

IEEE 802.11be 네트워크에서 협력 액세스 포인트를 위한 지능형 리소스 유닛 할당 기법

박주성, 박준희, 김민태, 안민기*, 이인규
고려대학교, 삼성전자*

{ps6983, pjh0585, wkd2749, inkyu}@korea.ac.kr, minki2.ahn@samsung.com*

Intelligent Resource Units Allocation for Cooperative Access Points in IEEE 802.11be Networks

Juseong Park, Junhee Park, Mintae Kim, Minki Ahn* and Inkyu Lee
Korea Univ., Samsung Elec.*

요 약

본 논문에서는 IEEE 802.11 be 기술 중 하나로 상정되는 동시전송기술을 사용할 때, sum rate 을 높이기 위한 resource unit(RU) 할당 방법에 대하여 심층강화학습을 사용한 기법을 제안한다. 겹쳐진 영역에 있는 station 들은 RU 를 공유한다는 제약을 고려하였을 때, 지능형 할당 방법은 다른 기법 대비 우수한 성능을 보임을 모의실험을 통해 보인다.

I. 서 론

IEEE 802.11 기술에 기반을 둔 wireless local area network(WLAN)의 보급률이 증가함에 따라, 현재의 WLAN 환경은 과거에 비해 네트워크의 access point(AP) 밀집도 및 단말 station(STA)이 요구하는 데이터 양이 급격히 증가하고 있다 [1]. 단일 AP 및 basic service set(BSS) 상황에서의 resource unit(RU)의 할당[2] 은 활발히 연구가 되었으나, 현재 차세대 WLAN 표준으로 논의되는 802.11be 의 새로운 요소 중 하나인 동시 전송 (AP coordination) 환경의 겹쳐진(overlapped) BSS 에 대한 RU 할당 문제는 아직 자세히 다루어지지 않았다. 기존 문헌에서 다른 RU 할당 문제와 달리 겹쳐진 BSS 상황에서 이용하는 주파수 대역대가 동일할 경우, 겹쳐진 영역에 있는 station(STA)는 IEEE 802.11be 에서 고려되고 있는 동시 전송을 위해 두 BSS 에서 동일한 RU를 사용하게 된다. 이는 전체적인 관점에서 보았을 때, 단일 BSS 환경의 RU 할당과 비교하여 문제를 더욱 복잡하게 만든다.

본 논문에서는 겹쳐진 BSS 상황에서 다운링크 RU 할당 문제에 대해 심층강화학습을 이용한 자원 할당 기법을 제안한다. Double deep Q-network(DDQN)을 마치 반복 알고리즘처럼 사용함으로써 적은 반복 수 내에 성능 이득을 얻을 수 있음을 시뮬레이션 결과를 통해 확인하였다.

II. 본론

1. 시스템 모델과 문제 형성

본 논문에서는 orthogonal frequency-division multiple access(OFDMA)에 기반한 Wi-Fi 네트워크를 고려한다. 네트워크에 할당된 전체 대역폭은 부반송파들로 이루어진 부분집합들로 나누어지며 이는 RU 라고 한다. RU 할당 문제는 업링크와 다운링크 상황에서 각각 따로 고려할 수 있으나 본 논문에서는 그림 1 과 같은 겹쳐진 BSS 구조에서 하향 링크 RU 할당



그림 1. 겹쳐진 BSS 상황에서의 시스템 구조

문제에 대하여 해결하고자 한다. IEEE 802.11be 의 대역 중 20MHz 를 가정하며, 이 때, RU 의 최소단위는 78.125kHz 의 부반송파와 26-tones 이다. 그림 1 과 같이 BSS 내 STA 이 4 개 일 때, RU configuration(RUC)는 (106, 52, 52, 26)로, RUC 내 RU 의 순서는 내림차순으로 가정한다. 또한 STA 간 공평성을 위해 각 RU 가 하나씩 할당 되는 시스템을 상정한다.

