

令和 5 年 5 月 19 日現在

機関番号：10101

研究種目：基盤研究(B)（一般）

研究期間：2019～2022

課題番号：19H04161

研究課題名（和文）バンディット問題の方策の実用化のための理論の深化

研究課題名（英文）Theory Deepening for Practical Applications of Bandit Problem Policies

研究代表者

中村 篤祥（Nakamura, Atsuyoshi）

北海道大学・情報科学研究院・教授

研究者番号：50344487

交付決定額（研究期間全体）：（直接経費） 7,500,000円

研究成果の概要（和文）：敵対的バンディットと確率的バンディットの両方において、実用性からヒントを得た問題を定式化し、効率的で高性能なアルゴリズムを提案し、理論的・実験的に性能を評価した。敵対的バンディットにおいては、最適腕が損失を被らないという設定における漸近最適なアルゴリズムを開発し、確率的バンディットにおいては、報酬が閾値以上の腕が閾値個以上あるか否かを、腕を選択して報酬を得ることを繰り返すことによって判定する、分類バンディット問題を定式化し、P-Trackingという効率的で漸近最適なアルゴリズムを開発した。これらの成果に関しては、主要な査読付き国際雑誌・国際会議で発表を行った。

研究成果の学術的意義や社会的意義

バンディット問題のアルゴリズムは、昔から効率的な治験を行うために研究され、現代ではインターネット広告配信、推薦システム、A/Bテストなどに用いられている。基本的に、能動的なサンプリングを行なって効率的に情報を得る方法の研究であり、様々な応用の可能性を秘めている。開発した分類バンディットアルゴリズムは、ラマン分光によるインタラクティブ計測による病理診断の高速化にも用いられることも可能であり、今後様々な分野の応用に発展することが期待される。

研究成果の概要（英文）：In both adversarial and stochastic bandit settings, we formalized problems that are inspired by practical utility, proposed their efficient and high-performance solution algorithms, and evaluated them theoretically and experimentally. In the adversarial bandit setting, we developed an asymptotically optimal algorithm under the condition that at least one arm does not suffer any loss. In the stochastic setting, we formalized the classification bandit problem, in which the player decides whether the number of arms with their expected rewards at least a given threshold is at least a given threshold or not by drawing arms iteratively, and developed the P-tracking algorithm that is efficient and asymptotically optimal. These results are published in major peer-reviewed international journals and conference proceedings.

研究分野：機械学習、データマイニング

キーワード：バンディット問題 オンライン学習

1. 研究開始当初の背景

探索と知識利用のトレードオフの問題を扱うバンディット問題の方策は、オンライン最適化の様々な問題に応用できるポテンシャルをもっているが、十分に実用化が進んでいるとは言い難い。

2. 研究の目的

バンディット問題の方策を実用化という観点から見直し、一般には情報量的または計算量的に困難だとされる問題をヒューリスティクスで解決するのではなく、現実的な制約を課して理論的に精度と効率性を保証する方策を開発することにより、バンディット問題の理論を深化させ、実用化を加速する。

3. 研究の方法

- (1) 敵対的多腕バンディット問題において、最適腕の累積損失が小さいことがわかっているとき、その事実を利用して既存の方策よりリグレットが少ない実用的な方策を開発する。
- (2) 確率的バンディット問題において、与えられた腕集合全体がある条件を満たすか否かを少ない能動的なサンプリングで判定する問題を定式化し、理論保証のある実用的なアルゴリズムを開発する。

4. 研究成果

- (1) ノイズフリー設定の漸近最適な敵対的バンディットアルゴリズムの開発

最適腕の損失が 0 であることが保証されている問題において、敵の最適な報酬決定アルゴリズムおよびプレーヤの最適な腕選択アルゴリズムを開発し、それらの漸近的最適性を証明した [1]。Table 1 [1]は、損失が 0 か 1 かの 2 値の場合、K 本の腕の内、損失 0 の腕が本ある時に、T 回腕を引いた際の最適期待損失を、フルインフォメーション設定(すべての腕の報酬が毎回公開される設定)、忘却型の敵のバンディット設定、適応型の敵のバンディット設定の 3 つの設定で比較してものであり、敵が強くなるに従って最適期待損失が大きくなっていることがわかる。

Table 1

Comparison of $L^*(K, T, \alpha)$, the minimax expected number of mistakes, for variants of the noise-free multi-armed game with K arms and T rounds.

