

Intelligent Network Management using Reinforcement Learning

M. Kim, [Y. Hong](#) (ETRI)

Y. Han (Koreatech.)

T. Ahn (KT)

K. Kim (ETRI)

NMRG Meeting@IETF 102 – Montreal

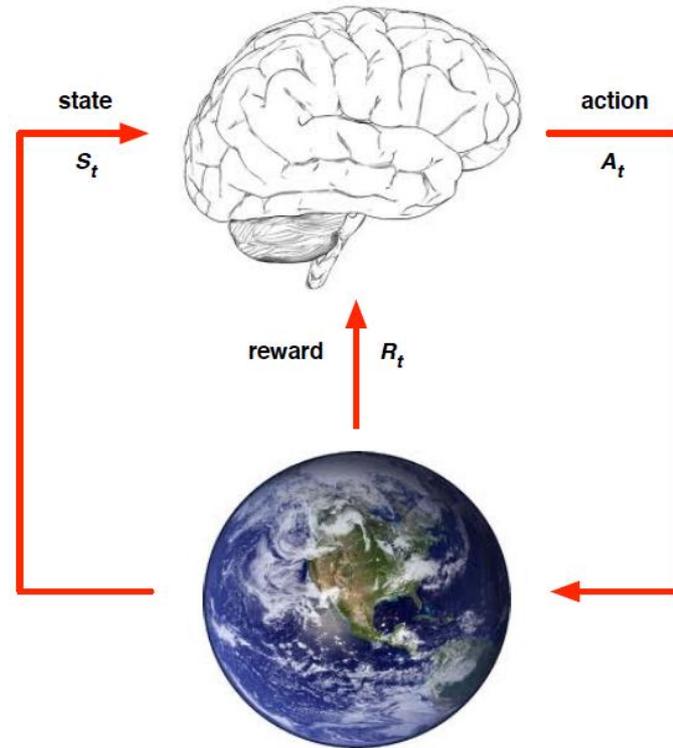
2017.07.19

draft-kim-nmrg-rl-03

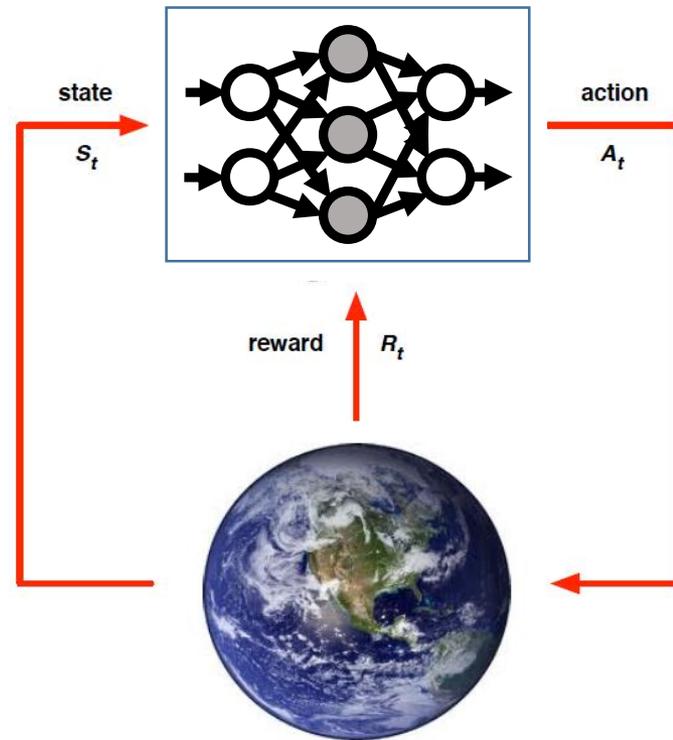
(<https://datatracker.ietf.org/doc/draft-kim-nmrg-rl/>)

