

Intelligent Management using Collaborative Reinforcement Multi- agent System

M. Kim, Y. Hong (ETRI)

Y. Han (Koreatech.)

NMRG Meeting@IETF 99 – Prague

2017.07.20

draft-kim-nmrg-rl-00

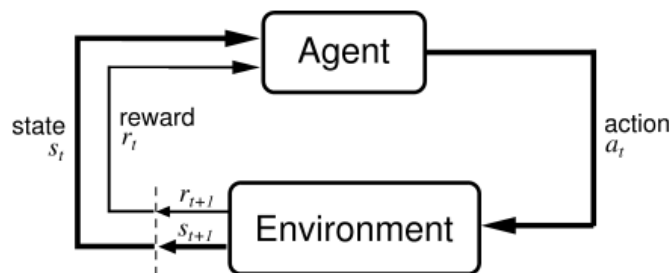
[\(https://datatracker.ietf.org/doc/draft-kim-nmrg-rl/?include_text=1\)](https://datatracker.ietf.org/doc/draft-kim-nmrg-rl/?include_text=1)

Table of Contents

| | |
|---|----|
| 1. Introduction | 3 |
| 2. Conventions and Terminology | 3 |
| 3. Motivation | 3 |
| 3.1. General Motivation for Reinforcement Learning | 3 |
| 3.2. Reinforcement Learning in networks | 4 |
| 3.3. Motivation in our work | 4 |
| 4. Related Works | 4 |
| 4.1. Autonomous Driving System | 4 |
| 4.2. Game Theory | 4 |
| 4.3. Wireless Sensor Network (WSN) | 5 |
| 4.4. Routing Enhancement | 5 |
| 5. Multi-agent Reinforcement Learning Technologies | 5 |
| 5.1. Reinforcement Learning | 5 |
| 5.2. Policy using Distance and Frequency | 5 |
| 5.3. Distributed Computing Node | 6 |
| 5.4. Agent Sharing Information | 6 |
| 5.5. Sub-goal Selection | 6 |
| 6. Proposed Architecture for Reinforcement Learning | 6 |
| 7. Use case of Multi-agent Reinforcement Learning | 8 |
| 7.1. Distributed Multi-agent Reinforcement Learning: Sharing Information Technique | 8 |
| 7.2. Use case of Shortest Path-planning via sub-goal selection | 9 |
| 8. IANA Considerations | 10 |
| 9. Security Considerations | 10 |
| 10. Acknowledgements | 10 |
| 11. References | 10 |
| 11.1. Normative References | 10 |
| 11.2. Informative References | 10 |
| Authors' Addresses | 12 |

Reinforcement Learning

- ◆ One of the machine learning algorithm
- ◆ An agent can explore an environment with **taking action** and **its current state** toward destination
- ◆ The **cumulative reward** to find an **optimal value** by agents
- ◆ The **policy from maximizing cumulative reward** for learning process



$$f(S_{k+1}, A_{k+1}) = \max D(S_k, A_k)$$
$$a = \operatorname{argmax}(R(S_k, A_k) \text{ where } a_k \in A$$

What is Motivation ?

◆ General Motivation for Reinforcement Learning (RL)

- ✓ Reinforcement Learning (RL) is **a system capable of autonomous acquirement and incorporation of knowledge**

◆ Reinforcement Learning (RL) in networks (Intelligent Management)

- ✓ In terms of **networking monitoring system**, to achieve **fair resource allocation** for nodes within the wire or wireless mesh setting