Full-information	
$L^*(K, T, \alpha)$	$\begin{cases} \leq T \left(1 - \left(\frac{\alpha}{K}\right)^{1/T}\right) & (T < K - \alpha) \\ = \sum_{i=\alpha+1}^K \frac{1}{i} = \Theta\left(\log \frac{K}{\alpha}\right) & (T \geq K - \alpha) \end{cases}$
Bandit with an oblivious adversary	
$L^*(K, T, \alpha)$	$\begin{cases} \frac{K-\alpha}{\alpha+1} \left(1 - \prod_{i=0}^{\alpha} \frac{K-T-i}{K-i}\right) & (T < K - \alpha) \\ \frac{K-\alpha}{\alpha+1} & (T \geq K - \alpha). \end{cases}$
Bandit with an adaptive adversary	
$L^*(K, T, \alpha)$	$\begin{cases} T \left(1 - \left(\frac{\alpha}{K}\right)^{1/T}\right) - \frac{\left(\frac{K}{\alpha}\right)^{(T-1)/T} - 1}{K \left(\left(\frac{K}{\alpha}\right)^{1/T} - 1\right)} & (T < K - \alpha) \\ K - \alpha - \frac{(K+\alpha-1)(K-\alpha)^2(2T+(K+\alpha+3)(K-\alpha))}{(2T+(K+\alpha-1)(K-\alpha))^2} & \left(T \geq \frac{(K+\alpha-1)(K-\alpha)}{2}\right) \end{cases} \leq$
$L^*(K, T, \alpha) \leq \min\{T, K - \alpha\} - \sum_{j=1}^{\min\{T, K-\alpha\}} Q(K - j + 1, T - j + 1, \alpha),$	
$\left(L^*(K, T, \alpha) = K - \alpha - \Theta\left(\frac{1}{T}\right)\right)$	
where $Q(K, T, \alpha) = \prod_{h=\alpha}^{K-1} \frac{h}{T+h} = \prod_{h=1}^T \frac{\alpha-1+h}{K-1+h}$. For $\alpha = K - 1$,	
$L^*(K, T, \alpha) = 1 - \frac{K-1}{T+K-1}$.	

- (2) 悪腕存在チェック問題のアルゴリズムの開発

確率的多腕バンディット問題において、期待報酬が閾値以上の腕の有無を、与えられた信頼度で少ないサンプル数で判定するアルゴリズムを開発した [2]。Table 5 [2]は実データ(Real-Time Bidding

Table 5 The average stopping times $\times 10^{-3}$ of the five algorithms and their 99% confidence intervals in the simulations based on real dataset. Note that $(\theta_L, \theta_U) = (\theta_{m'} - 0.01, \theta_{m'} + 0.01)$ for $m' = 0, 1, 5, 10, 19$