Table of Contents		
1.	Introduction	3
2.	Conventions and Terminology	4
3.	Motivation	4
3.1.	General Motivation for Reinforcement Learning	4
3.2.	Reinforcement Learning in networks	4
3.3.	Deep Reinforcement Learning in networks	4
3.4.	Motivation in our work	5
4.	Related Works	5
4.1.	Autonomous Driving System	5
4.2.	Network Defect Prediction	5
4.3.	Wireless Sensor Network (WSN)	6
4.4.	Routing Enhancement	6
4.5.	Routing Optimization	6
4.6.	Game Theory	6
5.	Intelligent Machine Learning Technologies	7
5.1.	Reinforcement Learning (RL)	7
5.2.	Deep Learning (DL)	7
5.3.	Deep Reinforcement Learning (DRL)	7
5.4.	Advantage Actor Critic (A2C)	8
5.5.	Asynchronously Advantage Actor Critic (A3C)	8
5.6.	Policy using Distance and Frequency	9
5.7.	Distributed Computing Node	9
5.8.	Agent Sharing Information	9
6.	Proposed Architecture	9
6.1.	Architecture for Reinforcement Learning	10
6.2.	Architecture for Deep Reinforcement Learning	11
7.	Use case of Reinforcement Learning	11
7.1.	Distributed Multi-agent Reinforcement Learning (RL): Sharing Information Technique	12
7.2.	Intelligent Edge Computing technique for Traffic Control using Deep Reinforcement Learning	13
7.3.	Edge computing system in a field of construction works using Reinforce Learning	14
7.4.	Fault prediction for core-network using Deep Learning	14
8.	IANA Considerations	15
9.	Security Considerations	15
10.	References	15
10.1.	Normative References	15
10.2.	Informative References	15
	Authors' Addresses	17
Kim, et al.	Expires January 3, 2019	[Page 2]
Internet-Draft	draft-kim-nmrg-rl-03	July 2018

Reinforcement Learning (RL)



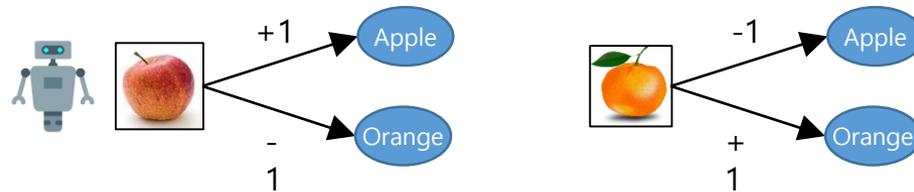
$DRL = RL + \text{Deep Learning}$



Decision Problem

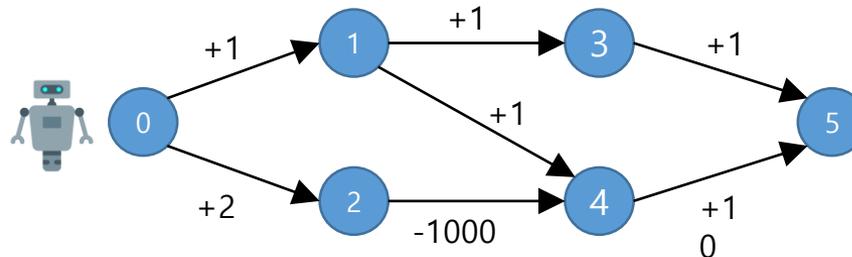
◆ Supervised Learning

- Single Decision



◆ Reinforcement Learning

- Sequential Decision



Robot Icon from <https://visualpharm.com/free-icons/robot%203-595b40b85ba036ed117da97e>

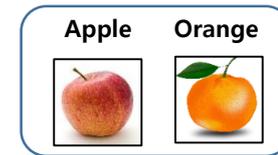
Apple and Orange from <https://pixabay.com>

Sequential Decision example from 'Reinforcement Learning: A User's Guide, Bill Smart, Washington University in St. Louis.'

