International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST

DRL Based Naïve Emergency Control for Complicated Power Systems

Alok Kumar Sriwastawa, Virendra Pratap Yadav
SHEAT College of Engineering, Varanasi, Uttar Pradesh, India

Article Info
Volume 9, Issue 1
Page Number : 309-321

Publication Issue
January-February-2022

Article History
Accepted : 15 Feb 2022
Published : 22 Feb 2022

ABSTRACT

The power system's operating point will be less predictable as a result of these adjustments. Designing emergency controls offline via lengthy simulations has long been the norm. New sophisticated “wide-area emergency control algorithms” are needed since power system for the future is likely to vary more. The final line of defense for grid security and resilience is emergency control of the power system. For the most part, existing emergency response plans are developed off-line, using a “worst-case” scenario or a handful of representative operating situations. As the level of uncertainty and variability in contemporary electrical grids rises, these systems face considerable challenges in terms of adaptability and resilience. “Deep reinforcement learning (DRL)” for complex power systems was used in this thesis to build unique adaptive emergency control techniques that make use of DRL's non-linear generalization capabilities and high-dimensional feature extraction. “Reinforcement Learning for Grid Control (RLGC)” is a new open-source platform that was created to aid in the advancement and assessment of “DRL algorithms” for controlling electricity systems. There is a description of the emergency control systems for dynamic generator braking as well as platform and DRL-based under-voltage load shedding. The created DRL approach is tested for its potential to manage a wide range of simulation situations, uncertainty in model parameters, and noise in data.

Keywords: Emergency Control, Deep Reinforcement Learning, Transient Stability Dynamic Breaking, Load Shedding

I. INTRODUCTION

A. Overview

Owing to rising uncertainty, complexity, and data dimensions in power systems, traditional approaches typically hit bottlenecks while trying to handle control and decision issues. Thus, data-driven strategies for fixing such challenges are being intensively investigated. “Deep reinforcement learning (DRL)” is one of these data-driven methodologies and is recognized as actual "artificial intelligence (AI)". DRL is a mix of "reinforcement learning (RL)" and "deep learning (DL)". This branch of study has been employed to handle a broad variety of complicated sequential decision-making challenges, including those in power systems. This study initially discusses the core principles, models, methods, and approaches of DRL. The applications in power systems are then examined, comprising the electricity market, demand response, operational control, and energy management. In addition, current breakthroughs in DRL include the merging of RL with other classical approaches and the potential and problems of applications in power systems are highlighted as well.

The second is the programming approach, such as mixed integer programming [1], [2], dynamic programming [3], power system is a dynamic, complex, large-scale network of electrical components. Power systems have gone through many decades of evolution. During this period, economic, technical, political, and environmental motivations have converted conventional grids into more sophisticated, resilient, sustainable, and efficient smart grids [4] – [6]. Smart grids leverage bi-directional energy flow accompanied by bi-directional details flow among all the players, including manufacturers, consumers, distribution and transmission system operators, and demand response aggregators [6], [8]. Such variables have caused issues to the electricity grid from diverse viewpoints. Firstly, the high penetration of renewable power provides higher volatility to a power system.

To tackle these difficulties, effective procedures are necessary for planning and controlling the grid. This continual change of networks leads to greater unpredictability and difficulty in both the commercial transactions and the actual physical flows of power [9].

B. Introduction Deep-Reinforcement Learning

DRL integrates deep learning's sensing capability with reinforcement learning's decision-making capability. It is a form of AI that is more akin to human thought and is widely recognized as true AI. Fig. 1.1 illustrates the basic framework of DRL. Deep learning gathers information about the target observation from the environment and offers state detail about the present environment.

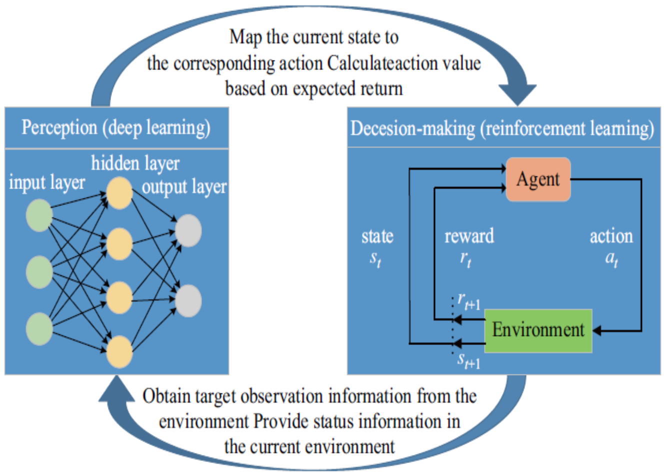


Fig. 1.1: DRL Framework

C. DRL Algorithms

DRL challenges may be phrased as optimization, planning, management, and control concerns. Solution methods RL, e.g. "Q-learning", the Q function's iteration procedure is as displayed in [10], whereas it will update as displayed in DRL [10]. The goal function may be described as [11] at this moment. The policy-based approaches directly optimize the quantity of interest while staying stable under the function approximations at each step by redefining the policy and calculating the value according to this new policy until the policy converges. At first, the objective function's gradient is derived as policy parameters as indicated in [11], and then the weight matrix will update using [12].

D. Applications in Power System

After years of study, several articles have been published regarding the uses of DRL in power systems, and most of them were published after 2018. These applications cover a broad variety of optimization issues, decisions, and control, in the power system, including electricity market, demand response, operational control, energy management, and many more.

