

Industrial General Reinforcement Learning Control Framework System based on Intelligent Edge

Kwihoon Kim*, Yong-Geun Hong*

* ETRI, 218 Gajeong-ro, Yuseong-gu, Daejeon, 34129, Korea

kwihoon@etri.re.kr, yghong@etri.re.kr

Abstract— *This paper is about the intelligent edge-based reinforcement learning control framework technology for manufacturing field solution and features large-scale learning, scalable edge distribution technology that can be applied to various task. In this paper, two items are proposed as features. The first proposes the General Reinforcement Learning Framework(GRLF) in manufacturing, and the second proposes the edge solution technology based on the GRLF in manufacturing. We apply the industrial solution such as grid sorter system for example. As a result of the industrial GRLF based grid sorter system, it was confirmed that when a total of 100 deliveries are randomly received into the grid sorter system by any emitter, all shipments are 100% accurate. It also classifies approximately 0.5 deliveries per step. This shows the efficiency of classifying around 30 deliveries per minute, assuming a step is performed per second.*

Keywords — reinforcement learning, deep learning, intelligent edge, supervised learning

I. INTRODUCTION

AI, especially Deep Learning, which is cited as a factor enabling the fourth industrial revolution, has developed rapidly. As a result, it has become possible to make relatively accurate judgments and predictions that could not be made before and is being applied to a variety of industries. For example, in the field of self-driving or unmanned mobile devices, AI technology is introduced to identify situations or predict movements by analyzing sensor data collected. In addition, the robot sector utilizes image-based person/thing recognition, voice recognition, etc. In particular, intelligent manufacturing that combines manufacturing with AI technologies is drawing attention[1].

There is a large amount of data generated by IoT (Internet of Things) such as sensors installed at smart manufacturing sites. Based on this, the optimal control operation technology required for the manufacturing process is critical.

For this optimal control, the deep reinforcement learning has been considered a lot recently. The deep reinforcement learning works by recognizing the present state of the environment that the agent wants to control and selecting actions, and the environment returns rewards for actions taken

by the agent. The agent repeats this process and learns a control policy that maximizes compensation[2-3].

In traditional manufacturing processes, sorting, stacking, and anomaly detection are mostly dependent on operator observations. Therefore, accuracy is inconsistent depending on the operator's condition and proficiency. It is also time-consuming depending on the operator's ability. We need a way to handle this automatically. Reinforcement learning technology automatically provides action and reward for optimum policy through the machine learning techniques[4-5].

This paper proposes to apply reinforcement learning on the basis of edge computing to the manufacturing sector. reinforcement learning is commonly used to solve specific problems. As a specific example, a number of services, such as smart cities, smart factories, and smart health care, are being developed around separate reinforcement learning technologies when providing enhanced learning, and the costs of developing and maintaining enhanced learning technologies are expected to increase more and more. Then, in order to solve a number of problems, we need to build a separate reinforcement learning solution. In order to improve this problem, this paper proposes an intelligent edge computing-based reinforcement learning control framework. In particular, intelligent edge-based reinforced learning control framework technology is proposed that can be real-time control in manufacturing and applied to a variety of tasks. In this paper, we propose the general reinforcement learning framework of technologies such as compatibility of the existing frame reinforcement learning and integration work that reflects the technology. It is effective to collect information from various sensors through intelligent edge computing equipment, analyze the situation at the site through reinforcement learning technology, and speed up work by classifying goods based on it[6].

In addition, this study proposes the grid sorter systems based on general reinforcement learning framework, which are available at smart manufacturing sites, defines problems to be solved, and describes the collaborative multi-agent reinforcement learning control structure and control behaviour elicitation process for such shipment classification based on reinforcement learning.