Example vs Experience

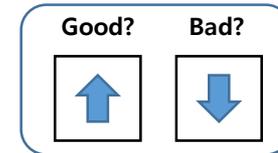
◆ Supervised Learning

- Learning by Example
- Labeled examples are given.
- Problems are defined by ‘examples’



◆ Reinforcement Learning

- Learning by Experience (or Interaction)
- Examples of ‘Good’ and ‘Bad’ (value-based) actions are not given
- Tasks are implicitly given through ‘**Reward**’



*Robot Icon from <https://visualpharm.com/free-icons/robot%203-595b40b85ba036ed117da97e>
Apple and Orange from <https://pixabay.com>*

Sequential Decision example from 'Reinforcement Learning: A User's Guide, Bill Smart, Washington University in St. Louis.'

With Possible Network Scenario with RL

◆ **Intelligent Traffic Management**

- **Edge-based traffic management system** for intelligent AI control using deep reinforcement learning to provide **fast response time, reliability and security**

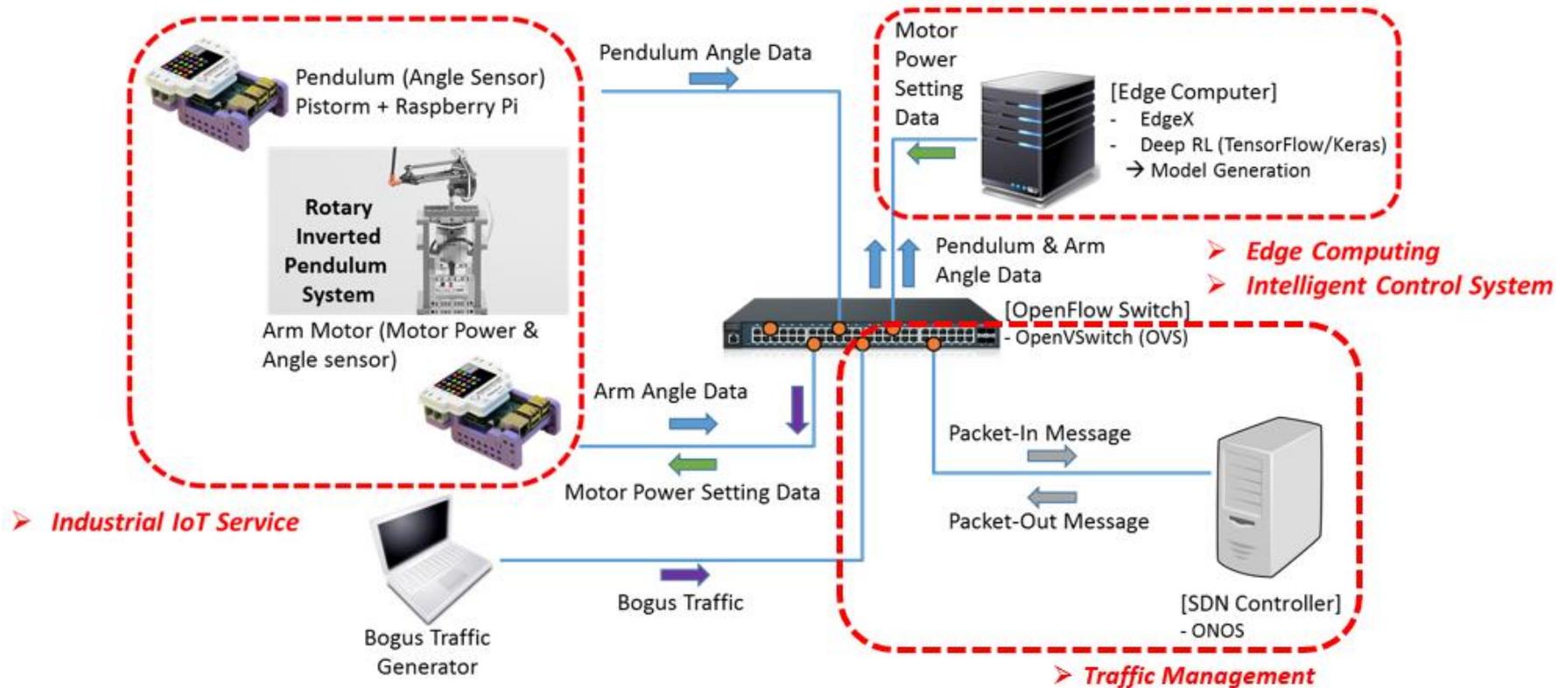
◆ **Network Defect Prediction**

- Deep reinforcement learning emerged **the preferred solutions to manage and monitor the networking equipment** (LTE core, router and switch) prevented by the **networking failure risk**

