

数理・情報のフロンティア
2021 年度採択研究代表者

2022 年度
年次報告書

土屋 平

京都大学 大学院情報学研究科
大学院生

高速かつ高性能な広範にわたる逐次的意思決定問題の方策開発と解析

研究成果の概要

本年度は主にバンディット問題における best-of-both-worlds (BOBW) 方策について研究を行った。バンディット問題では背後に決定される報酬列の生成のされ方によって、確率的設定と敵対的設定という2つの大きく異なる設定があり、それらで達成可能な性能は異なっている。しかし、ある実問題がどちらの設定に属するかを知るのは困難であり、両方の設定において同時に最適性を達成する方策が望まれる。このような性質を持つ方策は BOBW 方策と呼ばれ、近年盛んに研究されてきた。しかし、既存の BOBW 方策の確率的設定における best (最適性) の定義は、背後の報酬分布の期待値に注目した場合の最適性であり、これは報酬分布の高次の情報から定まる理論限界と乖離がある。さらに、実用的な側面では、広告配置問題やオンライン最短経路問題などの多くの実問題では、背後の報酬分布が非常に小さい分散を持つという特徴がある。以上を背景に、本年度は敵対的設定における性能を維持しつつ、確率的設定においては分散を考慮したリグレット上界を達成するような BOBW 方策の構築に取り組んだ。我々は、BOBW 方策を実現する上で最も代表的な方策の枠組みである follow-the-regularized-leader においてその学習率を漸近的に分散に収束する経験的推定量に基づき設計することで、これを実現できることを明らかにした。この提案方策は、数値実験において既存の分布の分散を考慮しない BOBW 方策と比較して、大幅に良い性能を示した。以上の成果の他に、構造を伴うバンディット問題における BOBW 方策や固定予算のもとでの最適腕識別問題において研究成果を挙げた。

【代表的な原著論文情報】

- 1) "Adversarially Robust Multi-Armed Bandit Algorithm with Variance-Dependent Regret Bounds,"
In Proceedings of Thirty Fifth Conference on Learning Theory (COLT 2022), pp.1421–1422, 2022.
- 2) "Minimax Optimal Algorithms for Fixed-Budget Best Arm Identification,"
In Advances in Neural Information Processing Systems 35 (NeurIPS 2022), pp.10393–10404, 2022.
- 3) "Best-of-Both-Worlds Algorithms for Partial Monitoring,"
In Proceedings of The 34th International Conference on Algorithmic Learning Theory (ALT 2023), pp.1484–1515, 2023.