DEEP REINFORCEMENT LEARNING FOR POWER GRID MULTI-STAGE CASCADING FAILURE MITIGATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Cascading failures in power grids can lead to grid collapse, causing severe disruptions to social operations and economic activities. In certain cases, multi-stage cascading failures can occur. However, existing cascading-failure-mitigation strategies are usually single-stage-based, overlooking the complexity of the multi-stage scenario. This paper treats the multi-stage cascading failure problem as a reinforcement learning task and develops a simulation environment. The reinforcement learning agent is then trained via the deterministic policy gradient algorithm to achieve continuous actions. Finally, the effectiveness of the proposed approach is validated on the IEEE 14-bus and IEEE 118-bus systems.

1 INTRODUCTION

The modern large power grid consists of thousands of generators, substations, and transmission lines, all intricately interconnected and interdependent, working together to maintain the stable transmission of electricity. However, during the operation of the power system, various events may occur, among which *cascading failures* is particularly complex and highly damaging Chen et al. (2019); Jyoti & Hayat (2023); Uwamahoro & Eftekharnejad (2023). Cascading failures in power systems are typically triggered by the failure of a single component, e.g., a transmission line. These faults can rapidly *propagate* through the tightly interconnected network, potentially causing severe disturbances across the entire power grid and even leading to a complete system collapse Li & Tse (2024); Zhang et al. (2023); Li et al. (2024). Such kind of events can pose a significant threat to the security of the power grids and result in severe social and economic consequences.

Cascading failures can lead to devastating outcomes Guo et al. (2017); Salehpour & Al-Anbagi (2024). For example, on June 19, 2024, at approximately 15:17, Ecuador experienced a nationwide blackout, resulting in a collapse of the nation's power grid, affecting around 18 million people, with the power outage lasting for approximately 3 hours. The direct cause of this incident was the failure of the Milagro-Zhoray transmission line, which triggered a series of cascading failures, ultimately resulting in a widespread outage. This severe outage underscores the importance of developing fast cascading failure mitigation strategies for complex power grids.

In the power system area, cascading failure mitigation refers to a series of control actions to prevent the chain reaction after the first fault (e.g., one-line tripping), thereby avoiding system-wide blackouts. In recent years, numerous studies have emerged in this field. For example, Guo et al. (2024) proposed a method combining transient stability analysis with interaction graphs to identify critical lines and mitigate cascading failures by reducing the fault probability of components on these critical lines. Li et al. (2023) applied network flow theory to study the process of power flow redistribution and proposed a cascading failure mitigation strategy based on adaptive power balance recovery and selective edge protection. Inspired by the propagation patterns of faults, Bhaila & Wu (2024) employed graph neural networks (GNNs) to model and analyze cascading failures in power grids using an end-to-end approach. Liu et al. (2024), on the other hand, utilized an improved percolation theory to analyze the survivability of nodes in power grids and proposed an effective mitigation strategy.

In this paper, a deep reinforcement learning (DRL) approach is developed for mitigating multi-stage cascading failures (MSCF) in power systems, with the following contributions: (1) A simulation environment for multi-stage cascading failure study is constructed; (2) The Deep Deterministic Policy Gradient (DDPG) algorithm is adopted to address the MSCF issue; (3) The proposed model is validated on the IEEE 14-bus and 118-bus systems, demonstrating its effectiveness.
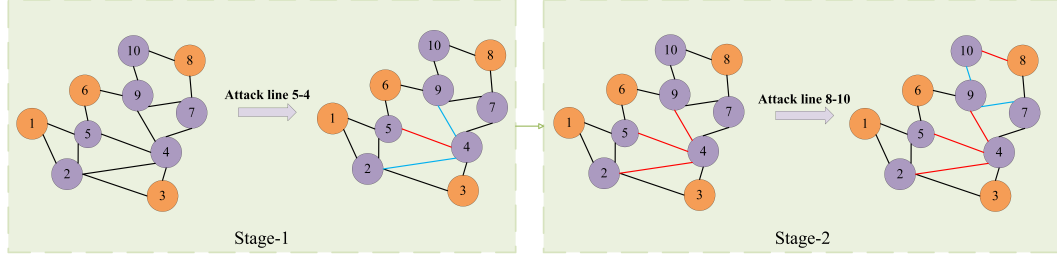
Figure 1: An example of a multi-stage cascading failure.

2 METHODOLOGY

2.1 MULTI-STAGE CASCADING FAILURE (MSCF) PROBLEM

Traditionally, single-stage cascading failure problems have been well studied Qi et al. (2017). However, in certain situations, multiple stages may occur Zhu (2021). For example, Fig. 1 depicts a multi-stage cascading failure example: an earthquake causes the loss of the power line 4-5, triggering the first stage of cascading failures (lines 2-4 and 4-9 are subsequently tripped due to the over-limit line power flow, *after* the loss of line 4-5). Suppose the (remaining) power grid does not collapse and enters a steady state. Then, after a short period, the aftershock may break another line, triggering *another stage* of cascading failures.