◆ **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**

Our Research: Edge-based Intelligent Traffic Control system using DRL



Step1: DRL Simulation for Edge-based Control system

◆ **Simulation Environment**

- OpenAI Gym (open source)
 - A toolkit for developing and comparing reinforcement learning algorithms (Environment)
- DRL Definition
 - State, Action and Reward
 - Step, Episode

◆ **Improvement of Performance (Simulation)**

- Deep-Q Network, A2C (Advantage-Actor-Critic), A3C (Asynchronous + A2C)

Simulation Environment for the System

◆ Pendulum Environment (Pendulum v0)

- Observation

Num	Observation	Min	Max
0	$\cos \theta$	-1.0	1.0
1	$\sin \theta$	-1.0	1.0
2	$\theta' (= \frac{d\theta}{dt})$	-8.0	8.0

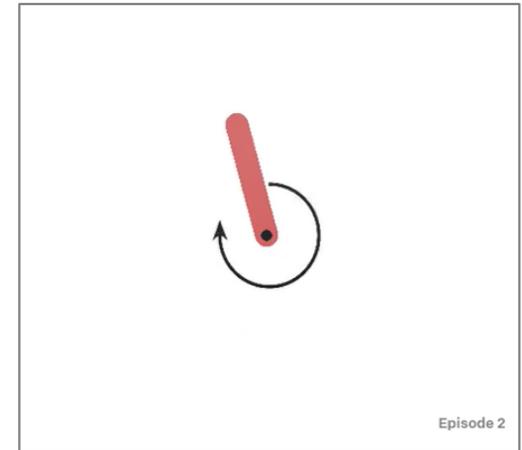
- Action

Num	Action	Min	Max
0	Joint effort	-2.0	2.0

- Reward (per step)

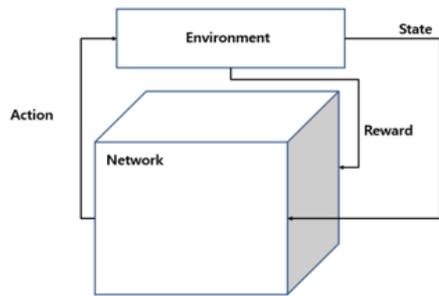
$$-(\theta^2 + 0.1 * \theta'^2 + 0.001 * action^2) = -(\pi^2 + 0.1 * 8^2 + 0.001 * 2^2) = -16.2736044 \rightarrow (** \text{ lowest cost})$$

** highest cost = 0

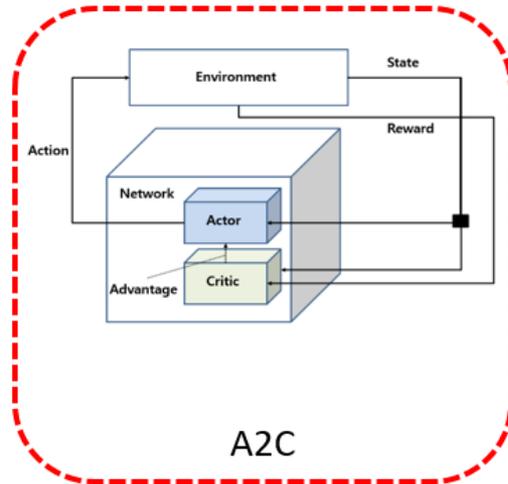


Simulation Result (render())

