

## 강화학습을 통한 시간에 엄격한 패킷 스케줄링

정현석<sup>0\*</sup>, 이태호<sup>\*</sup>, 이병준<sup>\*</sup>, 김경태<sup>\*</sup>, 윤희용<sup>\*\*</sup>

<sup>0\*</sup>성균관대학교 정보통신대학 전자전기컴퓨터공학과

<sup>\*\*</sup>성균관대학교 소프트웨어대학 소프트웨어학과

e-mail: {daun8928, leetaeho, byungjun}@skku.edu<sup>0\*</sup>, kyungtaekim76@gmail.com<sup>\*</sup>, youn7147@skku.edu<sup>\*\*</sup>

## Time Critical Packet Scheduling via Reinforcement Learning

Hyun-Seok Jeong<sup>0\*</sup>, Tae-Ho Lee<sup>\*</sup>, Byung-Jun Lee<sup>\*</sup>, Kyoung-Tae Kim<sup>\*</sup>, Hee-Yong Youn<sup>\*\*</sup>

<sup>0\*</sup>Dept. of Electrical and Computer Engineering, Sungkyunkwan University

<sup>\*\*</sup>Dept. of Software, Sungkyunkwan University

### ● 요약 ●

본 논문에서는 시간에 엄격한(Time critical) 산업용 IoT(Industrial IoT) 환경의 무선 센서 네트워크 시스템 상의 효율적인 패킷 전달과 정확도(Accuracy) 향상을 위해 강화학습과 EDF 알고리즘을 혼합한 스케줄링 기법을 제안한다. 이 방식은 다중 대기열(Multiple queue) 환경에서 각 대기열의 요구 정확도(Accuracy Requirement)를 기준으로 최대한 패킷 처리를 미룸으로써 효율적인 CPU자원 분배와 패킷 손실율(Packet Loss)을 조절한다. 제안하는 기법은 무선 센서 네트워크 상의 가변적이고 예측 불가능한 환경에 대한 사전지식이 없이도 요구하는 서비스의 질(Quality of service)를 만족할 수 있도록 한다. 또한 정확도를 요구조건으로 제시하여 마감시간이 중요시되는 작업에서도 효율을 최대화한다.

**키워드:** 강화학습(reinforcement learning), 패킷 스케줄링(packet scheduling), 큐러닝(q-learning), 무선 센서 네트워크(wireless sensor network), 최단 마감 우선 스케줄링(earliest deadline first scheduling)

### I. Introduction

실시간 시스템의 경우 주기적인 업무(Periodic Tasks)와 비주기적인 업무(Aperiodic Tasks)가 공존한다[1]. IIoT 환경의 무선 센서 네트워크 시스템에서는 수 많은 센서가 주기적으로 센서 데이터를 전송하고, 조종(Control)이나 알림(Notification) 목적으로 비주기적인 데이터가 전송된다. 또한 이 데이터들은 시간에 엄격한 방식으로 처리되기를 요구한다. 이들은 엄격한 마감시간을 설정하게 되고, 이를 벗어나지 않는 높은 정확도를 필요로한다. 이렇게 시간에 엄격하고, 주기적인 데이터와 비주기적인 데이터가 공존하는 환경에서 마감시간 이외에 따로 권장하는 딜레이 요구조건(Delay Requirement)을 설정하고, 이 시간을 기준으로 초과시 EDF 스케줄링을 적용하고 초과하지 않은 경우 큐러닝(Q-learning)에 기반한 스케줄링을 적용하여 유연하면서도 마감시간에 엄격하게 업무를 스케줄링한다.

본 논문은 1장 소개를 다음으로, 2장에서는 관련연구, 3장에서는 제안하는 기법, 4장에서는 결론으로 구성하였다.

### II. Preliminaries

#### 1. Related works

##### 1.1 EDF 스케줄링

이 스케줄링 기법은 시간에 엄격한 실시간 시스템에서 널리 사용되고 있다. EDF 스케줄링은 마감시간이 가까울수록 우선순위를 높게 부여하는 방식의 스케줄링 기법이다. EDF는 빠르고 안정적인 것이 장점이다.

##### 1.2 RM 스케줄링

주기적인 데이터의 경우 RM 스케줄링이 주로 사용되곤 한다. 이 기법은 패킷의 마감시간으로 주기를 설정하고, 주기가 짧을수록 높은 우선순위를 부여하여 더 높은 우선순위의 작업이 선점(Preemptive)되도록 한다.

