

Received September 1, 2019, accepted October 2, 2019, date of publication October 8, 2019, date of current version October 18, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2946282

# A Reinforcement Learning Approach to Solve Service Restoration and Load Management Simultaneously for Distribution Networks

LUCAS ROBERTO FERREIRA<sup>1,2</sup>, ALEXANDRE RASI AOKI<sup>1</sup>, (Senior Member, IEEE),  
AND GERMANO LAMBERT-TORRES<sup>3</sup>, (Fellow, IEEE)

<sup>1</sup>Electrical Engineering Department, Federal University of Parana, Curitiba 80060-000, Brazil

<sup>2</sup>Siemens Ltda, Curitiba 80215182, Brazil

<sup>3</sup>Research and Development Department, Gnarus Institute, Itajuba 37500-052, Brazil

Corresponding author: Germano Lambert-Torres (germanoltorres@gmail.com)

**ABSTRACT** Energy and economy are increasing the relationship over the years, where the energy becomes a significant resource to keep a country developing, and it supports its economy. Then, more reliable the energy should become, especially the distribution network, to keep the entire process running. In this level of energy distribution, where residential consumers and medium and small industries are supplied, the number of interconnections of the network is enormous. However, for economic and environmental aspects, these complex systems, which are operating close to their capacity, needs to increase the automation, appearing the concept of smart grids and the Advanced Distribution Management System (ADMS) and its methods to control. Inside of the ADMS, there are a lot of essential techniques. Among them, there are two techniques which are the most relevant for this paper: the self-healing and load management. In an ADMS system, these two techniques are treated separately, but the best solution occurs when they are computed together. In this paper, it is proposed an approach that can address both problems at the same time or individually, i.e., in place to have a sequential method to solve step-by-step the issues in the networks. The proposed algorithm, through reinforcement learning technique, can handle both problems together. The proposed approach is tested in a real urban distribution network with some created scenarios to compare the results with outages and overloads. Some comparisons with other methods are carried out.

**INDEX TERMS** Computational intelligence, load management, q-learning, power distribution, power operation, reinforcement learning, self-healing.

## I. INTRODUCTION

The energy is an essential resource in human life that affects directly in society, transports, security, life support, and all these points together, having a significant economic impact in a nation. Since years ago, researches tried to count how much the cost impact of an interruption is for the consumers, industrial, commercial, and residential [1]–[3]. With time, the energy is growing in importance, and nowadays, how much more power can be distributed, more development occurs, factories can keep their production, the commerce can maintain the trading, new ideas can emerge from the house garages, and many other uncountable good things can happen.

The associate editor coordinating the review of this manuscript and approving it for publication was Biju Issac<sup>1</sup>.

High quality and reliability in all chains should be kept to sustain energy distribution. This paper presents a contribution in this direction for distribution networks, increasing the operation quality. These networks have become complex throughout the last decades, due to the number of devices installed, the number of consumers, the environmental laws, new regulations, aging of the equipment, among others [4]. These points bring so many possibilities to solve an issue that the operator might take the wrong decision or take more time to resolve a problem. Thus, a lot of studies have been made to create autonomous systems, the major part of them with a certain level of intelligence, to minimize the error from the operator action, giving more instructions in a short period of time to his decision making in a viable time before that a blackout happens [5], [6].

One possible investment to improve the control is the Advanced Distribution Management System (ADMS), which is a software platform that includes many functionalities to guide the distribution system to be more resilient, in other words, to have the capacity to recover from any disaster, be reliable and efficient in its operations [7]. A complete ADMS is composed of algorithms for automated fault location, fault isolation, service restoration, conservation voltage reduction, peak demand management, volt/var control optimization, microgrid operation, and electric vehicle support [8], [9].

For each one of these features, there is a unique algorithm that will run independently when the correspondent trigger occurs. In general, it is not possible to reach an optimal global solution, resolving only a unique systemic problem; because the algorithm tries to find the best case to solve an isolated and unique issue. For instance, in the self-healing problem, there are some articles showing techniques to solve it [10], [11], by centralized, decentralized, or distributed systems, using graph theory or artificial intelligence; but, in all cases, they resolve an issue per time. Furthermore, other ADMS functionalities can have independent techniques for Volt/Var [12], microgrid [13], load management [14], and so on. Beyond the self-healing problem, there is another important trigger of outages, which is the overload, as pointed out in [3], so in this case, there are techniques to mitigate the issues related to the sudden growth of load in the system as showing in [15]–[17]. A combination of self-healing and load shedding can be seen in [18], [19], where more than one problem occurs to be resolved using the same technique, but it should address one single trouble first and after another.

Thus, this current paper brings a new approach to reconfigure the system, from two different triggers (fault and overload). Due to the complexity of this problem, it is hard to find a single strategy that can handle different approaches at the same time; the idea is to use the Machine Learning (ML) methods [20]. There are a lot of techniques that can work in complex environments, with a massive amount of data and in real-time, once the action to avoid permanent fault requires limited time to work. Some works can be seen in this area applied in the electrical power system [21]–[23].

The initial inspiration of this paper was based on these two papers [24], [25], which use Reinforcement Learning (RL) on Self-Healing for shipboard reconfiguration. In this paper, the approach is to use the core of RL, including environment, policy, reward, and penalty, changing points of convergence. The structure of this paper begins with the motivation of the proposed method. After, it is explained how the reality of distribution networks inside of computer program was modeled. In the sequence, a brief explanation about reinforcement learning is presented, showing the proposed approach. Following that, the application of the proposed method occurs in a small system to explain how the strategy works. And then, the proposed method is applied in a real distribution network. Some results for different scenarios and operating

points are carried out, and some comparisons with other sequential methods are made.

