

5G 네트워크에서의 에너지 효율을 고려한 DQN 기반 자원 할당 기법

문승일^o, 홍충선^{*}

경희대학교 컴퓨터공학과

moons85@khu.ac.kr, cshong@khu.ac.kr

DQN-based resource allocation scheme considering energy efficiency in 5G network

Seung Il Moon^o, Choong Seon Hong^{*}

Department. of Computer Science and Engineering, Kyung Hee University

요 약

최근 모바일 기기의 증가 및 모바일 네트워크 기술의 발전으로 인한 모바일 네트워크의 속도 증가로 인하여 매년 모바일 데이터로 인한 트래픽이 기하급수적으로 증가하고 있으며, 차세대 네트워크는 모바일 트래픽에 대한 보다 높은 처리능력이 요구되고 있다. 이를 위해 펌토 셀 기지국을 이용한 데이터 오프로딩 기법에 대한 연구가 활발히 진행되고 있다. 본 논문에서는 5G 기반의 Heterogenous Networks (HetNets) 환경에서 보다 효율적으로 자원을 할당하고 네트워크 사용량을 극대화하기 위한 강화학습 기반의 자원 할당 기법을 제안하고, 시뮬레이션을 통해 제안하는 기법의 성능을 평가한다.

1. 서 론

최근 모바일 기기의 기하급수적인 증가 및 다양한 서비스의 요구사항에 따라 매년 모바일 데이터로 인한 트래픽이 폭발적으로 증가하고 있다.[1] 하지만 기존의 네트워크 환경으로는 Radio Access Networks(RANs)상에 발생하는 병목현상을 해결하기에는 한계가 있다. 이에 이러한 문제를 해결하기 위한 패러다임으로 펌토 셀 기지국을 이용한 데이터 오프로딩, Device-to-Device (D2D) 통신, 네트워크 가상화 등이 제시되고 있으며 다양한 분야에서 관련연구가 수행되고 있는 추세이다. HetNet에서의 주요 연구 목표는 사용자 연결(User Association), 전력 제어, 무선자원 할당 문제를 해결하는 것이다. 하지만 이러한 서브 문제들은 동시에 해결해야만 하는 것이기에 쉽게 해결하기가 어렵다. 이러한 문제를 해결하기 위해 본 저자는 [2]에서 Markov approximation framework를 이용한 Self-Organizing Energy Efficient DL/UL 스케줄링 기법을 제안하였고 시뮬레이션을 통한 성능의 우수성을 입증하였다. [2]에서는 최적의 값을 찾기까지의 학습 시간이 오래 걸리기 때문에 Markov approximation framework를 이용하여 빠른 학습시간으로 근사 최적 값을 찾도록 접근하였다. 본 논문에서는 최적 값을 얻기 위한 학습시간을 단축시키고 성능 향상을 위해 Deep Q-Networks (DQN)[3]을 적용하여 HetNet에서의 사용자 연결, 전력 제어, 무선 자원 할당 문제를 해결하고 전체 네트워크의 사용률 최대화를 해결하기 위한 자원 할당 기법을 제안하였다.

