Multi-Armed Bandits for Optimizing New Peers in Peer-to-Peer Networks

Oscar Sandford [1] **Shawn Nettleton** [1]

Abstract

Write this last (fewer than 300 words). The completed document should be 5-9 pages.

1. Introduction

Peer-to-peer computer networks create a unique environment for content distribution wherein the integrity of the system is not compromised by the failure of a single, centralized node in the network. According to (Schollmeier, 2001), true peer-to-peer systems require peers to be mutually directly accessible (without intermediate entities), as well as the network state or quality of service being preserved in the advent of a peer being removed from the network, for any reason.

The requirements for peer-to-peer networks in different application domains vary. However, new peers that are directly accessing the server for the first time have no information on the network state. New peers therefore cannot be held accountable to preserve the network state and its content if other nodes disconnect. It is essential that this new peer is fed the relevant data as fast as possible in order to fulfill both the requirements of a true peer-to-peer environment, as well as any necessary quality of service targets. With the added volatility of a dynamic network setting, the rate at which a new peer can be brought ”up to speed” becomes far more crucial.

In this study, we abstract the new peer scenario described above as a reinforcement learning problem with multi-armed bandits. Various algorithms to solve the multi-armed bandit problem are considered, and a select few are implemented in order to evaluate their efficacy against this problem. Related literature is surveyed in order to compare our work with solutions to similar problems and verify the validity of our results.

[1]Department of Computer Science, University of Victoria, Victoria, Canada. Correspondence to: Oscar Sandford <oscarsandford@uvic.ca>, Shawn Nettleton <shawnnettleton@uvic.ca>.

RL Course Project Report, University of Victoria, Spring, 2022.

2. Related Work

2.1. Algorithms Used

2.2. Domain & Network Specific

Text here.

3. Problem Formulation

Consider the setting of a peer-to-peer network wherein a new peer joins with the intent to be brought ”up to speed” with the rest of the network as soon as possible (i.e. download all the data in the network from other peers). However, the new peer does not know the network speeds of its seeds, just how much data it receives over time when it chooses a peer and receives data from them for one time step. The reward is how many bytes received in that time slot.

We want to be careful about defining the reward, because we want the agent to choose the peer that is transmitting the fastest. However, consider that network speeds may change, and the optimal seed to leech from will not always be the best.

4. Approach

Various algorithms will be considered, starting with epsilon-greedy and UCB (upper confidence bound). More complex bandit algorithms are considered as well.

5. Results

Text here.

5.1. Theoretical Results

Text here.

5.2. Experiment Results

Text here.

6. Discussion

Text here.

7. Conclusion and Future Research

Text here.

References

Schollmeier, R. A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications. In *Proceedings First International Conference on Peer-to-Peer Computing*, pp. 101–102, 2001. doi: 10.1109/P2P.2001.990434.