## II. MOTIVATION

Currently, there are two processes to resolve the faults into the distribution grid. The first tries a simple service restoration where if a source is available to be used, a set of loads might be transferred to this feeder, and the process ends up at this moment. The second process considers the ADMS functionalities designed to resolve each problem per time, i.e., in a sequential way. For example, in a case where the self-healing solution does not have an alternative source to transfer the load without causing an overload, the service restoration will not be executed *a priori*. However, a second functionality might be triggered to avoid this overload. So, load shedding can cut off part of the load, before the self-healing actions. Then, self-healing is enabled after load shedding.

This second process is more complex, and it will require operational experience or automatic triggering of functionalities to develop it. So, the motivation of this paper is based on the idea of solving complex problems by several functionalities simultaneously. The problem in this approach is that the search space is increased exponentially, and simple techniques cannot find a solution or resolve in a short time. Then, the contribution of this paper is to call on Reinforcement Learning, which is prepared to handle complexity environments.

## III. THE ALGORITHM PROPOSED: MODIFIED Q-LEARNING AND THE ENVIRONMENT MODEL

The proposed method can be classified as Machine Learning (ML) techniques, which is a branch of Computational Intelligence field that treats with algorithms, mathematical, and statistical models. More specifically, the proposed method is a class (paradigm) of ML, named Reinforcement Learning (RL), in which the concept of the model is substituted by multiple agents (which are model-free). Each agent manages its interaction with the environment, and its actions trying to maximize the notion of cumulative reward.

The proposed method uses a Modified Q-Learning (MQL) adjusted to the power distribution environment. The original Q-Learning algorithm, where Q represents quality, which is also a class of RL, was proposed in [26], [27], that can't be used directly for the size of the real problem in distribution systems. The main goal of each agent of the MQL is to learn the operation policy of the distribution network, its issues, and constraints, the current operating point, and to act. Each action can produce changes in the current operating point of the distribution system, generating a reward (positive or negative) for the agent and modifying its internal behavior. The method proposed, in Figure 1, follows the self-healing workflow (fault location, isolation, and service restoration). The first step of the algorithm is to identify and isolate the fault, saving the switches that were isolated (list of actions). After, it is done the setup to run the training of

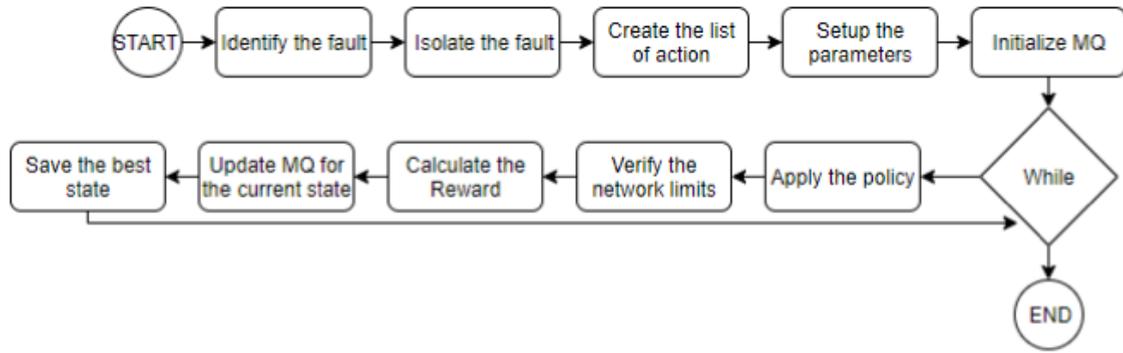


FIGURE 1. Proposed algorithm workflow.

MQL, so the method is initialized, and the algorithm can start to find the best actions to be executed and resolve the problem. Inside the loop, the greed policy, represented by equation (1), is used to determine the current state based on the values in MQ-matrix. Note that for each new state selected, the limits to avoid any new failure in the system are verified, as permanent parallelism, current, and voltage constraints. The reward calculation updates the MQ matrix values, and it determines if the current state is valid or not to be used, in case of the right solution, it is saved to be used:

$$\pi'(s) = \operatorname{argmax}(MQ^\pi(s, a)) \quad (1)$$

where  $\pi$  is the policy applied,  $s$  is the current state,  $s \in S$ ,  $Q$  is the value function for a pair variable and  $a$  is the current action,  $a \in A$ . The equation (1) is indirectly related with the equation (3), where this equation selects the next state to be verified by the algorithm.

Below is presented the relation between the MQL algorithm and the distribution system to process and resolve the problem.

- Agent: switch or load;
- Environment: Distribution System;
- Action: Open/Close switch, or Load Shedding [%];
- State: System topology;
- Reward: delta load, equation (2);
- MQ-matrix: Rows – the index represents the state in decimal that should be converted in binary, Columns – the actions to be performed, each index represents a possible device to commute or a perceptual of the shed, Figure 2 shows the general idea for MQ-matrix. Furthermore, to have a better comprehension, see the next section. The values are calculated based on equation (3). To initialize the matrix, it is used the load value for each topology configuration according to the index.

### A. LEARNING PROCESSING

The first state of the system is the device positions after the isolation, so the algorithm finds the best action based on the greed policy in the MQ-matrix. And then, a new action changes the state, if the selection is related to a device, or changes the load, whether a shedding action.

