

소프트웨어 정의 네트워크에서 강화학습을 활용한 원격 디바이스 제어 시스템 설계

임현교*, 김주봉**, 김민석***, 홍용근***, 한연희***

*한국기술교육대학교 창의융합공학협동과정

**한국기술교육대학교 컴퓨터공학과

***한국전자통신연구원

e-mail : {glenn89, rlawnqhd3, yhhan}@koreatech.ac.kr, {mskim16, yghong}@etri.re.kr

Design of A Remote Device Control System Using Reinforcement Learning in Software Defined Networks

Hyun-Kyo Lim*, Ju-Bong Kim**, Min-Suk Kim***, Yong-Geun Hong***, Youn-Hee Han**

*Interdisciplinary Program in Creative Engineering, KoreaTech University

**Dept. of Computer Engineering, KoreaTech University

***Electronics and Telecommunications Research Institute

요 약

최근, Industry 4.0 과 IoT 기기의 보급으로 인하여 수많은 센서와 액추에이터, 모바일 기기 등이 Cyber-Physical System 을 통해 네트워크와 연결되며, 더 효율적인 시스템을 요구한다. 이를 위하여, EdgeX 와 SDN 을 활용하여 빠르고 효율적인 네트워크 서비스를 제공한다. 따라서 본 논문에서는 CPS 기반의 Reinforcement Learning 을 활용한 Rotary Inverted Pendulum System 을 통해 실시간으로 빠르고 안전한 네트워크 서비스를 제공할 수 CPS 아키텍처를 구현한다.

1 서론

사이버 물리 시스템(Cyber-Physical System, CPS)은 현재의 물리 시스템과 사이버 시스템의 융합을 추구하는 새로운 패러다임 이다 [1, 2]. CPS 는 다수의 센서, 액추에이터 등 네트워크로 연결되어 있는 실제 물리 정보를 습득, 가공, 계산, 분석하여 그 결과를 액추에이터 시스템을 통해 적용한다.

본 논문에서는 Rotary Inverted Pendulum 시스템을 통해 pendulum 의 정보를 센싱하고, 액추에이터 를 통해 pendulum 을 세우는 작업을 수행한다. Rotary Inverted Pendulum 시스템을 통해 센싱된 정보는 네트워크를 통해 Reinforcement Learning (RL)의 에이전트가 있는 EdgeX 기반의 서버로 옮겨진다. Reinforcement Learning [7] 을 이용하여 학습된 정보는 다시 네트워크를 통해 Rotary Pendulum System 의 액추에이터 로 전달되어 Pendulum 을 세우는 것을 유지하게 된다. 이때, Rotary Inverted Pendulum 시스템과 EdgeX 를 연결하는 네트워



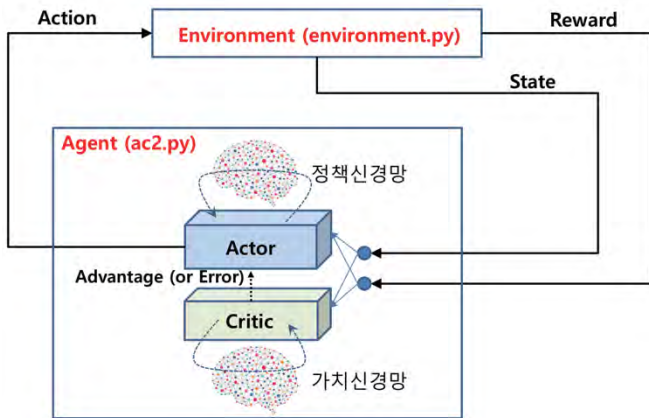
(그림 1) QUBE-Servo2

크는 SDN 기반의 네트워크로 구성되어 있다.

본 논문의 2 장에서는 전반적인 Cyber-Physical System 의 구성 요소들을 설명하며, 3 장에서는 제안하는 CPS 기반의 Rotary Inverted Pendulum system 을 소개한다. 마지막으로 4 장에서는 결론을 기술한다.

+ 교신 저자: 한연희(한국기술교육대학교)

이 논문은 2016 년도 및 2018 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 임 (No. NRF-2016R1D1A3B03933355, No. 2018R1A6A1A03025526)



(그림 2) Advantage Actor Critic 알고리즘 구조

2 Cyber-Physical System 구성

본 논문에서 제안하는 CPS 기반의 Rotary Inverted Pendulum 시스템은 실제 물리 시스템, Edge Computing 시스템과 SDN 기반의 네트워크 시스템이 결합된 사이버 물리 시스템이다.