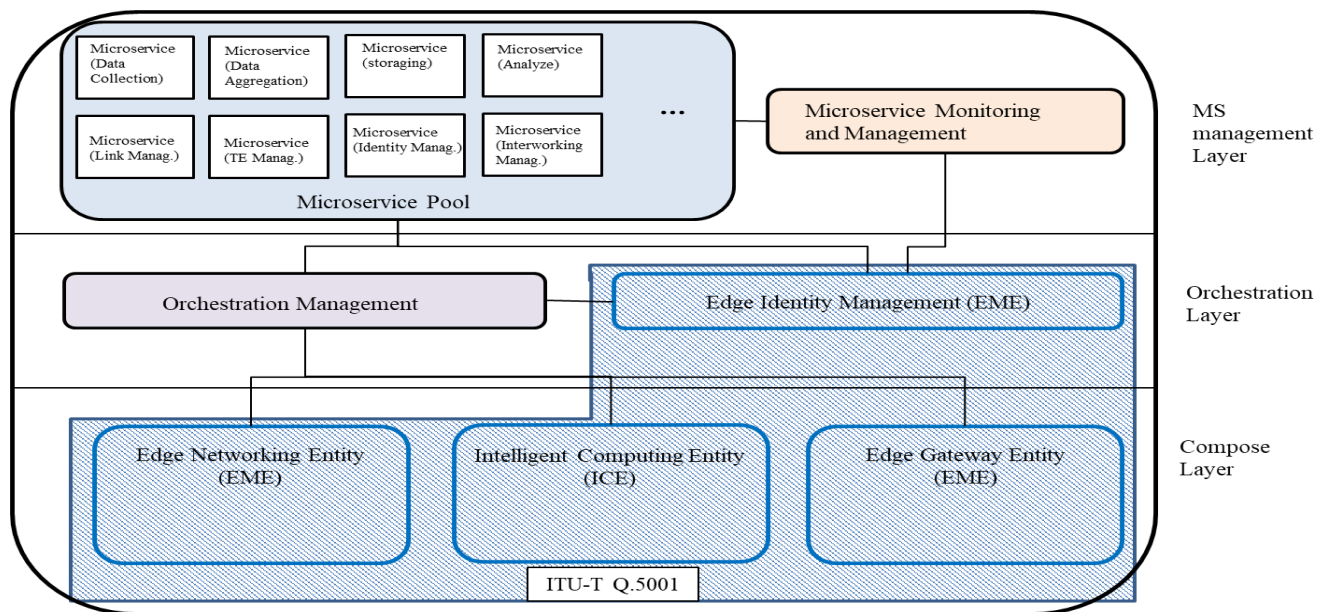


Figure 1. Core General Reinforcement framework based on Intelligent Edge Concept

The delivery classification system proposed in this study is similar to the GridSorter produced by the Karlsruhe Institute of Technology (KIT) University in Germany. However, GridSorter at KIT is a rule-based system that is controlled by predetermined rules for each state[7]. This rule-based approach has great limitations. The complexity of the system makes it very difficult to configure a sophisticated optimal rule, and the verification of the rules that you have configured also becomes difficult. Even if a well-modelled rule is created, changes to the system configuration are also very difficult.

On the other hand, applying reinforcement learning allows such rule configurations to be completed automatically and effectively with no human intervention. In addition, since enhanced learning-based artificial intelligence learns optimal behaviour and achieves its goals within a properly configured environment, it can be used universally by repeating only additional learning as the system configuration changes.

II. SYSTEM MODEL

Fig. 1 illustrates the technical composition of the core common reinforcement learning framework.

The framework is defined first by microservices, each of which is one or several workable functions. Then, perform microservices management. Finally, microservices are configured and implemented to develop IEC entities. Therefore, a framework can be defined as a protocol interface, procedure, and message format based on the compose layer, orchestration layer, and microservices management layer.