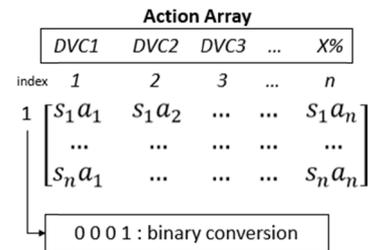


FIGURE 2. Understanding the MQ-matrix.

To determine the new state, topology, and loading, three constraints to avoid any new failure in the system are verified. The first constraint is the permanent parallelism, where it is executed a graph search for each feeder, and the same point (device) cannot appear twice. The second verification is the voltage limits, which is used the own software (similar to OpenDSS software [28]) to calculate the power flow and obtain the voltage profile. The last constraint is the maximum capacity of current on the equipment. In this case, the overall capacity for all devices and branches of the distribution system is used. If one of these constraints is exceeded the reward for the new state is minus one. Notice, all values are normalized, considering the total power calculated when the system is in normal operation mode.

If all constraints are right, in the step “verify the network limits” in Figure 1, the reward can be calculated according to equation (2), which represents the variation between the load from the previous state and the load for the current state. The values can be positive, that represents an improvement, or negative, where the new state has less load than the previous, this load is calculated based on the topology system from the current state ( $s$ ) and the next state ( $s'$ ) after the action ( $a$ ) defined by the policy in equation (1). Otherwise, the reward should receive -1 as a negative choice for that action.

$$R(s, a) = \Delta L = L(s') - L(s) \quad (2)$$

where  $R$  is the reward function,  $\Delta L$  load variation,  $L$  is the total load of the system based on state  $s$ ,  $s'$  is the future state when taking the action  $a$ ,  $s' \in S$ .

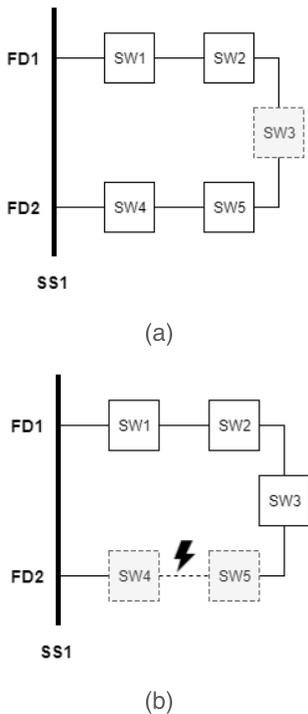


FIGURE 3. Example system with five reclosers. (a) the normal system, and (b) system after isolation.

Therefore, with a reward value for the current state, the MQ value can be obtained based on the equation (3), where the first term means how much the value function in position  $(s, a)$  will be decreased based on the learning factor  $\alpha$ , and after, in the second term, the value function is increased based on the reward and the maximum value for the next better action  $(a')$  from the next start  $(s')$ , so the result means if this position will be interesting for future states. Notice that the values for  $MQ$  are based on the load.

$$MQ(s, a) := (1 - \alpha)MQ(s, a) + \alpha[R(s, a) + \gamma \max_{a'} MQ(s', a')] \quad (3)$$

where  $\alpha$  is the learning rate,  $0 < \alpha \leq 1$ ,  $\gamma$  is the discount factor  $0 \leq \gamma < 1$ ,  $a'$  is the future action from  $s'$ . Both  $\alpha$  and  $\gamma$  are configured in the “Setup the parameters.”

To determine the best solution, the algorithm saves, in the end, the best state when there is no constraint violation. The algorithm stops when the number of iterations reaches the end, or there isn't more variation for the policy choice in 10 consecutive iterations.

### B. UNDERSTANDING THE MODEL

To have a better comprehension of how the method is executed; it is created a small hypothetical system to apply in a few steps, the complete idea of the algorithm. Figure 3 shows this system in a normal state (a) and in an isolated and restored state (b). The equation (3) is the initialization for matrix MQ after the isolation; notice that there are three possible simple

TABLE 1. The relation between MQ-matrix and system topology.

Q-matrix index	SW1 position	SW2 position	SW3 position
1	0	0	0
2	1	0	0
3	0	1	0
4	1	1	0
5	0	0	1
6	1	0	1
7	0	1	1
8	1	1	1



FIGURE 4. Real distribution network used for test of the proposed method.

actions, which are: to close or open switches SW1, SW2, and SW3, once SW4 and SW5 are out of scope. The initial stage in the algorithm, which is represented in Figure 3 (b), comprehends line 4 according to Table 1.

The first step of the algorithm is to determine the action using the greedy policy, so the maximum value is in line 4, inside of the matrix MQ, is 1 (60/60). This action is to close SW3, and this carries for a new state where all devices are closed and represent line 8 in Table 1.

$$Q_{initial} = \begin{bmatrix} 10/60 & 0 & 0 \\ -10/60 & 30/60 & 0 \\ 30/60 & 0 & 0 \\ -30/60 & -20/60 & 60/60 \\ 10/60 & 0 & 0 \\ -10/60 & 60/60 & 0 \\ 60/60 & 0 & 0 \\ -60/60 & -50/60 & -30/60 \end{bmatrix} \quad (4)$$

Once reached a new state, it should be verified all constraints, and, if any violation is reached, the reward receives a negative value, so the next step is to calculate the new value for MQ based on equation (3). The MQ-matrix is only

