid (in €/kWh). E_{export}: Total energy exported back to the main grid (in kWh).

4. Penalty for Grid Usage, P_{grid}:

$$P_{grid} = p_{grid} . \frac{E_{import}}{E_{total}} \quad (12)$$

Where p_{grid} : Penalty rate for grid reliance (unitless or in €/kWh). E_{total}: Total energy demand or consumption in the microgrid (in kWh). This term discourages grid usage by imposing a penalty proportional to the fraction of total energy demand met by imports from the grid.

3.2 DRL-GNN Integration for Dynamic Microgrid Optimization

The combination of DRL with a GNN addresses the computational challenges associated with large action spaces and varying topologies in microgrid management. The proposed method creates manageable clusters within the microgrid and one agent that can traverse these sub-microgrids. This method allows for the inclusion of distance minimization in cluster formation without expanding the action space excessively. Thereby, the findings by (Almasan et al., 2022) indicate that the DRL-GNN agent surpasses existing leading solutions in these diverse topologies.

The integration of DRL and GNNs represents a cutting-edge approach for solving complex, relational decision-making tasks. DRL uses deep neural networks to enable agents to learn optimal

https://doi.org/10.5194/isprs-annals-X-G-2025-427-2025 | © Author(s) 2025. CC BY 4.0 License.

policies through high-dimensional, dynamic environments, while GNNs are specifically designed to model data structured as graphs, capturing dependencies and interactions among entities. This integration allows DRL agents to process and learn from structured, graph-based data, improving their ability to generalize and adapt to complex, interconnected settings.

The embeddings from the GNN serve as inputs to the DRL agent's policy and/or value networks. In Q-learning (e.g., DQN), the GNN outputs are used to estimate Q-values, which represent expected future rewards for taking specific actions. This enables the DRL agent to learn policies that account for the underlying graph structure, effectively improving decision-making. Based on the GNN output, the agent takes actions that may be either node-level (acting on specific nodes) or graph-level (acting on the entire graph structure). After the agent takes an action, it receives a reward based on the effectiveness of its decision. This reward feedback is then used to update both the DRL and GNN components through backpropagation. The loss functions used in DRL propagate gradients not only through the policy/value network but also through the GNN layers. This adjusts the GNN parameters to better capture the structure needed for optimal decision-making. Over many episodes or iterations, the agent continually refines its policy. The GNN learns to produce embeddings that capture critical information about the graph structure, while the DRL agent learns an optimal policy based on these embeddings. The combined model eventually converges on a policy that leverages both the graph's structural information (through GNN) and reward-based adaptation (through DRL). Figure 3 presents the DRL-GNN integration concept. Here the interaction and optimization over the two neural networks is depicted within the reinforcement loop.

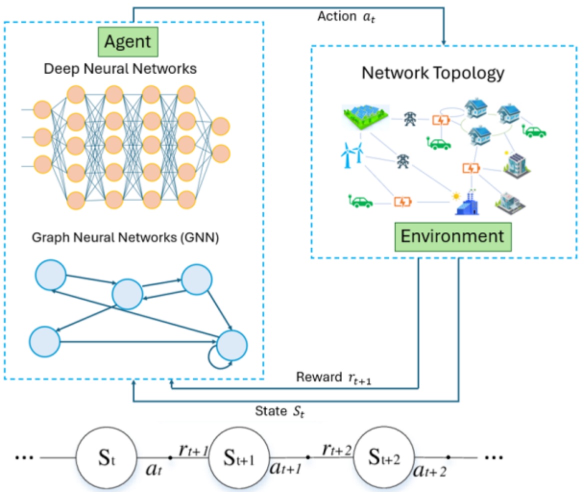

Figure 3. Interaction agent and model in DRL-GNN integrated model

The adaptation of DRL states and actions in GNN is presented in Figure 4.

