

대규모 MIMO 시스템에서 주파수 효율성 및 사용자 공정성 공동 최적화를 위한 다중목표 강화학습 기반 전력 할당 기법

오영우^o

최우열

조선대학교 컴퓨터공학과

Multi-Objective Reinforcement Learning-Based Power Allocation for Joint Optimization of Spectral Efficiency and User Fairness in Massive MIMO Systems

Youngwoo Oh^o

Wooyeol Choi

Dept. of Computer Engineering, Chosun University

ywo@chosun.kr

wyc@chosun.ac.kr

요 약

본 논문에서는 대규모 MIMO 시스템에서 요구되는 주파수 효율성과 사용자 공정성을 동시에 최적화하기 위한 multi-objective reinforcement learning (MORL) 기반의 송신 전력 할당 기법을 제안한다. 제안하는 방식은 다중 목표 사이의 우선순위를 결정하기 위해, 각 지표의 선호도 가중치를 적응적으로 조정한다. 이때, 주파수 효율성과 사용자 공정성은 보상 벡터로 수집되며, 각 목표의 선호도 가중치는 베이스 규칙에 의해 계산되는 사후확률을 바탕으로 결정된다. 실험 결과를 통해, 제안하는 MORL 기반 전력 할당 기법이 타 기법들을 능가하는 최적화 성능을 달성할 수 있음을 확인하였다.

1. 서론

Massive multiple-input multiple-output (MIMO) 기술은 다수의 안테나를 활용하여, 주파수 효율성을 개선할 수 있는 5G 네트워크의 핵심 기술이다. 그러나, 기지국에 배치되는 안테나의 개수 증가로 인해 사용자 간 간섭이 증가하며, 이는 사용자 공정성의 감소로 이어진다. 이를 해결하기 위해, massive MIMO 시스템에서의 다양한 지표를 동시에 최적화하기 위한 multiple objective optimization (MOO) 연구가 활발히 수행되고 있다. 그러나, 해당 연구의 경우, 여러 지표를 동시에 최적화해야 하므로 높은 계산 복잡도가 요구된다 [1]. 따라서, 본 논문에서는 계산 복잡도의 증가 없이, massive MIMO 시스템의 주파수 효율성과 사용자 공정성을 동시에 최적화할 수 있는 MORL 기반의 전력 할당 기법을 제안한다.

2. 본론

2.1 Multi-Cell Massive MIMO 시스템

본 논문에서는 L 개의 cell 내 M 개의 송신 안테나를 가진 massive MIMO 기지국과 K 명의 사용자가 존재하는 하향링크 시나리오를 고려한다. 임의의 cell l 내 기지국 j 의 전송 신호는 $\sum_{i=1}^K v_{j,k} s_{j,k}$ 로 표현되며, $v_{j,k}$ 는 프리 코딩 벡터, $s_{j,k}$ 는 사용자 k 에 대한 하향링크 신호를 의미한다. 따라서, 수신되는 하향링크의 신호는 기지국 j 의 전송된 신호, 인접 기지국 및 사용자에게 의해 발생하는 간섭과 잡음의 합으로 구성되어, 아래와 같이 표현할 수 있다.

$$y_{j,k} = \underbrace{v_{j,k} s_{j,k}}_{\text{Desired signal}} + \underbrace{\sum_{l \neq j}^L \sum_{i \neq k}^K v_{l,i} s_{l,i}}_{\text{Interference}} + \underbrace{n_{j,k}}_{\text{Noise}} \quad (1)$$

위의 수식에서의 $n_{j,k}$ 는 분산이 σ^2 인 additive white Gaussian noise (AWGN)를 의미하며, 기지국 j 와 사용자 k 의 SINR은 아래와 같이 정의된다.

$$\lambda_{j,k} = \frac{p_{j,k} \alpha_{j,k}}{\sum_{l=1}^L \sum_{i=1}^K p_{l,i} \beta_{l,i} + \sigma^2} \quad (2)$$