2.1 Rotary Inverted Pendulum System

Rotary Inverted Pendulum 시스템은 QAUNSER 에서 제공하는 QUBE-Servo2 [3] 이다. QUBE-Servo2 는 Arduino, myRIO 및 Raspberry Pi 와 같은 PC 또는 마이크로 컨트롤러를 사용하여 교육자에게 실험을 위한 환경을 제공한다.

QUBE-Servo2 는 위치 및 속도 피드백을 제공하는 광학 엔코더가 달려 있는 DC 모터, 진류 센서가 내장된 앰프, 데이터 수집 장치 및 QFLEX2 컴퓨팅 인터페이스로 구성되어 있다. QFLEX2 컴퓨팅 인터페이스의 USB 포트를 통해 PC 또는 마이크로 컨트롤러와 연결할 수 있으며, 모터의 각도, 앵글의 각도, 위치 정보를 센싱하여 받아들 수 있다.

2.2 EdgeX

최근 클라우드 컴퓨팅 [4]에 대한 관심이 많아지며, 클라우드를 통한 서비스를 제공하는 것이 많아 졌다. 하지만, IoT 장비 및 실제 물리 시스템과 클라우드 간의 물리적 거리로 인하여 속도나, 클라우드의 부하가 증가하는 문제가 발생한다. 이에 따라, Edge 컴퓨팅이 클라우드 컴퓨팅의 문제를 해결하는 방법으로 제시되고 있다. Edge 컴퓨팅 [5]은 분산된 개방형 아키텍처로서 분산된 처리 성능을 제공하여 클라우드를 거치지 않고 중간에서 데이터 수집 및 처리를 통하여 빠른 서비스를 제공하고, 클라우드의 부하를 줄인다. 다양한 Edge 컴퓨팅을 위한 솔루션들이 제공되고 있으며, 본 논문에서는 EdgeX Foundry 에서 제공하는 EdgeX를 활용한다.

EdgeX [6]는 EdgeX Foundry 에서 Linux Foundation 을 통해 호스팅하는 오픈소스 프로젝트로, IoT 엣지 컴퓨팅을 위한 공통 개방형 프레임워크를 구축할 수 있게

한다. EdgeX 는 사용하는 사용자의 필요에 따라 오픈 소스로 제공되는 마이크로서비스들을 묶음으로써 선택적 Edge 컴퓨팅을 위한 프레임워크를 구축 할 수 있다. 또한 개인이 필요에 따라서는 마이크로서비스를 만들어 배포할 수 있으며, 연동하여 사용가능 하다.

본 논문에서는 EdgeX 프레임워크 구축을 위해 필수적으로 필요한 core-data, core-metadata, core-command, export-client, export-distro, rules-engine 마이크로서비스들을 이용하였으며, 추가적으로 EdgeX 와 다른 장치들 간의 통신을 위한 device-MQTT 마이크로서비스를 추가로 설치하였다. device-MQTT 는 RL 학습 에이전트와 Rotary Inverted Pendulum 시스템 사이에 MQTT 통신을 할 수 있도록 도와주는 마이크로서비스이다.

