JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Survey on different techniques that use DDPGO with their implementation.

[1]Gajanan K. Bhoi, [2]G. A. Patil, [3]Umesh Kulkarni
[1]Student M. Tech (CSE) , [2]Professor, [3]Professor
[1]Department of Computer Engineering,
[1]D Y Patil College of Engineering & Technology, India

***Abstract*—In the last few years it has been brought into notice a complex issue of optimization systems that help track activities and provide outputs to analyze a certain pattern and predict behaviour and possibilities to understand any kind of system. Such kind of problem solving can be done with help of Deep Deterministic Policy Gradient Optimization (DDPGO). DDPGO is nothing but a technique that is made up by combining two different approaches i.e. Deep Q-Network and Deterministic Policy Gradient. It mainly consists of two major parts actor and critic. The Actor is an arrangement network that accepts the state as info and results the specific activity (persistent), rather than a likelihood appropriation over activities. The Critic is a Q-esteem network that makes an in state and move as info and results the Q-esteem. DDPG is an "off"- approach technique. DDPG is utilized in the consistent activity setting and the "deterministic" in DDPG alludes to the way that the Actor processes the activity straightforwardly rather than a likelihood circulation over activities.**

***IndexTerms*—DDPGO, Deep Learning Neural Network.**

1. I. INTRODUCTION

In today's progressive world it is a aid to have a powerful decision making systems to adapt to inputs passed to it and provide decisions and options to choose. To elaborate it lets consider examples like Artificial intelligence managed systems and environments which help us solve real time problems like scheduling, arranging, ordering which are required in each field of today's world like hospitals, traffic signals, hotels, public sector, airlines etc. Such fields need to take decision-based actions to accomplish tasks like traffic monitoring, detection of early diseases, scheduling flight departures, seat allocation. These all kinds of tasks can be easily accomplished by use of advanced neural network algorithms which have high capacity to take multiple inputs and provide better decisions. One of such kinds is DDPGO.

2. II. DIFFERENT SYSTEMS USING DDPGO TECHNIQUE

The proposed system mainly focused on the real time situation of traffic congestion caused at prime locations in the peak hours. We have to address the time and intensity of the traffic based on the distance between two prime points which have to be accurate to solve the issue of traffic congestion at given point to continue smooth passing of vehicles without affecting the remaining route as well. The author mainly compared two techniques i.e deep Q network(DQN) and deep deterministic policy gradient (DDPGO)[1] for solving the same issue of traffic congestion. DDPGO was applied to a variable network of traffic control to a single intersection points connecting multiple roads at a junction. The system used a self-learning approach as a plus point to train the network to efficiently pass the traffic. It is observed that DDPGO outperformed DQN in fix time control operation for optimization of traffic release. DDPGO self-learned and assessed the multipoint junctioning of traffic at single time. Value and Policy based approach was collaborated by DDPGO to fine tunes its performance. Moreover DDPGO operated in Dynamic real time situations rather than DQN.

Nowadays technology has taken up most of the sectors in the modern world, one of which is the ship operating sector. Smart ships have been one of the modern ways to live in an automated world where ships can be smartly controlled without much needed human intervention. The proposed system used two neural networks of Actor and critic each[2]. The actor network took care of the underlying robots controlling the ships while the critic network found out the error matrices and continuously improved the performance of the route assessment and stimulation with respect to other ships and schedules worldwide. Control over the ships and managing the complex paths and routes to improve the self-controlling mechanism for remote ships over oceans and continents was the achieved outcome of the DDPGO dual network.

Unmanned Aerial Vehicle (UAV) [3] is one of the progressive fields where DDPGO is used. The proposed work shows the implementation of DDPGO into such an important sector where UAV path programming and achievement of efficient path planning is to be achieved. In this field of UAV many obstacles are faced to plan the path for reaching the endpoint and one of it is the diverse and complex natural conditions and terrains which play and important part in path planning and management. To solve this issue Multi critic DDPGO(MCDDPGO)[3] was enforced to achieve accurate path analysis so as to trace and reach the endpoint in decided time frame. The actor network of MCDDPGO was the actual conditional input provided by the divide continuously to the multi critic network which had more than one critic networks. This multi critic network propagated the actual right path based on the dynamic continuous input from the actor network. Hence it was proved that MCDDPGO was more efficient to conventional DDPGO for path planning.

