

---

# Power Grid Cascading Failure Mitigation by Reinforcement Learning

---

Yongli Zhu <sup>\*</sup> <sup>1</sup>

## Abstract

This paper proposes a cascading failure mitigation strategy based on Reinforcement Learning (RL) method. Firstly, the principles of RL are introduced. Then, the Multi-Stage Cascading Failure (MSCF) problem is formulated, and its challenges are investigated. The problem is then tackled by the RL based on DCOPF (Direct Current Optimal Power Flow). Designs of the RL framework (rewards, states, etc.) are illustrated in detail. Experiments on the IEEE 118-bus system by the proposed RL method demonstrate promising performance in reducing system collapses.

## 1. Introduction

Increasing renewable sources (e.g., wind, solar) are integrated into modern power grids to reduce emissions. However, due to the intermittent natures of those renewable sources, the original power grid can become fragile, i.e., more easily affected by various risks, among which the cascading failure is one of the most challenging issues necessary to be addressed. A cascading failure is defined as a series of consecutive malfunctions of physical components (e.g., power transmission lines, power substations). Cascading failure is typically caused by an unexpected natural disaster such as a hurricane, typhoon and flood. Severe cascading failures can lead to a total system collapse (i.e., disintegrated into small energized pieces) or even a blackout event (i.e., loss of electricity for the entire city or country). To recover the power grid to its original healthy state, backup generator units need to be turned on. The relevance of the cascading failure problem to climate change is:

- 1) Given the increase in extreme weather events due to climate change, having a stable power grid is critical for renewable resources integration;
- 2) The backup generators are typically fossil-fuel (e.g., coal,

diesel) based units that emit greenhouse gases.

Therefore, it is meaningful to develop nuanced strategies to mitigate power grid cascading failures at its early stages with as little generation cost as possible.

The cascading failure mitigation can be regarded as a stochastic dynamic programming problem with unknown information about the risk of failures. Previous researches try to tackle this problem based on either mathematical programming methods or heuristic methods. For example, bi-level programming is used to mitigate cascading failures when energy storages exist (Du & Lu, 2014). In (Han et al., 2018), a percolation theory-based algorithm is employed for mitigating cascading failures by using UPFC (Unified Power Flow Controller) to redistribute the system power flow more evenly. In (Tootaghaj et al., 2018), a recovery plan for load-serving from cascading failures was put forward considering the uncertainty of failure locations. In (Cordova-Garcia et al., 2018), a cascade control algorithm using load shedding considering the communication network delays for power grids was proposed to reduce the failure of power lines. In (Shuvro et al., 2017), By characterizing the cascading-failure dynamics as a Markov chain model, it is found that the impact of human-operator action will have a significant impact on cascading failures.

However, some of the above research share the same limitation: unrealistic assumptions are often made, which yield impractical control strategies in terms of time or economic cost. Different than communication networks and social networks, the power grid is not a pure-mathematical *graph*, but a physical grid, i.e., both its node (called “bus” in power system terminology) and edge (called “branch” in power system terminology) are physical entities that can not be added or removed arbitrarily. On the other hand, most power system lines are equipped with automatic protection-relay devices, which can trip the line when its flowing current/power or temperature is above a certain threshold in a pre-defined time window. Thus, in this paper, the main focus is put on branch failure rather than node failure.

Meanwhile, some emerging artificial intelligence technologies, such as reinforcement learning (RL) and deep learning (DL), have nourished both the fields of power systems and control theory (Kiumarsi et al., 2018). In (Glavic et al., 2017), a holistic literature review is given on the recent

---

<sup>1</sup>Texas A&M University, College Station, USA.  
Correspondence to: Yongli Zhu <yongliz@tamu.edu, yzhu16@vols.utk.edu>.

development of RL applications in the power system area, though topics regarding cascading failure are not covered. In (Vlachogiannis & Hatziaargyriou, 2004), the RL method is used for reactive power control. In (Liu et al., 2018), voltage restoration for an islanded microgrid is achieved via a distributed RL method. In (Zhu et al., 2018), an application for disturbance classification is proposed based on image embedding and convolutional neural network (CNN). In (Yan et al., 2018), deep learning is applied in power consumption forecasting. However, the application of RL or DL for cascading failure study are less reported.

