Post-Disaster Microgrid Formation for Enhanced Distribution System Resilience

Mukesh Gautam[1], Michael Abdelmalak[1], Mohammed Ben-Idris[1], Eliza Hotchkiss[2]
[1]Department of Electrical & Biomedical Engineering, University of Nevada-Reno, Reno, NV 89557, USA
[2]National Renewable Energy Laboratory, Golden, CO 80401, USA
Emails: {mukesh.gautam, mabdelmalak}@nevada.unr.edu, mbenidris@unr.edu, eliza.hotchkiss@nrel.gov

***Abstract*—This paper proposes a deep reinforcement learning (DRL) based approach for post-disaster critical load restoration in active distribution systems to form microgrids through network reconfiguration to minimize critical load curtailments. Distribution networks are represented as graph networks, and optimal network configurations with microgrids are obtained by searching for the optimal spanning forest. The constraints to the research question being explored are the radial topology and power balance. Unlike existing analytical and population-based approaches, which necessitate the repetition of entire analyses and computation for each outage scenario to find the optimal spanning forest, the proposed approach, once properly trained, can quickly determine the optimal, or near-optimal, spanning forest even when outage scenarios change. When multiple lines fail in the system, the proposed approach forms microgrids with distributed energy resources in active distribution systems to reduce critical load curtailment. The proposed DRL-based model learns the action-value function using the REINFORCE algorithm, which is a model-free reinforcement learning technique based on stochastic policy gradients. A case study was conducted on a 33-node distribution test system, demonstrating the effectiveness of the proposed approach for post-disaster critical load restoration.**

***Index Terms*—Active distribution systems, microgrid formation, network reconfiguration, reinforcement learning, resilience.**

I. Introduction

Resilience enhancement strategies have gained significant interest during the last decade due to the high reliance on electricity access and availability, especially during and after disruptive events. The frequency and intensity of extreme weather- and cyber-related events have increased dramatically resulting in significant infrastructure damage and socio-economic losses [1], [2]. To reduce the negative impacts of disruptive events, fast and efficient restoration strategies provide a potential pathway for enhanced resilience in power systems operations [3], [4]. Restorative strategies aim to return the system to normal or semi-normal operations and performance through recovering failed components or executing redundant solutions [5]. Post-disaster load restoration can be achieved through microgrid formation [6], network reconfiguration [7], and allocation of distributed energy resources (DERs) [8]. However, a few challenges still exist to determine proper post-event actions in a fast-paced manner, including computational burdens, time constraints, and system size. Therefore, providing a load restoration strategy to improve the resilience of distribution systems considering the system operational constraints has become important.

Several studies have been conducted to enhance the load restoration mechanisms in distribution systems against severe outage events. Resilience-based microgrid formation frameworks have been proposed to enhance the restoration of critical loads in both radial and meshed networks [9]. An expert system-based approach has been developed to control tie-switches of distribution feeders for improved load restoration behavior [10]. A microgrid formation strategy has been used to retain critical loads after an extreme event, leveraging mixed-integer linear programming optimization methods [11]. An optimal network partitioning algorithm has been developed in [12] to improve system resilience through spectral clustering under $N-k$ (i.e., $k>1$) contingencies. A heuristic-based fuzzy multi-objective approach has also been developed to minimize the number of tie-switch operations of radial systems for reduced load curtailment [13]. An approximate dynamic programming methodology has been formulated in [14] to determine optimal switching decisions for network reconfiguration of distribution systems during hurricanes. Despite the significant contributions of these methods, the degree of approximation and linearization adopted in the problem formulation play a vital role in the obtained results. The studied approaches have focused mainly on analytical and heuristics-based techniques giving less importance to artificial intelligence-based approaches. Additionally, the capability of these methods to model large-scale systems decreases exponentially, with the system size imposing additional computational limitations.

The emerging advancement of reinforcement learning (RL) based approaches has encouraged their adoption in control-based resilience enhancement strategies. The integration of deep neural network structure to RL-based methods has shown a wide range of promising pathways to improve resilience of electric power systems [15]. A post-disaster RL-based optimization framework has been developed to enhance resilience of islanded microgrids through energy storage management and load shedding strategies [16]. Recent studies have leveraged model-free reinforcement learning algorithms to learn the value of an action in a particular application (Q-learning approaches) [17] and actor-critic approaches [18] to improve load restoration schemes of distribution systems. RL-based approaches have the capability to learn from experiences.