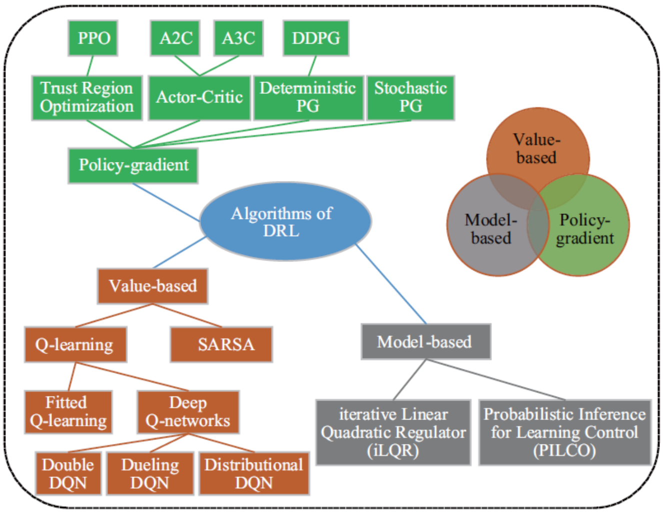


Fig. 1.2: Main DRL algorithms and their inter-relationship

In conclusion, DRL and its applications in the power system still confront numerous potential problems. These will provoke additional attention and inquiry, and there will certainly be more startling breakthroughs in the future.

II. LITERATURE REVIEW

A. Overview

In the literature, there are several examples of crises and preemptive management methods. It will be examined in this section which is the most crucial of all. The literature frequently makes use of one of two ways to choose remedial control measures. The first is sensitivity approaches, and the second is Optimal Power Flow (OPF) approaches. Emergency control measures are selected using sensitivity approaches, which are determined by the stability index's sensitivity when compared to the controls that are now accessible.

B. Sensitivity Approach

To find control actions, several linear methods have been proposed. From a power system model, it is possible to calculate the sensitivities of distinct measures of stability about the available control inputs, and then to implement these results to improve stability. These procedures are often extremely quick, which makes them excellent for use in emergencies. They are unsuitable for huge changes in configuration due to the failure of the linear assumption.

According to [3], voltage stability was improved by employing a control methodology that relied on descriptor load flow, Jacobian's least singular value. It relies on the sensitivity of reactive and active powers, and it is remedied utilizing a strategy that is based on the continuation of previous actions.

C. Optimal Power Flow Approach

The OPF may be utilized to reveal adequate emergency controls for a given situation. When compared to the sensitivity method, which relies on a linearization of the system, the OPF may encompass the entire model and thus take into consideration the nonlinear effects. Several stability indices may be applied to the “OPF formulation” to accommodate stability margins to achieve the desired level of resistance to fluctuation resistance. This is why the sensitivity methodologies have been largely replaced by the OPF technique in the scientific literature. The OPF is utilized to confront the economic challenges associated with the operation of a power system [15]. Consequently, “OPF formulations” such as “security-constrained OPF (SC-OPF),” “Integrate Security against Eventualities,” and “Security-Constrained OPF (SC-OPF)” is used. The SC-OPF is a computationally intensive device that is utilized to retain the power system operating in a normal state of operation. On the other hand, the OPF may also be utilized to restore the cohesion of a power system [16], if necessary. A set of stability constraints may be incorporated into the OPF to compute the corrective activity required to locate a new stable operating point.

According to [17], a method for revamping a system to offer damping to prevent small-signal instability without amending the closed-loop power oscillation damping controllers has been discovered and implemented. In [18], the topic of security against unforeseen events was also discussed. Because of the computational challenges involved, the OPF scheme for retaining small-signal stability is generally not adequate for emergency control in huge systems in general.

D. Operational Envelope

When managing power systems, it is possible to use an operational envelope that is estimated off-line, like that explained in [18], which addresses voltage and temperature stability. It is possible to implement emergency controls as controls that guide the operating point into the security region if the functional space of the power system has stability restrictions defined. With the help of specified contingencies, [19] developed a method for evaluating the boundaries of voltage, temperature, small-signal, and transient stability under specified conditions, and the results are displayed as a graph.

III. METHODOLOGY

A. Overview

All countries' national and economic security depends on reliable and resilient energy. Anticipated (e.g. N-1) threats have been well protected by a wide range of preventive management measures. In the preceding two decades, however, the United States, India, Brazil, and Europe all experienced multiple large-scale blackouts [1–3]. Emergency control has long been recognized as essential for limiting the scope and impact of power outages and other significant blackouts. Generation tripping or dispatch, dynamic braking, controlled system separation, and load shedding.

B. Problem Statement

To create new schemes that have great adaptiveness and resilience to handle the uncertainties and changes that occur in current electricity grids.

C. Research Objectives

To answer our issue statement, the present research has the below-mentioned objectives:

1. Developed an innovative and able to adapt “Emergency Control Techniques” utilizing “Deep Reinforcement Learning (DRL)” by using the non-linear generalization capabilities and high-dimensional feature extraction of DRL for “complex power systems.”
2. DRL's ability to withstand a wide range of simulation situations, the uncertainty of model parameters, and data noise is examined in the second phase of the research project.
3. Third, extensive case studies have shown that both the IEEE 39 bus and the two areas, four machine systems have excellent performance and resiliency.

D. Developing Algorithms For Grid Control

Reinforcement Learning for Grid Management (RLGC) is an open-source platform that has been created and released with the aim of designing, testing, and assessing RL algorithms for power system control [20]. Open-source benchmarks (like Image Net and Open AI Gym) are significant driving factors in machine learning improvement (including RL). RLGC's purpose is to provide a comparable open-source benchmark for RL in power grid management. Fig. 3.1 depicts the architecture of this open platform. It consists of two major sections:

1. The RL module; and
2. The power system simulation and control module.

Two configuration files are utilized to describe the settings for the dynamic simulation of the power system as well as the RL training parameters.

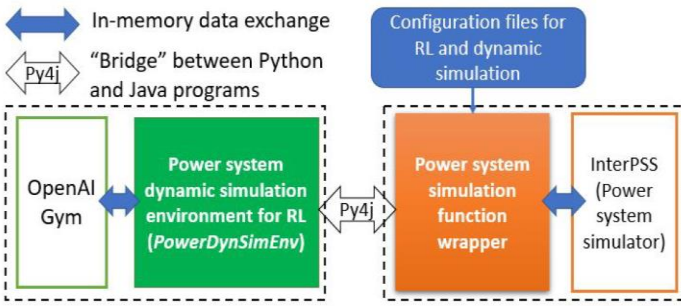


Fig. 3.1: An open platform for the development, skilling, and benchmarking of real-time control algorithms for power systems.

