

다중슬롯머신과 최신 통신 시스템 응용

이주현
한양대학교

joohyunlee@hanyang.ac.kr

Multi-armed Bandits and Recent Applications in Communication Systems

Joohyun Lee
Hanyang University

요 약

다중 슬롯 머신 (Multi-Armed Bandits)은 강화학습 (Reinforcement Learning) 문제의 가장 간단한 형태 중에 하나이다. 본 논문에서는 통신 시스템에서 다중슬롯머신을 이용한 최신 응용사례를 살펴본다.

I. 서 론

다중 슬롯 머신 (Multi-Armed Bandits)은 강화학습 (Reinforcement Learning) 문제의 가장 간단한 형태 중에 하나이다. [1] 간단한 문제의 형태 덕분에 다양한 수학적 분석이 이루어져왔으며 또한 통신 시스템을 포함한 많은 응용에 쓰이고 있다. 본 논문에서는 통신 시스템에서 다중슬롯머신을 이용한 최신 응용사례를 살펴본다.

II. 본론

대표적인 다중 슬롯 머신 알고리즘으로는 epsilon-greedy, upper confidence bound (UCB), EXP3 (Exponential-weight algorithm for Exploration and Exploitation), Thompson sampling 과 그 변형들이 있다. [1] UCB, EXP3, Thompson sampling 은 모두 일반적인 가정하에서 로그 함수로 증가하는 이론적인 최적 (order-optimal)의 리그렛 (regret)을 달성하지만 문제 환경에 따라 성능은 차이가 있다.

[2]에서는 802.11 프로토콜에서 최적 전송률 설정 (Rate adaptation) 문제에 변형된 UCB 알고리즘인 ORS (Optimal Rate Sampling)을 개발하여 적용하였다. [3]에서는 다중사용자 다중채널 (Multi-user Multi-channel) 상황에서의 인지 라디오 (Cognitive radio) 문제를 위해 복잡도가 낮은 GYRO (Greedy in Randomized Orders) 알고리즘을 개발하여 적용하였다. 환경이 시변하는 상황 (Non-stationary environment)에 적용하기 위한 알고리즘으로는 [4] 등이 있다.

III. 결론

본 논문에서는 통신시스템에서 다중슬롯머신을 이용한 최신 응용사례를 살펴보았다. 많은 통신시스템의 문제들이 다중슬롯머신을 이용하여 모델링될 수 있으며 알고리즘의 수학적 분석 등에 활용가능성이 높다.

참 고 문 헌

- [1] Tor Lattimore, and Csaba Szepesvári, "Bandit algorithms," Cambridge University Press, 2020.
- [2] Richard Combes, Alexandre Proutiere, Donggyu Yun, Jungseul Ok, and Yung Yi, "Optimal rate sampling in 802.11 systems," IEEE INFOCOM, 2014.
- [3] Sunjung Kang and Changhee Joo, "Low-complexity Learning for Dynamic Spectrum Access in Multi-User Multi-Channel Networks," IEEE INFOCOM, 2018.
- [4] Fang Liu, Joohyun Lee, and Ness Shroff, "A change-detection based framework for piecewise-stationary multi-armed bandit problem," Thirty-Second AAAI Conference on Artificial Intelligence, 2018.