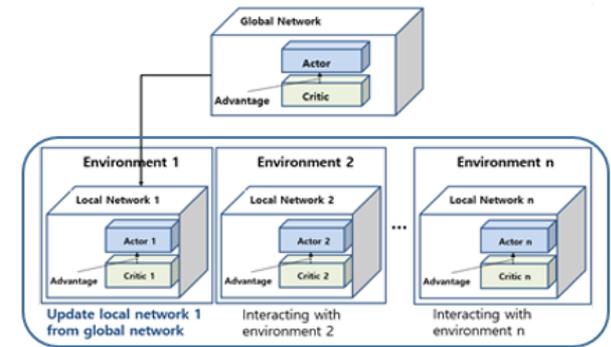
Simulation progress for Improvement of Performance



Deep-Q Network

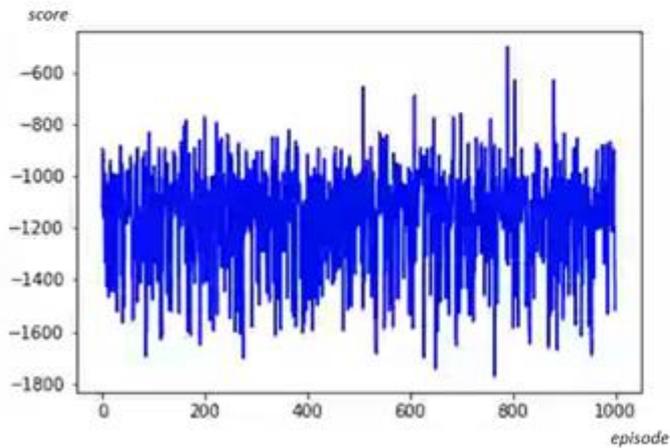


A2C

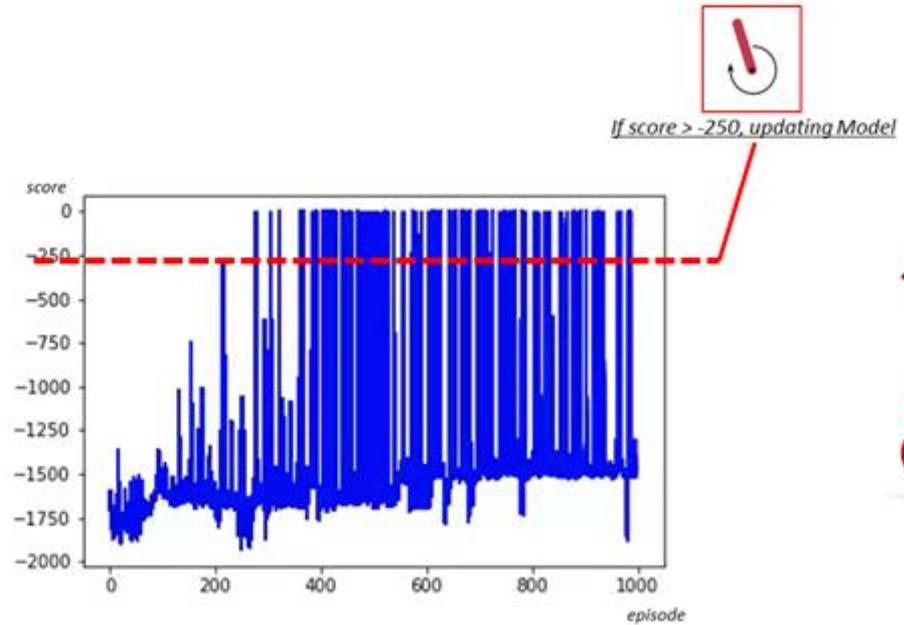


A3C

Simulation progress for Improvement of Performance



Deep-Q Network



A2C



A3C



More successful results than DQN

Next Steps for Contribution of NMRG :

- ◆ Adaptive open-flow network to intelligently manage and control traffic data
- ◆ Research and Implementation based on an edge-based platform
- ◆ AI engine in Edge-based platform using intelligent machine learning approach (DRL)

Thank you

Comment or Question?

Email:

mskim16@etri.re.kr

yghong@etri.re.kr