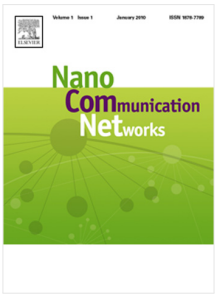
Accepted Manuscript



Deep reinforcement learning: Algorithm, applications, and ultra-low-power implementation

Hongjia Li, Ruizhe Cai, Ning Liu, Xue Lin, Yanzhi Wang

PII: S1878-7789(17)30085-6
DOI: https://doi.org/10.1016/j.nancom.2018.02.003
Reference: NANCOM 204

To appear in: *Nano Communication Networks*

Received date : 23 July 2017
Accepted date : 11 February 2018

Please cite this article as: H. Li, R. Cai, N. Liu, X. Lin, Y. Wang, Deep reinforcement learning: Algorithm, applications, and ultra-low-power implementation, *Nano Communication Networks* (2018), https://doi.org/10.1016/j.nancom.2018.02.003

Deep Reinforcement Learning: Algorithm, Applications, and Ultra-Low-Power Implementation

Hongjia Li[a,*], Ruizhe Cai[a,*], Ning Liu[a], Xue Lin[b], Yanzhi Wang[a]

[a] *Syracuse University, Syracuse*
[b] *Northeastern University, Boston*

Abstract

In order to overcome the limitation of traditional reinforcement learning techniques on the restricted dimensionality of state and action spaces, the recent breakthroughs of deep reinforcement learning (DRL) in Alpha Go and playing Atari set a good example in handling large state and action spaces of complicated control problems. The DRL technique is comprised of an offline deep neural network (DNN) construction phase and an online deep Q-learning phase. In the offline phase, DNNs are utilized to derive the correlation between each state-action pair of the system and its value function. In the online phase, a deep Q-learning technique is adopted based on the offline-trained DNN to derive the optimal action and meanwhile update the value estimates and the DNN.

This paper is the first to provide a comprehensive study of applications of the DRL framework on cloud computing and residential smart grid systems along with efficient hardware implementations. Based on the introduction of the general DRL framework, we develop two applications, one for the cloud computing resource allocation problem and one for the residential smart grid user-end task scheduling problem. The former could achieve up to 54.1% energy saving compared with baselines through automatically and dynamically distributing resources to servers. The latter achieves up to 22.77% total energy

*Hongjia Li and Ruizhe Cai contributed equally to this work.
Email addresses: hli42@syr.edu (Hongjia Li), rcai100@syr.edu (Ruizhe Cai), nliu03@syr.edu (Ning Liu), xuelin@coe.neu.edu (Xue Lin), ywang393@syr.edu (Yanzhi Wang)