

강화학습을 이용한 802.11 무선랜의 전력절감

임 태 현*, 이승형^o

Power Management of IEEE 802.11 WLAN Using Reinforcement Learning

Lim Tae Hyun*, Seung Hyong Rhee^o

요 약

IEEE 802.11 기반의 WLAN(Wireless Local Area Network) 기술은 무선으로 사용자들에게 높은 성능의 편의성을 제공하지만, 많은 전력을 소모하는 단점을 가지고 있다. IEEE 802.11 표준에서는 단말의 전력소모를 줄이기 위해 단말은 SLEEP 상태와 AWAKE 상태를 전환하며 에너지 소모량을 줄이고 있다. 본 논문에서는 현재 사용되고 있는 IEEE 802.11 전력절감방식에 대해 문제점을 제시하고, 강화학습 기술인 Q-learning을 이용하여 문제점을 해결한다. 네트워크 트래픽 환경에 따라 단말기는 유동적으로 SLEEP 상태 시간을 결정하며, 불필요하게 에너지를 소모하는 것을 최소화 한다. 네트워크 시뮬레이션인 NS-3에서 정의가 되어있지 않는 IEEE 802.11 전력절감방식의 구현방법과 에너지 소모량과 프레임 딜레이에 관한 결과를 나타낸다. 시뮬레이션 결과 기존방식에 비해 네트워크 성능 하락 없이 에너지 소모량을 약 27% 감소할 수 있다.

Key Words : IEEE 802.11 Power management, Power Save Mode, Reinforcement learning

ABSTRACT

The WLAN(Wireless Local Area Network) technology based on IEEE 802.11 provides users

with high performance but consumes a lot of power. So the IEEE 802.11 standard provides that the station can save the power consumption by switching the SLEEP state and AWAKE state. In this paper, we explain the problems that exist in conventional method and solve the problem by using the Q-learning model, a technology of reinforcement learning. Depending on the network traffic environment, the station determines the duration of the SLEEP state flexibly and minimizes unnecessary energy consumption. We explain how we implemented a series of IEEE 802.11 power management processes that were not defined in NS-3 and we show the energy consumption and frame delay and the simulation results.

1. 서 론

IEEE 802.11 기반의 무선통신 기술은 사용자들에게 이동성과 편의성을 제공하지만, 과거 노트북을 사용하기 위해 설계가 되었기 때문에, 많은 전력을 소모하는 단점을 가지고 있다.

^[1] IEEE 802.11 표준에서는 단말의 전력소모를 줄이기 위해 PSM(Power Save Mode)를 제공한다. PSM은 단말의 상태(State)를 SLEEP 상태와 AWAKE 상태로 정의한다. 단말의 SLEEP 상태는 외부와 데이터 송수신이 불가능하지만, 전력을 적게 소모한다. AWAKE 상태는 외부와 데이터 송수신이 가능하지만, 에너지소모량이 많은 상태이다. 단말은 SLEEP 상태와 AWAKE 상태를 전환하며, 에너지 소모를 절감하고 있다. 하지만 표준에서는 단말의 SLEEP 상태 기간에 대해서는 정확하게 정의가 되어있지 않고, 일반적으로 트래픽환경과 무관하게 일정한 주기로 상태를 전환하고 있다.

본 논문에서는 강화학습을 이용하여 네트워크 트래픽에 따라 단말의 SLEEP 상태 시간을 결정하는 전력절감방식을 제안한다. 2장에서 IEEE 802.11 전력절감방식과 강화학습, 제안하는 전력절감방식에 대해 설명한다. 3장에서 시뮬레이션 구현방법과 기존방식과 비교분석하고, 4장 결론으로 끝을 맺는다.

* 이 논문은 2017년도 광운대학교 교내 학술연구비 지원에 의해 연구되었음.