2. System Model 및 Problem Formulation

2.1. System Model

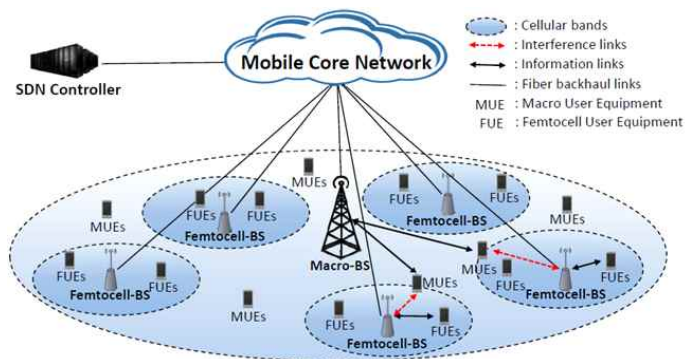


그림 1. System Model

본 논문에서는 그림 1에서와 같이 고정된 기지국들 (BSs)과 임의적으로 위치한 사용자 기기들(UEs)로 구성된 HetNets 환경을 고려하였다. 기지국은 하나의 매크로 BS와 여러 개의 펌토 셀 BS들로 구성되어 있으며 UE는 각각 하나의 기지국과 연결을 맺고 있다. 여기서 BS들의 집합을 \mathcal{B} 로 UE들의 집합을 \mathcal{U} 로 정의한다. 그리고 노드들의 집합인 $\mathcal{S} = \mathcal{B} \cup \mathcal{U}$ 로 표현할 수 있다. 또한 링크의 집합인 $\mathcal{E} = \mathcal{E}_{DL} \cup \mathcal{E}_{UL}$ 으로 표현할 수 있으며, 이를 통해 네트워크 $\mathcal{G} = (\mathcal{S}, \mathcal{E})$ 으로 나타낼 수 있다. 서브 채널의 집합은 \mathcal{F} , 전송 전력의 집합은 \mathcal{P} 로 정의한다. 해당 네트워크 환경에서의 SINR은 다음과 같이 정의할 수 있다.

$$\Gamma_{mi}^k = \frac{h_{mi}P_{mi}^k}{\sum_{n \in v^k/\{m\}} h_{ni}P_{ni}^k + WN_0} \quad (1)$$

이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2017R1A2A2A05000995). Dr. CS Hong is the corresponding author.

그리고 서브 채널 k 에서 링크 (m, i) 의 데이터 전송 속도는 다음과 같이 정의할 수 있다.

$$R_{mi}^k = W \log_2(1 + \Gamma_{mi}^k) \quad (2)$$

2.2. Problem Formulation

네트워크의 에너지 효율을 극대화하기 위하여 다음과 같이 utility function을 정의 하였다. utility function은 소비한 전력 대비 전송 속도이며, bit/s/W로 표현할 수 있다.

$$u_i(x, y, P) = \frac{R_{mi}}{P_{mi}} + \frac{R_{im}}{P_{im}} \quad \text{[Individual]} \quad (3)$$

$$U(x, y, P) = \frac{1}{|\mathcal{U}|} \sum_{i \in \mathcal{U}} u_i(x, y, P) \quad \text{[Global]} \quad (4)$$

그리고 최적화 문제를 해결하기 위하여 다음과 같은 제약사항을 두었다.

Unique association: UE i 는 타임 슬롯 t 에서 오직 하나의 BS와 association을 맺을 수 있다.

$$\sum_{m \in \mathcal{B}} x_i^m \leq 1, \quad \forall i \in \mathcal{U} \quad (5)$$

QoS: BS는 association 맺은 UE의 최소한의 QoS를 보장해야 한다.

$$R_{mi} \geq \psi_i^{DL}, \quad \forall (m, i) \in \mathcal{B} \times \mathcal{U} \quad \text{[DL]} \quad (6)$$

$$R_{im} \geq \psi_i^{UL}, \quad \forall (m, i) \in \mathcal{B} \times \mathcal{U} \quad \text{[UL]} \quad (7)$$

Interference: 먼저 downlink와 uplink에는 반드시 서로 직교 자원들이 할당되어야 한다.

$$y_{mi}^k + y_{im}^k \leq 1, \quad \forall (m, i) \in \mathcal{B} \times \mathcal{U} \quad (8)$$

그리고 conflict와 reuse 링크 쌍은 다음과 같이 정의할 수 있다.

$$y_{mi}^k + y_{nj}^k \leq 1, \quad \{(m, i), (n, j)\} \in \mathcal{E}_{\text{conflict}}, \quad \text{[DL]} \quad (9)$$

$$y_{im}^k + y_{jn}^k \leq 1, \quad \{(m, i), (n, j)\} \in \mathcal{E}_{\text{conflict}}, \quad \text{[UL]} \quad (10)$$

Resource: 네트워크에 사용 가능한 자원은 전체 자원량을 초과해서는 안 된다.

$$\sum_{i \in \mathcal{U}} x_i^m \sum_{k \in \mathcal{S}} (y_{mi}^k + y_{im}^k) \leq |\mathcal{S}|, \quad (11)$$

$$\sum_{i \in \mathcal{U}} P_{mi} \leq \hat{P}_m \quad (12)$$

Heterogeneous Network 환경에 에너지 효율을 고려한 downlink와 uplink 스케줄링에서 i) 사용자 연결, ii) 전력 제어, iii) 무선 자원 할당 문제를 해결하기 위하여 다음과 같이 Joint Optimization Problem(JOP)을 정의하였다.

$$\begin{aligned} \text{JOP : maximize: } & U(x, y, P) \\ & x, y, P \\ \text{subject to: } & (5), (6), (7), (8), \\ & (9), (10), (11), (12). \end{aligned} \quad (13)$$