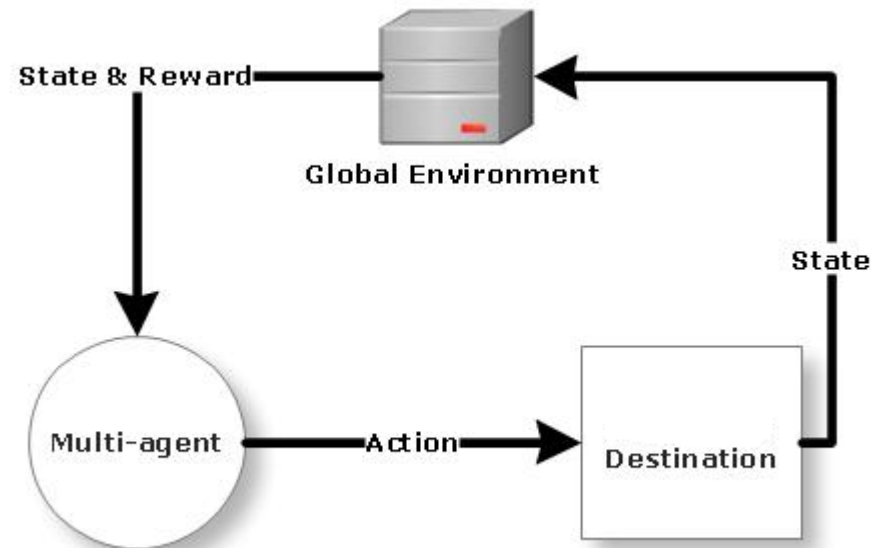
◆ Motivation in our work

- ✓ Networking issues such as **connectivity, traffic management, fast internet without latency and etc.**
- ✓ ML-based mechanisms such as reinforcement learning [RL] will provide **network solutions** with multiple cases **against human operating capacities**

Proposed Method

◆ Multi-agent Reinforcement Learning process cycle

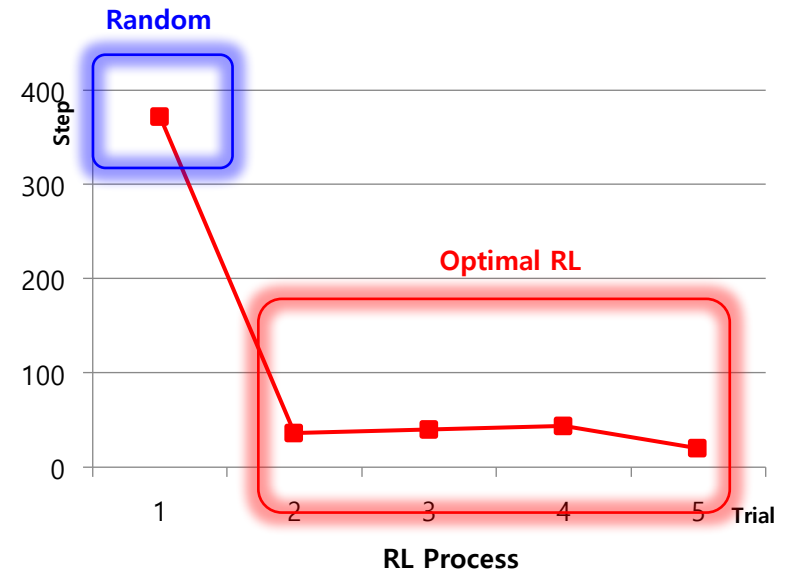
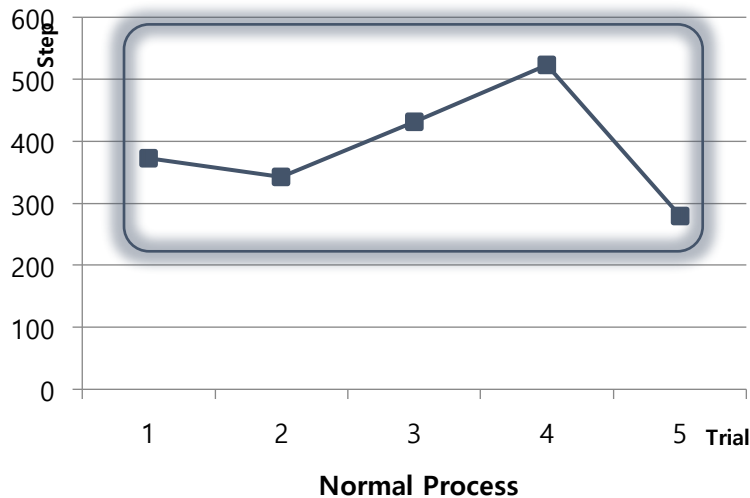
- ✓ Reinforcement Learning (RL)
- ✓ Action, State and Reward
- ✓ Policy for Monitoring and Management



Proposed Method 1 (cont'd)

