Reinforcement Learning-Based Active Distribution Management for Reducing the Risk of Cascading Failure

Oindrilla Dutta
Dept. of Electrical Engineering
The City College of New York
New York, USA
odutta@ccny.cuny.edu

Chin Chan
Dept. of Electrical Engineering
The City College of New York
New York, USA
cchan009@citymail.cuny.edu

Mahmoud Saleh
Dept. of Electrical Engineering
The City College of New York
New York, USA
mahmoudsalahem@gmail.com

Ahmed Mohamed
Dept. of Electrical Engineering
The City College of New York
New York, USA
amohamed@ccny.cuny.edu

***Abstract*—In this paper, a controller for reducing the risk of cascading line outages in an active distribution network (ADN), has been developed. The controller is based on adaptive critic design (ACD), which receives inputs from a distribution management system regarding the states of the distribution network. These states are then evaluated by the controller for critical contingencies in the network. Accordingly, the controller takes coordinated control actions by giving recommendations for adjusting the optimal power flow (OPF) set points for voltage and active power of the distributed energy resources (DERs) in the network. These control actions avoid considerable deviation from the optimal set points for generation cost and line losses. The effectiveness of the developed ACD controller is demonstrated with a test case of IEEE 30 bus system. The technique for formulation of this ACD algorithm, so as to attain a faster but accurate convergence, has been elaborately described in this work.**

***Index Terms*—Adaptive critic, Active Distribution Management, Cascading failure, Dual Heuristic Programming.**

I. Nomenclature

Variable	*Description*
$X(t)$	State of the plant at time *t*.
$A(t)$	Action output from Actor network at time *t*.
$J(t)$	Cost-to-go function at time *t*.
$U(t)$	Utility function at time *t*.
γ	Discount factor
t_0, t_f	Initial time, final time
N	Total number of trials

II. Introduction

Power system is one of the most critical infrastructures as it serves as the backbone of other critical infrastructures like transportation, communication, water purification, and so on. A fundamental property of such interdependent systems is that failure of nodes in one network leads to failure of dependent nodes in the other networks, resulting in a series of malfunctioning. A real-world example of such cascade of failures is the electrical blackout that affected much of Italy on 28 September 2003 [1], [2]. Node failures in a power network are more prevalent during severe weather conditions, which may result in over loading of lines and their eventual outage. This primarily occurs because the assets in a power network have often been lately operated near their capacity, in order to attain optimal generation cost and line losses. In this situation, the system elements have smaller margins of flexibility to handle minor changes in loading before failure. Hence, these traditional objectives of optimal line losses and cost may have to be adjusted to emphasize on the wider requirement of interdependent networks. An initial step towards preventing such concurrent failures in modern systems, is to avert one line outage cascading to the next (*N-1*) and resulting in node failures or load shedding.

Prevention of line outage in a distribution network (DN) during critical conditions is a real time control problem that requires coordinated control actions for managing load, generation and energy storage systems. Studies related to real time control problems in the ADN have primarily focused on isolation of a bus, which is connected to a microgrid [3]-[6]. These surveys have emphasized on changing protection strategies in an ADN and ensuring adequacy of reserve power in the DERs and ESSs when the load gets isolated. However, there is a dearth of research concerning real time prevention of node failure in a DN, which is a step ahead of node isolation.

In this paper, we have proposed a control technique that can adapt the actions to be taken in a ADN, based on its current states, so as to prevent an *N-1* line outage. The controller aims at better resource management by amending the voltage and generation set points of the DERs that are estimated by OPF. Care has to be taken for the controller to account for minimum adjustments to such traditional parameters, while also optimizing the power flow in the system, so that major deviation from optimal cost and line losses is avoided. Hence the controller has to solve an optimization problem by taking actions within critical constraints of time and resources.

Adaptive dynamic programming (ADP) based techniques have been found to be most efficient in taking such online and forward-in-time optimal control actions [7]. Hence, in this paper we have adopted an ADP based technique, called adaptive critic design (ACD), which can cope with a large number of variables in parallel, real-time and non-linear environment. Applications of robust ADP controllers in power system has received particular attention in the area of microgrid stability [8] – [10]. However, its application in line outage management has not been investigated, to the best of our

978-1-7281-4547-1/20/$31.00 ©2020 IEEE

knowledge. In this paper, we have demonstrated an effective application of ADP in tackling line outage in a DN. This work identifies the specific type of ACD algorithm which is best suited for this application and elaborates on the method of formulating the algorithm, so that it can converge effectively. The controller has been tested on a snapshot of the IEEE 30 bus system.