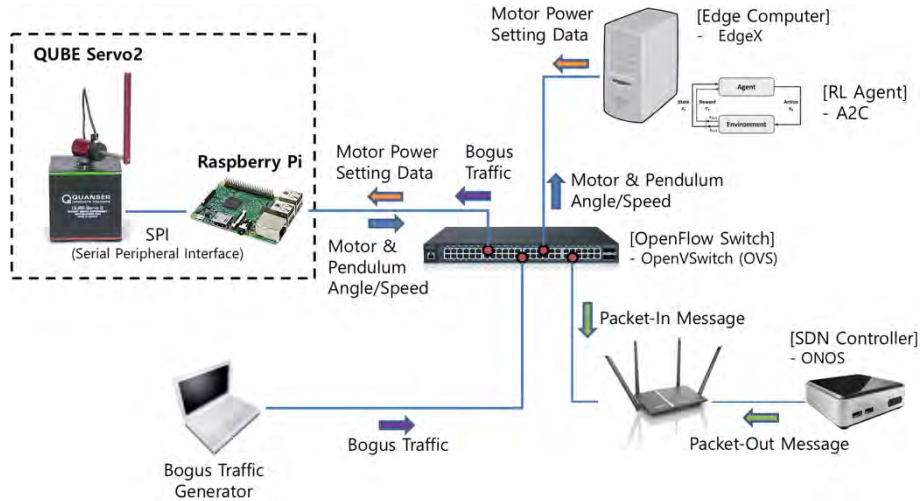
2.3 Reinforcement Learning

Rotary Inverted Pendulum 시스템의 추를 세우기 위해 사용되는 RL 의 알고리즘은 Advantage Actor Critic (A2C) 로써 알고리즘 [8, 9] 구조는 그림 2 와 같다. 물리적 시스템에서 들어온 정보 EdgeX 를 통해 받은 후 Environment.py 를 통해 state 로 정의하고 state 정보를 토대로 reward 를 계산한다. state 와 reward 정보는 정책 신경망의 역할을 하는 Actor 와 가치신경망으로 Error 값을 계산하는 Critic 으로 보내어 학습을 수행한다. 이후 학습된 결과를 토대로 action 을 만들고 다시 Environment.py 를 통해 EdgeX 에 보내어 물리적 시스템으로 전송하여 action 을 수행하게 한다. action 수행 후 다시 state 정보를 받으면, 이와 동시에 다음 state 에 대한 정보를 학습에 사용하고 이와 동시에 reward 를 계산하여 학습에 같이 사용하게 된다.

2.4 Software Defined Networking System

Software Defined Network (SDN)은 OpenFlow 를 통해 네트워크 트래픽 전달 동작을 소프트웨어 기반 컨트롤러에서 제어 및 관리 하는 방식이다 [10]. SDN 은 기존의 네트워크 시스템을 컨트롤 플레인과 데이터 플레인으로 분리함으로써 네트워크 세부 정보에 상관 없이 중앙의 컨트롤러를 통해 네트워크를 제어 및 관리가 가능하다.

본 논문에서는 SDN 시스템 구성을 위하여 OVS-Switch 와 네트워크 제어와 관리를 위한 ONOS 컨트롤러를 구성하였다. OVS-Switch 를 통해 물리적 시스템인 Rotary Inverted Pendulum 시스템 Reinforcement Learning 을 연결한다. ONOS 컨트롤러는 OVS-Switch 를 통해 전송되는 모든 트래픽의 flow-table 을 관리하고 제어 할 수 있다. 실시간으로 flow-table 의 정보를 받아 모니터링 할 수 있으며, 이를 이용하여 트래픽 flow 의 priority 를 조정할 수 있다. ONOS 컨트롤러 에서는 flow-table 관리를 위하여 REST API 를 이용해 관리정보 와 priority 정보를 관리자가 쉽게 조정 할 수 있도록 하며, 웹 서비스를 통해 각 트래픽의 플로우와 OVS-Switch 의 포트 정보들을 확인 할 수 있다.



(그림 3) 제안하는 CPS 기반의 Rotary Inverted Pendulum 구조

3 제안하는 CPS 기반 Rotary Inverted Pendulum System

본 논문에서 제안하는 CPS 기반의 Rotary Inverted Pendulum 시스템은 RL 을 통해 학습된 결과를 토대로 추를 세우는 것으로써, RL 학습 에이전트는 Edge 단에서 훈련이 된다. 또한 전체적인 네트워크는 SDN 기반의 네트워크로 전체적인 네트워크가 지능을 가짐으로써 모니터링과 관리가 용이하도록 하였다.

그림 3 은 본 논문에서 제안하는 CPS 기반의 Rotary Inverted Pendulum 시스템 구조를 나타낸 것이다. QUBE-Servo2 는 Raspberry Pi 와 SPI 통신을 이용하여 현재의 모터 앵글 값과 축의 앵글 값을 보낸다. Raspberry Pi 는 QUBE-Servo2 와 EdgeX 서버 간의 MQTT 통신을 위한 장치 역할을 수행한다. 또한 EdgeX 서버에서 수행되는 RL 의 action 인 모터 파워 값을 받아 QUBE-Servo2 로 SPI 통신으로 보내는 역할을 수행한다.

Raspberry Pi 에서 모터와 축의 앵글 값을 OVS-Switch 로 보내면 OVS-Switch 는 받은 트래픽 정보를 기 설정된 flow-table 에 따라 EdgeX 서버로 정보를 보내게 된다. flow-table 의 설정은 ONOS 컨트롤러에서 Packet_in 과 Packet_out 메시지를 통해 실시간으로 flow-table 의 상태와 조정을 하게 된다.