Moreover, once a deep reinforcement learning (DRL) model is fully trained, it can be easily integrated into the online decision making process for enhanced resilience. Due to the numerous DRL control-based approaches, further investigation is still required to validate the efficiency and effectiveness of these methods in resilience-based load restoration problems.

This paper proposes a DRL-based framework for post-disaster critical load restoration (PDCLR) in active distribution systems. The proposed approach is based on the training of a neural network that makes the best decision based on previous experience. Given a specific decision, the sequential decision process gives a reward value as a function of system outcome. The objective of the proposed agent is to minimize critical load curtailment. To ensure realistic representation of distribution system operations, system constraints, such as radiality and power balance limitations are considered. In the training phase of the proposed framework, an action probability distribution is generated using forward propagation of a deep neural network (DNN). Actions are randomly sampled from the action probability distribution, so that when actions are passed through the training environment, the DRL agent gets rewarded (or penalized) based on its performance. The discounted returns and policy gradients are calculated based on the stored values of states, actions, and rewards. The weights of DNN are then updated using the policy gradients. The trained DRL agent is then used to find the best network configuration. The proposed framework is validated through a case study on a prototypical 33-node system. The results show that the proposed framework can effectively find a network configuration, thereby minimizing critical load curtailment.

The remainder of the paper is organized as follows: the mathematical formulation of the PDCLR problem is explained in Section II; the proposed framework and solution approach are described in Section III; a case study on the 33-node system is used to validate the proposed work in Section IV; and finally, Section V, includes concluding remarks.

II. Mathematical Modeling

This section presents the graph theory-based modeling of the distribution network and the mathematical formulation of the PDCLR problem. Also, states, actions, and reward function are described in the context of PDCLR.

A. Graph Theoretic Modeling of Distribution Network

Distribution systems are equipped with sectionalizing switches (normally closed) and tie-switches (normally open). When all switches of a distribution network are closed, a meshed network is formed, and the meshed network can be represented by an undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ is a set of nodes (or vertices) and $\mathcal{E}$ is a set of edges (or branches). This graph is represented in spanning trees and forests, as described below. The post disaster critical load restoration (PDCLR) problem proposed in this paper changes the status of sectionalizing and tie switches in such a way that radiality is always maintained and microgrids are formed with DERs.

1) Spanning Tree: A spanning tree is defined as a subset of the undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with a minimal number of edges linking all vertices (or nodes). Each edge of the undirected graph $\mathcal{G}$ is assigned a specific weight based on the problem being studied. For example, the edge weights vary depending on the problem. The sum total of all edge weights of a spanning tree is minimized when establishing the minimum spanning tree. Fig. 1(a) shows a spanning tree of a hypothetical 12-node system. The spanning tree shown in the figure consists of all 12 system nodes and 11 closed edges or branches.

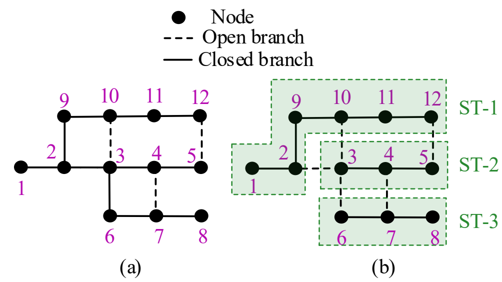

Fig. 1. (a) A spanning tree; and (b) a spanning forest of a hypothetical 12-node system

2) Spanning Forest: In graph theory, a forest is a disconnected union of trees. A spanning forest is a forest that covers all vertices of the undirected graph $\mathcal{G}$ and consists of a set of disconnected spanning trees [19]. Fig. 1(b) shows the spanning forest formed as a result of disconnection of two additional branches (2–3 and 3–6) in the spanning tree presented in Fig. 1(a). The spanning forest shown in Fig. 1(b) consists of three spanning trees (ST-1, ST-2, and ST-3).