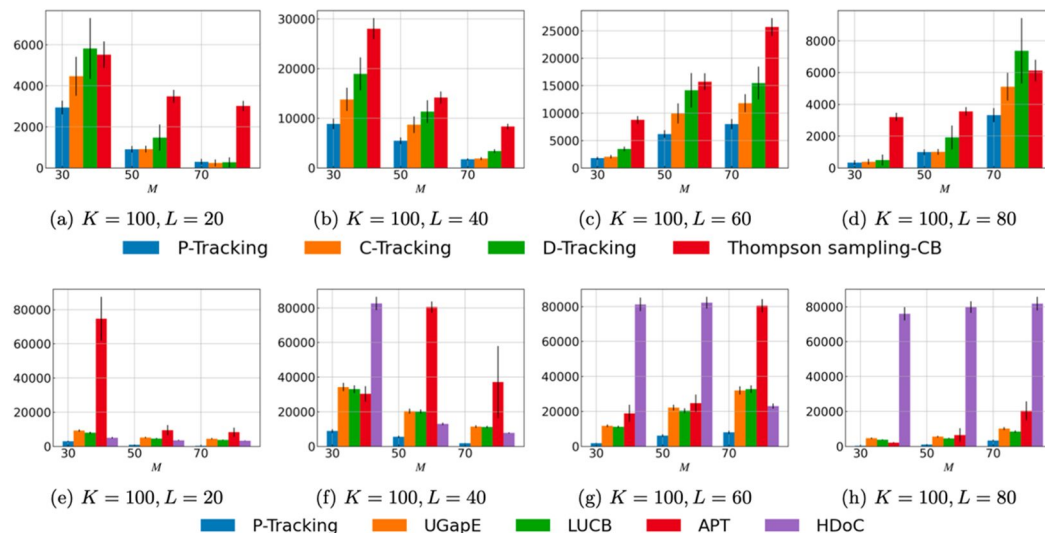
δ (K, T)	Policy	$\theta_0 = 0.06573$ ($m = 0$)	$\theta_1 = 0.05890$ ($m = 1$)	$\theta_5 = 0.04092$ ($m = 4$)	$\theta_{10} = 0.03409$ ($m = 10$)	$\theta_{19} = 0.01144$ ($m = 19$)
$\delta = 0.01$ (1776.1)	APT _P	149.9 ± 2.4	62.2 ± 4.3	24.9 ± 5.9	21.6 ± 3.9	9.5 ± 1.7
	UCB	149.9 ± 2.4	149.6 ± 6.0	52.5 ± 2.5	41.2 ± 2.1	21.5 ± 1.1
	LUCB	149.9 ± 2.4	123.8 ± 8.8	28.1 ± 2.4	19.3 ± 1.9	9.0 ± 0.8
	TS	149.9 ± 2.4	72.9 ± 3.8	21.9 ± 1.4	17.1 ± 1.3	9.9 ± 0.7
	MS	NA [†]	72.4 ± 3.6	21.9 ± 1.4	17.1 ± 1.3	9.9 ± 0.7
$\delta = 0.001$ (2021.0)	APT _P	174.2 ± 2.2	66.0 ± 3.4	29.0 ± 6.1	25.5 ± 4.5	9.9 ± 1.9
	UCB	174.2 ± 2.2	157.4 ± 5.6	56.5 ± 2.7	44.5 ± 2.4	22.8 ± 1.2
	LUCB	174.2 ± 2.2	131.6 ± 7.7	29.6 ± 2.6	21.4 ± 1.8	9.2 ± 0.8
	TS	174.2 ± 2.2	77.1 ± 3.1	22.9 ± 1.4	18.3 ± 1.4	10.6 ± 0.8
	MS	NA [†]	77.0 ± 3.1	22.9 ± 1.4	18.3 ± 1.4	10.6 ± 0.8

[†]For MS, we were not able to measure the stopping time due to the large computation time in the case with $m = 0$ in which the rejection probability in a posterior distribution of a loss-mean set approaches to one

dataset provided by iPinYou (Zhang et al. 2014))のクリック率を用いたシミュレーションの結果である。20 腕で 99%の信頼度で判定する設定において、提案方策 APT_p と 4 つの既存方策の停止するまでのサンプル数を比較したものであり、悪腕の数 m が 1 の時に提案方策が特に有効であるという実験結果が得られている。

(3) 分類バンディット問題のアルゴリズムの開発

確率的多腕バンディット問題において、期待報酬が閾値以上の腕が与えられた閾値個あるか否かを、与えられた信頼度で少ないサンプル数で判定するアルゴリズムを開発した[3][4]



上図[4]は 100 腕($K=100$)によるシミュレーション結果である。閾値以上の腕の数の閾値 L と実際の閾値以上の腕数 M を変えて、既存方策と停止までのサンプル数に関する比較実験をおこなった結果を示している。提案方策である P-Tracking の性能が既存方策と比べて安定して良いことが読み取れる。