III. Adaptive Critic Design (ACD)

The adaptive critic family of designs is used for finding a series of optimal (minimum/maximum) control actions that must be taken in sequence, where the quality of those actions is unknown until the end of that sequence [11]-[14].

A. *Component Networks of ACD*

The fundamental structure of ACD comprises three networks, whose functionalities have been outlined as follows:

1) *Model Network (MN)*

The Model network is a deterministic or stochastic representation of the plant that is to be controlled. The states of the plant are observable, discrete and deterministic within the time interval t_0 to t_f, and the initial state i.e. $X(t_0)$ is available. This network receives the states of the plant and the actions at current time step as inputs to provide estimated states of the plant at the next time-step as output.

2) *Actor Network (AN)*

The Actor network consists of a Backpropagation Neural Network (BpNN) or Recurrent Neural Network (RNN), which takes $X(t)$ as input and outputs successively improving control actions, $A(t)$. These control actions when applied to the plant will minimize/maximize the cost-to-go function, J.

3) *Critic Network (CN)*

Similar to Actor, the Critic network is also a BpNN or RNN. It receives input from the model and actor networks and generates improved value functions or its derivatives based on the improved control laws from the Actor.

B. *Types of ACD*

The ACD family is categorized into different types of strategies based on the method of approximating J. In this paper, Dual Heuristic Programming (DHP) has been found to be more relevant in defining the problem for analyzing and mitigating contingencies in power system. The fundamental structure of DHP is demonstrated here.

1) *Dual Heuristic Dynamic Programming (DHP)*

The working principle of DHP is shown in Fig. 1. DHP goes through two-levels of iterations, simultaneously. One level of iteration (denoted by k) represents the passage of time in the dynamic process, while the other (denoted by l) represents the number of trials in the search for the optimal solution. DHP starts its iteration through l and k by randomly initializing the weights in the AN and CN. AN is provided with the initial state of the plant, $X(k = t_0)$, as input and it generates the action, $A_l(k)$, for the next time step $(t = k+1)$. The MN also receives $X(k = t_0)$ from the plant and $A_l(k)$ from AN to generate $X(k+1)$. The outputs from plant, MN and AN are supplied to the derivative block, which estimates the derivatives required for evaluating the cost sensitivity of the algorithm to state perturbations on the optimal trajectory. The cost function is defined in (1) and its derivative with respect to the state perturbations is provided in (2).

$$J\big(X(t),A(t)\big) = U\big(X(t),A(t)\big) + \gamma\langle J(t+1)\rangle, \tag{1}$$

$$\begin{aligned}\lambda\big(X(t)\big) = \frac{\partial J\big(X(t),A(t)\big)}{\partial X(t)} = \\ \frac{\partial U(t)}{\partial X(t)} + \frac{\partial U(t)}{\partial A(t)}\frac{\partial A(t)}{\partial X(t)} + \gamma\left\langle \lambda\big(X(t+1)\big)\frac{\partial X(t+1)}{\partial X(t)}\right\rangle \\ + \gamma\left\langle \lambda\big(X(t+1)\big)\frac{\partial X(t+1)}{\partial A(t)}\frac{\partial A(t)}{\partial X(t)}\right\rangle\end{aligned} \tag{2}$$

The CN approximates the same derivative based on its random initial weights and the MN output. The error in this estimation by the Critic is expressed by (3).

$$\mathcal{E}_C = \lambda\big(X(t)\big) - \lambda^*\big(X(t)\big) \tag{3}$$

The mean square of this error (MSE), $\|\mathcal{E}_C\|^2$, is positive definite and is used to train the BpNN or RNN of CN, so as to obtain an improved value function. Now, the output of CN is used to update the AN parameters based on (4) and (5). Similar to CN, the AN BpNN or RNN is trained with the MSE of Actor, $\|\mathcal{E}_A\|^2$. The improved control law obtained from the training of AN is used to train the CN a second time. Hence, the weights of the Actor and Critic are not random anymore.