E. Implementation Details and Usage

RL module contains a Power DynSimEnv python class, which is created by widening the OpenAI Gym's standard basic environment Env class, which is named Power DynSimEnv. Control module and power system simulation developers are working on creating a wrapper for the Inter PSS simulation functions and capabilities, which will be used to interface with the PowerDynSimEnv environment in the Real-Time (RL) module. When applied to Algorithm 1, it consists of various key functions that represent the interactions between the environment and the learning agent (AL1).

A typical approach for testing algorithms of DRL and training NN models for grid control on the established platform consists of two stages: (1) the testing stage for verifying the taught NN and (2) the training stage for learning. The DRL will execute NN learning via training steps high in number throughout the training stage. It learns an optimum policy via exploitation and exploration and stores the best-performing NN settings automatically.

Fig. 3.2 illustrates the methodology for utilizing a grid control platform for testing and training the DRL model for grid control.

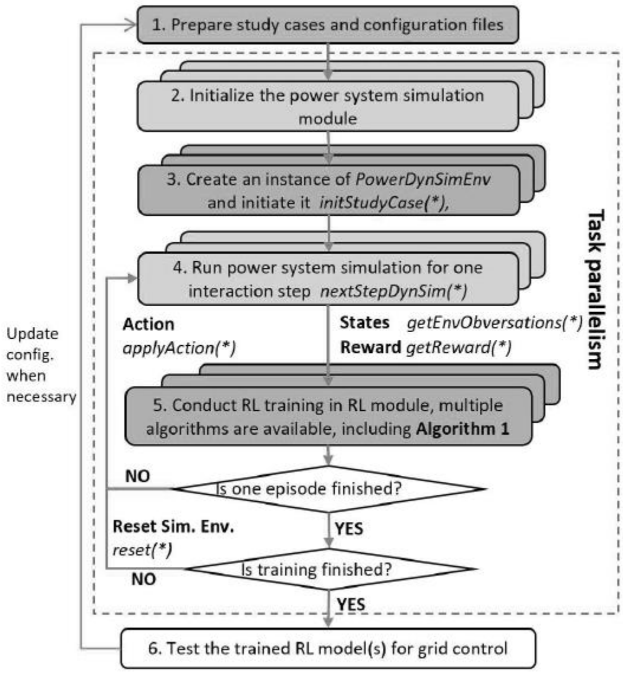


Fig. 3.2 shows a flowchart of a reminiscent approach for utilizing the platform to train and evaluate RL model(s) for grid control

F. DRL Algorithms For Grid Emergency Control

We explored and built control schemes based on DRL for two common forms of grid emergency control using the established platform stated in the preceding chapter:

1. Dynamic generator brake [8]; and
2. Low-voltage load shedding

Design and execution specifications for both emergency control systems' DRL algorithms will be covered in the following subchapters. These details will include neural networks, observations, actions, and reward systems, among other things.

IV. SIMULATION, TEST, AND RESULTS

Throughout the rest of the thesis, the same time steps are utilized in the test cases to ensure consistency. It took 9 hours to complete the training procedure on a Linux workstation with 32 AMD Opteron 1.44 GHz Processors and 64 Gigabit RAM, with no parallelism, on a computer. Our technique robustly learns effective policies when the parameters are properly tuned. Fig. 4.1 depicts the moving average of the reward over the course of the training. The decrease observed at the 3600th episode, as depicted in Fig. 4.1, is associated with a significant negative reward as a result of one "bad" excursion during training. The instability of the DQN algorithm is not indicated by this result, on the contrary. For the DQN algorithm to continue to be trained, the DQN model must learn to avoid the poor control actions that it has encountered during the training process. The DQN model eventually converges to the local optimum solution. We have done a lot of testing, and we have found that all of the local optimums that we found are excellent solutions.

T

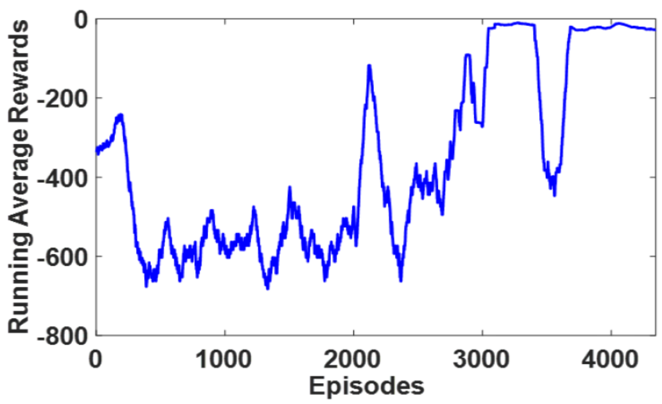


Fig. 4.1: The rewards' moving average throughout the DRL skilling

Using a new and much larger set of scenarios, including a variety of combinations of fault location, power flow condition, and fault duration, we test the structural rigidity of the culminating control policy (law) after the DRL model training.

1. A variety of power flow predicaments are tested, which would include (a) the original power flow case for learning and training, (b) each load in the system decreases or increases by 50 MW, 100 MW, and 180 MW, and (c) the tie-line power flow decreases or increases by 20 MW, 40 MW, 70 MW, and 100 MW between buses 7 and 10. Because the two tie-lines are the only means of connecting areas 1 and 2, it is possible to adjust the tie-line power flow by raising the generators' actual power output in one area while diminishing the generators' actual power output in the other area in the appropriate manner to achieve the desired result;
2. The problem location is picked for all the 10 buses;
3. The fault time is determined in a random manner between 0.3 s and 0.7 s.

