

Q-learning을 이용한 신뢰성 있는 패킷 스케줄링

김동현⁰, 유승연*, 김정태*, 윤희용*

⁰ 상균관대학교 정보통신대학 전자전기컴퓨터공학과

** 상균관대학교 소프트웨어대학 소프트웨어학과

e-mail: {kdh7263,seyoo90}@skku.edu⁰, kyungtaekim76@gmail.com*, youn7147@skku.edu**

Reliable packet scheduling using Q-learning

Dong-Hyun Kim⁰, Seung-Eon Yoo*, Kyung-Tae Kim*, Hee-Yong Youn**

⁰Dept. of Electrical and Computer Engineering, Sungkyunkwan University

**Dept. of Software, Sungkyunkwan University

● 요약 ●

본 논문에서는 무선 센서 네트워크 환경에서 신뢰성 있는 데이터 패킷 전송을 위한 효율적인 스케줄링 기법을 제안한다. 무선 네트워크는 수천 개의 센서노드, 게이트웨이, 그리고 소프트웨어로 구성된다. 큐러닝(Q-learning)을 기반으로 한 스케줄링 기법은 동적인 무선센서 네트워크 환경의 실시간 및 비실시간적인 데이터에 대한 사전 지식을 필요로 하지 않는다. 따라서 최종 결과 값을 도출하기 전에 스케줄링 정책을 구할 수 있다. 제안하는 기법은 데이터 패킷의 종류, 처리시간, 그리고 대기시간을 고려한 기법으로 신뢰성 있는 데이터 패킷의 전송을 보장하고, 전체 데이터 패킷에 공정성을 부여한다. 본 논문에서는 시뮬레이션을 통해 기존의 FIFO 알고리즘과 비교하여 제안하는 스케줄링 기법이 전체 데이터 패킷에 대한 공정성 및 신뢰성 측면에서 우수함을 증명하였다.

키워드: 강화학습(Reinforcement Learning), 패킷 스케줄링(Packet scheduling), 큐러닝(Q-learning), 무선 센서 네트워크(Wireless sensor network)

I. Introduction

최근 위치에 관계없이 컴퓨팅 환경에 접근할 수 있는 패러다임으로 무선 네트워크에 대한 연구가 활발히 진행되고 있다. 무선 센서 네트워크는 무선 네트워크에 연결된 수천 개의 센서 노드, 게이트웨이, 그리고 소프트웨어의 세 가지로 구성된다. 즉 컴퓨팅 기능과 무선 통신 기능을 갖춘 센서 노드를 배치하고, 자율적으로 네트워크를 구성하며 센서 노드로부터 감지된 물리적 데이터를 무선으로 수집하여 모니터링 및 제어하는 기술이다. 이러한 데이터 패킷을 스케줄링하는 기법으로 크게 선점(Preemptive) 및 비 선점(Non-preemptive) 기법이 있다. 비 선점 기법 스케줄링 기법에서, 중요한 작업은 중요도가 낮은 작업에 의해 대기할 수 있다. 따라서 이러한 알고리즘은 부적절하게 데이터 패킷을 Queue에 할당하기 때문에 높은 오버 헤드와 긴 End-to-End delay를 갖는다. 따라서 싱크노드에서 수집된 데이터는 데이터의 손실 및 가아 상태를 방지하기 위해 신뢰성 있고 공평하게 스케줄링 되어야 하며, 이를 위한 방법으로 큐러닝을 이용하였다. 본 논문은 1장 소개를 다음으로, 2장에서는 관련연구, 3장에서는

제안하는 기법, 그리고 4장에서는 결론으로 구성하였다.

II. Preliminaries

1. Related works

1.1 Q-learning

큐러닝은 강화학습(Reinforcement Learning)의 일종으로 스케줄링 기법에 대한 사전 지식을 필요로 하지 않기 때문에 동적인 네트워크 환경에 대한 스케줄링 기법에 적합하다. 또한 최종 결과 값이 도출되기 전에 큐러닝에서 최종적으로 구하고자 하는 State -Action pair 값을 구할 수 있으며 $Q^*(s,a)$ 을 구하는 식은 다음과 같다.

$$Q^*(s,a) = r_t + \gamma \max_{a'} Q^*(s', a') \quad [1]$$

위 수식에서 s 는 상태, a 는 행동, r 은 Action에 대한 보상 값을 나타내며, $Q^*(s,a)$ 는 다음에 실행할 패킷의 값으로 가장 큰 값을 찾는다.