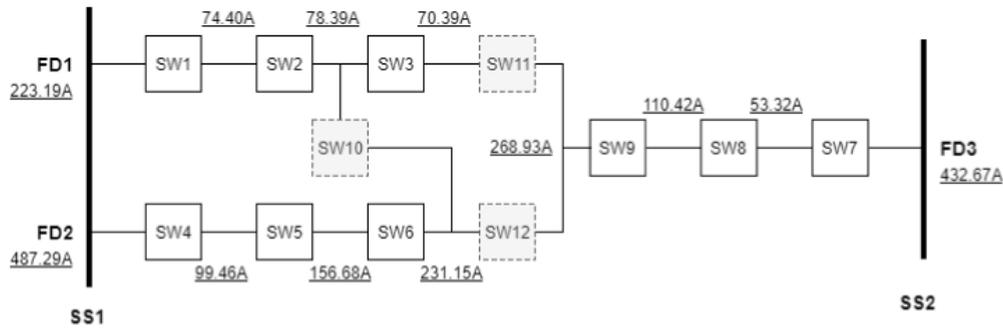


FIGURE 5. Distribution network example.

updated if the current state comprehends a better result from the previous iteration, the state is saved from being used as a solution at the end of the process. A new iteration starts with the new state (all devices closed), and a new action should be chosen again, in this case, the best solution is to open SW3, but the result return in the initial state, so there isn't a better solution than all devices closed, in this case, the algorithm can stop and show the best solution stored in the process.

IV. TEST ON A REAL DISTRIBUTION NETWORK

To illustrate the algorithm design, it is tested in a real urban grid with two substations interconnected by three feeders and disposed of 12 switches, where three of them are normally open and interconnects the three feeders. Figure 4 shows an aerial photo of the area supplied for this network. A snapshot of the normal topology and the current load, in Ampere, for each segment is demonstrated in Figure 5. Of course, in this figure, only the main devices are included to facilitate the representation. Table 2 shows some data from this network. The total number of transformers is 283; it means, around more than 5,000 consumers have their energy supplied by this part of the network. The number of branches is 79, with a total extension of about 18 km and the supplied power of approximately 15 MVA.

TABLE 2. Data of the real distribution network.

Name	Real Name	Total of Transformers	Number of Branches	Distance (km)	Total Power (MVA)
FD1	Reserban	57	15	3,76	2,9
FD2	Boulevard	118	30	6,65	6,33
FD3	Frambar	108	34	7,96	5,62

In Brazil, the distribution network has not yet a high level of automatism compared with North America and Europe. However, there is a specificity the differs Brazil with other countries, as the bigger amount of load concentrated and distance between the remote commanded devices. These points

are the reasons to increase the complexity to determine the best configuration. Furthermore, as the idea of this paper is to begin the discussion between the traditional way, where each problem is addressed to one technique, and the new approach, the size of the grid does not interfere with demonstrating the real purpose.

The limits for capacity and voltage used to simulate the scenarios were:

$$0.93 \leq V(i_b) \leq 1.05$$

$$I(sw) \geq cap$$

where  $V$  is the voltage calculated for each bus,  $i_b$  is the current bus,  $sw$  is the index for each switch,  $I$  is the current in each switch,  $cap$  is the capacity for the whole system.

Furthermore, the permanent parallelism between feeders is also checked to avoid any serious damage to the equipment. For the following tests, the maximum number of iterations is 2000, and the values for learning rate,  $\alpha$ , and discount factor,  $\gamma$ , are respectively 0.5 and 0.3. These values were defined after some initial tests and demonstrated better results than the other combinations.

The next subsections present two self-healing analyses with two different values for capacities for the system and one case when the system gets an overload.

A. SIMPLE FAULT ON FEEDER 1

In the first test case, the whole system has 500A of capacity, and a fault between the switches SW1 and SW2 in FD1 is simulated, where it is de-energized a total load in front of SW2 of 1.93MVA (148.79A). The possible solution is shown in Figure 6, where the SW11 is closed to transfer the loads to FD3, but as the resource feeder cannot sustain 148.79A, the shedding of 15% should be done to cut off 87.21A in FD1 and FD3. The option via FD2 is not possible, once the availability is just of 12.71A, far away from an ideal scenario to be considered in the analysis because should be considered a considerable shed to transfer the load that unfeasible the solution.

A second analysis can be done when the fault continuous the same, but the maximum capacity now is 560A,

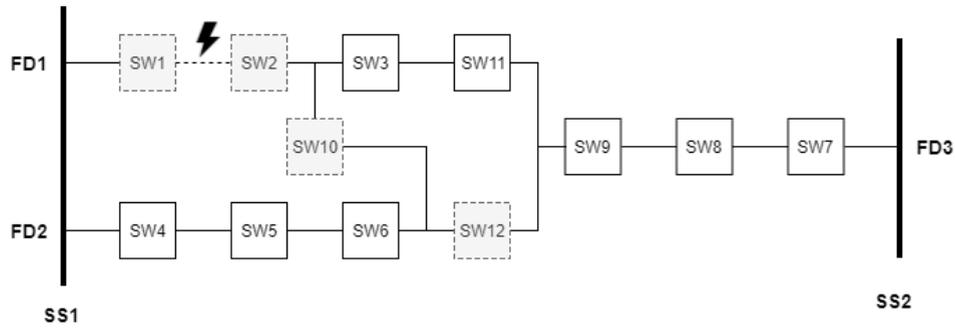


FIGURE 6. Case 1: capacity 500A, final topology.

