Network Management Research Group M-S.
Kim
Internet-Draft Y-G.
Hong
Intended status: Informational
ETRI
Expires: January 4, 2018 Y-H.
Han

KoreaTec
July 3,

2017

Intelligent Management using Collaborative Reinforcement Multi-agent System
draft-kim-nmrg-rl-00

Abstract

This document describes an intelligent reinforcement learning agent system to autonomously manage agent path-planning over a communication network. The main centralized node called by the global environment should not only manage all agents workflow in a hybrid peer-to-peer networking architecture and, but transfer and share information in distributed nodes. All agents in distributed nodes are able to be provided with a cumulative reward for each action that a given agent takes with respect to an optimized knowledge based on a to-be-learned policy over the learning process. A reward from the global environment is reflected to the next optimized action for autonomous path management in distributed networking nodes.

Copyright Notice

Table of Contents

1. Introduction

In large infrastructures such as transportation, health and energy systems, collaborative monitoring system is needed, where there are special needs for intelligent distributed networking systems with learning schemes. Agent Reinforcement Learning (RL) for autonomous network management, in general, is one of the challengeable methods in a dynamic complex cluttered environment over a network. The goal for autonomous network management using RL is self-management to manage optimized agent work-flow without minimal human dependency by learning process [RFC7575]. The system is needed by the development of computational multi-agents learning process in large distributed networking nodes, where the agents have limited and incomplete knowledge, and they only access local information in distributed networking nodes.

Reinforcement Learning can become an effective technique to transfer and share information among agents, as it does not require a priori knowledge of the agent behavior or environment to accomplish its tasks [Megherbi]. Such a knowledge is usually acquired and learned automatically and autonomously by trial and error.

Reinforcement Learning is Machine Learning techniques that will be adapted to the various networking environments for automatic networks[I-D.jiang-nmlrg-network-machine-learning]. Thus, this document provides motivation, learning technique, and use case for network machine learning.

2. Conventions and Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Motivation

3.1. General Motivation for Reinforcement Learning

Reinforcement Learning is a system capable of autonomous acquirement and incorporation of knowledge. It can continuously self-improve learning process with experience and attempts to maximize cumulative reward to manage an optimized learning knowledge by multi-agents-based monitoring systems[Teiralbar]. The maximized reward can be increasingly optimizing of learning speed for agent autonomous learning process.

3.2. Reinforcement Learning in networks

Reinforcement learning is an emerging technology in terms of monitoring and managing network system to achieve fair resource allocation for nodes within the wire or wireless mesh setting. Monitoring parameters of the network and adjusts based on the network dynamics can demonstrate to improve fairness in wireless environment Infrastructures and Resources [Nasim]. The fundamental goal for Reinforcement Learning is self-management, which is comprised of a couple of properties such as self-healing (adaptive function in the environment and heal problems automatically) and self-optimizing (function for automatically determine ways to optimize their behavior against a set of well-defined goals) [RFC7575].

3.3. Motivation in our work

There are many different networking management issues such as connectivity, traffic management, fast internet without latency and etc. We expect that ml-based mechanism such as reinforcement learning will provide solutions of networking issues with multiple cases against human operating capacities even if it is a challengeable area due to a multitude of reasons such as large state space search, complexity in giving reward, difficulty in agent action selection, and difficulty in sharing and merging learned information among the agents in a distributed memory node to be transferred over a communication network [Minsuk].

4. Related Works

4.1. Autonomous Driving System

Autonomous vehicle is capable of self-management for automotive driving without human supervision depending on optimized trust region policy by reinforcement learning that enables learning of more complex and special network management environment. Such a vehicle provides a comfortable user experience safely and reliably on interactive communication network [April] [Markus].

4.2. Game Theory

The adaptive multi-agent system, which is combined with complexities from interacting game player, has developed in a field of reinforcement learning. In the early game theory, the interdisciplinary work was only focused on competitive games, but Reinforcement Learning has developed into a general framework for analyzing strategic interaction and has been attracted field as diverse as psychology, economics and biology [Ann].