One approach to handling the MSCF problem is to decompose them into multiple sub-problems of single-stage and then solve each by each. However, this way might overlook the interdependence between stages. On the other hand, if we map the concept of "each *stage*" to the concept of "each *step*" in the RL context, then the MSCF problem can be investigated *holistically* under various mature frameworks of reinforcement learning, which is the motivation of this paper.

In this paper, the DDPG algorithm and the Actor-Critic framework are utilized Lillicrap et al. (2019), Mnih et al. (2016). The output of DDPG can be deterministic and real-number valued; hence, it performs well in solving problems with continuous actions Wang & Vittal (2023).

2.2 ENVIRONMENT IMPLEMENTATION

In our work, a simulation environment is developed for MSCF mitigation using Python and Matpower, which is a well-known MATLAB toolbox for AC power flow (ACPF) computation. Cross-tool interaction and data communication between Python and MATLAB have been achieved via a Python-MATLAB handler. Several key designs regarding this environment are described below.

2.2.1 DEFINITIONS OF STEP AND EPISODE

Step: a step means a **stage** when the power grid is attacked (e.g., by natural disasters), causing the grid to evolve into a new state (i.e., how many buses (i.e., nodes) and "lines" (i.e., edges) are still "available"; how many islands are formed; how large is the power flow on each remaining line; etc.).

Episode: an episode is one specific **set of steps** when the power grid is consecutively attacked. At the end of each episode, the final status is either "Win" or "Lose " (cf. definitions in later sections).

2.2.2 STATE DESIGN

For an *n*-bus power grid, our *state* is defined as follows:

$$state = [line\_status, P_1, Q_1, V_1, \theta_1, ..., P_n, Q_n, V_n, \theta_n]$$

where, $line\_status$ is the percentage value obtained via dividing the actual line power flow by its maximum limit; $P_i, Q_i, V_i, \theta_i, (i = 1, ..., n)$ denotes the active power injection, reactive power injection, voltage magnitude and angle of the i-th bus, respectively.
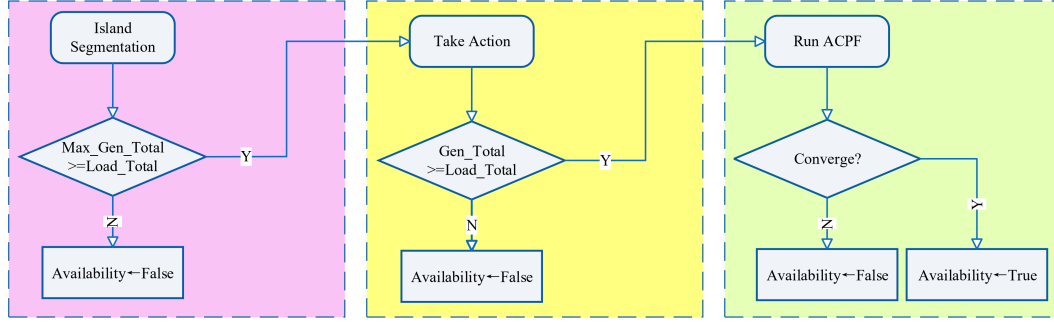
Figure 2: Island Availability assessment.

2.2.3 ACTION DESIGN

Cascading failure might be mitigated by adjusting the generator's power generation. Thus, the generation coefficients $[a_1, \ldots, a_m]$ of all m generators are considered as the *action*. The power output of the i-th generator is the product of a_i and its power capacity (i.e., the maximum power).

2.2.4 ISLAND DETECTION AND AVAILABILITY ASSESSMENT

The grid can become disconnected when lines are lost (due to an incident or line overload). Therefore, the first step is to assess the connectivity of the grid. To that end, we employ the union-find algorithm (c.f. Appendix A.1) to locate all the remaining islands.

The "availability" of an island means whether it is still *alive* at the end of a specific cascading failure stage; if not, it will be discarded in later stages. The availability assessment is carried out after the island detection. The criteria for island availability are described in Fig. 2. Max_Gen_Total and Gen_Total are respectively the total power capacity and the total actual power output of all the remaining generators in a specific island, and $Load\_Total$ is the total load demand in that island.

2.2.5 REWARD DESIGN

- **Total cost of generation:** $-c_1 \cdot cost$. Here, c_1 is a hyperparameter. *cost* means the total generation cost (\$) of all islands whose *availability* is true.
- **Loss of load penalty:** $-BaseReward_1 \cdot P_{loss}/P_{total}$. P_{loss} is the total load on *unavailable* islands at current stage, while P_{total} represents the original total load of initial power grid.
- **Convergence reward:** $BaseReward_2$. This reward is given when half or more of all the currently remaining islands have converged.
- **Win reward:** $BaseReward_3 \cdot (P_{available}/P_{total})^{c_2}$. This reward is given when the win conditions are met. $P_{available}$ is the total load of *available* islands.