In this paper, a reinforcement learning approach is designed for the mitigation of cascading failures with the following contributions:

- 1) Propose and formulate a novel problem called *Multi-Stage Cascading Failure* (MSCF) for the first time.
- 2) Present a systematic reinforcement learning framework to tackle the MSCF problem. Similar to AlphaGo (Silver et al., 2016), a “two-player game” idea is utilized in this study.
- 3) Unlike some previous study which treats power grid purely as a graph model with no or less physical background, this paper uses a professional power system simulator as the *environment* in the RL framework, to better reflect real characteristics of the power system. In this way, the learning result is more convincing, and the trained mitigation strategy will be more practical.

The remaining parts of this paper are organized as follows. Section 2 proposes an RL-based control framework for the mitigation of cascading failures. Section 3 presents the case study and results analysis. Finally, conclusions and future directions are given in Section 4.

## 2. Multi-Stage Cascading Failure Control

### 2.1. Multi-Stage Cascading Failure (MSCF) Problem

Firstly, the following definitions are given:

*Generation*: one event of the cascading failures within one stage, e.g., a line tripping.

*Stage*: after an attack (e.g., one line is broken by a natural disaster), the grid *evolves* with a series of potential *generations* (e.g., line tripping events if the line thermal limits are reached). Finally, the power system will either reach a new equilibrium point if it exists (i.e., the AC power flow (ACPF) converges, and all the branch flows are within secure limits), or the system collapses.

In conventional cascading failure analysis, only one stage is considered. However, in certain situations, succeeding

stages might follow shortly. For example, a wind storm results in one generation, in which certain lines are lost, and the system reaches a new steady state. Then, shortly, a new stage is invoked by tripping an important line due to the misoperation of human-operator or relay protection. As an example, Table 1 and 2 list the simulation results of the IEEE 118 system for a two-stage MSCF problem in two independent episodes.

A naïve way to handle such complicated multi-stage situations is to treat each stage independently by the existing control method. However, such kind of approach may not work well due to overlooking the correlations between any two consecutive stages. For example, the previous study (Zhu et al., 2014; Chen et al., 2007) found that attacking each line one by one sequentially can sometimes achieve a more severe effect (i.e., more components loss) than attack multiple lines simultaneously. Thus, the MSCF problem should be considered from a *holistic* perspective.

### 2.2. Mimicking the corrective controls by DCOPF

The following DC optimal power flow (DCOPF) is adopted to mimic the operator’s corrective control measures when a cascading failure event happens. i.e., re-dispatch the generators and do load shedding (when necessary).

$$\begin{aligned}
 \min_{p_i, p_j} \quad & \sum_{i \in G} c_i p_i + \sum_{j \in D} d_j (p_j - P_{dj}) \\
 \text{s.t.} \quad & \mathbf{F} = \mathbf{A} \mathbf{p} \\
 & \sum_{k=1}^n p_k = 0 \\
 & P_{dj} \leq p_j \leq 0, \quad j \in D \\
 & P_{gi}^{\min} \leq p_i \leq P_{gi}^{\max}, \quad i \in G \\
 & -F_L^{\max} \leq F_l \leq F_L^{\max}, \quad l \in L
 \end{aligned} \tag{1}$$

Where,  $n$  is the total bus number.  $G, D, L$  are respectively the generator set, load set and branch set;  $\mathbf{F} = [F_l](l \in L)$  represents the branch flow;  $p_i (i \in G)$  is the generation dispatch for the  $i$ -th generator;  $p_j (j \in D)$  is the load dispatch for the  $j$ -th load;  $\mathbf{p} = [p_k], k = 1 \dots n$  represents the (net)

Table 1. Result of Episode-1.

Stage-1	Generation-1	ACPF	OVER-LIMIT
		CONVERGE	LINES
		YES	0
Stage-2	Generation-2	ACPF	OVER-LIMIT
		CONVERGE	LINES
		YES	0
Result		WIN	

nodal power injections.  $\mathbf{A}$  is a constant matrix to associate the (net) nodal power injections with the branch flows.  $P_{dj}$  is the normal demand value for the  $j$ -th load;  $c_j$  is the given cost coefficient of generation;  $d_j$  is the given cost coefficient of load shedding.  $p_i (i \in G)$  and  $p_j (j \in D)$  are the decision variables for generators and load respectively.