이때, $p_{j,k}$ 는 송신 전력, $\alpha_{j,k}$ 와 $\beta_{l,i}$ 는 각각 채널, 간섭 이득을 의미한다. 달성 가능한 전송률 $C_{j,k}$ 는 $\log_2(1 + \lambda_{j,k})$ 와 같이 표현될 수 있으며, 전체 채널 용량 CP는 $\sum_{j=1}^J \sum_{k=1}^K C_{j,k}$ 와 같이 계산된다. 또한, 사용자 공정성은 모든 셀 L 에서의 모든 사용자 K 에 대한 공정성을 나타내며, 수식 3에 의해 계산된다.

$$FI = \frac{(\sum_{j=1}^J \sum_{k=1}^K C_{j,k})^2}{J \sum_{j=1}^J \sum_{k=1}^K (C_{j,k})^2} \quad (3)$$

2.2 다중목표 강화학습 기반 전력 할당 기법

본 연구에서는 다중 목표를 동시에 최적화하기 위한 multi-objective asynchronous advantage actor-critic (MOA3C) 알고리즘을 제안한다. 알고리즘을 적용하기 앞서 multi-objective Markov decision process (MOMDP)를 정의하는 과정을 수행한다. 상태 $S(t)$ 는 행동에 대한 SINR, 채널 용량, 사용자 공정성으로 구성된 유한 집합으로 정의된다.

$$S(t) = [\lambda_{j,k}(t), CP(t), FI(t)], \sim \forall j, k \quad (4)$$

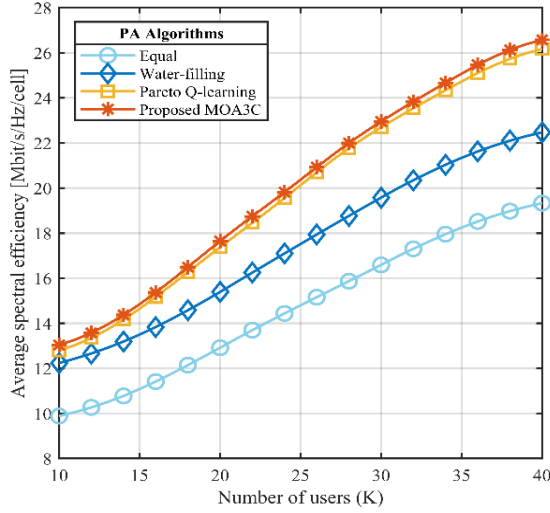


그림 1. 전력 할당 알고리즘별 주파수 효율성 분석

행동 $A(t)$ 는 학습 간 복잡도를 고려하여, 최소 전력 P_{\min} 과 최대 전력 P_{\max} 사이의 특정 단계로 나누어 선택되며, 주파수 효율성과 사용자 공정성에 대한 보상은 다음과 같이 벡터 형태로 구성된다.

$$R(t) = [CP(t), FI(t)] \quad (5)$$

마지막으로, 다중목표의 보상에 대한 우선순위를 결정하기 위해 선호도 개념이 추가되며 이때, 각각의 선호도는 0~1 사이의 값으로 수식 6에 의해 초기화된다.

$$p(\phi_o | D = 0) = \frac{1}{O}, \quad \forall o = \{1, 2, \dots, O\}, \quad (6)$$

2.3 베이지 규칙 기반의 선호도 업데이트 방안

본 논문에서는 주파수 효율성과 사용자 공정성 사이의 우선순위를 적응적으로 결정하기 위해 수식 7과 같은 베이지 규칙을 활용한다.

$$p(\phi_o | D) = \frac{p(D | \phi_o)p(\phi_o)}{\sum_{o=1}^O p(D | \phi_o)p(\phi_o)} \quad (7)$$

위의 수식은 목표 ϕ_o 에 대해 관찰된 보상 D 와 수식 6에 의해 초기화된 사전 확률을 통해 선호도 가중치를 업데이트하는 프로세스를 의미한다. 이때, $p(\phi_o | D)$ 는 관찰된 D 에서 목표 ϕ_o 의 사후 확률을 의미하고, 각각의 $p(D | \phi_o)$, $p(\phi_o)$ 는 주어진 목표 ϕ_o 에서 관찰되는 보상 D 의 우도(likelihood) 함수, 목표 ϕ_o 의 사전 확률을 의미한다. 따라서, 제안하는 선호도 갱신 기법은 목표의 보상과 사전 확률을 기반으로 선호도 가중치를 업데이트하여, 각 목표의 우선순위를 적응적으로 결정할 수 있다.