<引用文献>

- [1] A. Nakamura, D. Helmbold, M. Warmuth, “Mistake bounds on the noise-free multi-armed bandit game,” Information and Computation 269, 104453, 2019.
- [2] K. Tabata, A. Nakamura, J. Honda, T. Komatsuzaki, “A bad arm existence checking problem: How to utilize asymmetric problem structure?,” Machine Learning 109(2), 327-372, 2020.
- [3] K. Tabata, A. Nakamura, T. Komatsuzaki, “Classification Bandits: Classification Using Expected Rewards as Imperfect Discriminators,” Proceedings of Trends and Applications in Knowledge Discovery and Data Mining (PAKDD Workshops), 57-69, 2021.
- [4] K. Tabata, J. Komiyama, A. Nakamura, T. Komatsuzaki, “Posterior Tracking Algorithm for Classification Bandits,” Proceedings of the 26th International Conference on Artificial Intelligence and Statistics, PMLR 206:10994-11022, 2023.

5. 主な発表論文等

〔雑誌論文〕 計10件（うち査読付論文 10件 / うち国際共著 2件 / うちオープンアクセス 5件）

1. 著者名 Tabata Koji, Nakamura Atsuyoshi, Komatsuzaki Tamiki	4. 巻 12705
2. 論文標題 Classification Bandits: Classification Using Expected Rewards as Imperfect Discriminators	5. 発行年 2021年
3. 雑誌名 Lecture Notes in Computer Science	6. 最初と最後の頁 57 ~ 69
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-75015-2_6	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 -
1. 著者名 Chinone Kosuke, Nakamura Atsuyoshi	4. 巻 13151
2. 論文標題 An Explainable Recommendation Based on Acyclic Paths in an Edge-Colored Graph	5. 発行年 2022年
3. 雑誌名 Lecture Notes in Computer Science	6. 最初と最後の頁 40 ~ 52
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-97546-3_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 -
1. 著者名 Sugie Yuya, Yoshida Yuki, Mertig Normann, Takemoto Takashi, Teramoto Hiroshi, Nakamura Atsuyoshi, Takigawa Ichigaku, Minato Shin-ichi, Yamaoka Masanao, Komatsuzaki Tamiki	4. 巻 25
2. 論文標題 Minor-embedding heuristics for large-scale annealing processors with sparse hardware graphs of up to 102,400 nodes	5. 発行年 2021年
3. 雑誌名 Soft Computing	6. 最初と最後の頁 1731 ~ 1749
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s00500-020-05502-6	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 -
1. 著者名 Mitsuki Maekawa, Atsuyoshi Nakamura, Mineichi Kudo	4. 巻 129
2. 論文標題 Data-Dependent Conversion to a Compact Integer-Weighted Representation of a Weighted Voting Classifier.	5. 発行年 2020年
3. 雑誌名 Proceedings of Machine Learning Research	6. 最初と最後の頁 241, 256
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 Pelissier Aurelien、Nakamura Atsuyoshi、Tabata Koji	4. 巻 2019
2. 論文標題 Feature selection as Monte-Carlo Search in Growing Single Rooted Directed Acyclic Graph by Best Leaf Identification	5. 発行年 2019年
3. 雑誌名 Proceedings of the 2019 SIAM International Conference on Data Mining	6. 最初と最後の頁 450 ~ 458
掲載論文のDOI (デジタルオブジェクト識別子) 10.1137/1.9781611975673.51	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 該当する

1. 著者名 Nakamura Atsuyoshi、Helmbold David P.、Warmuth Manfred K.	4. 巻 269
2. 論文標題 Mistake bounds on the noise-free multi-armed bandit game	5. 発行年 2019年
3. 雑誌名 Information and Computation	6. 最初と最後の頁 104453 ~ 104453
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.ic.2019.104453	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 該当する