The two-area power system is estimated to be able to withstand a fault for up to 0.583 seconds without losing stability without the use of dynamic braking. As an alternative, the system may remain stable if it is used in conjunction with the control rule learned by DRL in the various situations described above (we test 220 distinct scenarios). To make the inputs to the DRL-based control system more pragmatic, we add a zero mean as well, one percent Gaussian-distributed noise to the data that is supplied to the trained NN. The trained control based on DRL was compared to the standard 2-dimension Q-table-based "Q-learning" approach in [8], which we found to be superior. When sound is added to the observations, the findings reveal that the control based on DRL outperforms the traditional "Q-learning"-based control for all of the testing situations.

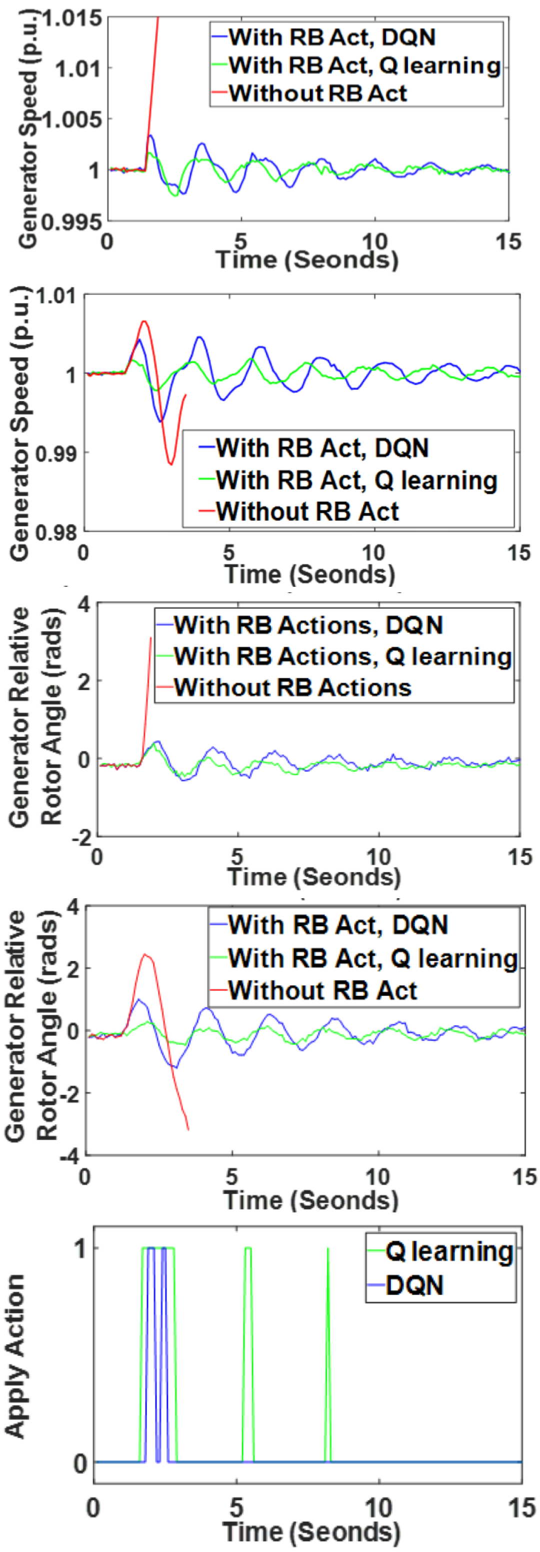
Fig. 4.2 (a) and (b) depict two illustrations of the RB actions for various power flow and faults conditions, for both DRL-based and traditional "Q-learning"-based control systems. As shown in Fig. 4.2 (a), the relative rotor angle and generator 3 speed (without and with RB actions) are depicted, as are the RB actions, for an intermittent defect at bus 4 with a period of 0.7 seconds, under the power flow condition that each load raises by 100 MW as compared to the power flow scenario used to train the operators. Under the original power flow situation for training, Fig. 4.2 (b) depicts the speed of generator 3 as well as the relative rotor angle and also the RB actions for a defect at bus 9 with a period of 0.6 seconds, and the RB actions for a defect at bus 9. Fig. 4.2 (a) and (b) show that if no RB actions are provided (red line), the system loses stability; however, when both DRL-based (blue line) and traditional "Q-learning"-based control (green line) provide RB actions, the system can maintain cohesion.

However, the control based on DRL offers unquestionably greater control actions as compared to the traditional "Q-learning"-based control, because the control based on DRL operates the RB in shorter time steps and, as a result, earns greater rewards. As illustrated in Fig. 4.2 (a) and (b), the control based on DRL would perform various RB actions at various times in each of the two circumstances. Every one of the outcomes displayed in Fig. 4.2 demonstrates the robustness, effectiveness, and adaptability of the DRL algorithm. However, it must be kept in mind that we also looked at different pre-fault periods; under normal circumstances, the control based on DRL does not apply any braking to the vehicle.

A. Under-Voltage Load Shedding

As shown in Fig. 4.3, an orchestrated UVLS scheme against FIDVR was constructed using the established platform and DRL algorithm, and the scheme was assessed on an amended IEEE 39-bus system [21], where step-down transformers were added to load buses 4, 7, and 18. A mix of constant impedance loads [23] and single-phase A/C motors [22] is used to model the original loads, which have been relocated to the transformers' low-voltage side.

It is necessary to use the OpenAI Baselines version of the DQN method to establish a closed-loop control strategy for implementing load shedding at buses 4, 7, and 18 to prevent FIDVR and attain the voltage recovery criteria displayed in Fig. 4 to prevent FIDVR. According to this study, the reward function coefficients (9) are as follows:c_1 = 260, c_2 = 150, and c_3 = 3. It is observed that the voltage magnitudes at buses 4, 7, 8, and 18 and the step-down transformers' low-voltage sides connected to them are greater than zero and that the proportions of loads served by buses 4, 7, and 18 are greater than one; therefore, N_m = 11. The most recent ten observation states are stacked and being utilized as a starting point.



