

네트워크 에너지 효율 최대화를 위한 불균일 트래픽 분포를 고려한 멀티 에이전트 큐러닝 기반 기지국 제어 기법

김은진(한경대학교), 이호원(한경대학교)

{gate1180, hwlee}@hknu.ac.kr

Non-Uniformly Distributed Traffic-Aware Base Station Control Based on Multi-Agent Q-Learning for Maximizing Energy Efficiency

Eunjin Kim(Hankyong National Univ.), Howon Lee(Hankyong National Univ.)

요약

최근 지구 온난화가 가속화됨에 따라 환경 친화적인 기술에 대한 수많은 연구가 진행 중이다. 또한, 도시화에 따른 인구 집중 등으로 인하여 지역마다 발생하는 트래픽이 매우 불균형이고 이에 대한 효과적인 지원이 필요하다. 이에, 본 논문에서는 고밀도 이기종 네트워크에서 각 스몰셀 기지국의 전송 전력을 효과적으로 조절함으로써 트래픽을 효율적으로 처리하는 멀티 에이전트 큐러닝 기반 기지국 제어 기법을 제안한다.

I. 서론

고밀도 이기종 네트워크에서 트래픽 발생량에 따라 기지국의 전력을 효율적으로 조절하는 것은 네트워크 에너지 효율 향상을 위해 매우 중요하다 [1][2]. 실제 네트워크에서 트래픽은 지역에 따라 매우 불균일하게 발생할 수 있으므로 사용자를 커버하기 위해 사용하는 전송 전력은 각 기지국마다 달라져야 한다 [3]. 이에, 본 논문에서는 고밀도 이기종 네트워크에서 사용자 분포에 따라 스몰셀 기지국이 전송 전력을 조절하여 에너지 효율적으로 사용자를 커버할 수 있는 강화학습 기반 기지국 전력 할당 기법을 제안한다. 특히, 강화학습 기반의 분산 Q-learning을 사용하여 각 스몰셀 기지국을 에이전트로 정의함으로써, 트래픽 패턴의 예측이 불가능한 복잡한 시뮬레이션 환경에서 환경에 대한 정보 없이도 자체 학습이 가능하도록 한다.

II. 멀티 에이전트 큐러닝 기반 기지국 전력 제어 기법

본 논문에서는 U 명의 사용자(user), N 개의 스몰셀 기지국(small-cell base station, SBS), M 개의 매크로셀 기지국(macro-cell base station, MBS)으로 구성된 네트워크에서 SBS j 에 대한 사용자 i 의 signal-to-interference-plus-noise ratio (SINR) $\gamma(i, j)$ 를 다음과 같이 계산한다.

$$\gamma(i, j) = \frac{P_r(i, j)}{\sum_{n \neq j, n \in N} P_r(i, n) + \sum_{m \in M} P_r(i, m) + \sigma_i^2} \cdot (2)$$

여기서, $P_r(i, j)$ 는 사용자 i 와 SBS j 의 reference signal received power (RSRP)로 정의되며, SINR 계산식에서 MBS와 다른 SBS는 간섭원으로 작용한다. σ_i^2 은 사용자 i 의 thermal noise power이다. 각 SBS j 는 사용 가능한 대역폭 W_j 와 자신과 연결된 사용자 수 $U_c(j)$ 을 기반으로 achievable data rate $\zeta(i, j)$ 와 에너지 효율 $\xi(j)$ 을 계산할 수 있다.

$$\zeta(i, j) = (W_j \div U_c(j)) \cdot \log_2(1 + \gamma(i, j)). \quad (2)$$

$$\xi(j) = \frac{\sum_{u \in U_c(j)} \zeta(u, j)}{P_t(j) + P_s(j) + P_c(j)}. \quad (3)$$

기지국의 에너지 효율은 기지국이 소모하는 총 전력량에 반비례한다. (3)에서 $P_s(j)$ 는 SBS j 가 커버하는 사용자 수 $U_c(j)$ 에 따라 달라지는 signal processing 전력이고, $P_c(j)$ 는 circuit power이다.