B. Problem Formulation

This subsection presents the objective function and the constraints of the PDCLR problem under consideration.

1) Objective Function: As a result of an extreme event, some or all parts of the system may lose power supply. Under such circumstances, sectionalizing and tie switches can be used to reconfigure the network for enhanced resilience. Therefore, the objective of the PDCLR problem is to minimize the critical load curtailment of the system. Mathematically, the objective function of the critical load curtailment minimization is expressed as follows.

$$\min \sum_{i=1}^{N} \omega_i \Delta P_i, \tag{1}$$

where ΔP_i is the load curtailment at node i; ω_i is the critical load factor at node i; and N is the total number of nodes in the system.

2) Constraints: The proposed PDCLR problem is subjected to various constraints, including nodal power balance and radiality. The power balance constraints at each node of the system can be expressed as follows:

$$\sum_{j \in \Omega_g(j)} P_{g,j} + \sum_{l \in \Omega_L(j)} P_{l,j} = P_{D,j} \tag{2}$$

where $\Omega_g(j)$ is the set of sources (including DER) connected to node j; $\Omega_L(j)$ is the set of lines connected to node j; $P_{g,j}$ is the power injected from the source j; $P_{D,j}$ is the load at node j; and $P_{l,j}$ is the line power flow from node l to node j.

A distribution system must always meet the radiality requirement. Therefore, each potential configuration should be radial (i.e., the radiality constraint should be met for each spanning tree of the network). Each spanning tree of the network is represented by a sub-graph $\mathcal{G}_s = (\mathcal{N}_s, \mathcal{E}_s)$, where $\mathcal{N}_s$ is a set of nodes (or vertices) and $\mathcal{E}_s$ is a set of edges (or branches) in the sub-graph. For the sub-graph, a node-branch incidence matrix should be constructed. If $n = |\mathcal{N}_s|$ denotes the number of nodes and $e = |\mathcal{E}_s|$ denotes the number of edges of a particular spanning tree, then the node-branch incidence matrix $A \in \mathbb{R}^{n\times e}$ is the matrix with element a_{ij} calculated based on (3). If the node-branch incidence matrix A is fully ranked, then the radiality constraint is satisfied.

$$a_{ij} = \begin{cases} +1 & \text{if branch } j \text{ starts at node } i \\ -1 & \text{if branch } j \text{ ends at node } i \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

C. States, Actions, and Reward Function

The choice of states, actions, and reward function plays a critical role for the proper training of a reinforcement learning (RL) agent. For the PDCLR problem under consideration, a vector of on/off status of network branches (or edges) is taken as the state. The action is a vector of on/off status of network switches and the cost function at time step t is expressed as follows:

$$C_t = LC_t^{RL} - LC_t^{min}, \tag{4}$$

where LC_t^{RL} is the critical load curtailment at time step t as a result of the action taken by the RL agent; and LC_t^{min} is the minimum critical load curtailment at time step t.

The total reward at time step t is computed as follows.

$$R_t = \begin{cases} \frac{100}{1+C_t} & \text{if all constraints are satisfied} \\ -\rho & \text{if any constraint is violated} \end{cases} \tag{5}$$

where ρ denotes the penalty factor.

III. Proposed Framework

This work leverages recently advanced reinforcement learning techniques to minimize the critical load curtailment following an extreme event. This section provides a brief overview of deep reinforcement learning, the REINFORCE algorithm, and the training attributes of the proposed approach.

A. Reinforcement Learning

Reinforcement learning (RL) is a branch of machine learning that studies how artificial intelligent algorithms should operate in a given scenario to maximize the cumulative reward. RL, along with supervised and unsupervised learning, is one of the three main machine learning techniques. The RL problem is frequently formally described as a Markov decision process (MDP), in which an agent remains in state S_t at timestep t, takes action A_t, obtains a reward R_t, and moves to the next state S_{t+1} based on environment dynamics guided by the transition probability $Pr(S_{t+1}|S_t, A_t)$. In order to maximize its cumulative reward, the agent learns a policy $\pi(A_t|S_t)$, which is a function that maps states to actions. The state and action spaces of the MDP are high-dimensional for the microgrid formation problems under consideration, and basic RL algorithms cannot solve them. Deep reinforcement learning (DRL) approaches are, therefore, utilized where deep neural networks (DNNs) are employed as function approximators for the policy and/or the value-function to solve MDPs.

