

レーザネットワークの同期現象を用いた競合バンディット問題における意思決定

Decision making for competitive bandit problem using synchronization of laser network

埼玉大¹、東大情理²

○巳鼻 孝朋¹、菅野 円隆¹、成瀬 誠²、内田 淳史¹

Saitama Univ.¹, Univ. Tokyo²

°Takatomo Mihana¹, Kazutaka Kanno¹, Makoto Naruse², and Atsushi Uchida¹

E-mails: t.mihana.229@ms.saitama-u.ac.jp, auchida@mail.saitama-u.ac.jp

はじめに: 近年、光の高速性や空間多重性を用いてコンピュータの処理を高速化させる光アクセラレータ[1]の研究が活発に行われている。また強化学習における課題の1つとして、競合バンディット問題[2]が挙げられる。これは複数人のプレーヤがスロットマシンを選択して報酬を受け取り、全プレーヤの総報酬を最大化する問題である。

競合バンディット問題を解くために、光による実装として量子もつれ光子を用いた実装が行われている[3]。光子がビームスプリッタを通過することで透過や反射をし、光検出器を置くことで進行方向を判別する。ここで、ビームスプリッタに通す前に偏光板で偏光を変えることで光検出器の検出割合を変更させて意思決定を行う。さらに、量子もつれ光子を用いることでプレーヤの選択競合を起こさないまま意思決定ができる。しかしながら、量子もつれ光子による意思決定では2台のスロットマシンにおける問題に限られている。

一方でレーザカオスによる遅延カオス同期に着目した意思決定[4]が報告されている。遅延カオス同期では先行振動するレーザ(リーダー)と遅れて同期するレーザ(ラガード)が観測される。また、結合強度を変化させることでリーダーとなる時間の長さを変えられる。そこでレーザとスロットマシンを対応させることで意思決定を達成でき、レーザネットワークを用いることでより多くスロットマシンに対するバンディット問題を解くことができる[4]。レーザネットワークのスケラビリティの良さにより、本手法を競合バンディット問題へ対応させることが可能であると考えられる。

そこで本研究ではレーザネットワークにおける同期現象を用いた競合バンディットの意思決定を行うことを目的とする。

方法と結果: 図1(a)には2人のプレーヤで3台のスロットマシンの競合バンディット問題を示している。また、複数のプレーヤが同じスロットマシンを選択した場合には報酬が山分けになる。従って、他プレーヤとの競合を避けながら良いスロットマシンを探す必要がある。本研究では、各プレーヤに対して1つの半導体レーザネットワークを設定する。それぞれのレーザネットワークは一方結合されたリング状のレーザネットワークを使用する。また、半導体レーザの個数は競合バンディットにおけるスロットマシンの台数に対応させる。図1(a)では3つのレーザのレーザネットワークにおいてリーダーとなるレーザを選択し、結合強度を変更していくことで意思決定を行う。本研究では競合バンディット問題による綱引きボンベ理論[5]を拡張した提案手法を導入する。綱引

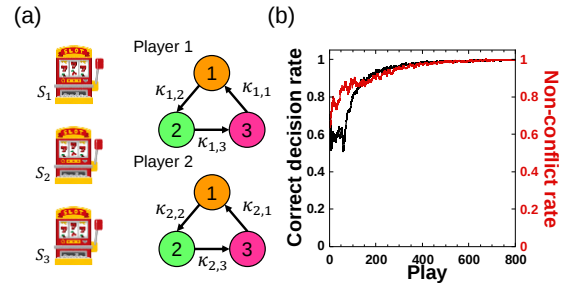


Fig. 1. (a) Schematic diagram for competitive bandit problem. (b) Evaluation for proposed method by (black) correct decision rate and (red) non-conflict rate.

きボンベ理論では意思決定初期に競合が起きるが、時間経過とともに競合を回避しながら、良いスロットマシンを選択できるアルゴリズムである。本研究ではあるプレーヤが先に最も良いスロットマシンを見つけた時に、そのプレーヤだけが一番良いスロットマシンを選択し、その他のプレーヤは一番良いスロットマシン以外から選択させるような専属的選択をアルゴリズムに追加した。

提案手法の評価指標として、平均正答率と非選択衝突率を用いる。2つの指標は意思決定を複数回試行することで求められる平均的な評価である。平均正答率では当たり確率の高いスロットマシンを選択できるかという指標であり、平均正答率が1のとき全プレーヤが当たり確率の高いスロットマシンを定常的に選択できることを示す。一方で非選択衝突率では、各回の選択においてプレーヤ同士が同じスロットマシンを選択していないかという指標になっており、非選択衝突率が1の時に全プレーヤが定常的に異なるスロットマシンを選択できることになる。本研究では2プレーヤ3台のスロットマシンに対する競合バンディット問題を Fig. 1(b) のように評価した。提案手法では平均正答率と非選択衝突率ともに1に到達していることから、選択競合を起こさずに、両プレーヤが良いスロットマシンを選択できていることが分かる。

まとめ: 本研究では一方向結合された半導体レーザネットワークによる遅延カオス同期を用いて、競合バンディット問題の解決を行った。提案手法を用いることで、プレーヤでの選択競合が起きず、良いスロットマシンを選択できる意思決定を達成した。

参考文献

- [1] K. Kitayama, et al., APL Photonics, **4**, 090901 (2019).
- [2] L. Lai, et al., IEEE Trans. Mob. Comput., **10**, 239 (2011).
- [3] M. Naruse, et al., Sci. Rep., **8**, 10890 (2018).
- [4] T. Mihana, et al., Opt. Exp., **27**, 40112 (2020).
- [5] S.-J. Kim, et al., Philosophies, **1**, 245 (2016).