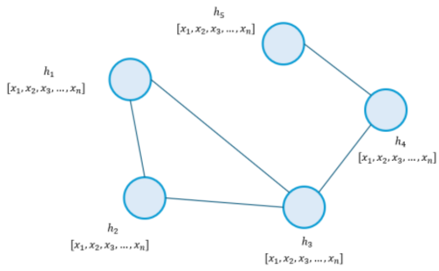

Figure 4. Representation of states and actions in GNN

Through DRL-GNN integration the microgrid topology becomes indifferent for the optimized model as it has optimized over the two dimensions. The GNN creates a graph representation where the connections between the nodes are represented as entities in the graph. In this representation, the hidden states of the connections are initialized based on the input link-level features and the action to be evaluated. The state and action space depicted in Figure 4. The state features x1-xn is defined for each DER component based on its operational status and characteristics. For instance, in the case of a solar panel, each node can encompass Current Power Output (kW), Capacity (kW), Operational Status (Online/Offline), Voltage (V) Efficiency (%) and Environmental Impact (gCO2/kWh).

3.3 Dynamic Clustering

To optimize load sharing within a microgrid, clustering techniques can be utilized to group buildings or parcels based on their energy consumption profiles (Rostami et al., 2020). This approach is frequently used in the formation of multiple microgrids but can also be applied within a single microgrid to dynamically create optimal energy balance at each time step. Dynamic clustering in microgrid optimization refers to the strategy of dynamically grouping or reconfiguring microgrid resources, such as distributed energy sources, storage units, and loads, to improve operational efficiency, stability, and adaptability. This concept plays a crucial role in optimizing the performance of microgrids, particularly under varying energy demands, environmental conditions, and grid interactions.
In the context of GNNs, dynamic clustering can be used to reduce the action space by simplifying the structure of the graph, thereby focusing the model's attention on the most relevant nodes and edges. This is particularly valuable in large or complex graphs, where the action space can quickly become overwhelming due to the exponential number of possible connections and actions at each node. Dynamic clustering groups nodes and edges into clusters based on relevance, connectivity, or certain features of interest. By replacing multiple nodes or edges with a single cluster node or edge, the GNN operates on a simplified version of the graph, effectively reducing the number of possible actions it has to consider.

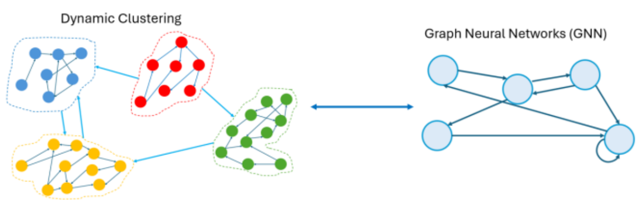

Figure 5. Schematic representation of dynamic clustering integration with GNN

Dynamic clustering can be used to create a hierarchical structure within the graph. Instead of having to select actions across all nodes and edges simultaneously, the GNN can first choose between clusters and then refine its focus within a selected cluster. This hierarchical reduction of the action space makes the decision-making process more efficient.
This clustering allows the GNN to make coarse-grained decisions at the cluster level first, before moving into finer details, which reduces the need to evaluate every individual node at the outset.