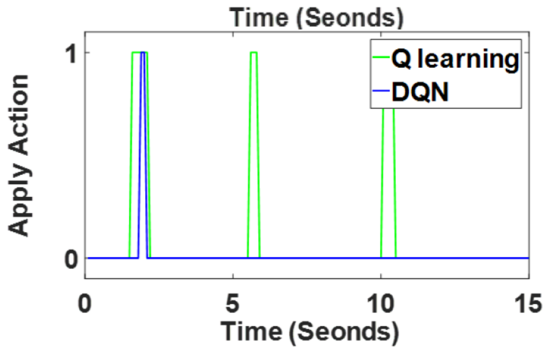


Fig. 4.2: The system's relative rotor angle and the generator speed evolution: Both buses 4 and 9 have a fault of 0.7 seconds, which means that N_i , the nodes in the input layer is 110 in number and N_r is 10

At each action time step, the control action for buses 4, 7, and 18 is either 1 (load shedding) or 0 (no-load shedding) (shedding 20 percent of the initial total load at the bus). This means that at each action step, there are a total of 8 possible combinations of potential discrete control actions, which corresponds to a total of 8 nodes in the output layer N_o. The following are some other critical hyperparameters to consider: 1,200,000 total interaction steps were used in the training; N_{h1} = N_{h2}= 256 nodes were used in the hidden layers; the learning rate was η = 0.00005; the learning rate was 0.00005; the minimum exploration rate was 0.02.

After a flat start of dynamic simulation, each episode begins with a short-circuit fault implemented randomly at bus 4, 15, or 21 with a randomly-chosen fault duration of 0.0 s (no-fault), 0.05 seconds, or 0.08 seconds; and the fault is self-cleared at the end of the simulation at 1.0 s of the simulation time. By selecting the fault location and duration at random, the training agent can ensure that the system interacts with it both with and without the presence of FIDVR conditions. No paralysis was experienced during the training procedures, which took 21 hours on the same Linux workstation that was utilized in the earlier case. Figure 4.4 depicts the rewards' moving average received throughout the training period.

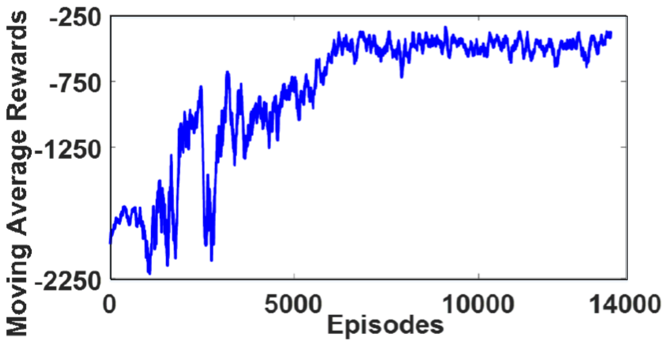
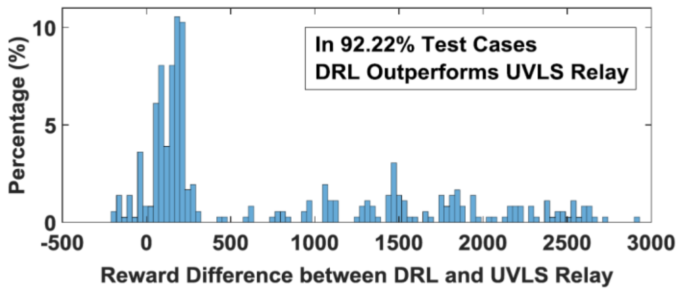


Fig. 4.3: The moving average of the awards received throughout the DRL load shedding management skilling for the 39-bus system.

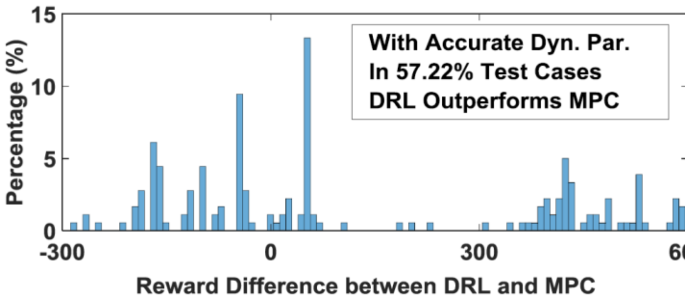
Using a set of 960 test situations containing various combo of power flow circumstances, fault locations, dynamic model parameters, and fault duration from the skilling contexts, we evaluated the structural rigidity and adaptiveness of the trained DRL agent. (1) four distinct load levels (80%, 90%, 110%, and 120%); (2) two distinct sets of critical dynamic parameters for the A/C motor model, one about (assumed) true values and the other incorporating a 10% rise in the A/C motor stalling performance parameters T stall and V stall[39]. Keep in mind that the A/C motor dynamic model is an aggregated model that displays a huge number of physical A/C in the real world, and as a result, its parameters may contain a variety of errors; (3) 30 distinct fault sites (corresponding to buses 1 through 30); and (4) four different fault length periods (corresponding to 0.02, 0.05, 0.08, and 0.1 s).

For the UVLS relay load shedding scheme, we have tested the previously trained load shedding control based on DRL, along with an MPC methodology that utilizes a mixed integer programming optimization to address the issue posed by (6). Each of the three control approaches has been assessed in terms of the reward and the execution time specified in the design (9). The reward differences (i.e., the DRL'reward minus the reward of a comparative approach) for each test situation are computed to demonstrate the comparison findings; a positive number indicates that the DRL technique is better for the corresponding test case, and a negative number indicates that the comparative methods are superior. 462 of the 960 test situations could result in FIDVR difficulties if nothing is done, necessitating the use of load shedding, according to the results.