$$\mathcal{E}_A = \frac{\partial J\big(X(t)\big)}{\partial A(t)} - 0, \tag{4}$$

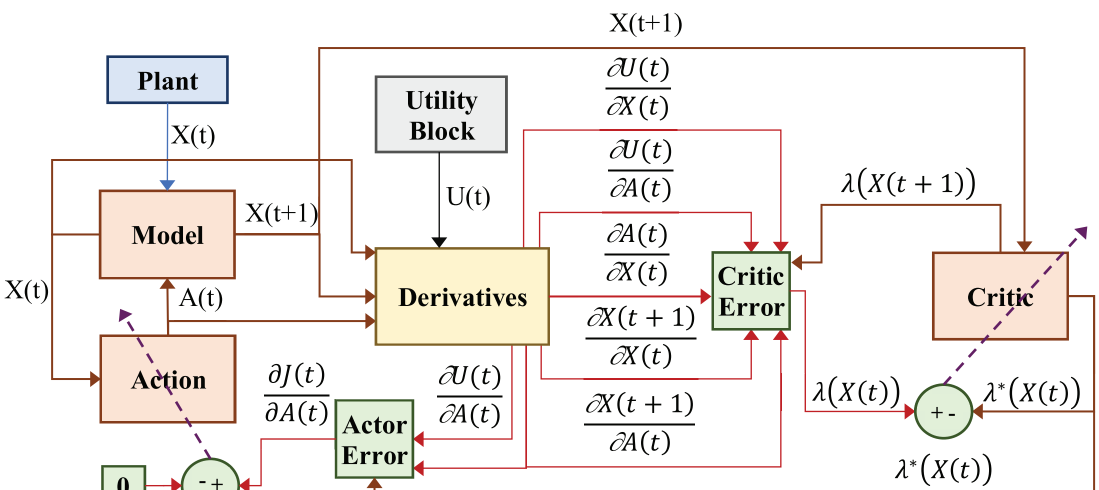


Fig. 1. Dual Heuristic Programing Process

Progress of the iteration in k continues in this manner until a stopping criterion is reached, after which the algorithm starts similar iterations for $l+n$, where $n = 1,2,\ldots,N\text{-}1$.

IV. Test Case

In this paper, the IEEE 30 bus system has been used for case study. This system consists of 6 generator buses (including one slack bus), 41 branches, and supplies a total load of 217.5 MVA. Five of these generators are considered as DERs and the slack bus is considered as connection to the transmission network. The system has been surveyed for critical contingencies that might induce a cascading failure when the generators operate at OPF set points. Fig. 2 shows one such cascading failure that starts with a step change in load at bus 15. The increased load results in critical loading of branches 29, 30 and 35 when OPF is used. Under these circumstances, a line outage at branch 30 causes ~15% overloading of branches 29 and 31, and 6% overloading of branch 35. This would cause outage at lines 29 and 31 first. In Fig. 2, only branch 29 has been disconnected, and this has resulted in 20% overloading of branch 28. That would cause fast disconnection of branch 28 and the OPF will no longer converge. In power system, disconnection of branches due to overloading happens within a span of 5-12 minutes [10], [11].

Cascading failure can be avoided by preventive measures [15], where the operator is sent warning if the system is operating near the thresholds. The operator must take prompt actions to mitigate the failure. A possible set of actions for mitigating the aforementioned failure could be adjustment of the optimal set points for active power and voltage of the five DERs prior to the disconnection of any branch, including branch 30, which is the most critical branch in this case. The primary task would be identification of critical branches and adjusting their line loadings, while avoiding violation of the line [illegible] the buses.

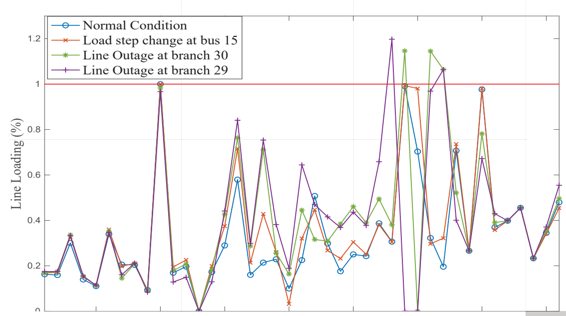


Fig. 2. Cascading line outage in IEEE 30 bus system.

