

### チーム体制

- 焔埜 晃平 (チームリーダー, 九大)
- Yaxiong Liu (Kyushu, ph.D student)

- 研究協力者
- 末廣 大貴 (客員, 九大)
  - 瀧本 英二 (九大)

### 研究

**動機:** 現在の深層学習は画像・音声など実数値データが対象であり、離散的・構造的な制約を持つデータには適用できない  
 例: 材料科学における化合物, DNA配列, 圧縮データ...

**目標:** 離散的・構造的な制約を持つ最適化・予測問題に対するアルゴリズム

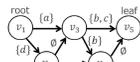
- **ボトルネック:** 候補が**指数的に大** (ナイーブな手法では計算爆発)
- **現状:** アドホック・発見的アプローチが主流

**本アプローチ:** 離散/オフライン/オンライン/バンディット最適化に基づく系統的な方法論

**成果例:** 圧縮データ上での機械学習 (ブースティング) [WALCOM 18]

- **キーアイデア:** 圧縮によるデータの構造化 + 構造制約下で**オンライン最適化**
- **省スペース性:** データを**圧縮したまま学習** (例: a9aデータ: 約70%圧縮)
- **位置づけ:** 従来圧縮型手法で解けなかった最適化問題の解決

**展望:** 離散構造制約下でのブラックボックス最適化技術の確立

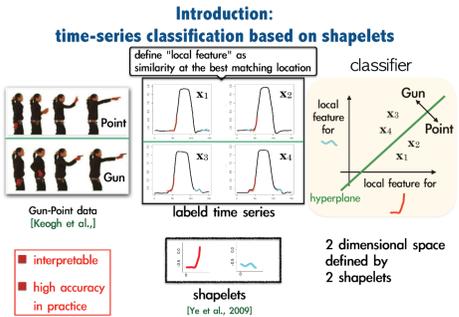


### 研究トピック

- Shapelet に基づくマルチインスタンス学習
- 構造化データ上での機械学習 (ongoing)
- 決定ダイアグラムに基づくスケジューリングの最適化 (離散最適化)
- 組合せ論的オンライン予測

### Shapelet に基づくマルチインスタンス学習

坂内 & 武田チームとの共同研究



#### Shapelet と MIL の関係

- Shapelet に基づく時系列分類  $\approx$  MIL (「部分時系列=インスタンス」を複数持つ「時系列=バッグ」)
- 実はMILにも似たようなアプローチがあった (MILES と呼ばれるアルゴリズムが有名)
- ただし, どちらの分野においても, 訓練データ内の部分時系列 (インスタンス) を Shapelet とし, 類似度に基づいて特徴写像を構成 (理論的にうまくいくか示されていない)

### Shapelet に基づくマルチインスタンス学習(2)

#### 貢献

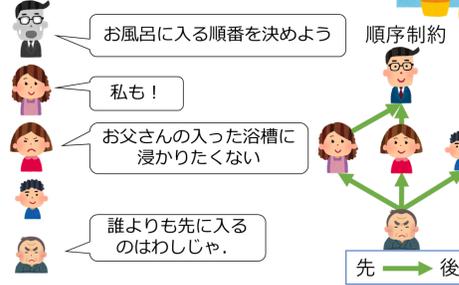
- (無限の) Shapelet に基づく仮説クラスの提案
- 汎化性能の保証
  - 既存の定理とは incomparable
- 無限のshapelet (無限次元) の最適化問題を, 有限次元の問題に帰着 (表現定理の拡張に基づく)
- 表現定理により, 既存手法 (shapeletに訓練バッグ内のインスタンスを用いる手法) の正当性を与えた
- 実験的にも高精度. ドメインに依存せず, 汎用性大

MIデータに対しても  
 既存の shapelet-型 手法に比べ  
 実験的に高精度

dataset	mi-SVM	MI-SVM	MILES	Ours
MUSK1	0.834 ± 0.043	0.831 ± 0.041	0.863 ± 0.070	0.850 ± 0.074
MUSK2	0.736 ± 0.040	0.840 ± 0.037	0.800 ± 0.070	0.858 ± 0.072
elephant	0.802 ± 0.028	0.822 ± 0.028	0.815 ± 0.056	0.826 ± 0.027
fox	0.618 ± 0.035	0.581 ± 0.045	0.636 ± 0.054	0.641 ± 0.069
tiger	0.765 ± 0.039	0.815 ± 0.029	0.787 ± 0.048	0.818 ± 0.051

### 決定ダイアグラムに基づくスケジューリングの最適化

例: お風呂スケジューリング問題



順序制約付きフロータイム最小化  
 単一機械スケジューリング問題

入力: 順序制約グラフ  $G$ , 処理時間ベクトル  $p \in \mathbb{R}^n$   
 出力: 順列  $\pi^* = \operatorname{argmin}_{\pi \in S_G} \pi \cdot p$