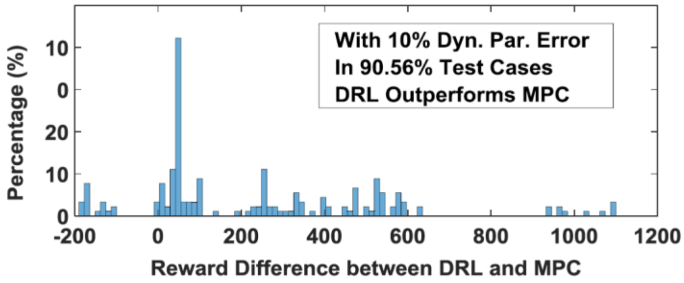
The reward difference' histogram between the UVLS relay and the DRL-based control is depicted in Fig. 8(a). This means that the control based on DRLoutperformed the UVLS relay in 92.22 percent of the 462 test cases. Test Set A contains 229 test scenarios with the same dynamic characteristics as the training scenarios, whereas Test Set B contains 233 test scenarios with dynamic load parameters Tstall and Vstall that are 10 percent higher than the training scenarios (Test Set B). It is the primary goal of Test Set B to replicate the modeling gaps (or uncertainties) that can occur in real-world applications. It should be noted that the DRL approach based on DQN is model-free, whereas methods based on MPC rely strongly on the correctness of the model; as a result, it is critical to address modeling errors in applications based on MPC.



(a)



(b)



(c)

Fig. 4.4. Histogram of the reward gap among (a) UVLS and DRL for the \s462 test cases that need load shedding; (b) MPC and DRL for the 229 test cases in \sTest Set A; and (c) MPC and DRL for the 233 test cases in Test Set B.

For Test Set A, Fig. 10 (b) portrays the reward difference' histogram between the DRL and the MPC, which reveals that DRL-based control outperforms the MPC in a slight majority of the test situations (the DRL outperforms the MPC in 57.22% of the test situations). The reward difference's histogram between the MPC and DRL techniques is shown in Fig. 10 (c) for Test Set B, and it reveals that the DRL approach outperforms the MPC technique in 90.56 percent of the test situations. Fig. 10 (b) and (c) illustrate considerable merit of the newly developed DRL approach over the MPC technique: the MPC method's effectiveness is highly dependent on the correctness of the system model, whereas DRL is model-independent and more resilient to modeling errors.

Calculation times for the DRL and MPC algorithms are summarized in Table 4.1. UVLS relays do not have a calculation time because they are either instantaneous or have a preset delay. The DRL method needs significantly less implementation time than the MPC technique, as the NN handling the complex mapping from observed states to actions is much more efficient in the DRL approach than the time-consuming, complex optimization solution procedures in the MPC technique. With a 0.13 second response time during an eight-second simulation event, the DRL technique can meet real-time operating prerequisites and enables grid operators to validate control actions as needed.

Table 4.1: Comparison of The MPC And DRL's Average Computation Time

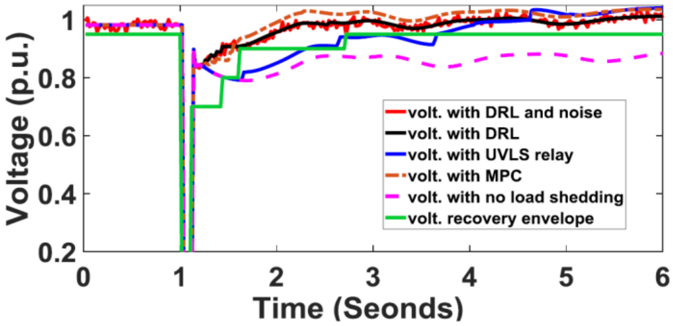
Average DRL Computation Time	Average MPC Computation Time
0.13 seconds	23.73 seconds

A new test scenario with a 120 percent load level is depicted in Figs. 11 and 12 to further demonstrate the advantages of the DRL technique. The effectiveness of the MPC, DRL, and UVLS relay control schemes are all compared for this new test scenario, which is shown in Fig. 13. There is a 0.1-second duration time for the fault to occur on bus 3, and the dynamic parameters V stall and T stall both rises by 10% as a result of the fault. The data is also contaminated with zero mean, one percent Gaussian-distributed noise to render the testing for the load shedding control based on DRL more realistically conducted This test scenario yields total rewards of -1271.61 points for the DRL, -1548.14 points for the MPC, and - 3778.80 points for the UVLS relay control, respectively. For various load shedding controllers, the voltage profiles at buses 4, 7, and 18 are shown in Fig. 11.

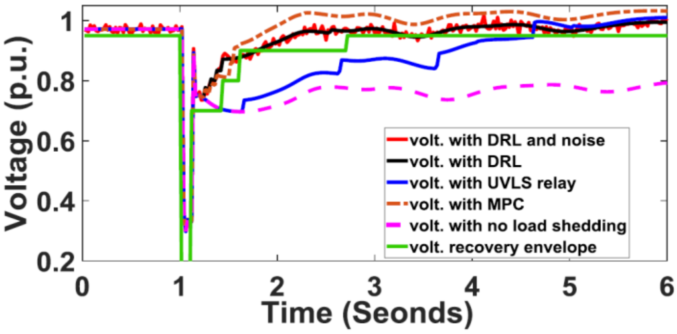
For various relay control schemes, the voltage profiles at buses 4, 7, and 18 are displayed in Fig. 12. The amount of load shedding at buses 4, 7, and 18 is shown in Fig. 12. It is considerable to keep in mind that the additional one percent noise does not affect the decision-making or the effectiveness of the DRL-based control. Both of the following factors contribute to the large reward difference (2507.19) between the UVLS relay and DRL: 1) Compared to the UVLS relay, the DRL sheds a considerably smaller amount of load. The figure illustrates that, when compared to the UVLS relay, the DRL sheds 60 percent (120 MW) less load for bus 4 (the DRL technique does not shed any load at bus 4) and 20 percent (14.64 MW) less load for bus 18; 2) when compared to the UVLS relay technique, the DRL method leads to a significantly better voltage recovery profile, as illustrated in Figure. With the control based on DRL, the voltages at all three load buses with the A/C motors quickly recover above the voltage recovery envelope permitted by the operating standard, allowing for faster voltage recovery.