Also, the capacity of the resources, in this case DERs, must be taken in consideration. Thus, modifications to optimal generator set points within specific limits of generation costs are primarily control actions, which when applied to the plant (30-bus system) will minimize a value function (branch loadings of critical branches). The number of control actions for each generator is two (active power and voltage set points), and five generators (excluding the slack) will have to be controlled in this system. Thus, there are 10 factorial combinations for each set of actions and there can be as many such sets as possible in the search for optimal solution. The dimensionality of the problem makes it difficult to be solved analytically in real time, where prompt decisions have to be made. Hence, a reinforcement learning based technique is more suitable for this situation. In this paper, the DHP algorithm has been chosen as means of solution for its ability to quickly and accurately solve the optimization problem, while evaluating the actions not just at current time step but also at future time steps. Hence, this algorithm provides a flexibility to analyze a system at both *N-1* and *N-2* contingencies.

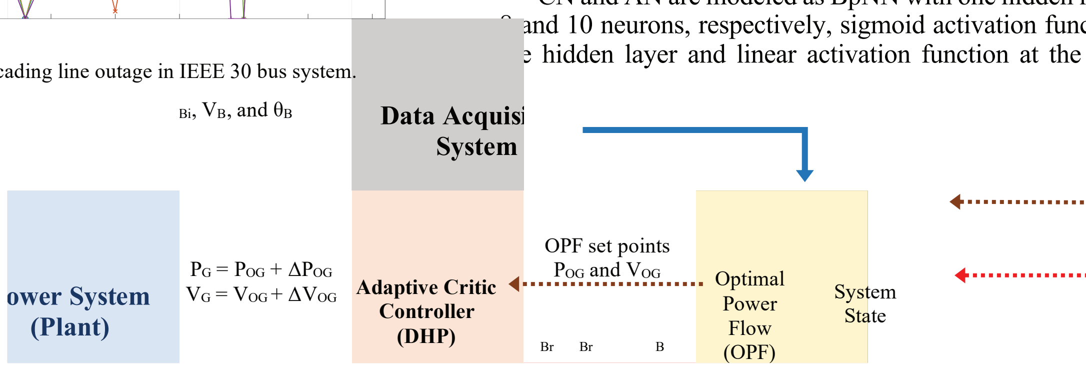
The flowchart for incorporating preventive actions is shown in Fig. 3, where P_{Br} and Q_{Br} are the active and reactive power flows in the branches, respectively; V_B and θ_B are the bus voltage and angle; P_{OG} and V_{OG} are the generator active power and voltage set points from OPF; P_G and V_G are the active power and voltage set points after modification by DHP.

The formulation of DHP for solving any optimization problem requires specific design of certain parameters that are critical to the fidelity of the solution. These parameters are: ratio between learning coefficients of Critic and Actor, discount factor in J, utility function, scaling factor, and reset criteria.

1) Design of Model, Critic and Actor Networks

The MN is designed as a Cascaded Feedforward Neural Network (CFNN) and has been trained offline so that it can generate accurate apparent power flow in the lines, based on active power and voltage set points for the five generators as inputs. This CFNN consists of one hidden layer with twenty neurons, a hyperbolic-tangent activation function for hidden layer and a linear activation function for the output layer. The stochastic design and offline training of the Model eliminates the requirement of performing load flow multiple times, thereby saving immense computation time. The state vector, which consists of apparent power flow in the branches, is scaled using the maximum allowable apparent power flow in each branch.

CN and AN are modeled as BpNN with one hidden layer of [illegible] and 10 neurons, respectively, sigmoid activation function at [illegible] hidden layer and linear activation function at the output



layer. The learning of CN has to be faster than that of AN, as the training of the latter depends on information from CN. Accordingly, CN learning rate has been chosen to be 0.01 and that of AN 0.001. The discount factor, γ in (1), which influences the speed of convergence, changes recursively with the iteration in k. It starts with 0.1 for lower values of k, and changes to 1 for higher values of the same.

2) *Design of Utility Function*

The utility block contains contingency screening algorithms that are trained offline. It is desirable for the utility function to be continuous, differentiable and quadratic in shape [7]. The utility function for this paper has been formulated as shown in (6), where U_1, U_2 and U_3 are the cost functions for line loadings, bus voltages and generation cost, respectively, with their corresponding weights $\mathcal{W}_1$ (1.0), $\mathcal{W}_2$ (1.0) and $\mathcal{W}_3$ (0.7). The expressions for U_1, U_2 and U_3 are shown in (7), (8) and (9). In (7), $X_i(t)$ represents the apparent power flow in

$$U\big(X(t), A(t)\big) = \mathcal{W}_1 U_1(t) + \mathcal{W}_2 U_2(t) + \mathcal{W}_3 U_3(t) \tag{6}$$