Here, c_1, c_2, $BaseReward_1$, $BaseReward_2$, and $BaseReward_3$ are constants related to a specific power grid. A basic idea in picking those constants is to make the above four parts in the same order of magnitude. Finally, the overall workflow for the RL-based MSCF study is shown in Fig. 3(a).

3 EXPERIMENTS AND RESULTS

The proposed approach is tested on the IEEE 14-bus and modified IEEE 118-bus systems. The IEEE 14-bus system has 5 generators and 20 lines, with its topology shown in Fig. 3(b). For other details about the experiment settings and hyperparameters, please refer to Appendix A.5.

For each power grid, a DRL model is trained for 300 episodes. After training, the model interacted with the environment for an additional 1000 episodes, during which the total reward in each episode is recorded, and the final win rate is computed.

The model is compared with three baseline strategies, as shown in Table 1 and Fig. 4. **Baseline 1** means each generator output a random power. **Baseline 2** means all generators output the maximum power. **Baseline 3** means all generators operate at half of their maximum power output. It can
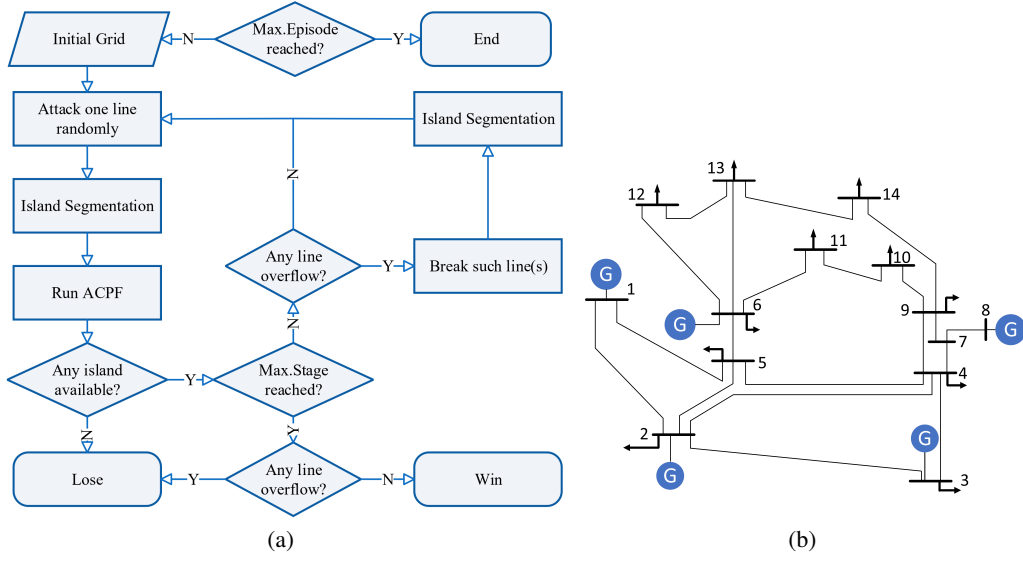
Figure 3: (a) The overall workflow of grid simulation for MSCF study; (b) The IEEE 14-bus system.

Table 1: Win rate comparison.

Method	IEEE 14-bus System	IEEE 118-bus System
DDPG	**95.5%**	**97.8%**
Baseline 1	52.0%	51.7%
Baseline 2	93.3%	8.40%
Baseline 3	85.6%	97.0%

be observed that the DRL achieves a good performance, with the highest win rate, large average rewards, and more stable behaviors.

4 CONCLUSION

This paper implements and validates a DRL-based solution for multi-stage cascading failure mitigation. One limitation of the current solution is that the differences in the states are relatively small, causing the majority of the model's actions to be similar. In future work, we will explore other state designs to improve the action's variability.
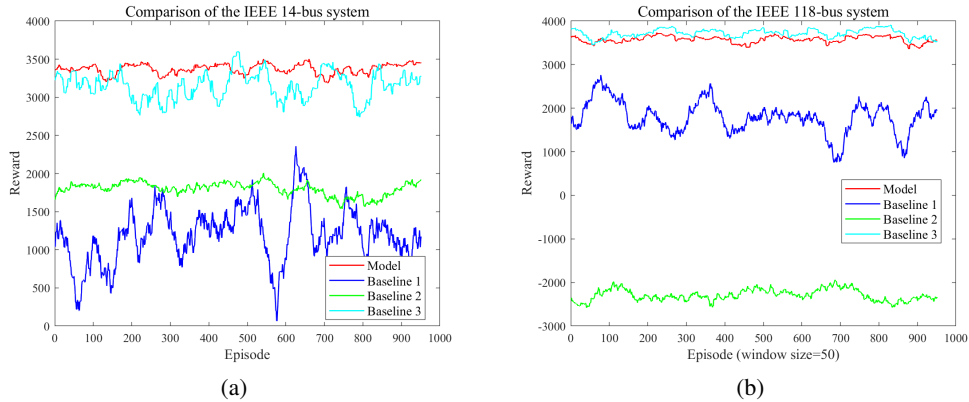

Figure 4: The moving-average reward comparison.

4