1.2 FIFO algorithm

가장 기본적인 비선점 스케줄링 기법으로 대기 Queue에 들어온 순서대로 CPU에 할당되는 방법이다. 하지만 한번 작업이 시작되면 다른 작업에 의해 교체되지 않기 때문에 중요한 작업의 처리가 늦어질 수 있으며, 처리시간이 짧은 작업들이 긴 작업에 의해 장시간 대기할 수 있으므로 동적인 WSN환경에서 스케줄링하기에 다소 제한이 있다.

패킷의 개수: 60

패킷번호	종류	처리시간	도착시간	대기시간
38	N	11.14	0	0.0
42	E	1.56	0	11.14
24	N	2.76	0	12.700001

18	N	6.61	5	10.460001
36	N	15.96	5	17.070002
1	N	4.97	5	33.030003

⋮				
3	R	12.09	95	444.9599
53	N	6.91	95	457.04993
60	N	13.75	95	463.9599

패킷 수: 60/평균대기시간: 7.732665

Fig. 2. Result of FIFO algorithm

III. The Proposed Scheme

제안하는 기법의 모든 실험은 Java로 실행되었다. 또한 실제 WSN 환경을 고려하기 위해 데이터 패킷의 종류는 Emergency data(E), Real time data(R), Non Real time data(N) 로 나누고 각 10%, 30%, 60%로 패킷의 숫자를 입력하여 생성되게 하였다.

```
Map<String, Double> w = new HashMap<String, Double>();
w.put("E", 100);
w.put("R", 300);
w.put("N", 600);
```

Fig. 1. Weighting according to packet type

여기서 E 는 데이터 패킷의 흐름에서 크게 벗어나는 데이터로 가장 우선시 처리되어야 하며, R 은 음성 및 영상 등 실시간 데이터로 데이터가 장시간 전송이 되지 못할 경우 데이터의 신뢰성 및 품질이 떨어질 수 있다. N 은 일반적인 데이터 패킷으로 우선순위가 가장 낮다. 도착시간은 센서 노드에서 데이터 패킷이 감지되어 싱크 노드로 전송 이후의 시간으로 5초마다 3개의 데이터 패킷이 감지된다고 가정하였다. 처리시간은 데이터 패킷을 처리하는데 걸리는 시간으로 1에서 15의 값을 설정하였다. 아래 Fig. 2는 기존의 FIFO를 이용한 스케줄링 결과이다.

위 결과에서 볼 수 있듯이 기존 FIFO 알고리즘은 데이터의 종류, 처리시간, 그리고 대기시간을 고려하지 않기 때문에 효율적이지 못하다. 제안하는 기법은 데이터 패킷의 대기시간 및 처리시간이 짧은 패킷으로 인해 처리 시간이 오래 걸리는 데이터 패킷이 기아 상태에 빠지는 것을 방지하기 위해 처리시간과 대기시간을 합하여 큐러닝의 r 값으로 부여 하였다. 따라서 처리시간이 짧고 대기시간이 길수록 r 값이 증가하여 $Q(s,a)$ 값이 증가하게 된다.

아래 Fig. 3은 제안된 기법의 처리 과정으로 현재 $Q(s,a)$ 에서 가장 큰 값을 가지는 $Q^*(s,a)$ 를 찾는 과정이다.

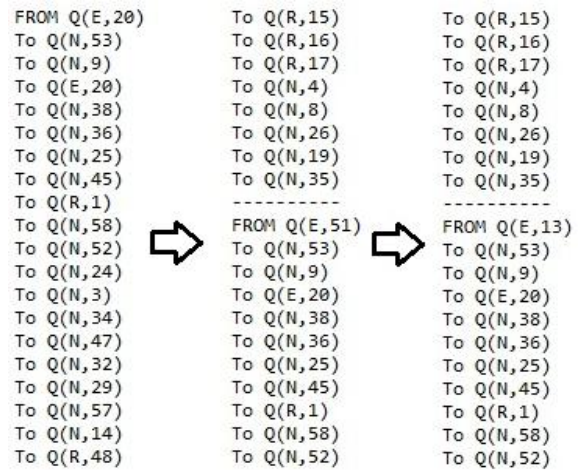


Fig. 3. Process of the proposed scheme

위 Fig. 3의 과정을 알고리즘으로 표현하면 다음과 같다.

Table 1. 제안된 기법의 알고리즘

알고리즘1. 제안된 기법의 처리과정	
1.	감지 된 데이터 패킷 중 가장 먼저 도착한 데이터 패킷의 $Q(s,a)$ 값을 구하고 처리한다.
2.	감지 된 모든 데이터 패킷에 대해 $Q(s,a)$ 값을 구하여 가장 큰 값이 다음 $Q^{n+1}(s,a)$ 로 선택된다.
3.	더 이상 처리할 데이터가 없을 때까지 1~2의 과정을 반복한다.

