

科学研究費助成事業 研究成果報告書

令和 3 年 6 月 23 日現在

機関番号：14301

研究種目：若手研究

研究期間：2018～2020

課題番号：18K17998

研究課題名(和文) 情報論的手法に基づく知識探索アルゴリズムの構築

研究課題名(英文) Construction of knowledge discovery algorithms based on information theoretic methods

研究代表者

本多 淳也 (Honda, Junya)

京都大学・情報学研究科・准教授

研究者番号：10712391

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：本研究では、限られた試行回数で探索すべき候補を適切に選択する問題である多腕バンディット問題において、情報論的手法に基づいた効率的な方策の開発を行った。特に、従来ヒューリスティックとして多く用いられてきたトムソン抽出とよばれる方策について、多くの複雑な設定において情報論的理論限界の観点からの性能保証を行った。さらに、累積報酬最大化ではなく期待値最大の候補の発見を目指す問題について、従来では本質的に非常に多くの試行数あるいは計算量が必要となっていた設定において、問題の困難さをそれぞれ情報論的に定量化することで現実的な試行数および計算量で実行可能な問題設定および方策の開発を行った。

研究成果の学術的意義や社会的意義

本研究の結果はトムソン抽出の適用可能性とその限界を明らかにしたものであるが、この方策は推薦システムなど既に実社会で多く用いられているものであり、その正当性を明らかにすることはバンディット方策を安全に社会で運用することに貢献するものである。また、この分野の発展に伴いこれらの方策を治験などより社会的に繊細な問題に対して適用しようとする試みが近年あるが、これらの設定では推薦システムといった設定に比べて可能な試行回数が大幅に少ないことが障害になっている。本研究はこういった設定に対しても意味のある保証が可能な枠組みを定式化した点で、より社会の広範な設定でバンディット方策を適用可能とする意義をもつ。

研究成果の概要(英文)：The multi-armed bandit problem is a problem of appropriately finding and choosing the candidates to be explored under a limited number of trials. In this research, we investigated policies for this problem based on the techniques of information theory. In particular, we established theoretical guarantees of the policy called Thompson sampling from the viewpoint of the information-theoretic lower bound, whereas this policy has been often used as an empirically promising heuristics. Furthermore, we also addressed the problem of finding the best candidate with the largest reward expectation rather than maximizing the cumulative reward in the multi-armed bandit problem. In this problem, existing formulations often required unrealistically large trials and heavy computation. In this research we formulate problems that are feasible under a realistic number of trials with practical algorithms by appropriately establishing the information-theoretic difficulty of the problem.

研究分野：機械学習

キーワード：機械学習 情報理論

1. 研究開始当初の背景

大規模なデータから有益な情報を取り出す機械学習の技術は最近では様々な場面で応用されているが、一方で現実では扱う対象の十分なデータがないまま意思決定が必要となる場面も数多く現れる。例えばウェブ上の広告配信では、クリック率（あるいは購入率）の高い広告を多く表示することが配信主の目標となるが、各広告のクリック率は事前には未知である。このように、ある行動を実際にとらない限りそれに関する情報が得られないという状況で逐次的に適切な行動を選択する問題は、機械学習において多腕バンディット問題あるいは単にバンディット問題とよばれる。この問題における代表的な目的設定は累積で得られた報酬を最大化するというものであり、これまで本研究代表者は、この設定に対して理論限界およびそれを達成するアルゴリズムの構成をおこなってきた。

一方、バンディット問題の実応用においては、累積報酬を最大化するのではなく報酬期待値が最大の候補をなるべく少ない試行回数で発見するという問題も数多く現れる。例えば、臨床試験を通過した複数の新薬のいずれを患者に割り振るべきかという問題は累積報酬最大化として定式化できるが、一方で臨床試験の段階では薬の成分の候補（あるいは用量）のうち最も効果が大きいものを発見することが重要となる。