AlphaGo is also one of the game theories using reinforcement learning, developed by Google DeepMind. Even though it began as a small learning computational program with some simple actions, it has now trained on a policy and value networks of thirty million actions, states and rewards for optimal management using learning process.

4.3. Wireless Sensor Network (WSN)

Wireless sensor network (WSN) consists of a large number of sensors and sink nodes for monitoring systems to manage event parameters such as temperature, humidity, air conditioning, etc. Reinforcement learning in WSNs has been applied in a wide range of schemes such as cooperative communication, routing and rate control. The sensors and sink nodes are able to observe and carry out optimal actions on their respective operating environment for network and application performance enhancements [Kok-Lim].

4.4. Routing Enhancement

Reinforcement Learning is used to enhance multicast routing protocol in wireless ad hoc networks, where each node has different capability. Routers in the multicast routing protocol are determined to discover optimal route with a predicted reward, and then the routers create the optimal path with multicast transmissions to reduce the overhead in Reinforcement Learning[Kok-Lim].

5. Multi-agent Reinforcement Learning Technologies

5.1. Reinforcement Learning

Agent reinforcement Learning is ml-based unsupervised algorithms based on an agent learning process. Reinforcement Learning is normally used with a reward from centralized node (the global environment), and capable of autonomous acquirement and incorporation of knowledge. It is continuously self-improving and becoming more efficient as the learning process from an agent experience to optimize management performance for autonomous learning process.[Sutton][Madera]

5.2. Policy using Distance and Frequency

Distance and Frequency algorithm uses the state occurrence frequency in addition to the distance to goal. It avoids deadlocks and lets the agent escape the Dead, and it was derived to enhance agent optimal learning speed. Distance-and-Frequency is based on more levels of agent visibility to enhance learning algorithm by an additional way that uses the state occurrence frequency.[Al-Dayaa]

5.3. Distributed Computing Node

Autonomous multi-agent learning process for network management environment is related to transfer optimized knowledge between agents on a given local node or distributed memory nodes over a communication network.

5.4. Agent Sharing Information

This is a technique how agents can share information for optimal learning process. The quality of agent decision making often depends on the willingness of agents to share a given learning information collected by agent learning process. Sharing Information means that an agent would share and communicate the knowledge learned and acquired with or to other agents using reinforcement learning.

Agents normally have limited resources and incomplete knowledge during learning exploration. For that reason, the agents should take actions and transfer the states to the global environment under reinforcement learning, then it would share the information with other agents, where all agents explore to reach their destination via a distributed reinforcement reward-based learning method on the existing local distributed memory nodes.

MPI (Message Passing Interface) is used for communication way. Even if the agents do not share the capabilities and resources to monitor an entire given large terrain environment, they are able to share the needed information to manage collaborative learning process for optimized management in distributed networking nodes.[Chowdappa][Minsuk]

5.5. Sub-goal Selection

A new technical method for agent sub-goal selection in distributed nodes is introduced to reduce the agent initial random exploration with a given selected sub-goal.

[TBD]

6. Proposed Architecture for Reinforcement Learning

The architecture using Reinforcement Learning describes a collaborative multi-agent-based system in distributed environments as shown in figure 1, where the architecture is combined with a hybrid architecture making use of both a master and slave architecture and a peer-to-peer. The centralized node(global environment), assigns each slave computing node a portion of the distributed terrain and an initial number of agents.

```
+-------------+                           +-----------------+
|             |<......>| node 1 |<.......>|    terrain 1    |
|             |                           +-----------------+
| Global env. |
|  (node 0)   |                           +-----------------+
|             |<......>| node 2 |<.......>|    terrain 2    |
+-------------+                           +-----------------+
```