3. HetNet에서의 Deep Reinforcement Learning

3.1 Q-function Approximation: Q-learning

[2]에서 제안한 downlink/uplink 스케줄링 기법은 최적의 값을 찾기 위한 학습 과정이 오래 걸리는 관계로 Markov Approximation Framework를 이용하여 빠른 시간 내에 근사 최적 값을 찾도록 하여 문제를 해결하였다. 학습 시간 및 accuracy 향상을 위해 본 논문에서는 DQN을 이용하여 [2]의 스케줄링 기법을 학습하였다. 먼저 에너지 효율을 고려하고 네트워크의 이용률을 최대화 하는 최적의 설정 값들을 찾기 위해 학습과정에서 최적의 설정 값을 찾을 확률 $\pi_f^*(U_f)$ 을 [2]에서 학습을 위해 사용하였던 최적의 설정을 찾기 위한 확률 변수를 이용하여 다음과 같이 정의하였다.

$$\begin{aligned} \pi_f^*(U_f) &= \frac{\exp(\beta U_f)}{\sum_{f' \in \mathcal{F}} \exp(\beta U_{f'})} \quad \forall f \in \mathcal{F} \\ &= \frac{\exp(\beta U_f)}{\exp(\beta U_f) + \sum_{f' \in \mathcal{F}} \exp(\beta U_{f'})} \\ &= \frac{1}{1 + \sum_{f' \in \mathcal{F}} \exp[\beta(U_{f'} - U_f)]} \end{aligned} \quad (14)$$

그리고 학습을 위한 Q-함수를 업데이트하기 위한 식을 다음과 같이 정의하였다. [4]

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)] \quad (15)$$

인공 신경망을 업데이트하기 위한 오차함수를 대표적인 오차함수 중 하나인 평균제곱오차 함수 MSE (Mean Squared Error)를 이용하여 다음과 같이 정의하였다.

$$\text{MSE} = (U_f - U_{f'})^2 = (R_{t+1} + \gamma \max_{a'} Q(S_{t+1}, a', \theta^-) - Q(S_t, A_t, \theta))^2 \quad (16)$$

3.2 Q-function Approximation: Q-Nets

아래 그림에서와 같이 HetNets에서 에너지 효율을 고려하고 전체 네트워크의 이용률 극대화를 위한 스케줄링 기법을 위하여 Q-함수 학습을 위한 Q-네트워크 구성하고 (15), (16)을 통해 네트워크를 학습 시킨다.

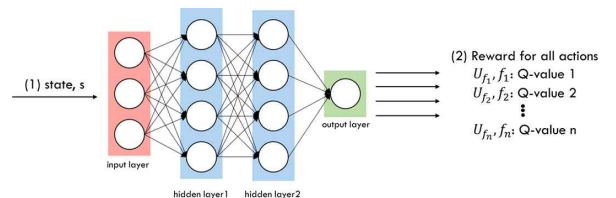


그림 3. Q-네트워크를 이용한 학습 과정

그리고 딥마인드가 [3]에서 제안한 개선된 DQN 알고리즘을 이용하여 downlink/uplink 스케줄링 알고리즘을 학습 시

킨다. [3]에서는 DQN을 이용한 학습을 수행할 때 mainDQN과 targetDQN을 분리시켜 운영하여 최대한 낮은 결과의 학습은 반영되지 않고 학습할 수 있도록 하였다.

4. DQN을 이용한 학습 알고리즘

다음은 DQN을 이용한 HetNets에서 에너지 효율성을 고려한 downlink/uplink 스케줄링 기법을 학습하기 위한 알고리즘이다. 학습을 위한 입력 값으로 학습시킬 Q-네트워크와 네트워크 구성 환경 정보를 이용한다. 초기화 과정을 통해 네트워크를 모델링 하고 학습하기 위한 네트워크 설정 값들을 생성하고 Q-함수를 통해 Q-값을 계산하여 메모리에 저장 시킨다. 그 후 메모리에 저장된 결과를 샘플링 하여 이를 Q-네트워크에 학습 시킨다.