### 2.3. Apply RL for MSCF problem

To apply RL to a specific power system problem, the first step is to map the physical quantities to corresponding concepts of RL theory, i.e., reward, state, and action.

#### 1) Reward design (of each Stage)

- – Total generation cost (i.e., the negative objective value of DCOPF), if DCOPF converge.
- – 1000, if DCOPF or ACPF diverge.
- +1000, an extra reward if system finally reaches a new steady-state **at the last stage**.

Those values ( $\pm 1000$ ) are based on by trial-and-error experiments.

#### 2) Action design

- If the line flow limit is too low, the DCOPF might not converge due to the narrow feasible region. On the contrary, if the line flow limit is too high, the feasible region also becomes large. However, the obtained optimal solution might lead to an operation point with *tighter* power flow status on each branch, resulting in cascading failures at the next stage of the MSCF problem. Thus, the “branch flow limit”  $F_l^{max}$  in the previous DCOPF formulation (2) is adopted as the action in the RL learning framework.

#### 3) State design

- Several quantities of each bus and the power flow of each branch are chosen and packed as the state, i.e.,  $state = [branch\_loading\_status, V_1, \theta_1, P_1, Q_1, \dots, V_n, \theta_n, P_n, Q_n]$ , where, *branch\_loading\_status* are the percentage values calculated by dividing each branch flow by its

loading limit for all the branches;  $V_i, \theta_i, P_i, Q_i (i = 1 \dots n)$  are respectively the voltage magnitude, voltage angle, active power injection, and reactive power injection of each bus.

#### 4) Environment

- In this study, the learning environment in the RL framework is just the power grid itself. Thus, a co-simulation platform based on the combination of DigSILENT and MATLAB is implemented. A commercial tool DigSILENT is used as the simulator (*environment*) to provide all needed information (*states* and *rewards*) to the RL network for training. Besides, the concept of *step* within one independent *episode* corresponds to one *stage* in the MSCF problem.

Finally, the overall workflow of the power grid simulation for the MSCF study is shown in Figure 1.

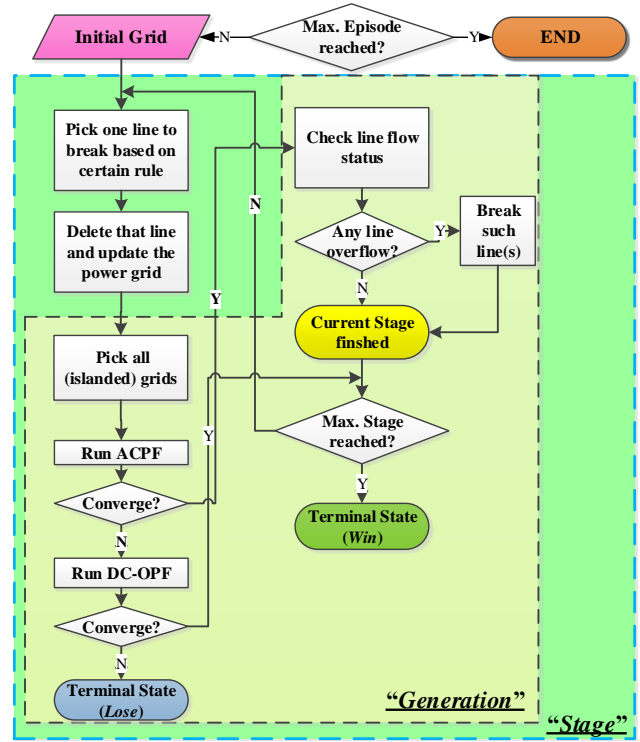


Figure 1. The overall workflow of grid simulation for MSCF study.

Table 2. Result of Episode-2.

