

アクティビティシミュレータにおける ベイジアンネットワークグラフ推定問題の高速・安定化

望月 陽介, 浦田淳司, 羽藤英二

Yosuke MOCHIZUKI, Junji URATA, Eiji HATO

本研究ではActivity-based Modelへの適用のため、Bayesian Networksのモデル構造学習を高速に、かつ安定して行う手法を提案する。Bayesian Networksは、変数間の説明関係をグラフ構造として学習することができ、近年Activity-based Simulatorとして研究が進んでいる。一方、グラフ構造の学習は膨大なグラフ構造からデータに適合する構造を探索する操作であり、計算時間と解の初期値依存が課題となっている。本研究では、これら課題への解決として挙げられるモデル制約とモデル平均化の枠組みの中で、それぞれObjectの導入と深層生成モデルによるグラフ構造の平均化を提案する。Objectの導入は探索するグラフ構造集合を大きく削減することで、短い計算時間で分散の小さい推定結果を出力する。深層生成モデルは学習された構造習合の中からグラフの部分構造を考慮した平均化を行い、スコアの高い部分構造を採用したグラフを安定して生成する。提案手法は、計算時間、Bayesian Networksモデルとしての評価指標に加え、Activity-based Modelに必要な変数感度を指標として、比較手法と共に検証を行なった。

Keywords: Activity-based Model, Bayesian Networks, Directed Acyclic Graph Learning, Deep Learning

1. はじめに

1.1 背景と目的

交通需要の分析では、エージェントごとの活動を予測するactivity-based model(ABM)の研究が行われている。ABMは活動の派生需要として交通需要を生成することで、旧来の四段階推定法に変わり交通施策の検討に貢献することが期待されている。これらモデルは、各個人の行動選択構造を変数間の計算関係として記述することで、個人の活動をシミュレーションする。変数間の計算関係は、そのモデルの活動生成精度や施策感度の再現性に大きく影響を持っており、これまで多くのモデルがヒューリスティクスや統計の知識に基づき提案されてきた^[1]。

その変数間の計算関係をデータから推定するために、Bayesian Networks(BN)^[2]をABMとして適用する研究が近年交通分野で行われている。BNは変数の説明関係を有効非循環グラフ(DAG)で記述するグラフィカルモデルの一種である。BNを用いたABMは、変数間の計算関係をデータから記述することで、活動生成精度や説明可能性を従来のモデルから向上することに成功している。

BNを活動生成モデルとして利用するためには、一般のモデルでパラメータ推定が必要な様に、グラフ構造の推定が必要である。グラフ構造の推定は、取りうるグラフ構造の中からスコア関数を最大化する構造を見つける離散最適化問題であり、変数の数が大きくなると厳密解を得るアルゴリズムには多大な時間を要する。このため、貪欲法を用いて局所解を探索するアルゴリズムの研究がBN分野で進んでおり^[3]、ABMへの適用の際も貪欲法により探索されたグラフ構造をモデルとして採用している。

しかしながら、こうした貪欲法により得られる局所解は初期値に大きく依存しており、学習毎に異なる構造が出力されることが多い。本研究の目的は、BNグラフの構造制約とモデル平均化の手法を導入することで、BNをABMとして用いる際に必要となる、素早く安定したグラフ構造推定を達成することである。グラフ構造の制約は、BNを利用する上で分野専門家の知識をグラフに反映させる形で設定されることが多く、特定の変数間に説明関係を強制的に設定することで探索するグラフ構造の解空間を狭めている。こうした制約はABMの多くの研究で採用されているが、制約がグラフ推定にもたらす計算時間短縮や推定結果の分散減少といった効果は検証されていない。同様に推定結果の分散を減少させる目的で導入されるモデル平均化は、初期値を変えた多数の異なる推定結果から「平均的な」構造を出力する手法である。しかしグラフ構造の平均化は、エッジの出現率を閾値処理する単純な方法が用いられており、変数の説明力やグラフの尤もらしさを考慮できていない。そこで、本研究ではBNグラフ構造推定問題に用いる構造制約の導入方法とモデル平均化手法を導入し、既存の推定手法と比較した上で、グラフ構造を高速に小さい初期値依存度で推定できることを示す。