EdgeX 서버에서는 device-MQTT 마이크로서비스를 통해 QUBE-Servo2 에 연결된 Raspberry Pi 와 RL 학습 에이전트의 통신을 가능하도록 해주며 모터와 축의 앵글 정보를 MongoDB 에 저장하며 모니터링 역할을 수행한다. RL 학습 에이전트는 MQTT 통신으로 받은 모터와 축의 앵글 정보를 Environment 에 전달하여 state 정보를 정의하고, 동시에 현재 각 앵글에서의 reward 값을 계산하게 된다. state 와 reward 값을 이용하여 각 Actor 와 Critic 신경망에서는 학습을 수행하게 된다. 학습된 결과를 토대로 action 인 모터의 파워 값을 생성하게 되고 이는 다시 Environment 를 통해 EdgeX 로 전달이 된다. 전달된 action 정보는 MQTT 통신을 통해 OVS-Switch 를 통해 Raspberry Pi 로 전달

되고 이는 다시 SPI 통신을 통해 QUBE-Servo2 로 전달되어 action 에 설정된 모터 파워 값 만큼 모터가 움직이게 되어 축을 세우게 된다.

이때, 실제 네트워크 환경과 동일하게 구성하기 위하여 Bogus Traffic 발생기를 OVS-Switch 에 연결하였다. Bogus Traffic 은 Raspberry Pi 에 일방적으로 많은 양의 트래픽을 보내어 모터와 축의 앵글 값을 보내거나 혹은 모터의 파워 값이 제대로 받지 못하도록 방해한다. 하지만, SDN 의 ONOS 컨트롤러를 이용하여 flow-table 의 각 트래픽을 확인하여 priority 를 조정함으로써, Bogus Traffic 에 대해서는 낮은 priority 를 주어 Bogus Traffic 이외의 통신에는 영향을 주지 않도록 한다.

4 결론

본 논문에서 제안하는 CPS 기반의 Rotary Inverted Pendulum 시스템은 Reinforcement Learning 과 EdgeX 그리고 SDN 기반의 네트워크 시스템이 융합된 구조이다. 이러한 Cyber Physical System 은 IoT 가 실제 세계에서 많이 사용됨에 따라 중요성이 강조되고 있다. 제안하는 CPS 구조는 실시간으로 빠르게 정보가 전달되어 축을 세워야 하는 Rotary Inverted Pendulum 시스템에 SDN 을 기반으로 일부 트래픽들을 제한하고, EdgeX 를 활용하여 디바이스 간의 MQTT 통신을 가능하게 했으며, RL 학습의 결과 빠르게 보내 제어 할 수 있는 CPS 기반의 시스템을 설계하였다.

참고문헌

- [1] R. Baheti, H. Gill, "The Impact of Control Technology: Cyber-physical System," IEEE Control Systems Society, pp. 1-6, Sep, 2018.
- [2] Jay Lee, Behrad Bagheri, Hung-An Kao, "A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems," Manufacturing Letter, Vol. 3, pp. 18-23, Jan, 2015.
- [3] <https://www.quanser.com/products/qube-servo-2/>

[4] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, "Above the Clouds: A Berkeley View of Cloud Computing". University of California, Berkeley, Feb, 2009

[5] Pedro Garcia Lopez, Alberto Montresor, Dick Epema, Anwitaman Datta, Teruo Higashino, Adriana Iamnitchi, Marinho Barcellos, Pascal Felber, and Etienne Riviere, "Edge-centric Computing: Vision and Challenges," ACM SIGCOMM Computer Communication. Rev. 45, No. 5, pp. 37-42, Oct, 2015.

[6] <https://www.edgexfoundry.org/>

[7] Sutton, S. Richard, Barto, G. Andrew, "Reinforcement Learning: An Introduction," MIT Press, 1998.

[8] Kai Arulkumaran, "A Brief Survey of Deep Reinforcement Learning", in IEEE Signal Processing Magazine Special Issue On Deep Learning For Image Understanding, 2017.

[9] Mnih, Volodymyr, "Human-level control through deep reinforcement learning", in Nature, 518(7540):529-533, 02 2015.

[10] Benzekki, Kamal, "Software-defined networking (SDN): a survey," Security and Communication Networks, Vol. 9, Issue. 18, pp. 5803-5833, Feb, 2017