• First Author : (ORCID:0000-0001-9832-4617)KwangWoon University, Department of Electronics Convergence Engineering, zlsuukd@naver.com, 정희원

o Corresponding Author : (ORCID:0000-0002-5925-1854)KwangWoon University, Department of Electronics Convergence Engineering, rhee@kw.ac.kr, 종신회원

논문번호 : 201912-343-B-LU, Received December 20, 2019; Revised December 27, 2019; Accepted December 27, 2019

II. 본 론

2.1 IEEE 802.11 Power Management

Infrastructure 네트워크 환경에서 하나 이상의 AP (Access Point)가 존재하며, 모든 트래픽은 이 AP를 통해 데이터 송수신이 이루어진다. 따라서 IEEE 802.11 전력절감방식에서 AP의 역할은 크게 2가지가 있다. 첫 번째로 외부로부터 온 데이터를 단말의 상태에 따라 데이터를 AP 버퍼에 저장할지, 단말에게 전송할지 결정한다. 단말이 SLEEP 상태일시 외부로부터 데이터를 수신하지 못하기 때문에, AP 버퍼에 데이터를 저장한다. 두 번째로 AP는 버퍼링 된 데이터의 존재 여부를 비콘 프레임 TIM(Traffic Indication Map) 필드를 통해 단말에게 알려준다. 이때 단말은 SLEEP 상태 기간을 LI (Listen Interval) 값을 통해 결정한다. LI는 BI(Beacon Interval) 단위로 얼마나 SLEEP 상태로 유지할지 결정하는 변수이다. 단말은 LI 값만큼 SLEEP 상태로 유지하고 AWAKE 상태로 진입한다. 비콘 프레임을 수신한 단말은 SLEEP 상태기간동안 AP 버퍼에 저장된 데이터를 PS-POLL(Power Save Poll) 프레임을 통해 하나씩 수신한다. 모든 데이터를 수신한 단말은 다시 SLEEP 상태로 진입한다.

2.2 강화학습(Reinforcement Learning)

^[2]강화학습은 머신러닝 기법 중 하나의 방법으로, 보상을 통해 에이전트가 가장 합리적인 행동을 할 수 있도록 학습한다. 강화학습의 상태는 현재 상황을 말해주는 지표로서, 상태 값에 따라 보상을 받게 되고, 이 보상을 최대화하는 행동을 결정한다.

Q-learning은 강화학습의 일종으로 각 상태에서 Q-value값이 가장 큰 행동을 선택하는 방법이다.

$$\pi^*(S) = \operatorname{argmax}_a Q(s,a) \quad (1)$$

Q-value값은 누적 보상의 합으로 표현되며 식 (2)과 같이 주어진다.

$$Q(s,a) = Q(s,a) + \alpha(r + \gamma \max_{a'} Q(s',a') - Q(s,a)) \quad (2)$$

Q(s,a)는 현재 상태의 행동에 대한 값이며, r은 현재 상태에 대한 보상이고, α는 학습 속도를 뜻한다. γ는 할인계수로 미래 보상에 대한 중요성을 결정한다. max(Q(s',a'))은 다음 상태의 가장 큰 Q-value 값을 말하며, 즉 현재 상태의 누적 보상을 계산하기 위해 다음 상태는 가장 큰 Q값을 통해 계산한다.

2.3 시스템 모델(System Model)

Q-learning에 적용하기 위해 강화학습의 상태(State), 행동(Action), 보상(Reward)을 정의한다.

강화학습의 상태 X는 어플리케이션층에서 측정된 평균 프레임 딜레이로 정의한다. 단말은 SLEEP 상태 동안 온 데이터를 PS-POLL 프레임을 통해 데이터의 수와 각 프레임의 딜레이를 알 수 있다.

$$X = \frac{1}{N} \sum_{i=1}^N D(i) \quad (3)$$

N는 단말기가 SLEEP 상태 동안 AP에 저장된 데이터 수를 뜻하고, D(i)은 i번째 프레임의 딜레이이다.