NP困難 [E.L Lawler, 78] → 緩和等を用いた多項式時間 2-近似解法 [Hall 9, 97] [Chekuri&Motwani, 99] [Chudak&Hochbaum, 99]

効率的な厳密解法 [提案手法, 18]

ゴール: 全員の待ち時間 & 処理時間の和を最小化する順番を求める

### 決定ダイアグラムに基づくスケジューリングの最適化(2)

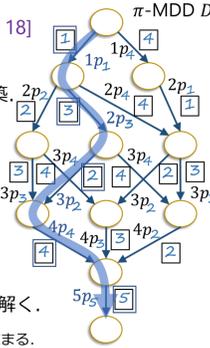
提案手法 [Matumoto+, 18]

1. 順序制約グラフ  $G$  から  $S_G^{-1}$  を表現する  $\pi$ -MDD  $D$  を構築

2. 処理時間ベクトル  $p$  を用いて,  $\pi$ -MDD  $D$  の各辺に重みを割り当てる.

$$1p_1 + 2p_2 + 3p_3 + 4p_4 + 5p_5 \rightarrow \text{フロータイム}$$

3.  $\pi$ -MDD  $D$  上の最短経路問題を解く.  
 → 最適解のインバース  $\pi^{*-1}$  が求まる.



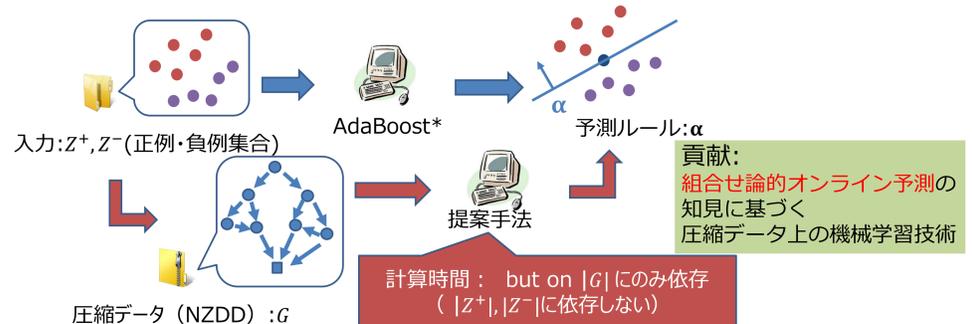
実験結果

- ジョブ数  $n = 30$
- ランダムな制約グラフ

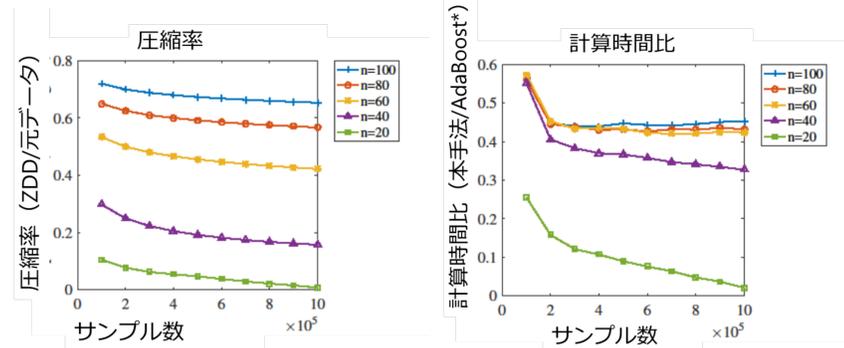


### ZDD上のブースティング

国際会議 WALCOM2018 best paper award



### ZDD上のブースティング(2)



### 発表成果

1. Kosuke Matsumoto, Kohei Hatano and Eiji 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.
2. Takahiro Fujita, Kohei Hatano, Shuji Kijima and Eiji Takimoto, "Online Combinatorial Optimization with Multiple Projections and Its Application to Scheduling Problem," IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences, Vol.E101-A No.9, pp.1334-1343, 2018.
3. Takahiro Fujita, Kohei Hatano and Eiji Takimoto, "Boosting over non-deterministic ZDDs," Theoretical Computer Science, 2018.
4. Ken-ichiro Moridomi, Kohei Hatano, and Eiji Takimoto, "Online Linear Optimization with the Log-Determinant Regularizer," IEICE Transactions on Information and Systems, Vol. E101.D Issue 6, pp. 1511-1520, 2018.
5. Ken-ichiro Moridomi, Kohei Hatano, and Eiji Takimoto, "Tighter Generalization Bounds for Matrix Completion Via Factorization Into Constrained Matrices," IEICE Transactions on Information and Systems, Volume E101.D Issue 8, pp. 1994-2004, 2018.
6. Daiki Suehiro, Kohei Hatano, Eiji Takimoto, Shuji Yamamoto, Kenichi Bannai and Akiko Takeda, "Multiple-Instance Learning by Boosting Infinitely Many Shapelet-based Classifiers," CoRR, abs/1811.08084, 2018.