

令和元年6月29日現在

機関番号：17102

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K00305

研究課題名(和文) 順序比較に基づくオンライン意思決定基盤技術

研究課題名(英文) Techniques for online decision making based on comparison

研究代表者

畑埜 晃平 (Hatano, Kohei)

九州大学・基幹教育院・准教授

研究者番号：60404026

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：本研究では、順序比較に基づくオンライン予測問題を、離散構造を有する最適化問題というより広い問題と位置づけ、関連する学習理論、最適化、オンライン予測手法の研究を行った。具体的には、(1) ランク制約付きの行列予測問題に対する汎化誤差解析(2) 順序制約つきスケジューリング問題に対するオンライン予測手法とオフライン最適化手法(3) 圧縮データ上の機械学習手法、等の研究を行った。特に、圧縮データ上の機械学習手法の研究はアルゴリズムに関する国際会議WALCOM2018においてベストペーパー賞を受賞した。

研究成果の学術的意義や社会的意義

本研究で得られた知見の1つは、離散構造上の最適化問題を広く捉えることにより、圧縮データ上の機械学習問題を定式化できたことである。この知見は新たな問題群を生み出し、より省スペース・高速な最適化手法の開発に大きく貢献している。実際、本成果の1つはベストペーパー賞という形で研究コミュニティからも興味深い結果と見なされている。また、圧縮データ上の最適化に対して、一見無関係であるオンライン予測技術の有用性を示した点も興味深いといえる。

研究成果の概要(英文)：We regard the online decision problem based on relative comparison as an optimization problem with discrete structures and study not only the problem itself but also related areas such as learning theory, optimization and online prediction. In particular, our project mainly focuses on (1) analyses of generalization ability for prediction of low-rank matrices (2) online/offline optimization methods for scheduling problems with precedence constraints, (3) machine learning algorithms over compressed data, and so on. Among them, our result on machine learning over compressed data won the best paper award at WALCOM2018, a conference on algorithms.

研究分野：機械学習

キーワード：離散構造 データ圧縮 オンライン予測 最適化 スケジューリング 機械学習 ブースティング

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

社会における意思決定には順序に関わるものが多い。例えば、(i) ランキング問題：野球やサッカーの勝敗予測、企業の倒産リスク予測 (ii) オンライン推薦問題：広告の提示、商品の推薦、(iii) ジョブスケジューリング問題が挙げられる。これらの問題の共通点は特定の集合（商品等）上の順序に関する意思決定問題と見なせる点にある。より厳密には、順序比較やスコア情報によって定義される目的関数に対するランキング（順列）の最適化問題と見なせる。特に、近年のデータの大規模化により、ユーザの嗜好やデータの重要度を反映したランキングの提示手法は必須となった。

例えば、機械学習・オンライン推薦の分野では、スコア付き訓練データからスコア関数を学習し、得られたスコアで順序づける事によりランキングを推定するというアプローチが主流である。しかし、スコアに基づく手法は訓練データのスコアのノイズに対して影響を受けやすい。例えば、商品推薦において、評価の甘い/辛いユーザが混在する場合、適切な商品のランキングを予測する事は困難となる（参考：次ページ予備実験）。

一方、近年スコアに代わるフィードバックとして、順序比較（例えば、「商品 A と商品 B どちらが好きか？」）が注目されている。とりわけ、複数の順序比較を統合することにより、ランキングを得る試みがある。比較の結果はスコアのノイズの影響を受けにくいいため、得られるランキングはノイズに対して頑健になる事が期待できる。しかし、比較に基づく順序の最適化には計算困難性のボトルネックが存在する：(a) 解空間が膨大である：実行可能解の個数は集合のサイズに対して指数的に大きい。したがってナイーブに実行可能解を列挙する事は現実的ではない。(b) 比較の統合は一般に困難：特に、与えられた複数の順序比較になるべく矛盾しない順列を求める問題は feedback arc set 問題と呼ばれ、近似解を求めることすら一般に困難である。そこで、既存手法はモデルに関して強い仮定を置く事で困難性を回避しているが、仮定の妥当性に問題がある。また、既存研究において、出力すべきランキング自体の理論性能評価は手付かずである。