RL algorithms can be broadly classified into value-based and policy gradient algorithms. A value-based algorithm learns the action-value function and takes actions based on the values of the best action-value function. Examples of value-based RL algorithms are Q-Learning and SARSA. A policy gradient algorithm learns a policy directly by utilizing a policy function and iteratively training to improve the likelihood of actions depending on cumulative rewards. Examples of policy gradient RL algorithms are REINFORCE and vanilla policy gradient.

REINFORCE is the most basic version of a policy gradient algorithm. It effectively maximizes the likelihood of an action based on the expected cumulative reward obtained after taking that action. The simplified version of the objective function of the REINFORCE algorithm is expressed as follows [20]:

$$J(\pi_\theta) = \mathbb{E}_{\tau\sim\pi_\theta}\left[G_0(\tau)\right] = \mathbb{E}_{\tau\sim\pi_\theta}\left[\sum_{t=0}^{T}\gamma^t R_t\right] \tag{6}$$

where $\mathbb{E}$ denotes the expectation operator; θ denotes DNN parameters; τ is the variable representing the full trajectory denoted $\tau = S_0, A_0, R_1, S_1, ..., S_{t-1}, A_{t-1}, R_t, S_t$; γ is the discount factor; T is the total time step; π_θ denotes the policy; and $G_0(\tau)$ denotes the return of the full trajectory starting from $t = 0$.

B. Training Attributes

The training of the proposed DRL framework is performed for a certain number of episodes (n_ep). Since the REINFORCE algorithm is a policy-based algorithm, the policy parameter (θ) is initiated for a DNN that serves as a policy network. The following maximization problem is solved using the REINFORCE algorithm:

$$\max_\theta J(\pi_\theta) = \mathbb{E}_{\tau\sim\pi_\theta}\left[G_0(\tau)\right] \tag{7}$$

For the maximization of the objective, the gradient ascent is performed on the DNN parameters θ and the parameters are updated using the update rule given by (8).

$$\theta \leftarrow \theta + \alpha\Delta_\theta J(\pi_\theta) \tag{8}$$

where α is the learning rate; Δ_θ is the gradient operator; and $\Delta_\theta J(\pi_\theta)$ is the policy gradient calculated as follows [20].

$$\Delta_\theta J(\theta) = \mathbb{E}_{\tau\sim\pi_\theta}\left[\sum_{t=0}^{T} G_t(\tau)\Delta_\theta \log \pi_\theta(A_t|S_t)\right] \quad (9)$$

In (9), $G_t(\tau)$ denotes the discounted return from time step t to the termination of trajectory, which can be expressed as:

$$G_t(\tau) = \sum_{t'=t}^{T} \gamma^{t'-t} R_{t'} \quad (10)$$

Algorithm 1 provides the procedure of training the proposed DRL-based approach.

Algorithm 1: Training of the proposed approach.

Input : System data including line data, load data, on/off status of edges, etc.

Initialize parameters θ of policy network with random values

for $episode \leftarrow 1$ **to** n_{ep} **do**

- Initialize the system with a random state (here, a vector of line/branch status)
- **for** $t \leftarrow 1$ **to** T **do**
  - Get the action probabilities based on current state and stochastically select an action
  - Calculate the reward function $R_{t+1}(S_t, A_t)$ after passing the state and action through reward generator
  - Store the transition $(S_t, A_t, R_{t+1}(S_t, A_t))$ consisting of state, action, and reward
- Calculate the discounted return based on (10)
- Calculate the policy gradient based on (9)
- Perform back-propagation to update parameters θ of the policy network

Output : On/off status of network switches

IV. Case Study and Discussion

A. System Description

To demonstrate the effectiveness of the proposed approach, a 33-node system is used for numerical simulation. The 33-node distribution test system is a radial distribution system with 33 nodes, 32 branches, and 5 tie-lines (37 branches) [21]. All branches (including tie-lines) are numbered from 1 to 37. The system's overall load is 3.71MW. The system is modified by adding five DERs each with 300kW capacity at nodes 4, 7, 14, 20, and 32. The lines 2, 7, 11, 19, and 28 are assumed to be equipped with sectionalizing switches. The locations and amounts of critical loads considered for the 33-node system are shown in Table I. The hyper-parameter settings of the policy neural network of the proposed framework for the 33-node system are shown in Table II.