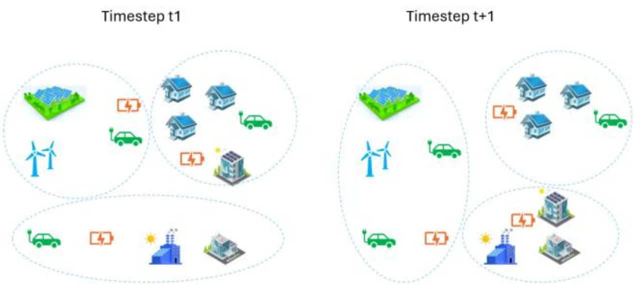

Figure 6. Microgrid dynamic clustering for time step *t* and timestep *t+1*

This clustering can be used to organize the microgrids strategically to optimize energy distribution. The resulting clusters can act as isolated units that manage their net load independently, which is particularly crucial for balancing the energy within the main microgrid system. Each cluster produces a specific net load that corresponds to the unbalanced demand within that cluster, connecting back to the main grid in various configurations. These configurations can range from fully interconnected systems where every cluster is connected to every other, to arrangements such as a circular or linear chain, depending on the specific energy and infrastructure needs (Ediriweera & Widanagama Arachchige, 2022). This clustering approach not only optimizes the physical distances between energy production and consumption points but also minimizes transmission losses, thereby enhancing the overall efficiency of the microgrid system.

3.4 Scalability and Computational Requirements

The dynamic clustering approach described in subsection 3.3 provides advantages in terms of scalability and computational efficiency. By managing the size of the sub-microgrids, the computational resources required to optimize the topology remain manageable. The clusters act as isolated units managing their net load independently, facilitating the addition of new clusters without significant impact on the overall computational load. The system can dynamically reconfigure clusters based on changing energy demands, ensuring that the microgrid remains optimized as it scales. This scalability allows for the efficient operation and optimization of higher number of sub-microgrids.

4. Conclusions

Combining the DRL-GNN agent with cluster techniques, can reduce the required computation and complexity, and increase the possibility of optimizing over multiple dimensions and topologies. With this novel approach it is possible to take in spatial relationships, and therefore to balance on a lower level.
The applied clustering can already impose strict spatial constraints necessary for efficient microgrid generation, while also limiting the size of the sub-microgrids. The microgrid is divided into several sub-microgrids that exchange energy, aiming to minimize distances and balance loads within the network. Constraints on the maximum number of loads per sub-microgrid ensure system efficiency and stability.

While the DRL-GNN-clustering triplet holds significant potential for scaling microgrid optimization, its performance should be measured in a follow-up study. The proposed framework should be applied to a case study to test the method and theory, and to compare the new results against traditional optimization methods, further validating its effectiveness.
The performance analysis should measure the computational load and convergence stability when scaling up the microgrid in terms of the number of elements and geospatial extent. Moreover, the analysis should include various microgrid and urban configurations to monitor the effectiveness of the DRL-GNN-clustering triplet in different settings. It is also important to consider the value of the clustering frequency and to use training data that includes both seasonal and weekly variations to test the approach's performance in various demand-generation settings. Exploring the framework in a multi-objective context would further extend the proposed method, enhancing its robustness.

5. Acknowledgements

5.1 Declaration of Generative AI and AI-assisted Technologies in the Writing Process

Statement: During the preparation of this work the authors used ChatGPT in order to reformat. After using this tool/service, the authors reviewed and edited the content as needed and takes full responsibility for the content of the publication.

5.2 Competing Interests

Competing interests: The authors declare that they have no competing interests. The research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Ali, Z., Saleem, K., Putrus, G., Marzband, M., Dudley, S., 2022. Control of PV and EV connected smart grid. *Microgrid Cyberphysical Systems*, Elsevier, p. 115–157.

Almasan, P., Sua´rez-Varela, J., Rusek, K., Barlet-Ros, P., Cabellos-Aparicio, A., 2022. Deep reinforcement learning meets graph neural networks: Exploring a routing optimization. use case. *Computer Communications*, 196, 184-194.

Bandeiras, F., Pinheiro, E., Gomes, M., Coelho, P., Fernandes, J., 2020. Review of the cooperation and operation of microgrid clusters. *Renewable and Sustainable Energy Reviews*, 133, 110311.

CBS Statline, 2023. Renewable energy; consumption by energy source, technology and application.

Chen, X., Qu, G., Tang, Y., Low, S., Li, N., 2022. Reinforcement Learning for Selective Key Applications in Power Systems: Recent Advances and Future Challenges. *IEEE Transactions on Smart Grid*, 13(4), 2935–2958.