このように、複数ある行動の候補のうちある要求を満たすものをなるべく少ない試行回数・なるべく小さい誤り確率で発見する問題は純粋探索問題とよばれ、それらのうち期待値最大の行動を発見する問題は最適腕識別とよばれる。この問題は最近では累積報酬最大化において構築されてきた手法を応用する形で少しずつ進展してきているものの、知られている理論限界と実際に達成可能な性能の間には未だ大きなギャップがあった。さらに、臨床試験といった状況では試行回数（被験者数）が非常に限られているが、そのような制約下で最適腕識別の枠組みで意味のある結論を導くこと（例えば、信頼度 95%で最適な候補を発見する）は理論限界に照らし合わせるとそもそも不可能であるという場合も多い。このような場合、例えば最適な候補を低い信頼度 60%で出力するよりも、報酬期待値が少なくとも一定の閾値を超える候補のいずれか 1 つを高い信頼度 90%で出力するといったようにそもそも別の純粋探索問題を考えたほうが実用上有益となりやすいが、これまで知られていた研究は最適腕識別に特化しており、必ずしも他の問題に対して自然に拡張可能な形となっていなかった。

2. 研究の目的

これらの背景を踏まえ、本研究課題では以下の 2 点を中心的な課題とした研究を行った。

- ・実問題の制約と理論限界との比較から望ましい性能が得られないことが分かる状況において、その制約下で意味のある結論を得るための別の純粋探索問題の枠組みを構築する
- ・最適腕識別、および一般の純粋探索問題において、実用的かつより精密な理論限界を達成するためのアルゴリズムを構成する

特に、バンディット問題における理論的基盤には主に情報論的理論限界とよばれる枠組みが多く用いられているが、本研究代表者は情報理論の研究にも長く従事しており、それらの手法をバンディット問題の研究にも多岐にわたり応用してきた。そこで、これらの情報論的手法をさらに発展させることで、さまざまな純粋探索問題においても適応可能な理論およびアルゴリズムを構築することを目指した。

3. 研究の方法

これらの目的を達成するため、まず現実的な試行数で扱うことのできる純粋探索問題の構築を行った。真に期待値最適な候補（あるいはそれから 以内の期待値をもつ候補）を発見することが求められる最適腕識別に対して比較的少ない試行数で実行可能な純粋探索問題としてはしきい値付きバンディットとよばれる設定が従来知られていた。これは各候補を指定された報酬期待値のしきい値を上回るかどうかについてそれぞれ判別を行う枠組みで、期待値最大に近い候補が複数ある場合でも比較的少ない試行数で実行可能であるのに対する一方で、しきい値に報酬期待値が非常に近い候補が 1 つでも存在すると大きく性能が悪化するという問題点があった。一方、例えば医療画像の診断や新薬の用量の選定といった設定では、全ての部位に対して以上の有無を判別する、あるいは全ての用量において効果の有無を判別するといったことは必ずしも求められず、異常のある箇所を 1 つでも発見する、あるいは効果のある用量を 1 つ発見する、といったことを可能な限り少ない試行回数で行うこと求められる。そこで、本研究課題では事前に定めたしきい値を上回る報酬期待値をもつ候補をなるべく少ない試行数で発見するという問題を良腕識別問題として定式化し、これらに対する情報論的理論限界の導出とそれを達成するアルゴリズムの構成を行った。

さらに、既存の最適腕識別に対するより理論的・実的に高性能なアルゴリズムの構成も併せて行った。試行ごとに何らかの組合せ制約を満たす候補集合を選ぶ設定は組合せバンディットとよばれるが、この問題に対するほとんどの研究はその各候補の報酬が個別に観測できる場合に対してのみに限られており、多媒体の広告候補の特定などで現れる報酬和のみが観測可能な設定に対するアルゴリズムはほとんど知られていなかった。一方、この設定は各候補が特徴量をもち報酬期待値がその線形和で与えられる線形バンディット問題とよばれる設定の特別なケースとして定式化できる。ここで組合せバンディットでは候補の組合せが指数個あることに起因して従来の線形バンディット問題のアルゴリズムは適用できないため、離散2次計画の近似解法といった各種の組合せ的アルゴリズムを適切に応用することで実時間で実行可能かつ理論限界に近い性能をもつアルゴリズムの構築を行った。さらに、より現実に近い設定への発展として、最終的な評価を行う報酬が線形和とは異なる設定についても併せて検討した。