Figure 1: Hybrid P2P and Master/Slave Architecture Overview

Reinforcement Learning actions involve interacting with a given environment, so the environment provides an agent learning process with the elements as followings:

o Agent actions, states and cumulative rewards

o One or more obstacles, and goals

o Initially, random exploration in a given node

o Next, optimal explorations under reinforcement learning

Additionally, agent actions with states toward its goal as below:

o Agent continuously actions to avoid an obstacle based on its policy and move to one or more available positions until it reaches its goal(s)

o After an agent reaches its destination, it can use the information collected by initial random learning process to next learning process for optimal management

o Agent learning process is optimized in the following phase and exploratory learning trials

In shown as Figure2, we illustrate the fundamental architecture for relationship of an action, state and reward, and each agent explores to reach its destination(s) under reinforcement learning. The agent does an action that leads to a reward from achieving an optimal path toward its goal. Our works will be extended depending on the architecture.

```
                                    +---------------------+
        ....state and reward........+ Global Environment +|
<............
        .                           +---------------------
+            .
 +------+------
+                                                        .
 | Multi-agent
|                                                        .
 +------+------+                    +---------------
+                 .
        ............action.........>+  Destiantion
+...................
                                    +---------------+
```

Figure 2: RL work-flow Overview

7. Use case of Multi-agent Reinforcement Learning

7.1. Distributed Multi-agent Reinforcement Learning: Sharing Information Technique

In this section, we deal with case of a collaborative distributed multi-agent, where each agent has same or different individual destination in a distributed environment. Since sharing information scheme among the agents is problematic one, we need to expand on the work described by solving the challenging cases.

Basically, the main proposed algorithm is presented by distributed multi-agent reinforcement learning as below:

```
   +--------------------------------------------------------------------
+
   | Proposed Algorithm
|
   +--------------------------------------------------------------------
+
   | (1) Let Ni denote the number of node (i= 1, 2, 3 ...)
|
   |
|
   | (2) Let Aj denote the number of agent
|
   |
|
   | (3) Let Dk denote the number of destination
|
   |
|
   | (4) Place initial number of agents Aj, in random position (Xm,
|
   | Yn)
```

```
|
   |
|
   | (5) Every Aj in Ni
|
   |
|
   | -----> (a) Do initial exploration (random) to corresponding Dk
|
   |
|
   | -----> (b) Do exploration (using RL) for Tx denote the number of
|
   | trial
|
   +--------------------------------------------------------------------
+
```

Table 1: Proposed Algorithm

Random Trial
(1) Let Si denote the the current state
(2) Relinquish Si so that the other agent can occupy the position
(3) Assign the agent new position
(4) Update the current state Si -> Si+1

Table 2: Random Trial

Optimal Trial
(1) Let Si denote the the current state
(2) Let ACj denote an action
(3) Let DRm denote discount reward
(4) Choose ACj <- Policy(Si, ACj)
(5) Move an available posiion
(6) Update learning process in the global environment
(7) Update the current state Si < Si+1

Table 3: Optimal Trial

Multi-agent reinforcement learning in distributed nodes can improve the overall system performance to transfer or share information from one node to another node in following cases; expanded complexity in RL technique with various experimental factors and conditions, analyzing multi-agent sharing information for agent learning process.

7.2. Use case of Shortest Path-planning via sub-goal selection

Sub-goal selection is a scheme of a distributed multi-agent RL technique based on selected intermediary agent sub-goal(s) with the aim of reducing the initial random trial. The scheme is to improve the multi-agent system performance with asynchronously triggered exploratory phase(s) with selected agent sub-goal(s) for autonomous network management.

[TBD]

8. IANA Considerations

There are no IANA considerations related to this document.

9. Security Considerations

[TBD]

10. Acknowledgements

David Meyerm, who chief scientist and VP in Brocade, has provided significant comment and feedback for the draft.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC7575] Behringer, M., Pritikin, M., Bjarnason, S., Clemm, A., Carpenter, B., Jiang, S., and L. Ciavaglia, "Autonomic Networking: Definitions and Design Goals", RFC 7575, DOI 10.17487/RFC7575, June 2015, <http://www.rfc-editor.org/info/rfc7575>.