The UVLS relay mechanism, on the other hand, is unable to recover the voltages at the three buses even 3 seconds after the fault has been resolved, causing the UVLS relays to shed a higher load at these three buses as a result. Because the DRL technique sheds less load than the MPC method while still meeting the operating standard criteria, the reward difference (276.53) between the two techniques is primarily due to this fact. The figure displays that the DRL technique reduces bus 7 load by 20 percent (26 MW) and bus 18 load by 20 percent (14.64 MW) when compared to the conventional technique. Because the MPC approach continues to suffer from fallacious crucial model parameters (10 percent divergence from the genuine values), the MPC approach results in higher load shedding (10 percent difference from the true values). Be aware that even though the Figure displays that the MPC technique's voltage recovery profiles are a little bit greater than those of the DRL method (at the expense of additional loads shed), this does not result in a higher reward because a voltage recovery profile that is higher than the standard voltage recovery standard is not rewarded as per the rules (9).

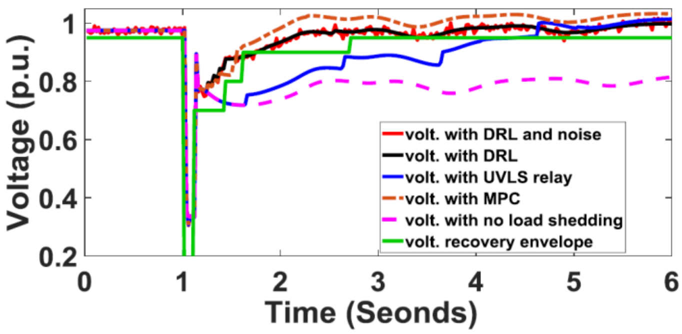
We conclude that this is appropriate given that the ultimate goal of UVLS controls is to recover the voltage over the envelope allowed by industry standards with the least amount of load shedding possible In summary, when comparing the DRL approach to the MPC control techniques and UVLS relay, the DRL approach displays considerable improvements in terms of resilience and adaptability. The DRL model may also deliver control actions incredibly quickly (0.13 s on average) in emergencies, making it a good candidate for use in real-time emergency control situations.



(a)



(b)

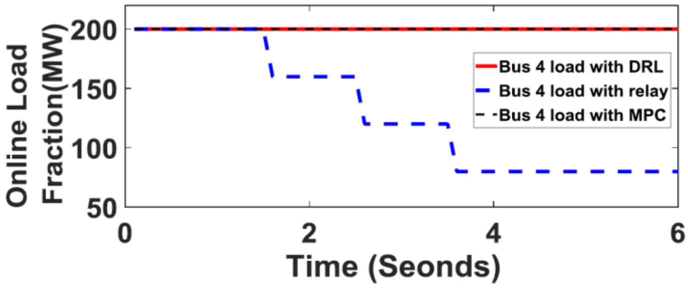


(c)

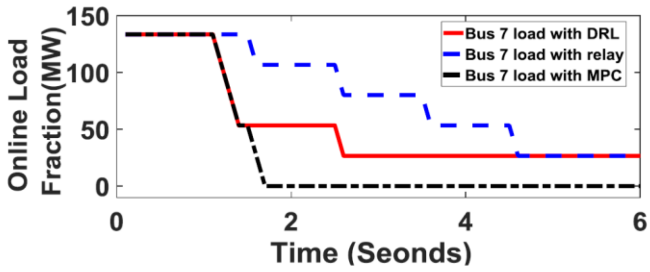
Fig. 4.5: Voltage profiles and required voltage recovery envelope for different types of loads shedding control methods: a) bus 4; b) bus 7; c) bus 18

V. DISCUSSIONS

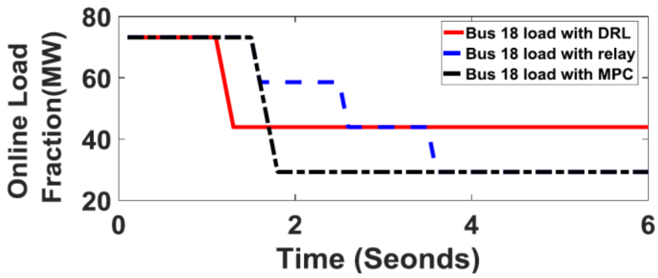
There are numerous key issues for DRL use in general, and notably in respect to its usage in power system emergency control.



(a)



(b)



(c)

Fig. 5.1: Load fractions online when load shedding is managed via MPC, DRL, or UVLS relay mechanisms: (a) bus 4; (b) bus 7; (c) bus 18. The Proportions are multiplied by the interim busload, expressed in megawatts.

VI. CONCLUSIONS AND FUTURE WORK

In the event of a large disruption or a severe emergency, a reliable emergency control system is very essential. Building "adaptive emergency control systems" using DRL is the focus of this dissertation. As part of an effort to speed up the development and testing of grid control algorithms, an open-source platform called RLGC has been created. We seek to serve as a starting point for future research in this field by releasing it as an open-source project. "Dynamic generator brake and UVLS" are two typical emergency control solutions built on the platform. It was found that both "DRL-based emergency control schemes" are adaptive and robust (to new scenarios, noise in observations, and model parameter uncertainty) as well as superior to MPC-based emergency control, conventional "Q-learning," and other prevailing protection mechanisms. To efficiently tackle control issues linked with increased uncertainties in power systems, recent innovations like deep meta-reinforcement learning and safe exploration are being used.

VII. REFERENCES

[1]. M. Zachar and P. Daoutidis, "Micro-grid/macro grid energy exchange: a novel market structure and stochastic scheduling, "IEEE Transactions on Smart Grid, vol. 8, no. 1, pp. 178–189, Jan. 2017.

[2]. C. Ordoudis, P. Pinson, and J. M. Morales, "An integrated market for electricity and natural gas systems with stochastic power producers, "European Journal of Operational Research, vol. 272, no. 2, pp. 642– 654, Jan. 2019.