B. Training

The proposed DRL-based approach is trained for 10,000 episodes. During each episode, the state is randomly initialized with an outage scenario. Based on the randomly initiated state, the probability distribution of actions is calculated. An action is sampled from the action probability distribution, then the state and action are passed into the reward generator to calculate the reward. The states, actions, and rewards are stored in a memory bank, which are later utilized to train the policy neural network.

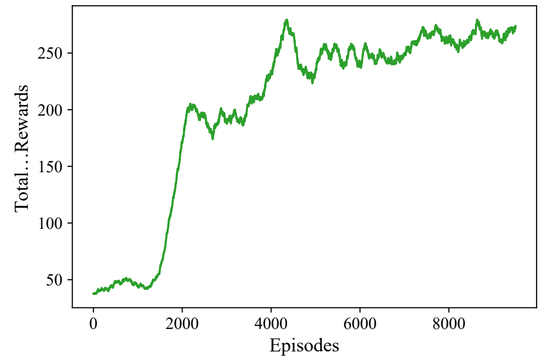

Fig. 2. Convergence process of Total Rewards

Fig. 2 shows the convergence process of Total Rewards during training of the proposed approach. The figure shows that the running mean (500-episode window) of the Total Rewards is initially very low until 2,000 episodes, but it increases as the training episode progresses and becomes relatively stable after 8,000 episodes. Since the policy neural network continues learning as the episodes increase, the Total Reward gradually increases and there is lower fluctuation after a certain number of episodes.

C. Testing and Implementation

For the testing and implementation of the proposed approach, three different test cases with different line outage scenarios were devised, which are as follows.

1) Test Case-I: For this test case, outages were assigned to lines 2, 3, 9, and 10. For this outage scenario, two microgrids (MG-1 and MG-2) and two isolates (IL-1 and IL-2) can be formed before the application of the proposed DRL approach, as shown in Fig. 3. The microgrids are energized by DERs, whereas the isolates are un-energized. The total critical loads in MG-1 are 940kW and the total generation is 900kW, which results in total critical load curtailment of 40kW in MG-1. In MG-2, the total generation is higher than the total critical load. In IL-1 and IL-2, the total critical loads are, respectively, 30kW and 45kW. Therefore, the total amount of curtailed critical load is 115kW for this baseline outage scenario, before implementing the proposed DRL-based approach.

After the implementation of the proposed approach, no changes were made in the status of sectionalizing switches (2, 7, 11, 19, and 28), whereas three tie-switches (35, 36, and 37) were closed as shown in Fig. 4. For this configuration, all system nodes are supplied through the substation, except an isolate (IL-1) consisting of node 10. The total critical load connected to node 10 is 30kW. Therefore, the total amount of curtailed critical load is only 30kW after the implementation of the proposed approach, recovering 85kW of the critical loads.

TABLE I
LOCATIONS OF CRITICAL LOADS FOR THE 33-NODE SYSTEM

Nodes	4	5	6	7	8	9	10	11	18	19	20	21	22	23	26	27	28	29	30	33
Critical Loads (kW)	60	30	60	200	200	60	30	25	45	45	45	45	45	45	60	60	60	60	60	30

TABLE II
HYPER-PARAMETER SETTINGS OF THE POLICY NEURAL NETWORK

Hyper-parameters	Values
Number of hidden layers	2
No. of neurons in hidden layers	10, 10
Learning rate	10^{-3}
Reward discount factor	0.95
Activation function of output layer	Linear
Activation function of hidden layers	ReLU
Optimizer	Adam