2. 研究の目的

本研究のねらいは、計算困難性を回避/克服しつつ、比較順序の統合に基づくオンライン意思決定技術を確立する事である。具体的には以下の目標を実現する：(1) 比較の統合に基づく協調フィルタリング技術とそのオンライン化、(2) 順序制約下でのオンラインジョブスケジューリング手法 (3) 順序に関する確率モデルのオンライン推定技術の開発とこれらの技術の裏付けとなる理論の確立である。

3. 研究の方法

当初のねらいであった比較順序の統合に基づくオンライン意思決定技術の確立には、離散最適化や学習理論といった関連分野からのアプローチが不可欠であると判断し、よりテーマを広げて研究を行った。結果として、新たな知見が得られた他、新たな問題発見につながった。

4. 研究成果

(1) メトリカルタスクシステム (MTS) 問題とは、意思決定の変更に関するコストを考慮したオンライン意思決定問題である。MTS 問題に対するアルゴリズムはいくつか知られており、ほぼ最適な競合比を達成する。しかし、意思決定集合が組合せ集合（グラフ上のパスの集合、順列の集合）の場合、従来手法をナイーブに適用すると各試行において指数時間かかってしまう。本研究では、従来手法の1つである Marking アルゴリズムを改良し、多くの組合せ集合に対して多項式時間で動作し、ほぼ最適な競合比を持つ手法を提案した。

(2) 低ランク行列の予測問題は協調フィルタリングなど多くの応用を持つ。本研究では、非負行列分解を含む低ランク行列のクラスに対する汎化誤差の理論解析を行った。本結果は、従来経験的に良い結果を持つ非負行列のクラスが理論的にもより厳密な汎化誤差の上界を持つことを示した。

(3) LogDet 正則化に基づくオンライン予測手法と 2 値協調フィルタリングへの応用：LogDet 正則化に基づくオンライン予測手法に対して厳密な理論的性能解析を行い、リグレットという指標の下で多くの問題に対していくつかの問題に対してほぼ最適である事を示した。また、解析結果は 2 値協調フィルタリングのオンライン予測問題に 응용可能である。特に、従来手法を LogDet 正則化に基づいた手法で置き換えることにより、最適な誤り回数の理論上界を導出した。また本結果を応用し、2 値行列補完問題に対して、“最適なカーネル” にほぼ匹敵する予測手法を導出した一連の成果は、LogDet 正則化手法に対するより深い理解を与え、頑健な最適化手法の開発に寄与する。

(4) 圧縮データ上でのブースティング：0,1 の系列で表現される事例 (bag-of-words 等) からなる訓練データを ZDD 形式に圧縮した上で、圧縮データ上でブースティングと呼ばれる機械学習手法を実行する、枠組みを提案し、非圧縮データ上と全く同様に動作することを示した。

この結果はビッグデータに対しても省メモリ・省スペースで動作可能であり，非常に有用な結果と言える．本成果は国際会議 WALCOM2018 において best paper award を受賞した．実際，人口データ上では，圧縮の度合いに応じて計算速度の向上も確認された．なお，実験を追加した結果は Theoretical Computer Science に採択された．本研究の興味深い点は，組合せ論的オンライン予測という（一見）全く異なる機械学習分野の知見が圧縮情報処理に活かされたことである．

（５）局所特徴に基づく学習理論とアルゴリズム時系列や画像，テキスト等の分類においては局所的なパターンが分類に大きく寄与する場合がある．局所的パターンに基づく分類手法はデータマイニングの分野でさかんに研究されているが，汎化性能の理論的な裏付けが知られていなかった．本研究では，局所的な特徴を組合わせた分類器の汎化誤差の理論上界を導出するとともに，分類器を生成するための最適化問題を定式化した．提案手法は多くの時系列データに対して良好な性能を示した．提案手法は従来手法と伍する性能を示しつつ，より少ない，調整が易しいパラメータで動作する．