두 AP 가 협력하는 시스템에서 각 AP 의 BSS 내에 4 개의 STA 들이 있다고 가정한다. 이 때 각각의 BSS 에 대응 되는 RUC 의 RU 개수는 4 개이다. 이 시스템에서 모든 STA 에 할당되는 파워가 동일하다고 가정할 때, STA 각각의 rate 은 다음과 같이 기술할 수 있다.

$$R_k = \sum_{m=1}^{n_k} b \log_2 \left(1 + \frac{P \tilde{h}_{k,m}^2}{\sigma^2} \right) \quad (1)$$

$$\text{where } \tilde{h}_{k,m} = \begin{cases} h_{k,\hat{m}} & , k=1,2,5,6 \\ \sqrt{\frac{1}{2}(|h_{k,\hat{m}}^1|^2 + |h_{k,\hat{m}}^2|^2)} & , k=3,4 \end{cases}$$

식 (1) 에서 b 는 tone 당 대역폭, P 는 송신파워, $n_k \in \{106, 52, 26\}$ 는 STA k 에 할당된 RU 에 따라 결정되는 tone 수, 그리고 $h_{k,\hat{m}}$, $h_{k,\hat{m}}^1$ 그리고 $h_{k,\hat{m}}^2$ 는 각각 STA k 의 부반송파 \hat{m} 의 주파수 영역 채널, 겹쳐진 영역에 있는 STA k 의 AP 1 로부터의 주파수 영역 채널

그리고 AP 2로부터의 주파수 영역 채널을 의미한다. 수식 (1)에서 겹쳐진 영역의 STA 들에 대해 동시 전송을 반영하기 위해 $\tilde{h}_{k,m}$ 을 정의하였다. 그리고 m 은 할당된 RU 에 따라 재배열된 인덱스를 의미한다.

시스템 전체 sum rate R 은 모든 STA 들에 대해 $R = \sum_{k=1}^N R_k$ 으로 얻을 수 있는데, 이때 N 은 겹쳐진 BSS 시스템 전체의 STA 수를 나타낸다. 할당 벡터를 사용하여 RU 할당 문제를 식 (2) 같이 sum rate 최대화 문제로 표현할 수 있다

$$\max_{\{n_1, \dots, n_N\}} \sum_{k=1}^N \sum_{m=1}^{n_k} b \log_2 \left(1 + \frac{P \tilde{h}_{k,m}^2}{\sigma^2} \right). \quad (2)$$

설명의 편의를 위해 본 논문에서는 그림 1 의 STA 배치 모델을 가정하였다.

2. DDQN 기반 RU 할당 알고리즘

식 (2)로 표현되는 문제에 DDQN [3] 기법을 적용하기 위해 Markov decision process(MDP) 모델링을 해야한다. State 벡터 $\mathbf{s}(i)$ 는 식 (3)과 같이 STA 들의 채널과 이전 action $a(i-1)$ 로 구성한다.

$$\mathbf{s}(i) = \{\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3^1, \mathbf{h}_3^2, \mathbf{h}_4^1, \mathbf{h}_4^2, \mathbf{h}_5, \mathbf{h}_6, a(i-1)\} \quad (3)$$

여기서 $\mathbf{h}_k, \mathbf{h}_k^l$ 은 주파수 영역 채널 벡터로, 각각

$$\mathbf{h}_k = [h_{k,1}, \dots, h_{k,234}]^T, \mathbf{h}_k^l = [h_{k,1}^l, \dots, h_{k,234}^l]^T \text{ 로 정의된다.}$$

Action $a(i)$ 은 가능한 RUC 에 따른 RU 할당방법들로 정의한다. Reward $r(i)$ 는 식 (2)로부터 계산한 sum rate $R(i)$ 와 이전 iteration 에서의 $R(i-1)$ 의 비교를 통해 $-1, 0, 1$ 세가지 값 중 하나를 가지도록 다음과 같이 정의한다

$$r(i) = \begin{cases} 1 & , R(i) > R(i-1) \\ 0 & , R(i) = R(i-1) \\ -1 & , R(i) < R(i-1) \end{cases}$$

위와 같이 정의된 MDP 로부터 강화학습 기법을 적용할 수 있다. 기존의 deep Q-network(DQN) 기법의 문제점으로 제기된 overestimation 을 완화하기 위해 [3]에 target 을 아래와 같이 계산하도록 제안되었다.