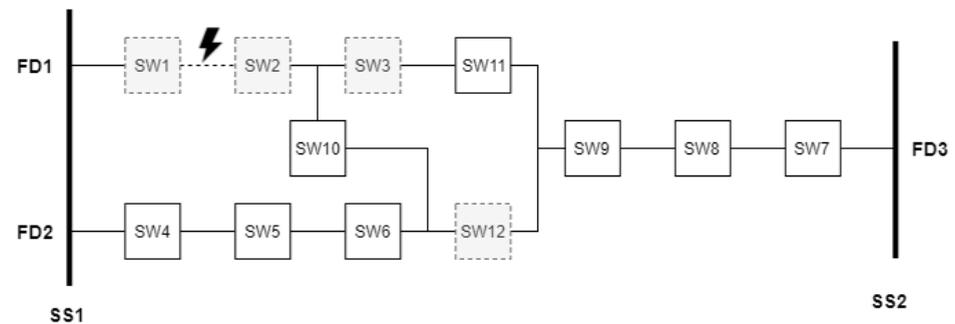


FIGURE 7. Case 1: capacity 560A, final topology.

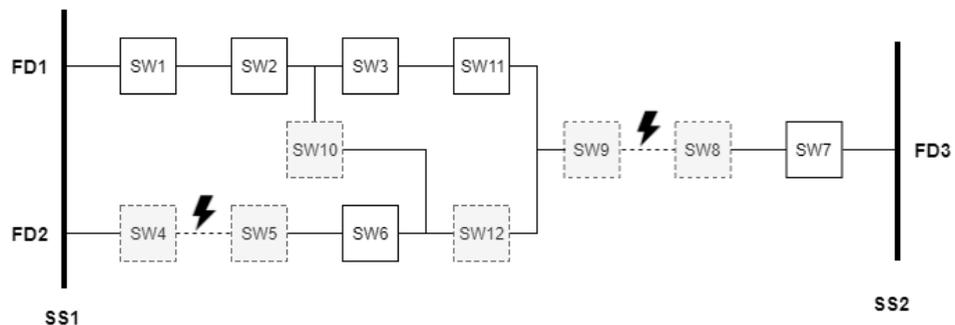


FIGURE 8. Case 2: capacity 500A, final topology.

according to Figure 7, the algorithm can split the blocks in FD1 to be transferred one by FD2 via SW10 and the another to FD3 via SW11, and a small shed of 5% is executed on FD2 to alleviate 28.28A. When the capacity was 500A, the total of power was 12.76MVA, less than 13.52MVA, when the capacity was 560A. In this case, the algorithm could determine that the better solution was to open the SW3 and transfer the FD1-block1 to FD2 and FD1-block3 to FD3 in place of just commuting the normally open switch.

### B. SIMULTANEOUS FAULT

In this test case, two simultaneous faults are simulated. The first fault is between the switches SW4 and SW5, where

the de-energized segment downstream of SW5 has 5.04MVA (387.83A), and the second fault is between SW8 and SW9, with 3.49MVA (268.93A) de-energized. The reconfiguration when the capacity is 500A is demonstrated in Figure 8, where SW11 closes to absorb the FD3-block3, the load limit in FD1 is 276.81A, and the block transferred has 268.93A, so FD2 stays completely de-energized because FD1 is in maximum capacity.

Now, the second case is simulated with a capacity equal to 560A. The final configuration, demonstrated in Figure 9, tries to restore the maximum load from FD2, but a shed of 10% should be done to mitigate 61.10A, and more load should be transferred instead of just the FD3-block3, the final result increases the load in 11%.

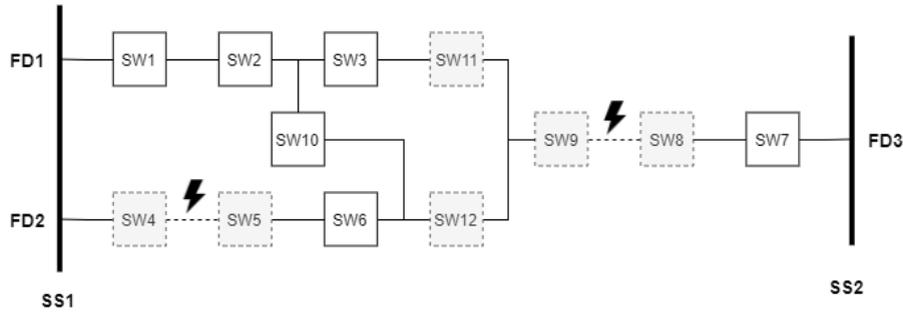


FIGURE 9. Case 2: capacity 560A, final topology.

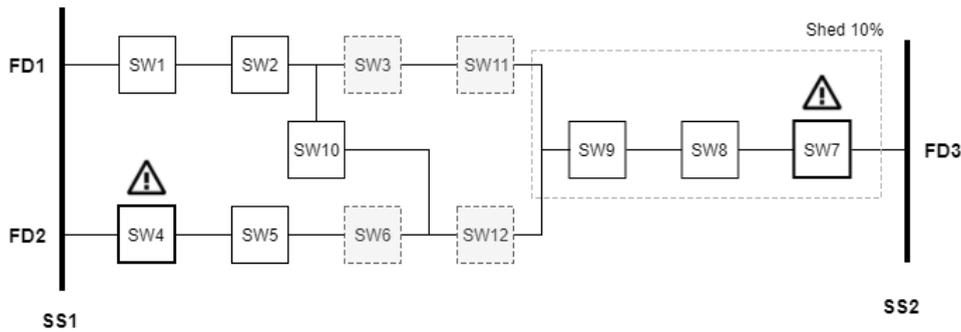


FIGURE 10. Case 3: capacity 400A, final topology.