（６）圧縮された決定集合上の離散最適化：単一のプロセッサ，複数のジョブとその実行時間，ジョブ間の順序制約が与えられた下で，全ジョブの待ち時間と実行時間の和を最小化するジョブの順序を見つける問題は古典的なスケジューリング問題の１つである．この問題は NP 困難であり，最悪時には指数時間かかるおそれがある．本研究では (i) 順序制約を満たすジョブの集合を決定ダイアグラム的一种で表現した上で (ii) ダイアグラム上の最短経路問題を解く，という枠組みにより，実時間内で効率的に解く方法を示した．提案手法は，順序制約が密な場合に，既存の整数計画ソルバーに基づく解法を凌駕する性能を示した．本成果は国際会議 SEA2018 に受理された．

（７）その他：順列に対するバンディット問題やスパース正則項付きの最適化問題に対する効率的なスクリーニング手法の研究，逐次射影に基づく組合せ論的オンライン予測手法，スパースな超平面に基づくランキング手法等の研究を行った．

5．主な発表論文等

〔雑誌論文〕(計 14 件)

1. K. Hatano, "Combinatorial Online Prediction," Proceedings of the 2018 International Symposium on Information Theory and Its Applications (ISITA), 40-44, 2018.
2. T. Fujita, K. Hatano and E. Takimoto, "Boosting over non-deterministic ZDDs," Theoretical Computer Science, 2018.
3. K. Moridomi, K. Hatano, and E. Takimoto, "Online linear optimization with the log-determinant regularizer," IEICE Transactions on Information and Systems, vol. E101.D, No.6, pp. 1511-1520, 2018.
4. T. Fujita, K. Hatano, and E. Takimoto, "Online Combinatorial Optimization with Multiple Projections and Its Application to Scheduling Problem," IEICE Transactions on Information and Systems, vol.E101-A no.9, pp. 1334-1343, 2018.
5. D. Suehiro, K. Hatano, and E. Takimoto, "Efficient Reformulation of 1-norm Ranking SVM," IEICE Transactions on Information and Systems, to appear, vol. E101-D, No.3, pp. 719-729, 2018.
6. K. Moridomi, K. Hatano, and E. Takimoto, "Tighter generalization bounds for matrix completion via factorization into constrained matrices," IEICE Transactions on Information and Systems, vol. E101.D no. 8, pp. 1997-2004, 2018.
7. K. Hatano, "Can Machine Learning Techniques Provide Better Learning Support for Elderly People?," Proceedings of the 6th International Conference on Distributed, Ambient and Pervasive Interactions: Technologies and Contexts (DAPI 2018), LNCS, vol 10922, 178-187, 2018.
8. T. Fujita, K. Hatano and E. Takimoto, "Boosting over non-deterministic ZDDs," Theoretical Computer Science, 2018.
9. D. Suehiro, K. Hatano, and E. Takimoto, "Efficient Reformulation of 1-norm Ranking SVM," IEICE Transactions on Information and Systems, to appear, vol. E101-D, No.3, pp. 719-729, 2018.
10. K. Matsumoto, K. Hatano and E. Takimoto, "Decision Diagrams for solving a job

- scheduling problem under precedence constraints,” Proceedings of the 17th International Symposium on Experimental Algorithms (SEA 2018), LIPICS, vol. 103, 2018.
11. T. Fujita, K. Hatano and E. Takimoto, “Boosting over non-deterministic ZDDs,” Proceeding of the 12th International Conference and Workshops on Algorithms and Computation (WALCOM 2018), LNCS vol.10755, 195-206, 2018.
 12. N. Ailon, K. Hatano, E. Takimoto, “Bandit online optimization over the permutahedron,” Theoretical Computer Science, vol.650, 92-108, 2016.
 13. T. Nakazono, K. Moridomi, K. Hatano, E. Takimoto, “A Combinatorial Metrical Task System Problem Under the Uniform Metric,” Proceedings of the 27th International Conference on Algorithmic learning theory (ALT 2016), LNCS 9926, 276-287, 2016.
 14. A. Shibagaki, M. Karasuyama, K. Hatano, and I. Takeuchi, “Simultaneous Safe Screening of Features and Samples in Doubly Sparse Modeling,” Proceedings of The 33rd International Conference on Machine Learning (ICML 2016), JMLR W&CP, vol.48, pp.1577-1586, 2016