제안된 기법을 통해 얻은 결과는 아래의 그림과 같다.

패킷번호	종류	처리시간	도착시간	대기시간
42	E	1.56	0	15.32
47	E	13.84	20	0.0
44	E	9.03	65	0.0
55	E	11.69	65	0.0
57	E	1.48	90	0.0
39	R	14.79	80	0.0
59	R	12.93	50	2.3899
:				
23	N	13.85	15	526.07!
31	N	7.49	70	484.92!
7	N	8.46	85	477.41!
17	N	1.83	35	535.87!
패킷 수 : 60 / 평균대기시간 : 8.931333				

Fig. 4. Result of the proposed scheme

제안된 기법은 기존의 FIFO와 다르게 데이터 패킷의 종류, 처리시간, 그리고 도착시간을 고려하여 Queue에 할당하였다.

본 연구의 정확성을 검증하기 위해 같은 패킷의 수를 입력하여 10번의 반복 및 기존의 FIFO 알고리즘을 이용한 스케줄링 결과와 비교하였다. 아래 Fig. 5는 제안된 기법과 FIFO 알고리즘의 비교 결과이다.

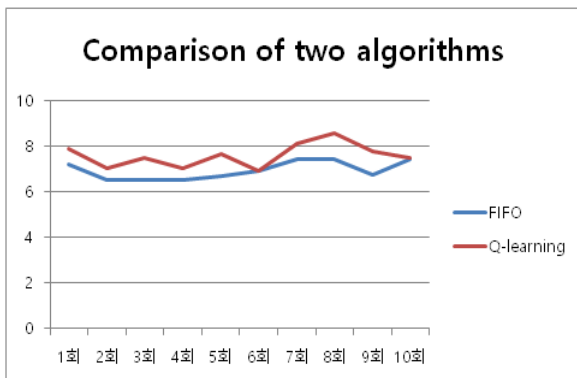


Fig. 5. Comparison of two algorithms

비교 결과 제안된 기법이 기존의 FIFO 알고리즘에 비해 평균 대기시간이 증가한 결과를 보였다. 하지만 제안된 기법은 기존의

기법과 달리 데이터의 종류를 고려하여 우선순위에 따라 처리하였다.

IV. Conclusions

본 논문에서는 실시간으로 변화하는 동적인 WSN 환경에서 센서 노드에서 감지된 전체 데이터 패킷에 공정성을 부여하고 신뢰성 있는 데이터 패킷의 전송을 위한 방법으로 큐러닝을 이용하였다. 연구의 효율성을 검증하기 위해 기존의 FIFO 알고리즘과 비교 실험하였고, 실험결과 제안된 스케줄링 기법은 기존의 FIFO 알고리즘을 이용한 스케줄링 기법에 비해 평균 대기시간이 조금 증가하였지만, 데이터 패킷의 종류, 처리시간, 그리고 대기시간을 고려한다는 점에서 효율적으로 데이터 패킷을 스케줄링 할 수 있음을 보인다. 앞으로의 연구 계획으로 데이터의 처리시간 및 대기시간에 따른 reward 값을 부여하여 평균대기시간을 더욱 줄이는 방법에 대해 연구할 예정이다.

Acknowledgment

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 정보통신방송연구 개발 사업(No.B0717-17-0070, 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] N. Nasser, L. Karim, T. Taleb "Dynamic Multilevel Priority Packet Scheduling Scheme for Wireless Sensor Network," IEEE transactions on wireless communications, Vol. 12, No.4, 2013.
- [2] H. Ferra, Ken Lau, C. Leckie, and A. Tang, "Applying Reinforcement Learning to Packet Scheduling in Routers," American Association for Artificial Intelligence, 2003.
- [3] V. K. Sharma, S.S.P. Shukla, V. Singh, "A Tailored Q-Learning for Routing in Wireless Sensor Networks," IEEE International conference on parallel, distributed and grid computing, 2013.
- [4] C. Cirstea, R. Davidescu, A. Gontean. "A Reinforcement Learning Strategy for Task Scheduling of WSNs with

Mobile Nodes,” IEEE International Conference on Telecommunications and Signal Processing(TSP), pp.348-353, 2013..

- [5] <https://www.slideshare.net/ssuser06e0c5/q-learning-cnn-object-localization>