1.2 本研究の特徴

本研究の貢献は三点ある。この内二つは、ベイジアンネットワークのグラフ推定問題を高速にかつ小さい初期値依存度で解く手法の導入である。これらの導入により、BNをABMとして利用する際に、より高い精度で活動生成や施策分析、行動選択解釈を行うことが可能になる。

またこれら手法の検証を行うことで、提案手法によるグラフ推定問題の高速・安定化に加え、BNを用いて活動生成を行う際のモデル性能の分散の確認と分析も、本研究が行なったABMとしてのBNの利用への貢献である。

本研究の特徴の一つは、Objectを導入することによるグラフ構造推定の解空間削減である。Objectの概念は、変数を集合として上位概念の元に集約させるものであり、Object-Oriented Bayesian Networks(OOBN)^[4]によりBNに導入された。本研究では、変数の順序配列を探索するグラフ構造推定手法^[5]において、Objectによる変数の集約を導入することで、従来のエッジに付与するグラフ制約よりも大きく、かつ変数間の関係に直接制約を設けることなく解空間の削減を達成した。

二つ目の特徴は、変数ごとの説明能力を考慮したグラフ構造のモデル平均化手法である。グラフ構造では、親変数からの変数に対する説明能力の指標が変数ごとに与えられる。変数ごとの指標をモデル平均化に導入することで、グラフの部分構造を学習済みの結果から最良なものに近づけて出力することに成功した。具体的には、真相学習モデルであるConditional VAE^[6]を用いて、各変数の指標が高い値でグラフ構造を条件付けしてサンプルする。この手法により、有効非循環の制限の中で各変数が最良な部分構造を選ぶ場合のグラフ構造を得ることに成功した。

これら二つの手法を導入した後、本研究ではグラフ構造推定の検証を行なった。BN分野では、真のモデルを仮定した場合のエッジ検出率やBICスコアを指標として、グラフ構造推定手法の検証が行われている。本研究ではこれら指標に加え、交通分野のABMへ適用する際に重要となる変数感度を加えて、提案手法とABMの既往研究で採用されている手法とを検証した。提案手法を導入することで、推定されるグラフ構造は高速に得られ、かつ性能の分散が小さいことを確認した。また、グラフ構造に制約を課さない場合や、既存のモデル平均化手法で推定されたグラフ構造では、性能に大きく分散があることを確認した。

2. 既往研究

本章では、ABMとしてBNを利用した既往研究の中で置かれているグラフ構造推定問題の設定について、BN分野の推定手法研究を参照しながら整理する。

BNのグラフ構造は、各変数 X_i に対して最適な親変数の集合 $\Pi_i \subset \{X_j \text{ for } j \neq i\}$ を決定する最適化問題を解くことで得られる。Ma(2015)^[7]をはじめに近年複数の研究で行われているBNのABMへの導入では、スコア関数、特にBayesian Information Criteria(BIC)(式(1))を最大化する定式化が採用される。

$$BIC(\Pi) =$$

$$\sum_{i=1}^I \left\{ \sum_{\pi \in \Pi_i} \sum_{x \in X_i} N_{x,\pi} \log \hat{\theta}_{x|\pi} - \frac{\log N}{2} (|X_i| - 1)(|\Pi_i|) \right\} \quad (1)$$

BICを最大化するようなグラフ構造の探索方法としては、大きな分類で貪欲法が用いられる。多く採用されるHill-climbing Search^[3]は、 $\Pi_i.append(X_j)$ 又は $\Pi_i.remove(X_j)$ の操作により最もBICが改善される変数の組 $\{i, j\}$ を探索し、操作を行うことでグラフを更新していく。Ordered-based Search^[8]は変数の順序を定義する配列 $order[] = (X_1, \dots, X_i, X_{i+1}, \dots, X_I)$ に対して、配列内の順序を入れ替える処理 $swap(i, i+1)$ (式(2))により最もBICが改善される変数の組 $\{i, i+1\}$ の探索を行い、 $swap(i, i+1)$ を繰り返してグラフを更新する。この手法は、配列 $order[]$ を定めると最適なグラフ構造が容易に得られることに基づいており、探索効率の高い貪欲法である。

$$swap(i, i+1): \quad (2)$$

$$(X_1, \dots, X_i, X_{i+1}, \dots, X_I) \mapsto (X_1, \dots, X_{i+1}, X_i, \dots, X_I)$$