$$Y_i^{\text{target}} = r(i+1) + \gamma Q\left(\mathbf{s}(i+1), \underset{a}{\operatorname{argmax}} Q(\mathbf{s}(i+1), a; \boldsymbol{\theta}(i)); \boldsymbol{\theta}^-(i)\right)$$

여기서 $\boldsymbol{\theta}(i), \boldsymbol{\theta}^-(i)$ 는 각각 Q-network 와 target network 의 파라미터들을 의미한다. 또한 $Q(\mathbf{s}(i), a(i))$ 는 action-value 함수를 나타낸다. 심층강화학습은 감가율 γ 을 고려한 모든 i 에서 예상되는 reward 의 함수인 return 을 최대화하도록 Q 함수를 학습한다. 학습된 Q 함수에 대하여 충분히 큰 i 에서 선택된 $a(i)$ 에 따라 RU 를 할당한다.

3. 모의실험 결과

그림 1 과 같이 AP 가 2 개, 각 BSS 당 STA 가 4 개인 상황을 설정하였다. 또한 겹쳐진 지역에 STA 이 두 개가

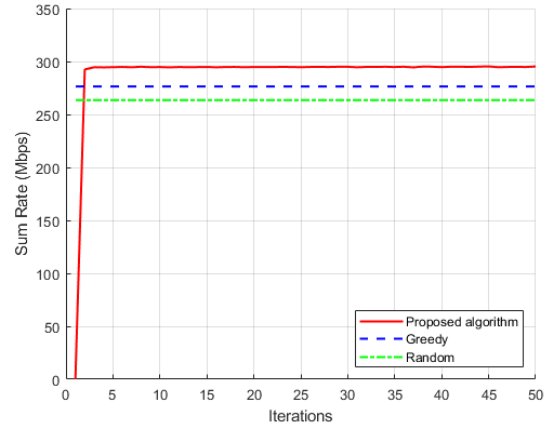


그림 2. Iteration 에 따른 다른 기술과의 sum rate 비교

있는 상황을 고려하였다. 그림 2 는 iteration 수에 따른 sum rate 을 5000 번 실험하여 그 평균을 나타내었다. 본 논문에서는 강화학습을 반복적인 알고리즘처럼 사용하였기에 기존 강화학습에서 매 시간에 해당되는 것을 iteration 이라 표현하였다. 비교하는 다른 기법으로 'Greedy'는 겹쳐진 지역에 있는 STA 에 먼저 RU 를 할당하고 남은 STA 들에 할당하는 방식으로 하였고, 'Random'은 무작위로 RU 를 할당하는 방식을 의미한다. 5 번째 iteration 기준 제안하는 기법은 다른 기법 대비 성능 이득을 확인할 수 있다. 따라서 제안하는 강화학습 기법을 사용할 경우, 고려하는 BSS 내 STA 의 sum rate 관점에서 효과적인 RU 할당이 가능함을 확인하였다.

III. 결론

본 논문에서는 IEEE 802.11 be 에서 상정하는 동시전송상황에서의 RU 할당 문제에 관하여 연구하였다. 동일한 대역대의 두 BSS 가 공유하고 겹쳐진 영역에 있는 STA 는 동일한 RU 를 사용하는 상황에서 DDQN 기법을 통해 반복적인 RU 할당 방법을 제안하였다. 모의실험을 통해 다른 기술에 비해 우수한 성능을 보임을 확인하였고, 또한 빠르게 수렴함을 보였다.

ACKNOWLEDGMENT

본 연구는 삼성전자의 지원(과제번호 IO210203-08373-01)을 받아 수행된 연구임.

참 고 문 헌

- [1] D. Lopez-Perez, et al., "IEEE 802.11 be extremely high throughput: The next generation of Wi-Fi technology beyond 802.11 ax," *IEEE Commun. Mag.*, vol. 57, no. 9, pp. 113-119, Sep. 2019.
- [2] K. Wang and K. Psounis, "Scheduling and resource allocation in 802.11ax," in *Proc. INFOCOM IEEE Conf. Comput. Commun.*, Apr. 2018, pp. 279-287.
- [3] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in *Proc. 30th AAAI Conf.*, Feb. 2016, pp. 2094-2100.