- **Compose Layer:** The IEC function blocks for each IEC entity are allocated and performed by organizing microservices from the microservice pool.
 - ENE, ICE, and EGE entities are implemented by organizing microservices.
 - Each microservices communicate with each other to perform the IEC function.
 - Microservices are managed as security by using the identification manager.
- **Orchestration layer:** Each microservices is identified and orchestration is performed securely.
 - Orchestration management configures and manages microservices for each IEC entity.
 - EME identifies all microservices before they are constructed.
 - This defines various IEC services.
- **Microservice(MS) Management Layer:** Each microservices is continuously monitored and updated.
 - It manages microservice pools.
 - It monitors each microservice condition.
 - It registers and updates microservices.
- **The central node is as follows.**
 - Information-oriented IoT networking node (ENE), information processing and analysis node (ICE), data transmission gateway node (EGE), and name management system (EME)
- **Typical microservice in microservice pools are as follows:**
 - Data collection, aggregation, storing, analysis, link management, TE management, reinforcement management

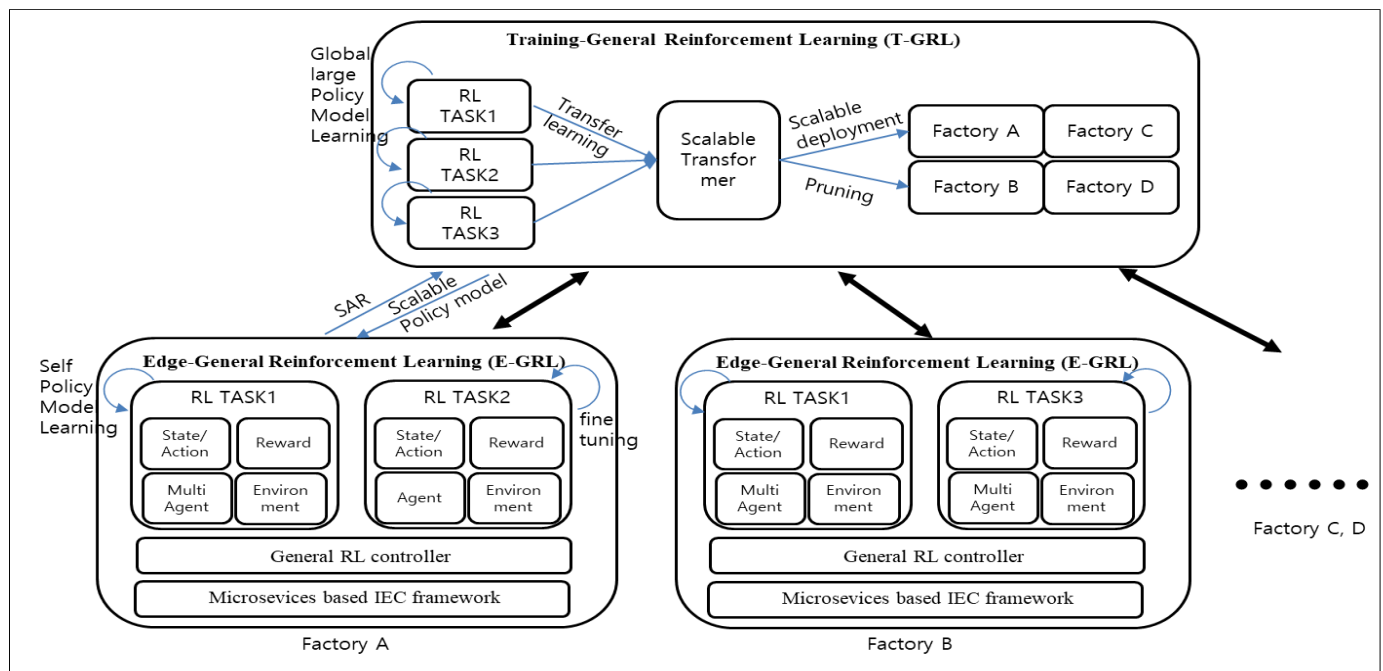


Figure 2. The system configuration of Industrial General Reinforcement Learning Framework

III. PROPOSED GENERAL REINFORCEMENT LEARNING CONTROL FRAMEWORK

A. Industrial General Reinforcement Learning Framework (GRLF) System concept & configuration

Fig.2 describes the technical composition of the common reinforcement learning framework in the manufacturing sector

It is largely divided into Training-General Reinforcement Learning (T-GRL) and Edge-General Reinforcement Learning (E-GRL). E-GRL is installed in the local area of each factory, and T-GRL is used for learning in the center as a central equipment. E-GRL of each factory collects sensing information and learning results and posts them on T-GRL, which controls the entire organization. The results learned in T-GRL will be made into a model that is suitable for each factory and distributed to each E-GRL. Models distributed in each E-GRL will perform fine tuning to learn self policy model for each factory.

