

LITERATURE SURVEY PROPOSAL ON REINFORCEMENT LEARNING AND HPC SYSTEMS ADAPTIVE ROUTING

Ram Sharan Chaulagain, Tushar Chandra Mondol

Florida State University

1 BACKGROUND

The performance of Network interconnect and HPC applications are interdependent. Performance of network depends on traffic patterns coming from different HPC applications where as performance of HPC applications depend on network throughput and latency of messages. Routing is one of the key factor in interconnect performance. Sensing the traffic pattern and adaptive routing the packets are challenging. It requires approximating global traffic information using local router level information.

Reinforcement learning is a subarea of machine learning that automatically learns to maximize cumulative reward through interaction with the environment[1]. So, we want to explore more into intersection of reinforcement learning and HPC systems optimization problems, hoping that will help us to give some insights on designing AI-based routing technology in HPC systems. From this survey we want to acquire what reinforcement learning is, how HPC scholars are using this tool for similar problems.

Following are the some list of papers but not limited:

2 Papers considered

- a) Q-adaptive: A Multi-Agent Reinforcement Learning Based Routing on Dragonfly Network
- b) Packet Routing in Dynamically Changing Networks: A Reinforcement Learning Approach
- c) Predictive Q-routing: A memory based reinforcement learning approach to adaptive traffic control
- d) Richard S Sutton and Andrew G Barto. 2018. Reinforcement learning: An introduction. MIT press.

- e) Reinforcement Learning for Adaptive Routing
- f) Deep Reinforcement Agent for Scheduling in HPC

References

1. Richard S Sutton and Andrew G Barto. 2018. Reinforcement learning: An introduction. MIT press.