2.4 MOA3C 알고리즘 설계 및 구현

MOA3C 알고리즘은 수식 7을 통해 각 목표의 선호도를 조정함으로써, trade-off에 의한 sub-optimal 수렴을 극복하고, 주파수 효율성과 사용자 공정성을 동시에 최적화하는 방향으로 학습할 수 있으며, 정책 θ 는 아래 수식 8과 같은 형태로 업데이트된다.

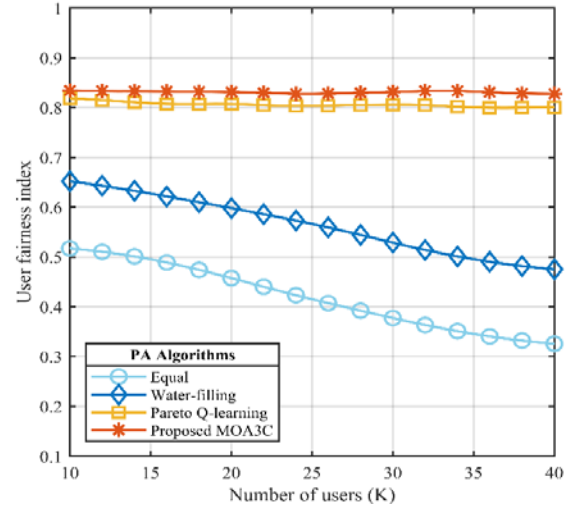


그림 2. 전력 할당 알고리즘별 사용자 공정성 분석

$$\Sigma_{t=1}^T \left[\left[\sum_{o=1}^O w_o (R_o(S(t)) - V_o(S(t))) \right] \nabla \log_{\pi}(A(t) | S(t)) \right] \quad (8)$$

이때, $R_o(S(t))$, $V_o(S(t))$ 는 각각 목표에 대한 보상과 가치함수를 의미하고, 선호도 가중치 w_o 는 수식 6, 7에 의해 얻어지는 각 목표의 우선순위를 결정한다. 또한, 각 목표에 대한 보상과 가치 함수 간의 차이를 계산하여, 선택한 행동이 예상과 비교하여 더 나은지를 판단함으로써, 주파수 효율성과 사용자 공정성을 동시에 개선하도록 정책 θ 를 업데이트할 수 있다.

그림 1, 2는 $L = 4$, $M = 64$ 일 때, 사용자 수 증가에 따른 주파수 효율성과 사용자 공정성을 분석한 결과를 나타낸다. 제안하는 MOA3C 기반의 전력 할당 기법이 주파수 효율성과 사용자 공정성 모두 우수한 최적화 성능을 달성할 수 있음을 보인다.

3. 결론

본 실험에서는 주파수 효율성과 사용자 공정성을 동시에 최적화하기 위한 MOA3C 알고리즘과 베이지 규칙 기반의 선호도 조정 기법을 활용한 전력 할당 기법을 제안하였다. 실험 결과를 통해, 제안하는 기법이 기존의 최적화 알고리즘 대비 높은 최적화 성능을 달성할 수 있음을 확인하였다. 추후 후속 연구에서는 차세대 무선 통신에서 발생하는 다중 목표 간의 trade-off 문제를 최적화할 수 있는 MORL 기반의 자원할당 연구를 수행할 예정이다.

4. 감사의 글

본 연구는 정부(교육부)의 재원으로 한국연구재단의 지원으로 수행되었음(No. NRF-2021R1I1A3A050535).

5. 참고 문헌

- [1] Z. Liu, W. Du, and D. Sun, "Energy and Spectral Efficiency Tradeoff for Massive MIMO Systems with Transmit Antenna Selection", in *IEEE Transaction on Vehicular Technology*, vol. 66, no. 5, pp.4453-4457, May. 2018.