Stage-1		ACPF	OVER-LIMIT
		CONVERGE	LINES
Generation-1	Generation-1	YES	2
	Generation-2	YES	0
Stage-2		ACPF	OVER-LIMIT
		CONVERGE	LINES
Generation-1	Generation-1	YES	4
	Generation-2	YES	2
	Generation-3	YES	2
	Generation-4	YES	3
	Generation-5	YES	10
	Generation-6	YES	20
	Generation-7	NO	–
Result		LOSE	

## 3. Case Study

In this section, a modified IEEE 118-bus system is adopted as the test-bed for the proposed MSCF mitigation strategy. The maximum stage number is set to 3. It contains 137 buses and 177 lines (parallel lines included), 19 generators, and 91 loads. The system topology is shown in Figure 2, where the red dot stands for generators.

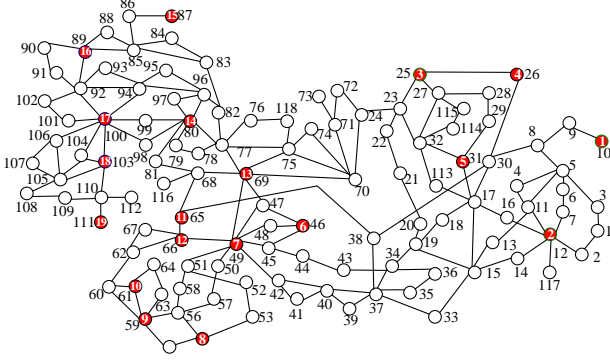


Figure 2. The topology of IEEE 118-bus power systems.

### 3.1. Shallow Neural Network

The architecture for the shallow neural network is that: one input layer, one output player, and one hidden layer with 10 neuron units. Its input is a 1-D vector with 753 ( $=137 \times 4 + 177 + 28$ ) elements; the output is the action in the RL framework (i.e., the line flow limit, c.f. Section 3).

Since both the hidden-layer dimension and output-layer dimension of the shallow network are one, the *SARSA* (On-policy TD) method is employed. During the training, the *action* is bounded by the range  $[0.80, 1.25]$ .

### 3.2. Deep Neural Network

#### 1) Feature engineering

To create an image-like input to utilize the CNN, the original input (length = 753) is extended to a length of  $784 = 28 \times 28$  by appending extra zeros.

2) Network structure Typically, a deeper network and more layers might lead to over-fitting in practice. Thus, the network structure used in this paper is shown in Figure 3.

3) The *Q*-learning (Off-policy TD) method is applied on it. The output of the 2nd-last layer (dimension  $1 \times 10$ ) will be used in both  $\epsilon$ -greedy policy and greedy policy. The last-layer output (dimension  $1 \times 1$ ) will be finally used to update the *Q*-network parameters. The candidate set of *action* is  $[0.8, 0.85, 0.9, 0.95, 1.0, 1.05, 1.1, 1.15, 1.20, 1.25] \in \mathbf{R}^{10}$ .

### 3.3. Experiments and Results

For both networks, the maximum episode number is 10000, the learning rate is  $10^{-4}$ , and the discount rate  $\gamma$  is 0.7. The learning performance is shown in Table 3. The plot of moving average *reward* (window size = 1000) for deep network case is shown in Figure 4. It can be observed that both RL and Deep RL have achieved satisfactory results in terms of winning rates (i.e., lower cascading risks). In both cases, the average return per episode is more than half of the

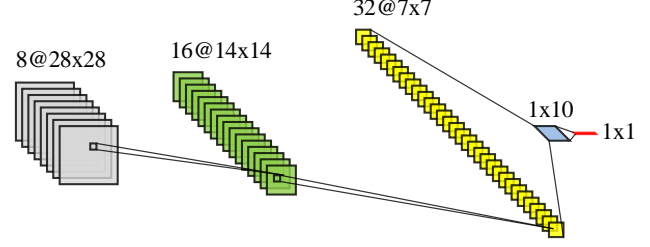


Figure 3. The network structure used in Deep RL.

Table 3. Learning Performance.

PERFORMANCE	SHALLOW NETWORK	DEEP NETWORK
Winning rate	78.00%	78.07%
Avg. reward	640.08	630.46

maximum possible value (i.e.,  $500 = 1000/2$ ), which shows a positive learning ability of the RL agent in mitigating cascading failures.