$$U_1(t) = \begin{cases} \sum_{i=1}^{41} (X_i(t) - \delta_{Ci})^2 \lceil X_i(t) - \delta_{Ci} \rceil, & X_i(t) > 0 \\ \sum_{i=1}^{41} (X_i(t) - \delta_{Ci})^2 \lfloor -X_i(t) + \delta_{Ci} \rfloor, & X_i(t) < 0 \end{cases} \tag{7}$$

$$U_2(t) = \sum_{i=1}^{41} \left(V_{OG,i} - \vartheta\right)^2 \tag{8}$$

$$U_3(t) = (C_{OG} - \varsigma)^2 \tag{9}$$

critical branches. δ_{Ci} represents a threshold value for $X_i(t)$, corresponding to a particular contingency, C. The value of δ_{Ci} can be estimated from the Line Outage Distribution Factor, based on the contingency that one wishes to study. The first term in (7) gives a quadratic shape to the utility function, the second term consists of a ceiling function for positive state vector and a floor function for the negative state vector, where the sign of the state vector is related to the direction of active power flow in the branches. The ceiling and floor functions are critical in reducing the computation of the algorithm as they eliminate branches that are less than the threshold value, thereby making the algorithm scalable. (8) and (9) are the constraints on bus voltage and deviation from optimal generation cost, respectively.

3) *Design of Reset Criteria*

The reset criterion for iteration over k is based on the violation of limits for voltage, line loading and generation cost. During a particular trial in l, the algorithm does not try to optimize the utility function once any of the aforementioned limits are violated. Instead, it starts a new trial in l. The stopping criterion over l has been set to ten iterations, as satisfactory results were obtained within these number of trials.

V. Results & Discussion

The DHP algorithm, as designed above, has been utilized for line outage prevention in the IEEE 30 bus system, by modifying the OPF set points of the DERs. The algorithm has been provided with *a-priori* knowledge for these modifications or 'actions', such that ΔP_{OG} and ΔV_{OG} lie between 1.0-5.0 and 0.01-0.06, respectively. The actions are shown in Figs. 4a and 4b for ΔP_{OG} and ΔV_{OG}, respectively. Although the number of iterations in k start at a higher value and gradually reduce as the iteration in l increases, only three iterations in k are shown here for simplicity. The actions in the three time steps (or iterations in k) change considerably between $k=1$ and $k=2$, for most cases. However, the actions between $k=2$ and $k=3$ are more in consensus with one another. This is because the AN and CN

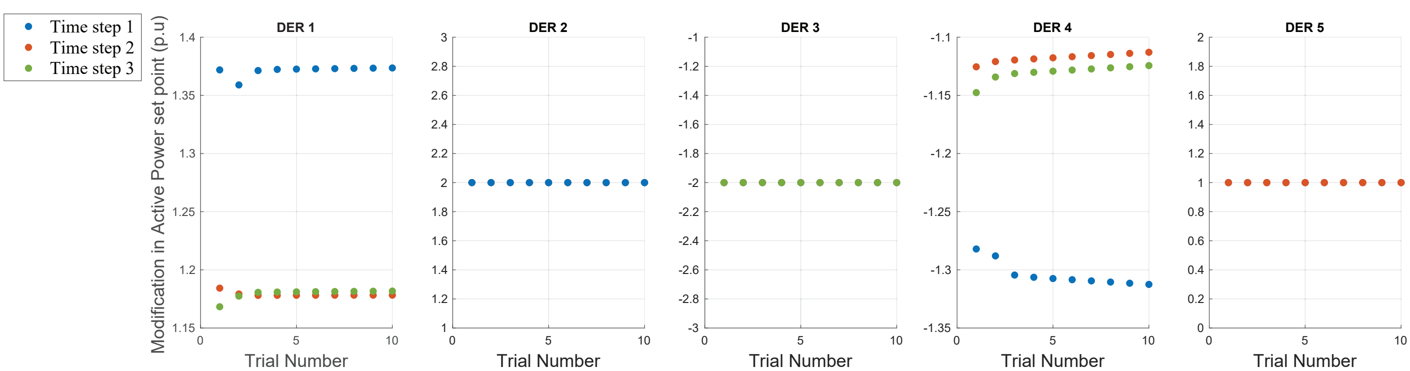


Fig. 4a. Actions for modifying active power set points of the five DERs for $k = 1$ to 3 and $l = 1$ to 10.