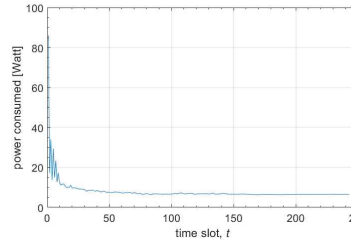


그림 4. 전력 소비량

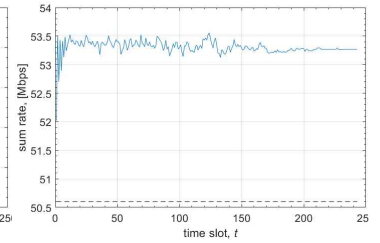


그림 5. 데이터 전송 속도

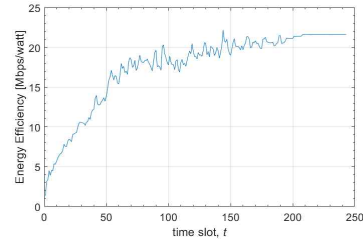


그림 6. 에너지 효율

그림 4는 시뮬레이션을 통해 확인한 제안하는 기법을 통해 학습한 스케줄링 기법을 이용했을 때의 전력 소비량으로 빠른 시간 내에 전력 소비량이 수렴함을 알 수 있다. 그림 5는 데이터 전송 속도에 대한 결과 값으로 그래프 하단의 점선은 요구되는 대역폭을 의미한다. 그림에서와 같이 요구 대역폭을 넘는 수치로 수렴함에 따라 QoS를 만족함을 알 수 있다. 마지막으로 그림 6은 이전 두 그래프 결과를 통해 구한 시간 당 에너지 효율을 나타낸 것으로 모든 결과들이 타임 슬롯 t 가 200일 때 안정적으로 수렴함을 알 수 있다.

6. 결론

본 논문에서는 Heterogenous Networks 환경에서 에너지 효율을 고려하고 전체 네트워크의 사용을 극대화를 위한 스케줄링 기법의 학습을 위하여 DQN을 이용하는 알고리즘을 제안하였다. 또한 시뮬레이션을 통해 DQN 기반의 학습 알고리즘을 이용하여 학습된 스케줄링 기법의 성능을 검증하였고, 빠른 시간 내에 안정적으로 수렴하며, 에너지 효율적으로 네트워크가 운영됨을 확인할 수 있었다.

7. 참고 문헌

[1] “Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016-2021 White Paper”, Cisco, 2017.02.07
 [2] Moon, S., Oo, T. Z., PARK, B. J., & HONG, C. S. (2017). SDN-Based Self-Organizing Energy Efficient Downlink/Uplink Scheduling in Heterogeneous Cellular Networks. IEICE Transactions on Information and Systems, 100(5), 939-947.
 [3] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Petersen, S. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529.
 [4] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602.

알고리즘 1 : DQN을 이용한 학습 알고리즘

1. **Procedure** Training
2. **Input:** Q-network structure, environment information
3. **Output:** Q-network
4. **Start:**
 Random initialize the policy π
 Initialize the network model : generate BSs, UEs, conflict and reuse links
5. **Loop:**
 - Random sample links in the system
 - Generate a set of user info (UE with required data rates) from network model
 - Save the data item {state, reward, action, post-state} into memory
 - Sample a mini-batch of data from the memory
 - Train the deep Q-network using the mini-batch data
 - Update the policy: chose the action with maximum Q-value
6. **End Loop**
7. **Return:** Return the deep Q-network

5. 성능평가

표 15. 시뮬레이션 파라미터

Quantity	Values
Area of region (A)	500 m \times 500 m
Static traffic: # of UE ($ \mathcal{U} $)	100
UE traffic demand (ψ_i^{DL}, ψ_i^{UL})	[0.1, 1] Mbps
# of BS ($ \mathcal{B} = \mathcal{B}_m \cup \mathcal{B}_f $)	25 = 1 + 24
Total transmit power of BSs	{46, 26} dBm
Antenna gain of BSs (G)	{12, 6} dBi
Reference distance of BSs (d_0)	{1000, 20} m
Transmit antenna height of BSs (h_t)	{30, 3} m
# of sub-channels ($ \mathcal{S} $)	12 \times 100
Bandwidth of each sub-channel (W)	15 kHz
Thermal noise for 1 Hz at 20 °C	-174 dBm

성능평가를 위해 MATLAB을 이용하여 다음과 같이 실험을 수행했다. 먼저 아래 표는 실험에 사용된 파라미터에 대한 정보이다. 이를 기반으로 네트워크를 생성하고 시뮬레이션을 통해 성능을 평가하였다.