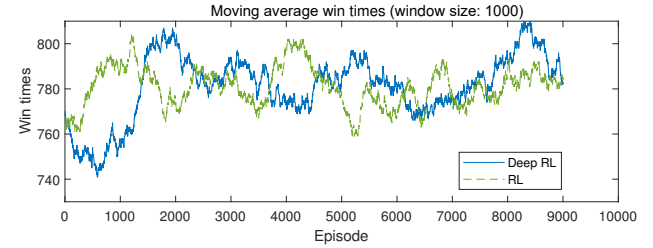


Figure 4. The learning performance by the DRL (10000 episodes) (best viewed in color).

## 4. Conclusions

In this paper, a reinforcement learning-based mitigation strategy for the Multi-Stage Cascading Failure problem is proposed. The trained RL agent works effectively on the IEEE 118-bus system under both shallow and deep architectures with an approximately 78% chance to avoid the power grid collapse. Potential benefits of the proposed idea in this paper include 1) enhanced resilience to extreme weather events and 2) increased penetration level of renewable sources. Investigating the effects of hyper-parameters (e.g., layer numbers, hidden neuron units, learning rate, reward amount, discount factor) of the RL network on the mitigation performance will be the next step.

## References

- Chen, X., Sun, K., Cao, Y., and Wang, S. Identification of vulnerable lines in power grid based on complex network theory. In *2007 IEEE Power Engineering Society General Meeting*, pp. 1–6, June 2007.
- Cordova-Garcia, J., Wang, X., Xie, D., Zhao, Y., and Zuo, L. Control of communications-dependent cascading failures in power grids. *IEEE Transactions on Smart Grid*, pp. 1–1, 2018.
- Du, P. and Lu, N. *Energy Storage for Smart Grids*. Academic Press, 1st edition, 2014.
- Glavic, M., Fonteneau, R., and Ernst, D. Reinforcement learning for electric power system decision and control: Past considerations and perspectives. *IFAC-PapersOnLine*, 50(1):6918 – 6927, 2017. 20th IFAC World Congress.
- Han, Y., Guo, C., Ma, S., and Song, D. Modeling cascading failures and mitigation strategies in pmu based cyber-physical power systems. *Journal of Modern Power Systems and Clean Energy*, 6(5):944–957, Sep 2018.
- Kiumarsi, B., Vamvoudakis, K. G., Modares, H., and Lewis, F. L. Optimal and autonomous control using reinforcement learning: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, 29(6):2042–2062, June 2018.
- Liu, Z., Luo, Y., Zhuo, R., and Jin, X. Distributed reinforcement learning to coordinate current sharing and voltage restoration for islanded dc microgrid. *Journal of Modern Power Systems and Clean Energy*, 6(2):364–374, Mar 2018.
- Shuvro, R. A., Wangt, Z., Das, P., Naeini, M. R., and Hayat, M. M. Modeling cascading-failures in power grids including communication and human operator impacts. In *2017 IEEE Green Energy and Smart Systems Conference (IGESSC)*, pp. 1–6, Nov 2017.
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., and Hassabis, D. Mastering the game of go with deep neural networks and tree search. *Nature*, 529:484–503, 2016.
- Tootaghaj, D. Z., Bartolini, N., Khamfroush, H., He, T., Chaudhuri, N. R., and Porta, T. L. Mitigation and recovery from cascading failures in interdependent networks under uncertainty. *IEEE Transactions on Control of Network Systems*, pp. 1–1, 2018.
- Vlachogiannis, J. G. and Hatziargyriou, N. D. Reinforcement learning for reactive power control. *IEEE Transactions on Power Systems*, 19(3):1317–1325, Aug 2004.
- Yan, K., Wang, X., Du, Y., Jin, N., Huang, H., and Zhou, H. Multi-step short-term power consumption forecasting with a hybrid deep learning strategy. *Energies*, 11(11), 2018.
- Zhu, Y., Yan, J., Tang, Y., Sun, Y., and He, H. The sequential attack against power grid networks. In *2014 IEEE International Conference on Communications (ICC)*, pp. 616–621, June 2014.
- Zhu, Y., Liu, C., and Sun, K. Image embedding of pmu data for deep learning towards transient disturbance classification. In *2018 IEEE International Conference on Energy Internet (ICEI)*, pp. 169–174, May 2018.