TABLE 3. Results of the presented cases.

Case	Functionality	Capacity [A]	Trigger	Switches changed	Best Shed [%]	Feeder Selected to Shed
1	Self-Healing	500	SW1	SW11 - CLOSED	15	3
	Self-Healing	560	SW1	SW3 - OPENED SW10 - CLOSED SW11 - CLOSED	5	2
2	Self-Healing	500	SW4 and SW8	SW11 - CLOSED	0	0
	Self-Healing	560	SW4 and SW8	SW10 - CLOSED	10	1
3	Load Management	400	SW4 and SW7	SW3 - OPENED SW6 - OPENED SW10 - CLOSED	10	3

C. OVERLOAD TRIGGER

When the capacity is set to 400A in the entire system, it is triggered an overload on FD2 and FD3, shown in Figure 10, which has respectively 22% and 8% more than the actual capacity. In this case, the algorithm will process the best reconfiguration and shedding to keep the maximum load possible and inside of the constraints (voltage and current). The reconfiguration is done in FD1 and FD2, where a whole block in FD1 should be shed via SW3 to transfer the entire block 3 of FD2 through SW10. The FD1 has 176.81A available, and the FD2 block3 has 231.15A, so when cutting off the last block in FD1, the availability goes to 247.20A,

enough to shift the feeder. The second overload is resolved with a simple shedding of 10% on the customer that represents 4.26MVA of saved load on FD3, and it is enough to keep the system stable. Without this method, the final topology should open SW6 and SW9 to alleviate 25% of the system; instead of the 10% obtained by this method, it was saved 171.47A.

V. DISCUSSION ABOUT THE RESULTS AND COMPARISON WITH OTHER METHODS

In this article, it is shown five different cases using the proposed MQ reinforcement learning methodology applied

TABLE 4. Results of other methods for comparison.

Case	Functionality	Capacity [A]	Trigger	Switches changed	Best Shed [%]	Feeder Selected to Shed
<b>BPSO + FPO Method</b>						
1	Self-Healing	500	SW1	SW10 - CLOSED	23	2
2	Self-Healing	500	SW4 and SW8	SW11 - CLOSED SW12 - CLOSED	47	1
<b>Distributed Method</b>						
1	Self-Healing	500	SW1	SW11 - CLOSED	14	3
2	Self-Healing	500	SW4 and SW8	SW10 - CLOSED SW11 - CLOSED	47	1

to resolve the self-healing and load management simultaneously. Table 3 resumes all results for comparisons. Notice that according to the problem, the functionality triggered was different, but the technique used was the same.

Another critical point to highlight is the use of the load shedding included in the self-healing approach, the cases 1 (500 and 560A), and 2 (560A) have the reconfiguration after the fault and also the load shedding to avoid a new problem after the reconfiguration and the final solution could bring more load restored when compared with a complete shed.

Table 4 shows the results of two other methods: one centralized and another distributed. The strategy of the first one is based on a Binary Particle Swarm Optimization (BPSO) for switching and an Optimum Power Flow (OPF) for load shedding [19]. The second one, the decentralized solution, each agent, installed in the switches, has operational rules-of-thumb [10]. Verifying the best solutions and comparing the results, the following conclusions can be taken:

Case 1: MQL and Distributed methods cut less load than BPSO+OPF method, 15% and 14% against 23% of load shedding, respectively; and

Case 2: MQL has less load than BPSO+OPF and Distributed methods, 0% against 47% and 47%, respectively.

TABLE 5. Results of other methods for comparison.

Case	Training Time [s]	Learning Time [s]	MQ size
1-500A	15.6	4.3	1024x25
560A		6.2	
2-500A	2.8	17.0	256x21
560A		17.9	
3	111.2	172.0	4096x29

The training and learning time can be seen in Table 5 for each case. It is also compared with the MQ dimension.

The training and learning time are directly related to the MQ size, but when the system has more availability, the process increases in some seconds because there are more possibilities to go through the MQ matrix.

According to Table 5, in cases where the number of reclosers increases to more than 20, the search space of MQ matrix falls in the concept known as, curse of dimensionality. To avoid this problem, the idea of modified Q-learning algorithm shown in this paper is from the fault location, select the feeders nearby the unhealthy feeder to consider in the solution, so it is not necessary to look the complete distribution system, but just the relevant border, which can resolve the majority of the cases of fault.

Another important aspect of the result of MQL is related to the number of switching to restore the service during a simultaneous fault. It can be observed that MQL demands only one switching while other methods require two switchings.

## VI. CONCLUSION

The reinforcement learning method modeled, to comprehend the electrical environment based on a reward that represents the load variation, represents satisfactory results, better than the operator can handle alone or with sequential programs. The proposed Modified Q-Learning shows that it is possible to execute different actions (switching and shedding) from an initial problem, as overcurrent and overload. The greedy policy was enough; in this case, to help the algorithm to choose the best action in the next iteration.

The novelty of this article is to treat more than one problem in a single technique, simultaneously and not sequentially. Instead of having multiple algorithms to solve the various issues of distribution networks, a centralized method is proposed to solve the problems at the moment of fault and overload. The base of the article comes of learning by reinforcement learning, in particular, Q-learning, which was modified, generating the Modified Q-learning method, producing a simultaneous solution of the issues for the distribution network problems. Furthermore, the learning matrix (matrix MQ), in addition to the modeling of the positions of the keys is also included the percentage of load shedding to

be selected for the final system back to a normal operating state.