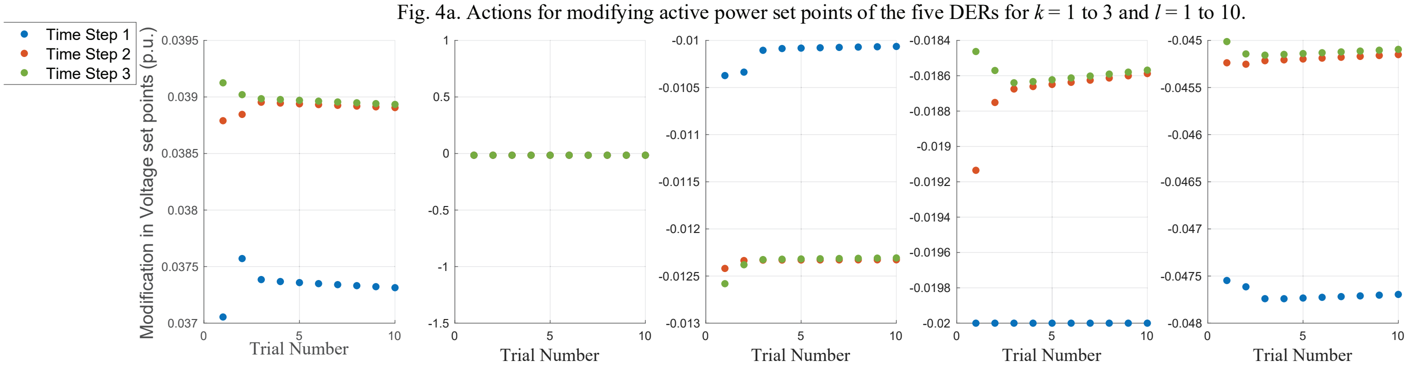


Fig. 4b. Actions for modifying voltage set points of the five DERs for $k = 1$ to 3 and $l = 1$ to 10.

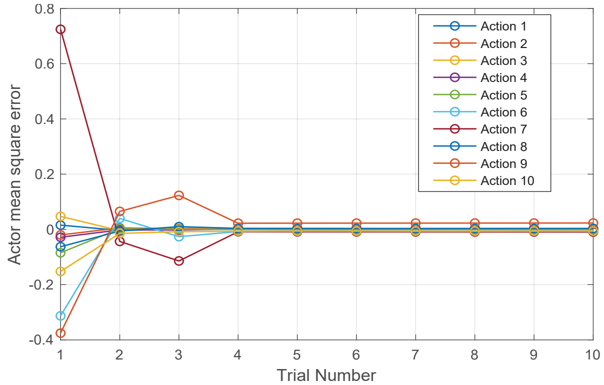


Fig. 5. Mean square error of actor over 10 trials.

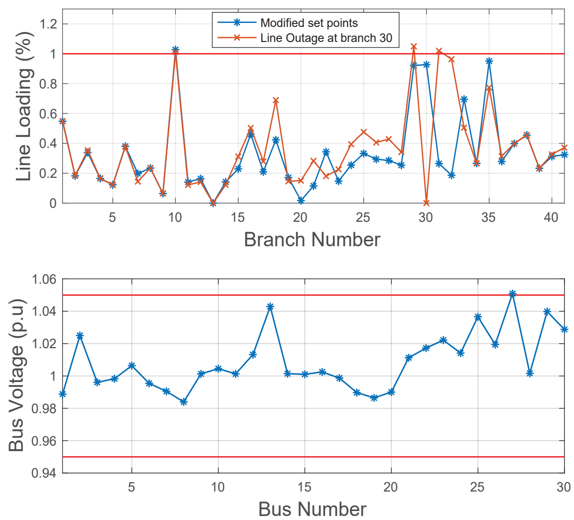


Fig. 6. Line loading and bus voltahe after modifications in set points.

weights for $k = 1$ and $l=1$ are random and hence, the error margin is high. This error is used to train the weights for the next trials in k and l. The convergence of the algorithm is demonstrated through Fig. 5, which shows that the value of the error in (4) is approaching zero. The convergence time of the algorithm is 413.23 seconds that is less than 10 mins. The final actions from the optimization at $k = 3$ and $l=10$ have been applied to the plant and the resulting line loading as well as bus voltages are shown in Fig. 6. A comparison between Figs. 2 and 6 shows that a failure in branch 30 during a step change in the load at bus 15, no longer causes overloading of the other branches. Besides, the bus voltages are within their tolerance limit. Also, the generation cost has changed from \$639.60 to \$642.21 for OPF set points and modified set points, respectively. Thus, the resources in the network have been effectively managed without causing significant deviation from optimal cost or load shedding. However, it should be noted that other critical scenarios might require load shedding capabilities. In that case, a load shedding function must be incorporated in the utility function.