### III. The Proposed Scheme

제안하는 방식은 EDF 스케줄링과 큐리닝을 혼합한 방식이다. 현재 타임슬롯(Time Slot)이  $t$ 일 때, 각 대기열의 딜레이 요구시간  $R_i(t)$ 과 각 대기열의 현재 평균 지연시간(Mean Delay)  $M_i(t)$ 에 따라 현재의 상태(State)에 대해 모델링 한다. 이 때 딜레이 요구시간  $R_i(0)$ 은 다음 식과 같이 초기 값이 설정된다.

$$R_i(0) = \frac{D_i}{2}$$

현재 상태를 나타내는 값을  $s_i$ 로 정의한다.

다음으로 현재  $Q$ -table에서 가장  $Q$ -value를 가진 대기열을 다음 대기열로 선택한다. 다음으로 보상 함수(Reward Function)  $r_i(t)$ 를 통해 서비스 이후 결과에 기반하여 각 대기열이 상대적으로 어느 정도의 수치만큼 좋은 상태에 있는지를 점수로 채점한다. 이는 다음 식과 같이 표현된다.

$$r_i(t) = \begin{cases} \frac{C_1 M_i(t)}{R_i(t)} & \text{if } M_i(t) < R_i(t) \\ -C_2 & \text{if } M_i(t) > R_i(t) \end{cases}$$

여기서  $C_1$ 과  $C_2$ 는 각각 보상과 처벌 계수이다. 이렇게 얻어진 보상을 통해 MDP를 정의할 수 있게 된다.

만약 평균 지연시간이 딜레이 요구시간보다 높아질 경우, 큐값에 의해서가 아니라 각 대기열의 첫 번째 패킷들을 이용하여 EDF 스케줄링 정책을 적용한다. 요구되는 정확도(Accuracy Requirement) 값을  $\alpha_i$ 라고 하고, 혼합 정책을 사용하여 얻어지는 평균 정확도 값을  $A_i(t)$ 로 정의한다. 둘 사이의 차이를  $\beta_i$ 로 정의한다. 요구되는 정확도와 평균 정확도 값에 따라 다음 타임슬롯에서의 평균 지연시간  $M_i(t+1)$ 을 새롭게 업데이트한다. 이는 다음 식과 같이 표현된다.

$$M_i(t+1) = M_i(t) + a\beta_i(t)$$

여기서  $a$ 와  $b$ 는 각각 보상과 패널티(penalty) 계수이다.

### IV. Conclusions

본 논문에서는 IIoT 환경이 요구하는 시간에 엄격한 환경에서 강화학습의 일종인 큐리닝을 이용한 패킷 스케줄링 방식을 제안하였다. 기존 큐리닝을 이용한 알고리즘의 시간에 엄격하지 않은 방식의 문제점을 EDF 스케줄링과 혼합하여 딜레이 요구시간을 학습을 통해 변경함으로써 시간에 엄격하면서도 다양한 환경에 적응하는 알고리즘을 제안하였다. 하지만 적용과정이 단순화되었기 때문에 강화학습의 특성을 더욱 활용할 수 있도록 더 많은 요소와 더 많은 상태공간(State Space)을 도입하는 방식에 대해 추가 연구가 이루어질 예정이다.

### ACKNOWLEDGEMENT

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 정보통신-방송연구 개발 사업(No. 2016-0-00133, 초연결 IoT 노드의 군집 자동화를 통한 Edge Computing 핵심 기술 연구), SW중심대학지원사업(2015-0-00914), 한국연구재단 기초연구사업(No.2016R1A6A3A11931385, 실시간 공공안전 서비스를 위한 소프트웨어 정의 무선 센서 네트워크 핵심기술 연구, 2017R1A2B2009095, 실시간 스트림 데이터 처리 및 Multi-connectivity를 지원하는 SDN 기반 WSN 핵심 기술 연구), 삼성전자, BK21PLUS 사업의 일환으로 수행되었음.

### REFERENCES

- [1] J. Hall, and P. Mars, "Satisfying QoS with a Learning Based Scheduling Algorithm," IEEE Sixth International Workshop on Quality of Service, pp. 171-173, May 1998.
- [2] J. Yin, and G. Guo, "An Algorithm for Scheduling Aperiodic Real-Time Tasks on a Static Schedule," IEEE Second International Conference on Information and Computing Science, Vol. 1, pp. 70-74, May 2009.