〔学会発表〕(計 16 件)

1. 畑埜晃平, “圧縮データ上のブースティング,” AIMAP ワークショップ「非ノイマン型計算、理論と応用」, 2019.
2. Liu Yaxiong, 森富賢一郎, 畑埜晃平, 瀧本英二 “Improved Mistake Bounds for Binary Matrix Completion,” 第 21 回情報論的学習理論ワークショップ (IBIS 2018), 2018.
3. 三宅 郁人, 瀧本 英二, 畑埜 晃平, “MDD の拡張による線形順序拡大集合の省領域表現とスケジューリング問題への応用,” 第 170 回アルゴリズム研究会, 2018.
4. 三宅 郁人, 瀧本 英二, 畑埜 晃平, “MDD の拡張による線形順序拡大集合の省領域表現とスケジューリング問題への応用,” 夏の LA シンポジウム, 2018.
5. 畑埜晃平, “ブースティング:最適化の視点に基づくサーベイ,” AIMAP ワークショップ「非ノイマン型計算、理論と応用」, 2018.
6. D. Suehiro, K. Hatano, E. Takimoto, S. Yamamoto, K. Bannai, A. Takeda, "Learning theory and algorithms for shapelets and other local features," NIPS 2017 Time Series Workshop, 2017.
7. 森富 賢一郎, 畑埜 晃平, 瀧本 英二, “LogDet 正則化を用いた二値行列補完,” IBIS2017, 2017.
8. Ryohei Nagaura, Kohei Hatano, Eiji Takimoto, “Combinatorial bandit prediction with relaxation-based approximation algorithms,” 情報処理学会 第 167 回アルゴリズム研究会, 2017..
9. 森富 賢一郎, 畑埜 晃平, 瀧本 英二, “ノルム制約付き行列分解に基づいた行列補完問題に対する汎化誤差の導出,” 電子情報通信学会 第 28 回 IBISML 研究会, 信学技報, vol. 116, no. 500, IBISML2016-100, 1-8, 2017..
10. 森富 賢一郎, 畑埜 晃平, 瀧本 英二, “FPL を用いた組合せ集合に対するオンライン予測,” 冬の LA シンポジウム, 2017.
11. D. Suehiro, K. Kuwahara, K. Hatano and E. Takimoto, “Time Series Classification Based on Random Shapelets,” NIPS 2016 Time Series Workshop, 2016.
12. 永浦 良平, 畑埜 晃平, 瀧本 英二, “k-集合に対する射影と分解に基づく敵対的バンディット予測,” 電子情報通信学会 第 25 回 IBISML 研究会, 信学技報, vol. 116, no. 121, IBISML2016-7, 225-229, 2016.
13. 桑原健吾, 末廣大貴, 畑埜晃平, 瀧本英二, “Shapelet の統合に基づく時系列分類,” 第 19 回情報論的学習理論ワークショップ (IBIS2016), 2016.
14. 松本晃輔, 畑埜晃平, 瀧本英二, “制約付きスケジューリング問題に対する DD を用いた解法,” 夏の LA シンポジウム, 2016.
15. 柴垣篤志, 烏山昌幸, 畑埜晃平, 竹内一郎, “スパースモデルのための特徴と標本の同時セーフスクリーニング,” 電子情報通信学会 第 25 回 IBISML 研究会, vol.116, no. 1, pp.201-208, 2016.
16. 森富賢一郎, 畑埜晃平, 瀧本英二, “ノルム制約付き行列分解に基づいた行列補完問題に対する汎化誤差の導出,” 第 19 回情報論的学習理論ワークショップ (IBIS2016), 2016.

〔図書〕(計 0 件)

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

〔その他〕

ホームページ等: <https://sites.google.com/view/koheihatano/home>

6. 研究組織

(1) 研究分担者

(2) 研究協力者

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。