4. 研究成果

まず、事前に定めたしきい値を上回る期待報酬をもつ候補の発見を目指す良腕識別問題に対して、情報論的理論限界を導出し、それを達成するアルゴリズムを構成した。特に、この問題が純粋探索問題でありながら、それを達成するためのアルゴリズムが累積報酬最大化と非常に近い性質をもつことを明らかにし、報酬最大化と純粋探索問題の新たな関係性を発見した。さらに発展研究として、しきい値を上回る累積報酬をもつ候補を複数個発見する問題も新たに考え、これが単一候補の発見を目指す際に用いたアルゴリズムとは別の枠組みで優れた性能を達成できることも明らかにした。これらの研究成果は機械学習のトップ論文誌 *Machine Learning* に2報が採録された。

さらに、組合せバンディット問題の設定においても、適切な近似アルゴリズムを設計することで組合せ的な候補数がある場合にも多項式時間で実行可能かつ計算量を考慮しない指数時間アルゴリズムに近い性能を達成可能なことを示した。また、密グラフ抽出を雑音のある重み観測クエリから行う問題を線形でない報酬をもつ純粋探索問題として新たに定式化し、こちらについても多項式時間で実行可能なアルゴリズムを構築した。これらの研究成果は機械学習トップの国際会議 *ICML* ならびに論文誌 *Neural Computation* に採録された。

また、上述のようにバンディット問題においては純粋探索問題と累積報酬最大化問題は相補的な関係にあり、いずれかの理論やアルゴリズムの発展がもう一方にも間接的に寄与するが、報酬がノンパラメトリックな分布に従う場合や報酬に関する直接的な情報が得られない場合など、多様な設定に対する理論限界ならびにアルゴリズム構築も行った。さらに、公平性の考慮や事前知識の存在する場合など近年バンディット問題との組合せが考えられつつある周辺分野の研究も併せて行い、トップ国際会議 *NeurIPS* をはじめとして上述の成果のほか計8報が採録された。

5. 主な発表論文等

〔雑誌論文〕 計13件（うち査読付論文 13件 / うち国際共著 2件 / うちオープンアクセス 11件）

1. 著者名 Kuroki Yuko, Xu Liyuan, Miyauchi Atsushi, Honda Junya, Sugiyama Masashi	4. 巻 32
2. 論文標題 Polynomial-Time Algorithms for Multiple-Arm Identification with Full-Bandit Feedback	5. 発行年 2020年
3. 雑誌名 Neural Computation	6. 最初と最後の頁 1733~1773
掲載論文のDOI (デジタルオブジェクト識別子) 10.1162/neco_a_01299	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Yuko Kuroki, Atsushi Miyauchi, Junya Honda, Masashi Sugiyama	4. 巻 119
2. 論文標題 Online Dense Subgraph Discovery via Blurred-Graph Feedback	5. 発行年 2020年
3. 雑誌名 The 37th International Conference on Machine Learning (ICML2020)	6. 最初と最後の頁 5522-5532
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Tabata Koji, Nakamura Atsuyoshi, Honda Junya, Komatsuzaki Tamiki	4. 巻 109
2. 論文標題 A bad arm existence checking problem: How to utilize asymmetric problem structure?	5. 発行年 2019年
3. 雑誌名 Machine Learning	6. 最初と最後の頁 327-372
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s10994-019-05854-7	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Hideaki Kano, Junya Honda, Kentaro Sakamaki, Kentaro Matsuura, Atsuyoshi Nakamura, Masashi Sugiyama	4. 巻 -
2. 論文標題 Good Arm Identification via Bandit Feedback	5. 発行年 2019年
3. 雑誌名 Machine Learning	6. 最初と最後の頁 1-25
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s10994-019-05784-4	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------