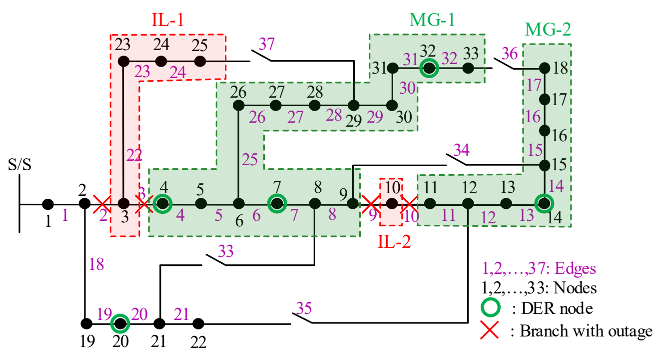

Fig. 3. Test Case-I before implementing the proposed DRL-based approach

2) Test Case-II: In the second test case, an outage scenario to recover all critical loads was used. Here, outages were assigned to lines 4, 8, 25, and 32. As a result, three microgrids (MG-1, MG-2, and MG-3) and an isolate (IL-1) are formed as shown in Fig. 5. The DERs in MG-1 and MG-2 can supply all of their critical loads. Microgrid MG-3 is energized by a DER of capacity 300kW connected to node 7. However, the total critical load in MG-3 is 490kW, of which 190kW is left unrecovered. The isolate IL-1 consists of node 33 with 30kW critical load, which is devoid of power supply. Therefore, the total amount of curtailed critical load for this case is 220kW before implementation of the proposed approach.

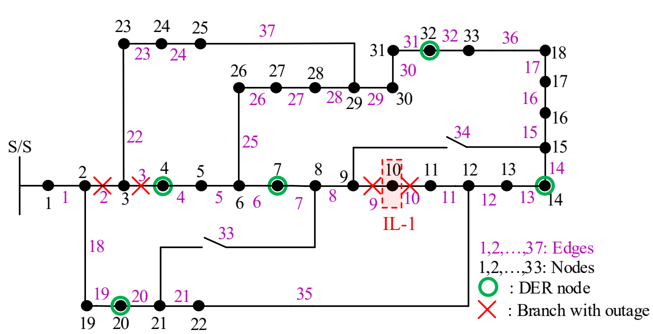

Fig. 4. Test Case-I after implementing the proposed DRL-based approach

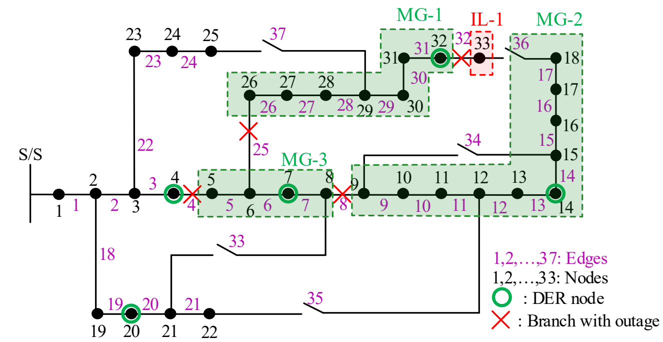

Fig. 5. Test Case-II before implementing the proposed DRL-based approach

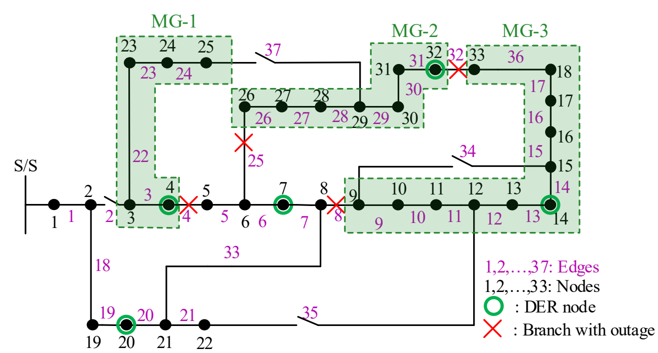

Fig. 6. Test Case-II after implementing the proposed DRL-based approach

When the proposed DRL-based approach is implemented, sectionalizing switch 2 is opened and tie-switches 33 and 36 are closed as shown in Fig. 6. As a result of these changes, three microgrids (MG-1, MG-2, and MG-3) are formed without any isolates as shown in Fig. 6. Each of the microgrids are energized by DERs and the total generation exceeds the total critical loads in each of them. Therefore, all critical loads are recovered in this test case.