본 논문에서는 각 SBS를 에이전트로 정의한다. 각 에이전트는 자신의 전송 전력 $P_t(j)$ 를 상태로 정의하고, 전송 전력을 조절하여 에너지 효율을 최대화하는 방향으로 학습하게 된다. 에이전트 j 의 보상 값은 다른 에이전트의 에너지 효율 값 활용 여부에 따라 R_j^c , R_j^d 로 구분된다. R_j^c 는 다른 에이전트의 에너지 효율을 사용하여 $R_j^c = \sum_{n \in N} \xi(n)$, R_j^d 는 에이전트 자신의 에너지 효율만을 사용하여 $R_j^d = \xi(j)$ 로 정의된다. 또한, 에이전트의 충분한 exploration을 위해

서 감쇠 입실론 탐욕 정책(decayed epsilon greedy policy)를 행동 선택 정책으로 사용하였다. 매 타임 스텝 t 에서 에이전트 j 의 상태 $s_j(t)$ 에 따른 행동 $a_j(t)$ 에 대한 Q-table 업데이트 식은 감가율 α , 학습률 η 에 의해 다음과 같이 정의된다.

$$Q_j(s_j(t), a_j(t)) = (1 - \alpha) Q_j(s_j(t), a_j(t)) + \alpha [R_j + \eta \max_{a_j(t)} Q_j(s_j(t+1), a_j(t))]. \quad (4)$$

III. 시뮬레이션 결과 및 분석

제안 방안과 'no transmit power control (No TPC)', 그리고 'random action' 기법을 비교하였다. 그림 1(a)는 특정 지역에 불균일하게 발생하는 트래픽 밀도(0~1의 값)에 따라 색상을 지정하여 각 SBS의 커버리지를 플로팅한 것으로 1에 가까울수록 많은 사용자를 커버하고 있음을 보여준다. 또한, 그림 1(b)는 제안 방안의 수렴성을 확인하기 위해 각 에피소드에서 총 T 번의 행동 후, 계산된 모든 보상 값을 평균화한 것으로 학습이 진행됨에 따라 제안 방안은 SBS의 전송 전력이 2W로 고정된 'No TPC' 값에 비해 높은 값으로 수렴하는 것을 확인할 수 있다. 또한, 'Proposed(Centralized)'는 네트워크 전체의 에너지 효율 값을 사용하여 학습하기 때문에 각 에이전트의 에너지 효율 값을 사용하여 학습하는 'Proposed(Distributed)'에 비해 비교적 높은 값에 수렴한다.

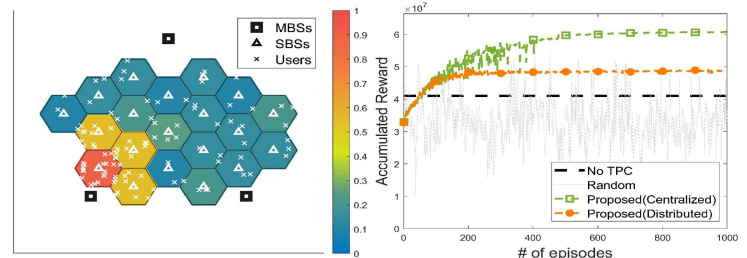


그림 1. (a) 네트워크 트래픽 분포, (b) Accumulated Reward

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1F1A1063606).

참고 문헌

- [1] H. Yu, H. Lee, and H. Jeon, "What is 5G? emerging 5G mobile services and network requirements," MDPI Sustainability, vol. 9, pp. 1 - 22, Oct. 2017.
- [2] W. Lee, B. C. Jung, and H. Lee, "DeCoNet: Density Clustering Based BS Control Algorithm for Energy-Efficient Ultra-Dense Networks," IEEE Access, vol. 8, pp. 120881-120891, Jul. 2020.
- [3] P. Frenger, C. Friberg, Y. Jading, M. Olsson and O. Persson, "Radio network energy performance: Shifting focus from power to precision", Ericsson Rev., vol. 2, pp. 1-5, Feb. 2014.