Cheong, D. L. K., Fernando, T., Iu, H., Reynolds, M., Fletcher, J., 2017. Review of clustering algorithms for microgrid formation. 2017 *IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia)*, IEEE, Auckland, 1–6.

Ediriweera, S., Widanagama Arachchige, L., 2022. Design and protection of microgrid clusters: A comprehensive review. *AIMS Energy*, 10, 375–411.

European Commission, 2023. Renewable energy directive.

European Council, 2022. REPowerEU: Council agrees on accelerated permitting rules for renewables.

European Council, 2023. Fit for 55.

Fathinezhad, F., Adibi, P., Shoushtarian, B., Chanussot, J., 2023. Graph Neural Networks and Reinforcement Learning: A Survey. *Deep Learning and Reinforcement Learning. IntechOpen.*

Israr, A., Yang, Q., 2021. Resilient and sustain- able microgeneration power supply for 5G mobile networks. *Renewable Energy Microgeneration Systems*. Academic Press, p. 213-228.

Ji, M., Wu, Q., Fan, P., Cheng, N., Chen, W., Wang, J., Letaief, K. B., 2024. Graph Neural Networks and Deep Reinforcement Learning Based Resource Allocation for V2X Communications. *IEEE Internet of Things Journal*.

Klemm, C., Wiese, F., 2022. Indicators for the optimiza- tion of sustainable urban energy systems based on energy system modeling. *Energy, Sustainability and Society*, 12(1), 3.

Li, S., Oshnoei, A., Blaabjerg, F., Anvari-Moghaddam, A., 2023. Hierarchical Control for Microgrids: A Survey on Clas- sical and Machine Learning-Based Methods. *Sustainability*, 15(11), 8952.

Nakabi, T. A., Toivanen, P., 2021. Deep reinforcement learning for energy management in a microgrid with flexible demand. Sustainable Energy, *Grids and Networks*, 25, 100413.

Philipo, G. H., Chande Jande, Y. A., & Kivevele, T. 2021. Clustering and Fuzzy Logic-Based Demand-Side Management for Solar Microgrid Operation: Case Study of Ngurudoto Microgrid, Arusha, Tanzania. *Advances in Fuzzy Systems*, 2021(1), 6614129.

Rostami, Z., Ravadanegh, S. N., Kalantari, N. T., Guerrero, J. M., Vasquez, J. C., 2020. Dynamic Modeling of Multiple Microgrid Clusters Using Regional Demand Response Programs. *Energies*, 13(16), 4050.

Shahzad, S., Abbasi, M., Ali, H., Iqbal, M., Munir, R., Kilic¸, H., 2023. Possibilities, Challenges, and Future Opportunities of Microgrids: A Review. *Sustainability,* 15, 6366.

Subramanya, R., Sierla, S., Vyatkin, V., 2022. Exploiting Battery Storages With Reinforcement Learning: A Review for Energy Professionals. *IEEE Access*, 10, 54484-54506.

Sutton, R. S., Barto, A. G., 2018. Reinforcement learning: an introduction. Adaptive computation and machine learning series, second edition edn, *The MIT Press*, Cambridge, Massachusetts.

Tajjour, S., Singh Chandel, S., 2023. A comprehensive re- view on sustainable energy management systems for optimal operation of future-generation of solar microgrids. *Sustainable Energy Technologies and Assessments*, 58, 103377.

Yang, Q., Yang, T., Li, W., (eds), 2020. Renewable Energy Microgeneration Systems, Academic Press, 213–228.

Yu, H., Niu, S., Shao, Z., Jian, L. 2021. A scalable and reconfigurable hybrid AC/DC microgrid clustering architecture with decentralized control for coordinated operation. *International Journal Of Electrical Power & Energy Systems*, 135, 107476.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-G-2025-427-2025 | © Author(s) 2025. CC BY 4.0 License.