Energy internet is one of the aroused issue in the coming times where we need to create and generate networks that do not form grids which overconsume unnecessary amount of energy. Here we have to manage and created techniques to optimize the issue and find ways to achieve it and way to solve both the needs. In this paper ,Combined cooling-heating and power (CCHP) microgrid[4] was proposed by author to achieve the goal of energy saving by managing the cycle and delays to form the grid and pass energy. To monitor this circulation and delay of energy DDPGO was used. DDPGO critic and actor network worked hand-in-hand to handle the same. The multi energy and multi device energy consumption was solved by the DDPGO. The system was implemented on the dummy source and it worked as expected, the future implementation was to be carried onto the actual complex real time network.

A controlled optimal control strategy of active distribution networks (ADN) [5] formulated on Deep Deterministic Policy Gradient (DDPG) is proposed by the author to determine the fluctuation of energy. The Active Distribution Network (ADN) is used to actively track the minimal threshold voltage to come to a point where fluctuation can be avoided. This calculation of dynamically changing threshold depending on the huge spread network of energy is difficult and needs to be achieved in a manner where the minimal threshold voltage has to be continuously calculated. To get this dynamic voltage calculated DDPGO is induced which monitors the network by tracing and learning the energy network and periodically calculated the threshold voltage and sets it to avoid the fluctuations. This paper implemented this approach on a small section of the energy network and was successful. Hence it was to be implemented on a wide real time network.

The proposed system proposed a simulation to control a bicycle based on the predefined plan and achieve the decided path through a derived plan. In this approach a active disturbance rejection control (ADRC)[6] mechanism was formulated with help of DDPGO. Basically, ADRC handled the uncertain things that were caused while controlling the bicycle and bypassing those to stabilize it. A learning method i.e. composite was imposed which involved DDPGO to learn and analyse the plan formulated for the bicycle to follow. The actor network was the one which raced the course of plan while the critic network stabilized the bicycle by throwing the unwanted actions that caused instability. Both balancing based on motion and tracking of path both were achieved successfully through the ADRC by using DDPGO.

3. III. CONCLUSION

The above paper mainly focused on different approaches where the DDPGO technique was successfully implemented to enhance the decision making. Hence it can be concluded that the DDPGO is extensively used in large networks where tracing and understanding of the nature of the system is required. DDPGO mainly works in two network i.e. Actor and critic. Each network involves analysis and understanding of feed network while the critic network has to propagate the uncertain change or findings from the actor network. the actor and critic pair work in a manner where actors continuously analyse and feed data to critic which calculates the error or uncertain pattern and helps train the actor network to take appropriate actions. Basically, DDPGO was extensively used in all sectors like traffic management, Aviation systems, Ship industries, Voltage management and health care systems where it has worked in a efficient manner and provided best results.

4. IV. REFRENCES

[1] H. Pang and W. Gao, "Deep Deterministic Policy Gradient for Traffic Signal Control of Single Intersection," 2019 Chinese Control And Decision Conference (CCDC), 2019, pp. 5861-5866, doi: 10.1109/CCDC.2019.8832406.

[2] W. Wang, F. Ma and J. Liu, "Course Tracking Control for Smart Ships Based on A Deep Deterministic Policy Gradient-based Algorithm," 2019 5th International Conference on Transportation Information and Safety (ICTIS), 2019, pp. 1400-1404, doi: 10.1109/ICTIS.2019.8883840.

[3] R. Wu, F. Gu and J. Huang, "A multi-critic deep deterministic policy gradient UAV path planning," 2020 16th International Conference on Computational Intelligence and Security (CIS), 2020, pp. 6-10, doi: 10.1109/CIS52066.2020.00010.

[4] X. Fang, Y. Huo, Y. Han and J. Wang, "Energy Scheduling and Decision Learning of Combined Cooling, Heating and Power Microgrid Based on Deep Deterministic Policy Gradient," 2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2), 2020, pp. 476-481, doi: 10.1109/EI250167.2020.9346953.