3) Test Case-III: In the third test case, a more extreme outage scenario was considered, where an outage of the line connected to the substation (i.e., line 1) occurs, in addition to an outage of three lines (2, 12, and 23). Three microgrids (MG-1, MG-2, and MG-3) and an isolate (IL-1) can be formed as a result of the outage, as depicted in Fig. 7. The total critical load curtailment in MG-1 is 95kW, whereas MG-2 and MG-3 don't have any curtailed critical loads. The isolate IL-1 consists of nodes 24 and 25 which don't have any critical loads. Therefore, the total amount of curtailed critical loads is 95kW in this test case before implementing the proposed approach.

After the implementation of the proposed DRL-based approach, the status of sectionalizing switches were kept unchanged while tie-switches 33 and 36 were closed. This results in the formation of a microgrid MG-1 and an isolate IL-1

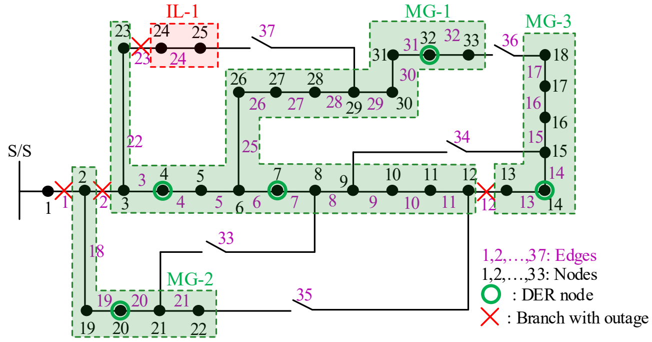

Fig. 7. Test Case-III before implementing the proposed DRL-based approach

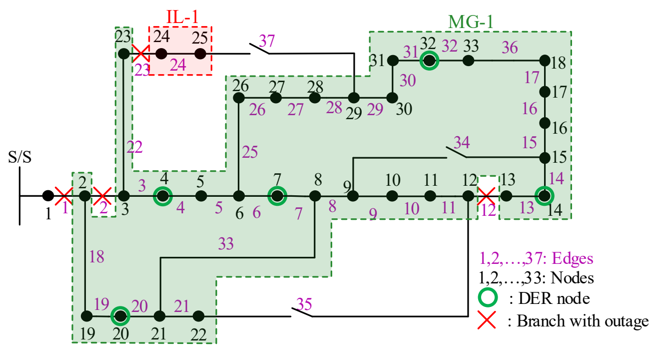

Fig. 8. Test Case-III after implementing the proposed DRL-based approach

as shown in Fig. 8. Since the total generation in MG-1 is higher than the total critical loads, there is no critical load curtailment. Also, the isolate IL-1 doesn't have any critical loads. Therefore, the implementation of the proposed approach leads to the recovery of all critical loads in this test case.

V. Conclusion

This paper has proposed a DRL-based approach for critical load restoration in active distribution systems against extreme events and multiple line outages. The proposed PDCLR utilizes DERs to form microgrids by changing the status of sectionalizing and tie switches present in the network. The distribution network was represented by an undirected graph and an optimal spanning forest was formed. REINFORCE, a policy gradient reinforcement learning algorithm, was used to train the policy neural network of the proposed DRL-based model. The convergence process during the training showed that the proposed approach continues learning as the training episode progresses and becomes relatively stable after a certain number of episodes. The proposed approach was tested and implemented on a 33-node distribution test system. The three test cases exhibit the effectiveness of the proposed approach to recover critical loads of the system by utilizing DERs and forming microgrids thereby enhancing the distribution system resilience. Applying this research to real-world examples, using historic outage events, and to other types of nodal distribution test systems would be valuable for future research.

Acknowledgement

This work was supported by the U.S. National Science Foundation (NSF) under Grant NSF 1847578.