11.2. Informative References

[I-D.jiang-nmlrg-network-machine-learning]
Jiang, S., "Network Machine Learning", ID draft-jiang-nmlrg-network-machine-learning-02, October 2016.

[Megherbi]
"Megherbi, D. B., Kim, Minsuk, Madera, Manual., "A Study of Collaborative Distributed Multi-Goal and Multi-agent based Systems for Large Critical Key Infrastructures and Resources (CKIR) Dynamic Monitoring and Surveillance", IEEE International Conference on Technologies for Homeland Security", 2013.

[Teiralbar]
"Megherbi, D. B., Teiralbar, A. Boulenouar, J., "A Time-varying Environment Machine Learning Technique for Autonomous Agent Shortest Path Planning.", Proceedings of SPIE International Conference on Signal and Image Processing, Orlando, Florida", 2001.

[Nasim] "Nasim ArianpooEmail, Victor C.M. Leung, "How network monitoring and reinforcement learning can improve tcp fairness in wireless multi-hop networks", EURASIP Journal on Wireless Communications and Networking", 2016.

[Minsuk] "Dalila B. Megherbi and Minsuk Kim, "A Hybrid P2P and Master-Slave Cooperative Distributed Multi-Agent Reinforcement Learning System with Asynchronously Triggered Exploratory Trials and Clutter-index-based Selected Sub goals", IEEE CIG Conference", 2016.

[April] "April Yu, Raphael Palefsky-Smith, Rishi Bedi, "Deep Reinforcement Learning for Simulated Autonomous Vehicle Control", Stanford University", 2016.

[Markus] "Markus Kuderer, Shilpa Gulati, Wolfram Burgard, "Learning Driving Styles for Autonomous Vehicles from Demonstration", Robotics and Automation (ICRA)", 2015.

[Ann] "Ann Nowe, Peter Vrancx, Yann De Hauwere, "Game Theory and Multi-agent Reinforcement Learning", In book: Reinforcement Learning: State of the Art, Edition: Adaptation, Learning, and Optimization Volume 12", 2012.

[Kok-Lim] "Kok-Lim Alvin Yau, Hock Guan Goh, David Chieng, Kae Hsiang Kwong, "Application of reinforcement learning to wireless sensor networks: models and algorithms", Published in Journal Computing archive Volume 97 Issue 11, Pages 1045-1075", November 2015.

[Sutton] "Sutton, R. S., Barto, A. G., "Reinforcement Learning: an Introduction", MIT Press", 1998.

[Madera] "Madera, M., Megherbi, D. B., "An Interconnected Dynamical System Composed of Dynamics-based Reinforcement Learning Agents in a Distributed Environment: A Case Study", Proceedings IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, Italy", 2012.

[Al-Dayaa]
"Al-Dayaa, H. S., Megherbi, D. B., "Towards A Multiple-Lookahead-Levels Reinforcement-Learning Technique and Its Implementation in Integrated Circuits", Journal of Artificial Intelligence, Journal of Supercomputing. Vol. 62, issue 1, pp. 588-61", 2012.

[Chowdappa]
"Chowdappa, Aswini., Skjellum, Anthony., Doss, Nathan, "Thread-Safe Message Passing with P4 and MPI", Technical Report TR-CS-941025, Computer Science Department and NSF Engineering Research Center, Mississippi State University", 1994.

Authors' Addresses

Min-Suk Kim
ETRI
161 Gajeong-Dong Yuseung-Gu
Daejeon 305-700
Korea

Phone: +82 42 860 5930
Email: mskim16@etri.re.kr

Yong-Geun Hong
ETRI
161 Gajeong-Dong Yuseung-Gu
Daejeon 305-700
Korea

Phone: +82 42 860 6557
Email: yghong@etri.re.kr

Youn-Hee Han
KoreaTec
Byeongcheon-myeon Gajeon-ri, Dongnam-gu
Choenan-si, Chungcheongnam-do
330-708
Korea

Phone: +82 41 560 1486
Email: yhhan@koreatec.ac.kr