As demonstrated in the results, with an application in a real network, depending on the load capacity in the system, the algorithm could determine which are the better sequences of switching to execute. Moreover, in all cases, the limits were respected, keeping the grid safe and reliable to avoid any new problem caused by the switching. The proposed MQL method was also compared with the other two approaches (centralized and distributed), which deal with self-healing and load shedding sequentially, and the results were better.

A disadvantage of this method is that the utility should have complete information for the grid to run the power flow. For further works should be enhanced the process to analyze the MQ matrix (the policy), in order to direct the searches for the best solution in less time. This enhancement might help to speed up the solution for more interconnected distribution networks. The proposed method takes into account only two functionalities, so when increasing more applications, some other enhancement should be done.

## ACKNOWLEDGMENT

The authors would like to thank the National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES), and the Brazilian Electricity Regulatory Agency (ANEEL R&D) for supporting this project.

## REFERENCES

- [1] G. Wacker and R. Billinton, "Customer cost of electric service interruptions," *Proc. IEEE*, vol. 77, no. 6, pp. 919–930, Jun. 1989.
- [2] M. J. Sullivan, T. Vardell, and M. Johnson, "Power interruption costs to industrial and commercial consumers of electricity," *IEEE Trans. Ind. Appl.*, vol. 33, no. 6, pp. 1448–1458, Nov. 1997.
- [3] K. Yamashita, S.-K. Joo, J. Li, P. Zhang, and C.-C. Liu, "Analysis, control, and economic impact assessment of major blackout events," *Eur. Trans. Elect. Power*, vol. 18, no. 8, pp. 854–871, 2008.
- [4] J. A. Momoh, *Smart Grid: Fundamentals of Design and Analysis*. Piscataway, NJ, USA: IEEE Computer Society Press, 2012.
- [5] Z. Ma, F. Liu, C. Shen, Z. Wang, and S. Mei, "Fast searching strategy for critical cascading paths toward blackouts," *IEEE Access*, vol. 6, pp. 36874–36886, 2018.
- [6] F. Liu, J. Guo, X. Zhang, Y. Hou, and S. Mei, "Mitigating the risk of cascading blackouts: A data inference based maintenance method," *IEEE Access*, vol. 6, pp. 39197–39207, 2018.
- [7] D. E. McNair, D. J. Phelan, and L. Coleman, "Voices of experience: Insights into advanced distribution management systems," US Dept. Energy, Washington, DC, USA, Tech. Rep. DOE/GO-OT-6A42-63689, 2015.
- [8] J.-W. Kang, L. Xie, and D.-H. Choi, "Impact of data quality in home energy management system on distribution system state estimation," *IEEE Access*, vol. 6, pp. 11024–11037, 2018.
- [9] K. Mahmud, M. J. Hossain, and G. E. Town, "Peak-load reduction by coordinated response of photovoltaics, battery storage, and electric vehicles," *IEEE Access*, vol. 6, pp. 29353–29365, 2018.
- [10] B. S. Torres, L. R. Ferreira, and A. R. Aoki, "Distributed intelligent system for self-healing in smart grids," *IEEE Trans. Power Del.*, vol. 33, no. 5, pp. 2394–2403, Oct. 2018.
- [11] T. D. Sudhakar and K. N. Srinivas, "Restoration of power network—A bibliographic survey," *Eur. Trans. Elect. Power*, vol. 21, no. 1, pp. 635–655, 2011.
- [12] D. Salles, A. C. Pinto, and W. Freitas, "Integrated volt/Var control in modern distribution power systems based on support vector machines," *Int. Trans. Elect. Energy Syst.*, vol. 26, no. 10, pp. 2216–2229, 2016.
- [13] F. O. Resende, N. J. Gil, and J. A. P. Lopes, "Service restoration on distribution systems using multi-microgrids," *Eur. Trans. Elect. Power*, vol. 21, no. 2, pp. 1327–1342, 2011.
- [14] W. Zhang, Y. Xu, S. Li, M. Zhou, W. Liu, and Y. Xu, "A distributed dynamic programming-based solution for load management in smart grids," *IEEE Syst. J.*, vol. 12, no. 1, pp. 402–413, Mar. 2018.
- [15] H. Mortaji, S. H. Ow, M. Moghavvemi, and H. A. F. Almurib, "Load shedding and smart-direct load control using Internet of Things in smart grid demand response management," *IEEE Trans. Ind. Appl.*, vol. 53, no. 6, pp. 5155–5163, Nov./Dec. 2017.
- [16] Q. Xu, Y. Ding, Q. Yan, A. Zheng, and P. Du, "Day-ahead load peak shedding/shifting scheme based on potential load values utilization: Theory and practice of policy-driven demand response in China," *IEEE Access*, vol. 5, pp. 22892–22901, 2017.
- [17] T. S. P. Fernandes, J. R. Lenzi, and M. A. Mikilita, "Load shedding strategies using optimal load flow with relaxation of restrictions," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 712–718, May 2008.
- [18] C.-H. Lin, C.-S. Chen, T.-T. Ku, C.-T. Tsai, and C.-Y. Ho, "A multiagent-based distribution automation system for service restoration of fault contingencies," *Eur. Trans. Elect. Power*, vol. 21, no. 1, pp. 239–253, 2011.
- [19] L. R. Ferreira, L. C. Siebert, A. R. Aoki, and T. S. P. Fernandes, "Load shedding through optimal power flow to support self-healing actions in distribution feeders," in *Proc. IEEE PES Transmiss. Distrib. Conf. Expo.-Latin Amer. (PES T&D-LA)*, Medellin, Colombia, Sep. 2014, pp. 1–6.
- [20] L. Cheng and T. Yu, "A new generation of AI: A review and perspective on machine learning technologies applied to smart energy and electric power systems," *Int. J. Energy Res.*, vol. 43, no. 6, pp. 1928–1973, 2019.
- [21] Y. Zhao, R. Ball, J. Mosesian, J. F. D. Palma, and B. Lehman, "Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2848–2858, May 2015.
- [22] H. Livani and C. Y. Evrenosoğlu, "A fault classification and localization method for three-terminal circuits using machine learning," *IEEE Trans. Power Del.*, vol. 28, no. 4, pp. 2282–2290, Oct. 2013.
- [23] X. Chen, Z. Y. Dong, K. Meng, Y. Xu, K. P. Wong, and H. W. Ngan, "Electricity price forecasting with extreme learning machine and bootstrapping," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2055–2062, Nov. 2012.
- [24] S. Pal, S. Bose, S. Das, C. Scoglio, B. Natarajan, and N. Schulz, "Shipboard power system reconfiguration using reinforcement learning," in *Proc. North Amer. Power Symp.*, Arlington, TX, USA, Sep. 2010, pp. 1–7.
- [25] S. Das, S. Bose, S. Pal, N. N. Schulz, C. M. Scoglio, and B. Natarajan, "Dynamic reconfiguration of shipboard power systems using reinforcement learning," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 669–676, May 2013.
- [26] C. J. C. H. Watkins, "Learning from delayed rewards," Ph.D. dissertation, King's College, Univ. Cambridge, Cambridge, U.K., 1989.
- [27] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, nos. 3–4, pp. 279–292, 1992.
- [28] *OpenDSS Software*, EPRI, Palo Alto, CA, USA, 2013.