[3]. L. Ze´phyr and C. L. Anderson, "Stochastic dynamic programming approach to managing power system uncertainty with distributed storage, " Computational Management Science, vol. 15, no. 1, pp. 87–110, Jan. 2018.

[4]. D. X. Zhang, X. Q. Han, and C. Y. Deng, "Review on the research and practice of deep learning and reinforcement learning in smart grids, " CSEE Journal of Power and Energy Systems, vol. 4, no. 3, pp. 362–370, Sep. 2018.

[5]. M. L. Tuballa and M. L. Abundo, "A review of the development of Smart Grid technologies," Renewable and Sustainable Energy Reviews, vol. 59, pp. 710–725, Jun. 2016.

[6]. R. C. Qiu and P. Antonik, Smart Grid and Big Data: Theory and Practice, New York: Wiley Publishing, 2017.

[7]. X. He, Q. Ai, R. C. Qiu, W. T. Huang, L. J. Piao, and H. C. Liu, "A big data architecture design for smart grids based on random matrix theory, " IEEE Transactions on Smart Grid, vol. 8, no. 2, pp. 674–686, Mar. 2017.

[8]. T. Yu, B. Zhou, and W. G. Zhen, "Application and development of reinforcement learning theory in power systems," Power System Protection and Control, vol. 37, no. 14, pp. 122–128, Jul. 2009.

[9]. X. He, L. Chu, R. C. M. Qiu, Q. Ai, and Z. N. Ling, "A novel data-driven situation awareness approach for future grids—Using large random matrices for big data modeling, " IEEE Access, vol. 6, pp. 13855–13865, Mar. 2018.

[10]. L. Chu, R. Qiu, X. He, Z. N. Ling, and Y. D. Liu, "Massive streaming pmu data modeling and analytics in smart grid state evaluation based on multiple high-dimensional covariance test, " IEEE Transactions on Big Data, vol. 4, no. 1, pp. 55–64, Mar. 2018.

[11]. R. Qiu, L. Chu, X. He, Z. N. Ling, and H. C. Liu, "Spatiotemporal big data analysis for smart grids based on random matrix theory, " in Transportation and Power Grid in Smart Cities: Communication Networks and Services, H. T. Mouftah, M. Erol-Kantarci, and M. H. Rehmani, Eds. John Wiley & Sons Ltd, 2018, pp. 591–633.

[12]. R. C. Qiu, X. He, L. Chu, and Q. Ai, "Big data analysis of power grid from random matrix theory, " Institution of Engineering and Technology (IET) in Smarter Energy: From Smart Metering to the Smart GridSmarter Energy: From Smart Metering to the Smart Grid, 2016, pp. 381–425, doi:10.1049/pbpo088e ch13.

[13]. M. J. Han, R. May, X. X. Zhang, X. R. Wang, S. Pan, D. Yan, Y. Jin, and L. G. Xu, "A review of reinforcement learning methodologies for controlling occupant comfort in buildings, "Sustainable Cities and Society, vol. 51, pp. 101748, Nov. 2019.

[14]. J. L. Duchaud, G. Notton, C. Darras, and C. Voyant, "Power ramp-rate control algorithm

with optimal State of Charge reference via Dynamic Programming, " Energy, vol. 149, pp. 709–717, Apr. 2018.
[15]. D. P. Bertsekas. Reinforcement learning and optimal control. [online]. Available: http://www.athenasc.com/.
[16]. F. Agostinelli, G. Hocquet, S. Singh, and P. Baldi, "From reinforcement learning to deep reinforcement learning: an overview, " in Braverman Readings in Machine Learning. Key Ideas from Inception to Current State, L. Rozonoer, B. Mirkin, and I. Muchnik, Eds. Cham: Springer, 2018, pp. 298–328.
[17]. B. V. Mbuwir, M. Kaffash, and G. Deconinck, "Battery scheduling in a residential multi-carrier energy system using reinforcement learning, " in Proceedings of 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids, 2018.
[18]. J. D. Wu, H. W. He, J. K. Peng, Y. C. Li, and Z. J. Li, "Continuous reinforcement learning of energy management with deep Q network for a power-split hybrid electric bus, " Applied Energy, vol. 222, pp. 799–811, Jul. 2018.
[19]. Z. Q. Wan, H. P. Li, H. B. He, and D. Prokhorov, "Model-free realtime EV charging scheduling based on deep reinforcement learning, " IEEE Transactions on Smart Grid, vol. 10, no. 5, pp. 5246–5257, Sep. 2019.
[20]. Y. Hu, W. M. Li, K. Xu, T. Zahid, F. Y. Qin, and C. M. Li, "Energy management strategy for a hybrid electric vehicle based on deep reinforcement learning, " Applied Sciences, vol. 8, No. 2, pp. 187, Jan. 2018.
[21]. H. W. Wang, C. J. Li, J. Y. Li, X. He, and T. W. Huang, "A survey on distributed optimization approaches and applications in smart grids, " Journal of Control and Decision, vol. 6, no. 1, pp. 41–60, Nov. 2019.
[22]. H. Li, D. Yang, W. Z. Su, J. H. Lu¨, X. H. Yu, "An overall distribution particle swarm optimization MPPT algorithm for photovoltaic system under partial shading, " IEEE Transactions on Industrial Electronics, vol. 66, no. 1, pp. 265–275, Jan. 2019.
[23]. H. J. Gu, R. F. Yan, and T. K. Saha, "Minimum synchronous inertia requirement of renewable power systems, " IEEE Transactions on Power Systems, vol. 33, no. 2, pp. 1533–1543, Mar. 2018.
[24]. Y. X. Li, "Deep reinforcement learning, " arXiv: 1810.06339 (2018).

Cite this article as :

Alok Kumar Sriwastawa, Virendra Pratap Yadav, "DRL Based Naïve Emergency Control for Complicated Power Systems", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 9 Issue 1, pp. 309-321, January-February 2022.
Journal URL : https://ijsrst.com/IJSRST229158