1. 著者名 Koji Tabata, Atsuyoshi Nakamura, Junya Honda, Tamiki Komatsuzaki	4. 巻 109(2)
2. 論文標題 A bad arm existence checking problem: How to utilize asymmetric problem structure?	5. 発行年 2020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 オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Nakamura Atsuyoshi、Hayashi Tatsuya	4. 巻 15
2. 論文標題 Efficient alignment-based average delay time estimation in fluctuating delayed propagation	5. 発行年 2022年
3. 雑誌名 Array	6. 最初と最後の頁 100240 ~ 100240
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.array.2022.100240	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 Hayashi Tatsuya, Nakamura Atsuyoshi	4. 巻 12
2. 論文標題 Propagation graph estimation from individuals' time series of observed states	5. 発行年 2022年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 6078-6078
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41598-022-10031-3	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 Koji Tabata, Junpei Komiyama, Atsuyoshi Nakamura, Tamiki Komatsuzaki	4. 巻 PMLR 206
2. 論文標題 Posterior Tracking Algorithm for Classification Bandits	5. 発行年 2023年
3. 雑誌名 Proceedings of The 26th International Conference on Artificial Intelligence and Statistics	6. 最初と最後の頁 10994-11022
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 -

[学会発表] 計8件 (うち招待講演 0件 / うち国際学会 1件)

1. 発表者名 茅根宏介, 中村篤祥
2. 発表標題 ユーザ-アイテム間の関係パスによる複数の説明が可能な推薦
3. 学会等名 第24回情報論的学習理論ワークショップ
4. 発表年 2021年

1. 発表者名 Gong Linghua, 中村篤祥
2. 発表標題 データ分布のクラスタ構造適合による転移学習
3. 学会等名 第14回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2022年

1. 発表者名 中村 篤祥、櫻田 健斗
2. 発表標題 決定木の分岐条件の共有化の効果と応用
3. 学会等名 人工知能学会全国大会
4. 発表年 2020年

1. 発表者名 中村篤祥, ペリシエ オレリアン, 田畑 公次, 小松崎 民樹
2. 発表標題 バンディット問題の方策を用いたモンテカルロ木探索による最適属性集合 探索
3. 学会等名 第19回日本蛋白質科学会年会第71回日本細胞生物学会大会合同年次大会
4. 発表年 2019年

1. 発表者名 白川稔, 中村篤祥, 工藤峰一
2. 発表標題 モンテカルロ木特徴探索に基づく非線形グラフ分類回帰
3. 学会等名 第22回情報論的学習理論ワークショップ
4. 発表年 2019年

1. 発表者名 Ryo Shirakawa, Atsuyoshi Nakamura, Mineichi Kudo
2. 発表標題 Learning a Nonlinear Model of Subgraph Features Using Monte Carlo Tree Search
3. 学会等名 ACML 2019 Workshop on Statistics & Machine Learning Researchers in Japan (国際学会)
4. 発表年 2019年

1. 発表者名 中村篤祥
2. 発表標題 与えられたデータに無矛盾なコンパクトな多出力二分決定グラフの質問学習
3. 学会等名 第123回人工知能基本問題研究会
4. 発表年 2023年

1. 発表者名 田畑 公次, 中村 篤祥, 高見 亮佑, Joshua Arenson, 和田 弥生, Walker Peterson, 合田 圭介, 園下 将大, 小松崎 民樹
2. 発表標題 単調増加制約のあるレベルセット推定
3. 学会等名 第124回人工知能基本問題研究会
4. 発表年 2023年

〔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
研究 分担者	田畑 公次 (Tabata Koji) (20814445)	北海道大学・電子科学研究所・准教授 (10101)	
研究 分担者	工藤 峰一 (Kudo Mineichi) (60205101)	北海道大学・情報科学研究院・教授 (10101)	

	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
研究 協力者	小宮山 純平 (Komiya Junpei)		Assistant Professor of Technology, Operations and Statistics (TOPS) at NYU Stern School of Business

6. 研究組織（つづき）

	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
研究協力者	ヘルムボルド デイヴィッド (Helmbold David)		Professor of UC Santa Cruz
研究協力者	ワームス マンフレッド (Warmuth Manfred)		Professor of UC Santa Cruz

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		
米国	UC SANTA CRUZ	New York University	