◆ Normal Process vs RL Learning Process

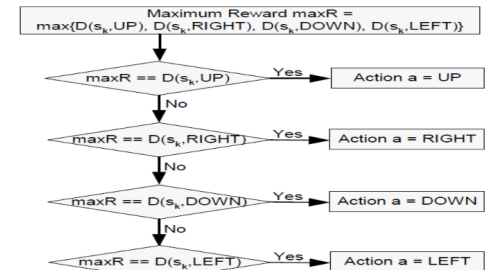
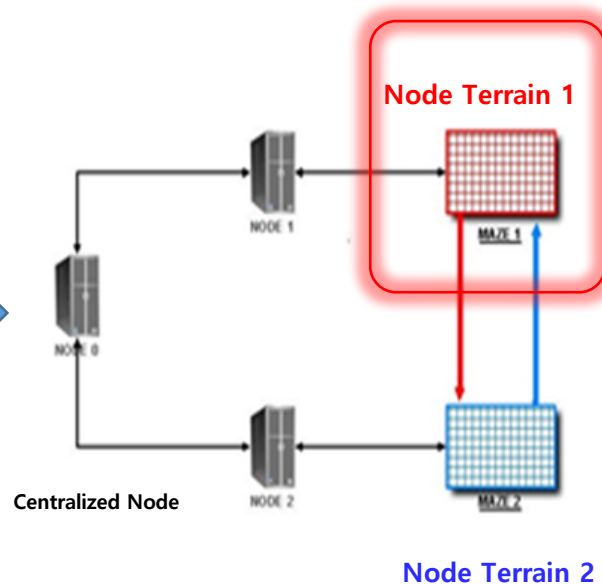
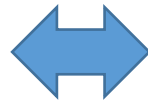
- ✓ Initial random exploration for learning process
- ✓ Random for the 1st processing
- ✓ Updated policy after random process



Proposed Method 1 (cont'd)

◆ Multi-agent Reinforcement Learning (RL) Technologies

- ✓ Reinforcement Learning (RL)
- ✓ Policy using Distance and Frequency
- ✓ Distributed Computing Node

