Expert Systems With Applications 213 (2023) 119192

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

A theoretical demonstration for reinforcement learning of PI control dynamics for optimal speed control of DC motors by using Twin Delay Deep Deterministic Policy Gradient Algorithm

Sevilay Tufenkci [a,*], Baris Baykant Alagoz [b], Gurkan Kavuran [c], Celaleddin Yeroglu [d], Norbert Herencsar [e], Shibendu Mahata [f]

[a] Malatya Turgut Ozal University, Department of Computer Technology, 44100 Malatya, Turkey
[b] Inonu University, Department of Computer Engineering, 44100 Malatya, Turkey
[c] Malatya Turgut Ozal University, Department of Electrical-Electronics Engineering, 44100 Malatya, Turkey
[d] Inonu University, Department of Computer Engineering, 44100 Malatya, Turkey
[e] Brno University of Technology, Faculty of Electrical Engineering and Communications, Dept. of Telecommunications, Technicka 12, 616 00 Brno, Czechia
[f] Dr. B. C. Roy Engineering College, Department of Electrical Engineering, Durgapur, West Bengal 713206, India

ARTICLE INFO

Keywords:
Deep reinforcement learning
DC motor
PI controller
Twin-delayed deep deterministic policy gradient
Metaheuristic optimization

ABSTRACT

To benefit from the advantages of Reinforcement Learning (RL) in industrial control applications, RL methods can be used for optimal tuning of the classical controllers based on the simulation scenarios of operating conditions. In this study, the Twin Delay Deep Deterministic (TD3) policy gradient method, which is an effective actor-critic RL strategy, is implemented to learn optimal Proportional Integral (PI) controller dynamics from a Direct Current (DC) motor speed control simulation environment. For this purpose, the PI controller dynamics are introduced to the actor-network by using the PI-based observer states from the control simulation environment. A suitable Simulink simulation environment is adapted to perform the training process of the TD3 algorithm. The actor-network learns the optimal PI controller dynamics by using the reward mechanism that implements the minimization of the optimal control objective function. A setpoint filter is used to describe the desired setpoint response, and step disturbance signals with random amplitude are incorporated in the simulation environment to improve disturbance rejection control skills with the help of experience based learning in the designed control simulation environment. When the training task is completed, the optimal PI controller coefficients are obtained from the weight coefficients of the actor-network. The performance of the optimal PI dynamics, which were learned by using the TD3 algorithm and Deep Deterministic Policy Gradient algorithm, are compared. Moreover, control performance improvement of this RL based PI controller tuning method (RL-PI) is demonstrated relative to performances of both integer and fractional order PI controllers that were tuned by using several popular metaheuristic optimization algorithms such as Genetic Algorithm, Particle Swarm Optimization, Grey Wolf Optimization and Differential Evolution.

1. Introduction

Reinforcement Learning (RL) is an effective machine learning method that is designed for learning from experience (Kaelbling, Littman, & Moore, 1996; Mnih, Kavukcuoglu, Silver, Graves, Antonoglou, Wierstra, & Riedmiller, 2013). In recent years, it has been utilized for intelligent control of real systems (Mnih et al., 2015; Lillicrap, Hunt, Pritzel, Heess, Erez, Tassa, Silver, & Wierstra, 2016; Rabault, Kuchta, Jensen, Réglade, & Cerardi, 2019). Today, many systems in daily use include DC motors to convert electrical energy to mechanical energy. Therefore, performance improvement in DC motor control contributes to many innovative application areas such as electric vehicles (Wu, Cheng, & Cui, 2004), Unmanned Aerial Vehicles (UAV) (Solomon et al., 2006; Solomon, 2007), etc.

DC motors have been frequently utilized in many application areas (Cui, Gong, & Xu, 2012; Berahim, 2014). Their lower price, ease of use,

* Corresponding author.
E-mail addresses: sevilay.tufenkci@ozal.edu.tr (S. Tufenkci), baykant.alagoz@inonu.edu.tr (B. Baykant Alagoz), gurkan.kavuran@ozal.edu.tr (G. Kavuran), c.yeroglu@inonu.edu.tr (C. Yeroglu), herencsn@vut.cz (N. Herencsar), shibendu.mahata@gmail.com (S. Mahata).

https://doi.org/10.1016/j.eswa.2022.119192
Received 20 May 2022; Received in revised form 29 October 2022; Accepted 29 October 2022
Available online 7 November 2022
0957-4174/