VI. Conclusion

In this paper, a DHP algorithm has been used for taking real-time control decisions in an ADN consisting of the IEEE 30 bus system. The system has been studied for cascading failure that results from operating the resources in the network at their critical limits. The algorithm has successfully optimized the loadings in the branches with the help of proper resource management, by modifying the OPF set points for active power and voltage of the distributed generators. The actions from DHP are potentially preventive measures so that critical contingencies do not culminate in a sequential line outage. Besides, the control decisions have avoided large scale deviation from optimal generation cost within tolerable limits for bus voltages.

References

[1] Rosato, V. et al., "Modelling interdependent infrastructures using interacting dynamical models", *Int. J. Crit. Infrastruct*. 4, 63–79, 2008.

[2] M. Parandehgheibi, E. Modiano and D. Hay, "Mitigating cascading failures in interdependent power grids and communication networks," *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Venice, 2014, pp. 242-247.

[3] C. Chen, W. Wu, B. Zhang, and C. Singh, "An analytical adequacy evaluation method for distribution networks considering protection strategies and distributed generators," *IEEE Trans. Power Del.*, vol. 30, no. 3, pp. 1392–1400, Jun. 2015.

[4] Z. Bie, P. Zhang, G. Li, B. Hua, M. Meehan, and X. Wang, "Reliability evaluation of active distribution systems including microgrids," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2342–2350, Nov. 2012.

[5] H. Farzin,M. Fotuhi-Firuzabad, and M. Moeini-Aghtaie, "Reliability studies of modern distribution systems integrated with renewable generation and parking lots," *IEEE Trans. Sustain. Energy*, vol. 8, no. 1, pp. 431–440, Jan. 2017.

[6] X. Xufeng, J. Mitra, W. Tingting, and M. Longhua, "Evaluation of operational reliability of a microgrid using a short-term outage model," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2238–2247, Sep. 2014.

[7] D. Wang, H. He and D. Liu, Adaptive Critic Nonlinear Robust Control: A Survey, in *IEEE Transactions on Cybernetics*, vol. 47, no. 10, pp. 3429-3451, Oct. 2017.

[8] J. Han, S. Khushalani-Solanki, J. Solanki and J. Liang, "Adaptive Critic Design-Based Dynamic Stochastic Optimal Control Design for a Microgrid With Multiple Renewable Resources," in *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 2694-2703, Nov. 2015.

[9] T. Bian, Y. Jiang, and Z.-P. Jiang, "Decentralized adaptive optimal control of large-scale systems with application to power systems," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2439–2447, Apr. 2015.

[10] Y. Jiang and Z.-P. Jiang, "Robust adaptive dynamic programming for large-scale systems with an application to multimachine power systems," IEEE Trans. Circuits Syst. II, Exp. Briefs, vol. 59, no. 10, pp. 693–697, Oct. 2011.

[11] P. J. Werbos, Neurocontrol and Supervised Learning: An Overview and Evaluation. In *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches*. Van Nostrand Reinhold, 1992, pp. 65-89.

[12] D. V. Prokhorov and D. C. Wunsch, Adaptive critic designs, in *IEEE Transactions on Neural Networks*, vol. 8, no. 5, pp. 997-1007, Sept. 1997.

[13] S. N. Balakrishnan and V. Biega, *Adaptive critic based neural networks for control (low order system applications)*, *Proceedings of 1995 American Control Conference - ACC'95*, Seattle, WA, USA, 1995, pp. 335-339 vol.1.

[14] Si, Jennie. *Handbook of Learning and Approximate Dynamic Programming*. IEEE Press, 2004.

[15] W. Qiao and R. G. Harley, Indirect Adaptive External Neuro-Control for a Series Capacitive Reactance Compensator Based on a Voltage Source PWM Converter in Damping Power Oscillations, in *IEEE Transactions on Industrial Electronics*, vol. 54, no. 1, pp. 77-85, Feb. 2007.