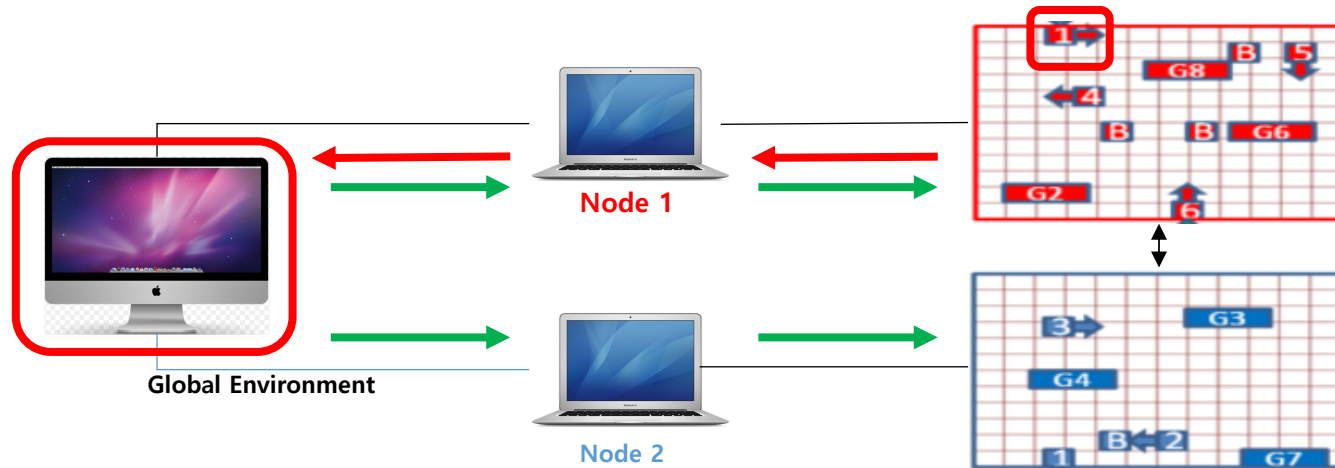


Agent → Goal

Proposed Method 2

◆ Agent Sharing Information

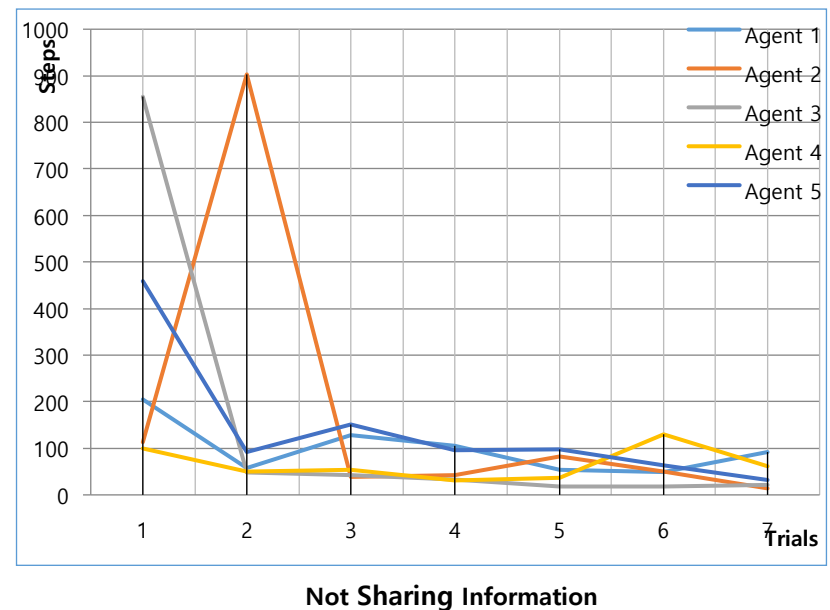
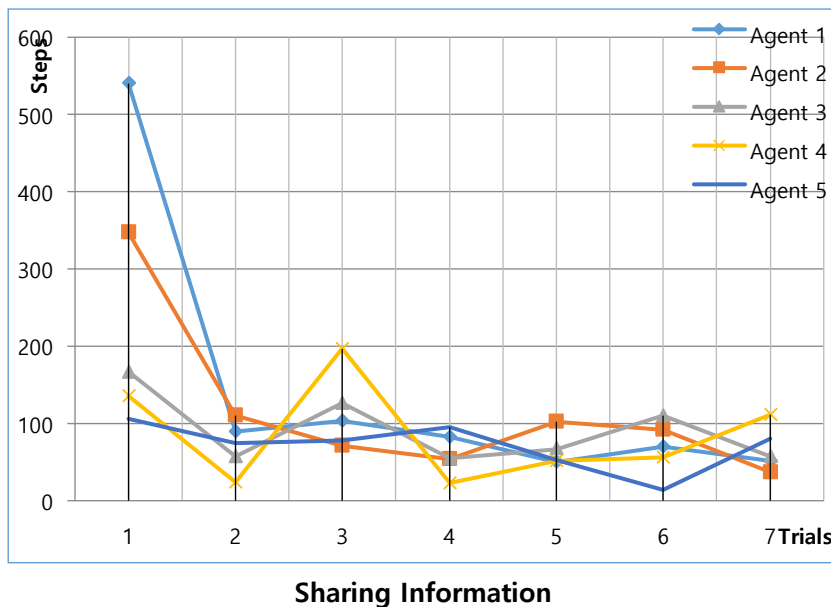
- ✓ Agents should **take actions and transfer the states** to the global environment under reinforcement learning (RL), then **it would share the information with other agents**



Proposed Method 2 (cont'd)

◆ Agent Sharing Information

- ✓ Effect of sharing information vs not sharing information
- ✓ Initially Random process for the 1st trial
- ✓ Efficient performance with shared information between agents



With Possible Network Scenario with RL

◆ Autonomous Driving System

- ✓ Self-automotive driving without human supervision depending on **optimized trust region policy**

◆ Wireless Sensor Network (WSN)