**LUCAS ROBERTO FERREIRA** received the B.S. degree in automation and control engineering from the Catholic University of Parana, in 2012, and the M.Sc. degree in electrical engineering from the Federal University of Parana, in 2015. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, Federal University of Parana, Curitiba, Brazil. He is a Software Development Engineer with Siemens, Curitiba. His research interests include

the machine learning applied in smart grids, self-healing, load shedding, and power system operation.



**ALEXANDRE RASI AOKI** (M'12–SM'17) received the B.S., M.Sc., and Ph.D. degrees in electrical engineering from Itajuba Federal University (UNIFEI), Itajuba, Brazil, in 1996, 1999, and 2003, respectively. He was a Senior Researcher and a Manager with the Institute of Technology for Development (Lactec Institutes), from 2003 to 2017. He was a Professor with the Development of Technology from the Lactec Institutes, from 2004 to 2017, and a Coordinator,

from 2010 to 2014. Since 2008, he has been a Professor with the Federal University of Paraná, Curitiba, Paraná, Brazil. He serves on a number of committees related to distributed generation and smart grids, including the IEEE and CIGRÉ. His research interests include planning and operation studies for distribution systems and smart grid, including renewable energy sources and storage systems.



**GERMANO LAMBERT-TORRES** (S'87–M'90–SM'13–F'14) received the B.S. and M.Sc. degrees in electrical engineering from Itajuba Federal University (UNIFEI), Itajuba, Brazil, the B.S. degree in economics from the South Minas Gerais Economic and Social Sciences Faculty, Itajuba, the B.S. degree in mathematics from the Itajuba Faculty of Sciences and Letters, Itajuba, and the Ph.D. degree in electrical engineering from the École Polytechnique de Montreal, Montreal, Canada, in 1990.

From 1983 to 2012, he was a Professor with the Electrical Engineering Department, UNIFEI. From 2000 to 2004, he was the Dean of the Research and Graduate Studies, UNIFEI. From 1995 to 1996, he was a Visiting Professor with the University of Waterloo, Waterloo, ON, Canada. He is currently the Director of Research and Development at PS Solutions, Itajuba, Brazil. He also serves as a Consultant for many power industries in Brazil and South America, and he has taught numerous IEEE tutorials in the USA, Europe, and Asia. He has completed more than 80 M.Sc. and Ph.D. thesis supervisions and published more than 600 journal and conference technical articles. He is also the author/editor or coauthor of nine books, more than 30 book chapters, and 50 transactions articles on intelligent systems and nonclassical logic.

Dr. Lambert-Torres is a member of the International Conference on Intelligent System Applications to Power Systems (ISAP) international board. He serves on a number of committees related to intelligent systems, including IEEE and the International Council on Large Electric Systems (CIGRÉ). He was a recipient of several awards, including the Technical Committee Working Group Recognition Award on New Technologies and Practical Applications from the IEEE Power and Energy Society (PES), in 2006, Outstanding Leadership as Member of the ISAP Board of Directors from the International Council of ISAP, in 2007, and Technical Committee Working Group Recognition Award on Multi-Agent Systems from the IEEE-PES, in 2008. He served as the General Chair for ISAP, in 1999 and 2009, respectively, the Vice-General Chair for ISAP, in 2001, and for the Congress on Logic Applied to Technology, in 2003 and 2007, respectively.

• • •