References

[1] E. L. Hotchkiss and A. Dane, "Resilience roadmap: a collaborative approach to multi-jurisdictional resilience planning," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2019.
[2] R. J. Campbell, "Weather-related power outages and electric system resiliency," Congressional Research Service, Tech. Rep., 2012.
[3] A. Kavousi-Fard, M. Wang, and W. Su, "Stochastic resilient post-hurricane power system recovery based on mobile emergency resources and reconfigurable networked microgrids," *IEEE Access*, vol. 6, pp. 72 311–72 326, 2018.
[4] A. Gholami, T. Shekari, and S. Grijalva, "Proactive management of microgrids for resiliency enhancement: An adaptive robust approach," *IEEE Trans. on Sust. Energy*, vol. 10, no. 1, pp. 470–480, Jan 2019.
[5] M. Abdelmalak and M. Benidris, "Enhancing power system operational resilience against wildfires," *IEEE Transactions on Industry Applications*, vol. 58, no. 2, pp. 1611–1621, 2022.
[6] H. Gao, Y. Chen, Y. Xu, and C.-C. Liu, "Resilience-oriented critical load restoration using microgrids in distribution systems," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2837–2848, 2016.
[7] J. Liu, Y. Yu, and C. Qin, "Unified two-stage reconfiguration method for resilience enhancement of distribution systems," *IET Generation, Transmission & Distribution*, vol. 13, no. 9, pp. 1734–1745, 2019.
[8] J. Wu, X. Chen, S. Badakhshan, J. Zhang, and P. Wang, "Spectral graph clustering for intentional islanding operations in resilient hybrid energy systems," *arXiv preprint arXiv:2203.06579*, 2022.
[9] K. S. A. Sedzro, A. J. Lamadrid, and L. F. Zuluaga, "Allocation of resources using a microgrid formation approach for resilient electric grids," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2633–2643, 2018.
[10] C.-C. Liu, S. Lee, and S. Venkata, "An expert system operational aid for restoration and loss reduction of distribution systems," *IEEE Transactions on Power Systems*, vol. 3, no. 2, pp. 619–626, 1988.
[11] C. Chen, J. Wang, F. Qiu, and D. Zhao, "Resilient distribution system by microgrids formation after natural disasters," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 958–966, 2016.
[12] R. Rocchetta, "Enhancing the resilience of critical infrastructures: Statistical analysis of power grid spectral clustering and post-contingency vulnerability metrics," *Renewable and Sustainable Energy Reviews*, vol. 159, p. 112185, 2022.
[13] D. Das, "A fuzzy multiobjective approach for network reconfiguration of distribution systems," *IEEE Trans. Power Del.*, vol. 21, no. 1, pp. 202–209, 2006.
[14] C. Wang, P. Ju, S. Lei, Z. Wang, F. Wu, and Y. Hou, "Markov decision process-based resilience enhancement for distribution systems: An approximate dynamic programming approach," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2498–2510, 2020.
[15] J. Xie, I. Alvarez-Fernandez, and W. Sun, "A review of machine learning applications in power system resilience," in *2020 IEEE Power Energy Society General Meeting (PESGM)*, 2020, pp. 1–5.
[16] H. Nie, Y. Chen, Y. Xia, S. Huang, and B. Liu, "Optimizing the post-disaster control of islanded microgrid: A multi-agent deep reinforcement learning approach," *IEEE Access*, vol. 8, pp. 153 455–153 469, 2020.
[17] M. Gautam, M. Abdelmalak, M. MansourLakouraj, M. Benidris, and H. Livani, "Reconfiguration of distribution networks for resilience enhancement: A deep reinforcement learning-based approach," in *IEEE Industry Applications Society Annual Meeting*. IEEE, 2022, pp. 1–6.
[18] M. Abdelmalak, M. Gautam, S. Morash, A. Snyder, E. Hotchkiss, and M. Benidris, "Network reconfiguration for enhanced operational resilience using reinforcement learning," in *2022 International Conference on Smart Energy Systems and Technologies*. IEEE, 2022, pp. 1–6.
[19] A. Gibbons, *Algorithmic graph theory*. Cambridge university press, 1985.
[20] L. Graesser and W. L. Keng, *Foundations of deep reinforcement learning: theory and practice in Python*. Addison-Wesley Professional, 2019.
[21] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Power Engineering Review*, vol. 9, no. 4, pp. 101–102, 1989.