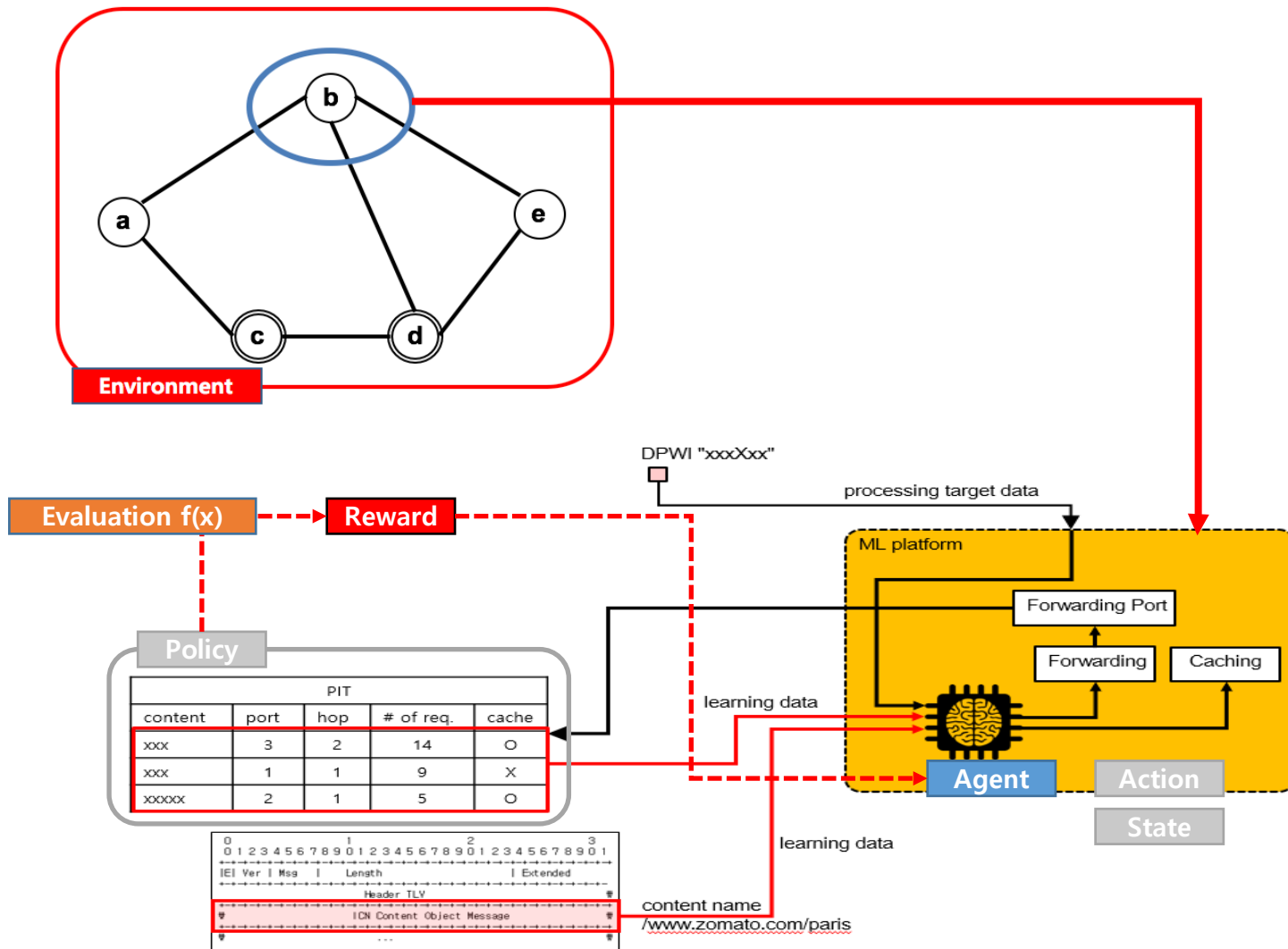
- ✓ RL in WSNs has been applied in a wide range of schemes such as **cooperative communication, routing and rate control for intelligent monitoring and management**

◆ Routing Enhancement

- ✓ 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**

On-going work

◆ Adaptive Networking model on Routing using RL



Next Step

- ◆ **We create adaptive intelligent networking management Model with RL**
- ◆ **We also set up RL scenario and preprocessing for Dataset**

Thank you

Comment or Question?

Email: mskim16@etri.re.kr