강화학습의 행동(action) A(x)는 단말의 LI이다. 즉 SLEEP상태 기간을 결정한다.

$$A(x) = \{ak \mid ak : LI = k \ k \in \{0,1,2,\dots,n\}\} \quad (4)$$

강화학습의 보상은 그림 1과 같이 정의한다. N는 단말기가 SLEEP 상태 동안 AP에 저장된 데이터 수를 의미하고, D는 상수 값이다. 평균 프레임 딜레이가 D값보다 작을시, 양수의 보상 또는 0의 보상이 주어진다. LI값은 에너지 소모량을 상대적으로 확인할 수 있는 지표이다. LI이 클수록 단말의 SLEEP 상태 기간이 늘어나기 때문에 에너지 소모량은 적게 소모하는 것을 의미한다. 즉 에너지 소모량이 작을수록 보상을 최대화 한다. N이 0인 경우 불필요하게 AWAKE 상태로 진입했기 때문에 보상을 0으로 주어진다. 평균 프레임 딜레이가 D값보다 클시, 제한적인 AP 버퍼 때문에 프레임 손실이 발생할 수 있기 때문에 보상은 음수로 정의한다. 즉 제안하는 전력절감방식은 평균 프레임 딜레이가 D보다 작으면서, 최소한의 에너지를 소모하는 단말의 LI값을 결정하는 방식이다.

Algorithm 1 Reward Function

```

if State < D then
  if N==0 then
    R(s,a)=ListenInterval
  else
    R(s,a)=0
  end if
else
  R(s,a)=-1
end if
    
```

그림 1. 보상함수
Fig. 1. Reward Function

III. 시뮬레이션 결과

3.1 NS-3 IEEE 802.11 전력관리

3.1장에서 구현되어 있지 않는 NS-3 IEEE 802.11 전력관리 시뮬레이션 구현방법을 설명한다.

NS-3의 trace 기능은 시뮬레이션 도중 이벤트가 발생하는 동시에 특정 데이터에 접근권한을 부여할 수 있다. 이 기능을 이용하여 특정 이벤트가 발생할 시, 내가 원하는 함수나 스케줄을 정의할 수 있다. 따라서 AP는 외부에 데이터를 받을 때 혹은 비콘 프레임을 전송할 때 다음과 같은 역할을 수행한다.

- 1) 외부로부터 데이터를 받을 시 단말의 상태에 따라 데이터를 전송할지 버퍼에 저장할지 결정한다.
- 2) 비콘 프레임을 전송할 때 단말의 상태를 확인하여 버퍼에서 데이터를 하나씩 전송한다.

단말은 AP의 비콘 프레임을 수신할 때, AP버퍼의 데이터 유무에 따라 다음 데이터를 기다리거나 단말의 LI만큼 SLEEP 상태로 진입한다. 이때 단말은 2장에서 정의한 강화학습 식을 통해 Q-value값을 업데이트 하며, 단말의 LI값을 결정한다.

3.2 제안하는 전력절감방식 시뮬레이션

3.1장에서 구현한 NS-3 전력절감방식 시뮬레이션을 이용하여, 기존 전력절감방식과 비교분석한다. 자세한 파라미터는 표1에서 나타낸다.

에피소드는 100s씩 20회 진행하였으며, 그림 2, 3은 각 에피소드마다 에너지소모량과 평균 프레임 딜레이의 그래프를 나타낸다. 표준 전력절감방식은 단말의 LI 값은 1인 경우이다.

표준 전력절감방식(LI=1)인 경우는 매 비콘 프레임마다 AWAKE 상태로 진입하기 때문에 네트워크 성능은 우수하지만, 불필요하게 AWAKE상태로 진입하

표 1. 시뮬레이션 파라미터
Table 1. Simulation Parameter

Parameter	Value
Network	802.11ac
Distance	5(m)
Beacon Interval	100ms
Packet Size	1448(Byte)
α, γ, D (reward function)	0.2, 0.6, 0.05
Packet Interval	0.33s
Listen Interval	1~9(Beacon Interval unit)
Simulation Time(Episode)	100(s)