B. Industrial GRLF based Edge Solution

Fig.3 shows the configuration of edge solution technology based on GRLF in industry. This architecture consists of many functional blocks distributed to each IEC entity, such as ENE, ICE, and EGE. General RL controller performs common env, agents, and PLC functions to perform reinforcement lessons. The General RL controller can be configured with one or multiple microservice in the microservice pool.

A typical service is the Multi-stacker or Grid Sorter.

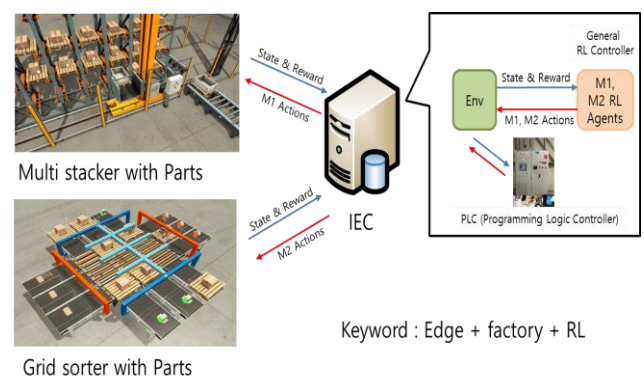


Figure 3. The configuration of edge solution technology based on GRLF in industry

C. Industrial GRLF based Grid Sorter system example

Sortation is the process of delivering specific items in a conveyor system to a specific destination. A grid sorter is a device that can move an object left or right or up and down. The performance figures for Grid server are largely two measured Accuracy and Efficiency at the destination. Accuracy is measured by the number of items classified/unit time. Efficiency is measured as a correctly classified item/all item. Sensing event information is transferred from the physical environment to the cyber environment and collected by converting it into state and return information that each agent can learn. Based on that information, each agent performs an actuator command.

IV. CONCLUSIONS

This paper is about intelligent edge based general reinforcement learning control framework technology for manufacturing field solution and proposes large capacity learning, scalable edge distribution technology that can be applied to various task. As a specific example, a number of services, such as smart cities, smart factories, and smart health care, are developed around separate reinforcement learning technologies when providing enhanced learning. As a result, the cost of developing and maintaining the enhanced learning technology is expected to increase as a result of numerous reinforcement learning technologies. However, it is effective to collect information from various sensors through intelligent edge computing equipment, analyse the situation at the site through enhanced learning technology, and speed up the work by classifying goods based on it. In addition, the following four expected effects of this paper are as follows: The first is expected to enable the replacement of the existing expensive control engineering advanced workforce for the delivery classification system control area through automatic optimal control through artificial intelligence. Second, import substitution, export increase and new market preoccupation effect through grid sorter technology that can be applied to smart manufacturing sectors are expected. Third, expansion of the market through new logistics systems is expected, and the effect of job creation is expected. Fourth, it is expected to gain market dominance due to technological superiority with global companies such as China, Japan and the United States, and thus increase exports.

ACKNOWLEDGMENT

This work was supported by the National Research Council of Science and Technology (NST) grant by the Korea government (MSIP) (No.CRC-15-05-ETRI).

REFERENCES

- [1] N. Kato, et al., "The Deep Learning Vision for Heterogeneous Network Traffic Control: Proposal, Challenges, and Future Perspective," IEEE Wireless Communication, 2017.
- [2] V. Mnih et al., "Playing Atari with Deep Reinforcement Learning," arxiv:1312.5602, 2013.
- [3] R.S. Sutton et al., "Policy Gradient Methods for Reinforcement Learning with Function Approximation," in Proc. Int. Conf. Neural Inf. Process. Syst., Denver, CO, USA, 2000, pp. 1057-1063.
- [4] R. Lowe et al., "Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments," in Conf. Neural Inf. Process. Syst., Long Beach, CA, USA, 2017.
- [5] M. Tan, "Multi-Agent Reinforcement Learning: Independent vs. Cooperative Agents," in Int. Conf. Mach. Learning(ICML), Amherst, MA, USA, 1993.
- [6] K. Kim et al., "signalling requirement and architecture of intelligent edge computing," ITU-T Q.5001, Oct. 2018.
- [7] Z. Seibold et al., "Layout-optimized sorting of goods with decentralized controlled conveying modules", 2013 IEEE International Systems Conference, Apr. 2013