[5] J. Gong, G. Mei and Y. Liu, "The real-time optimization of active distribution system based on deep deterministic policy gradient," 8th Renewable Power Generation Conference (RPG 2019), 2019, pp. 1-6, doi: 10.1049/cp.2019.0545.

[6] K. He, C. Dong, A. Yan, Q. Zheng, B. Liang and Q. Wang, "Composite deep learning control for autonomous bicycles by using deep deterministic policy gradient," IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, 2020, pp. 2766-2773, doi: 10.1109/IECON43393.2020.9254787.

[7] T. Dao and C. Chen, “Sliding-mode control for the roll-angle tracking of an unmanned bicycle,” Vehicle System Dynamics, vol. 49, no. 6, pp. 915–930, 2011.

[8] L. Guo, Q. Liao, S. Wei, and Y. Huang, “A kind of bicycle robot dynamic modeling and nonlinear control,” in The 2010 IEEE International Conference on Information and Automation. IEEE, 2010, pp. 1613–1617.

[9] H. M. Kim, Y. J. Lee, and P. K. Taik, “Robust back-stepping control of vehicle steering system,” in 2012 12th International Conference on Control, Automation and Systems. IEEE, 2012, pp. 712–714.M.

[10] Defoort and T. Murakami, “Sliding-mode control scheme for an intelligent bicycle,” IEEE Transactions on Industrial Electronics, vol. 56,no. 9, pp. 3357–3368, 2009.

[11] Nikolaos C. Koutsoukis ; Pavlos S. Georgilakis ; Nikos D. Hatziargyriou, Multistage Coordinated Planning of Active Distribution Networks[J]. IEEE Transactions on Power Systems, 2018, 33(1): 32 – 44.

[12] M. Usman, M. Coppo, F. Bignucolo, R. Turri,Losses management strategies in active distribution networks: A review[J].Electric Power Systems Research, Volume 163, Part A, October 2018, Pages 116-132.

[13] Liu, Yali ; Zhao, Xin; Li, Guodong; Research on the Optimization Plan Method based on Electric Vehicles' Scheduling Strategy in Active Distribution Network with Wind/Photovoltaic/Energy Storage Hybrid Distribution System[C]. 3rd Asia Conference on Power and Electrical Engineering, ACPE 2018, 366(1), June 13, 2018.

[14] Wang, Dongxiao; Meng, Ke; Luo, Fengji, etal, Coordinated dispatch of networked energy storage systems for loading management in active distribution networks[J]. IET Renewable Power Generation, 2016, 10(9): 1374-1381.

[15] A. Q. Huang, M. L. Crow, G. T. Heydt, J. P. Zhang, S. J. Dale, “The future renewable electric energy delivery and management (FREEDM) system: the energy Internet,” Proceedings of the IEEE, vol. 99, pp. 133–148, January 2011.

[16] W. Gu, Z. Wu, R. Bo, W.Liu, et al, “Modeling, planning and optimal energy management of combined cooling, heating and power microgrid: A review,” International Journal of Electrical Power & Energy Systems, vol. 54, pp. 26–37, January 2014.

[17] C. Marino, M. Marufuzzaman, M. Q. Hu, M. D. Sarder, “Develop-ing a CCHP-microgrid operation decision model under uncertainty,”Computers & Industrial Engineering, vol. 115, pp. 354–367, January 2018.

[18] L. Guo, W. J. Liu, B. W. Hong, C. S. Wang, “A two-stage optimal planning and design method for combined cooling, heat and power microgrid system,” Energy Conversion and Management, vol. 74, pp. 433–445, October 2013.

[19] Pappis, C., Mamdani, E. A Fuzzy Logic Controller for a Traffic Junction. IEEE Transactions on Systems, Man, and Cybernetics 7, 707–717 (1977)

[20] Srinivasan, D., Choy, M.C., Cheu, R.L. Neural Networks for Real-Time Traffic Signal Control. IEEE Transactions on Intelligent Transportation Systems 7(3), 261–272 (2006)