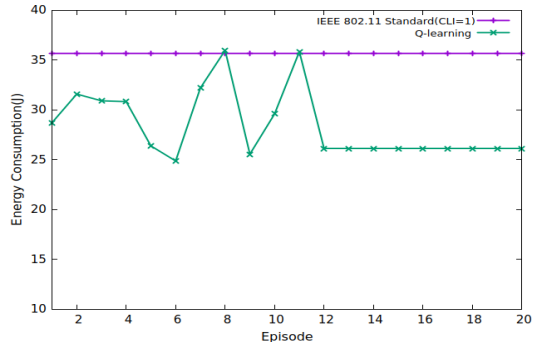


그림 2. 에너지 소모량
Fig. 2. Energy Consumption

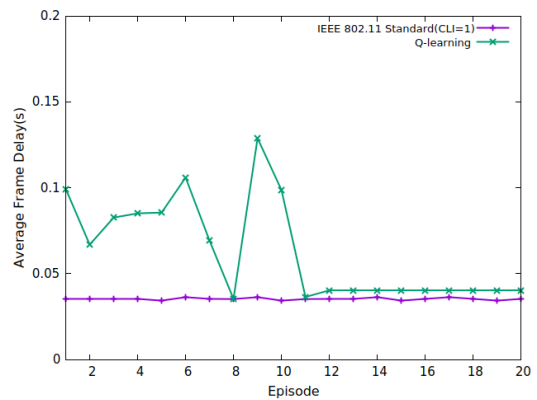


그림 3. 평균 프레임 딜레이
Fig. 3. Average Frame Delay

기 때문에 많은 에너지를 소모한다. 제안하는 전력절감방식은 에피소드가 진행하면서 단말의 평균 프레임 딜레이에 따라 양수의 보상 또는 음수의 보상을 받게 된다. 따라서 단말은 평균 프레임 딜레이가 상수 D보다 작은 성능을 내도록 학습을 진행하며, 적은 에너지 소모를 할수록 높은 보상을 받기 때문에 최소한의 에너지를 소모하는 최적의 LI값을 결정한다. 초기에는 학습을 하는 과정으로 에너지소모량과 평균 프레임 딜레이가 불규칙하지만, 에피소드가 10회 이상 진행된 후부터는 표준 전력절감방식과 유사한 딜레이 성능에 에너지 소모량은 약 27% 감소할 수 있었다.

IV. 결론

본 논문에서는 IEEE 802.11 표준 전력절감방식을 설명하고, 네트워크 트래픽환경에 상관없이 고정된 LI 값을 사용하는 방식에 대해 문제점으로 제시한다. 제안하는 전력절감방식은 최근에 많은 분야에 사용되고

있는 강화학습 기법인 Q-learning을 이용하여 트래픽 환경에 따라 최적의 LI(Listen Interval)값을 결정한다. 네트워크 시뮬레이터로 NS-3을 사용하였으며, 3.1장에서 구현 되어있지 않는 NS-3 전력절감방식에 대한 시뮬레이션 구현방법을 설명한다. 제안하는 전력절감 성능은 기존 전력절감방식과 비교분석 한 결과 평균 프레임 딜레이 성능은 거의 유사하지만, 에너지 소모량은 약 27% 감소한 것을 확인 할 수 있다.

References

- [1] *IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 2007.
- [2] R. S. Sutton and A. G. Barto, *Introduction to reinforcement learning* Cambridge, MA: MIT press, 1998.
- [3] Y. Li, X. Zhang, and K. L. Yeung, "DLI: A dynamic listen interval scheme for infrastructure based IEEE 802.11 WLANs," *IEEE 26th Annu. Int. Symp. PIMRC*, Hong Kong, China, 2015.
- [4] R. Karmakar, "Online learning-based energy-efficient frame aggregation in high throughput WLANs," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 712-715, Apr. 2019.
- [5] H. Chen and C.-W. Huang, "Power management modeling and optimal policy for IEEE 802.11 WLAN systems," *IEEE VTC2004-Fall*, Los Angeles, CA, USA, Sep. 2004.
- [6] W. Jung and S. Rhee, "Traffic-aware dynamic algorithm for power saving of IEEE 802.11 mobile access point," *J. KICS*, vol. 42, no. 8, pp. 1647-1653, Aug. 2017.