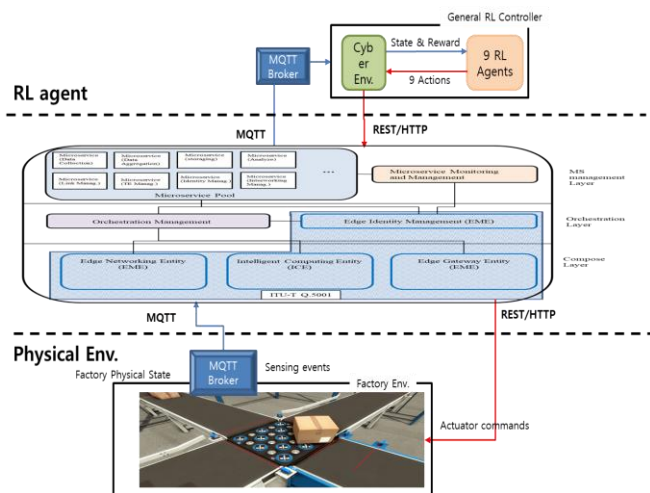


Figure 4. The configuration example of grid sorter technology based on GRLF in industry

Fig.4 shows the configuration example of grid sorter technology based on GRLF in industry. In this paper, using Factory I/O, the 3D Simulation Software provided by Real Games Inc. of Portugal. As shown in Fig.4, N-grid sorter was designed and implemented, and reinforcement learning was applied. In the proposed N-grid sorter, N stands for the number of sorter comprising one side of the grid model, and the study was started with N=3. Therefore, there are a total of nine sorter systems in this system. On the other hand, there are six emitters that ship arbitrary shipments to N-grid system and six removers that receive fixed shipments from N-grid system. Each shipment received in the system has its own type (or destination), and each remover has its own type (or destination) of shipment. The optimal control goal of this system is to receive as many deliveries as possible in a limited time and to classify them into remover, which is set to 100% accuracy through 9 sorter's classification collaborative control. In other words, the system requires high performance in terms of efficiency and accuracy.

The status information to be used in the reinforcement learning is determined by the location and type of shipment placed on the nine sorter and emitter. If any shipment is classified correctly, the agent will receive +2 compensation, or -2 compensation if it is misclassified or if it is not classified more than a certain step. In addition, an agent who owns the shipment at every step in order to learn to quickly categorize it as an optimal route will receive a -0.5 reward

As a result of learning, it was confirmed that when a total of 100 deliveries are randomly received into the N-grid system by any emitter, all shipments are 100% accurate. It also classifies approximately 0.5 deliveries per step. This shows the efficiency of classifying around 30 deliveries per minute, assuming a step is performed per second.



Kwihoon Kim studied in KAIST, M.S. degree and Ph.D. degree in 2000 and 2019, respectively. He worked in LG DACOM 2000~2005 and is a principle researcher in ETRI since 2005. He is a principal research engineer of intelligent IoE networking research team, ETRI now. He is an editor and rapporteur of ITU-T SG11 since 2006. His interested fields are Fog/edge computing, Internet of Things, 5G/IMT2020, deep learning, machine learning, reinforcement learning, GAN and knowledge-

converged intelligent service.



Yong-Geun Hong is a director of the Electronics and Telecommunications Research Institute (ETRI). He is a project leader of IoT Network and Intelligent Network R&D projects in ETRI. He received his B.S., M.S. and Ph.D in computer engineering from the Kyungpook National University, Daegu, Korea. He is also working for the IoT related standardization at IETF (Internet Engineering Task Force) and ITU-T. His research interests include IPv6, Internet Mobility, M2M/IoT, and

network intelligence.