以上に示した貪欲法は局所解を導出する手法であり、初期値依存が大きいこと、また変数が多くなると計算に長い時間を要することが知られている。そこでABMの利用では、グラフ構造推定前に、専門家知識を利用してグラフに制約を設けることで解空間の削減を行なっている。特定の変数間にエッジを固定して導入又は除外する方法である($X_j \in \Pi_i$ 又は $X_j \notin \Pi_i$)。エッジを固定する方法は、局所的な因果関係の方向を直感と一致させることでグラフ推定結果を改善する見込みはあるものの、特定の変数間の説明関係をデータの考慮せずに与えることで、活動生成精度や変数感度を低下する可能性が指摘されており^[8]、またその解空間の削減効率は高くない。本研究では、構造探索効率の高いOrdered-based Searchの変数順序配列 $order[]$ に対して制約を設けることで、変数間の説明関係を恣意的に与えることなく、大幅な解空間の削減を達成する。

Ma et al. (2017, 2018)^{[9],[10]}では、モデル平均化によるグラフ構造推定の安定化の導入提案されている。モデル平均化では、初期値の異なるグラフ構造推定で得られた複数のグラフ構造集合 $(\Pi^1, \Pi^2, \dots, \Pi^K)$ から平均的なグラフ構造 $\Pi^{average}$ を出力する。パラメータを推定する一般的な機械学習モデル学習とは異なり、グラフ構造の推定問題ではモデル平均化として採用される手法は多くは提案されていない。一つは最も学習スコアが最大のグラフ構造を選択する方法であり、もう一つは各エッジの出

現率を集計し、閾値により採択を決定する方法^{[9],[10]}である。式(3)で集計される変数 X_j から X_i へのエッジがグラフ構造集合 $(\Pi^1, \Pi^2, \dots, \Pi^K)$ に含まれる割合 \hat{P}_{ij} を、式(4)のように閾値 β で(0,1)値の \tilde{P}_{ij} に落とし込む。 $\tilde{P}_{ij} = 1$ ならば $X_j \in \Pi_i$ とする。この時、閾値 β は式(5)のように、離散化による情報量の損失が最小となるように選ば⁽³⁾れる。

$$\text{calculate } \hat{P}_{ij} = \frac{1}{n} \sum_{k=1}^K \delta(X_j \in \Pi_i^k) \quad (4)$$

$$\text{if } \hat{P}_{ij} > \beta \Rightarrow X_j \in \Pi_i^{\text{average}}, \tilde{P}_{ij} = 1 \quad (5)$$

$$\beta = \underset{\beta}{\operatorname{argmin}} \sum_i \sum_j |\hat{P}_{ij} - \tilde{P}_{ij}|$$

本手法では、親変数の集合としての説明能力を考慮せずグラフ構造を確率的な処理のみで選択しており、初期値に依存し分散の大きい複数の構造から、高スコアな構造を選択する手法にはなっておらず、学習スコア最大のグラフ構造を選択する場合にエッジ検出率等の指標で劣ることが示されている^[11]。本研究では、深層生成モデルを用いて、各変数の親集合に対して計算されるスコア $\text{BIC}(\Pi_i^k)$ を考慮するモデル平均化を行うことで、グラフの部分構造に対して高スコアな構造を選択する手法を提案する。

3. 提案手法

本章では、グラフ構造推定に導入する提案手法の記述を行う。はじめに3.1節でObjectの導入による解空間削減手法を述べた後、3.2節では深層生成モデルによるモデル平均化手法について述べる。

3.1 解空間を削減したグラフ推定手法

本節で導入する解空間削減手法では、Object-Oriented Bayesian Networks(OOBN)^[4]でBNに導入されたObjectの概念を用いる。グラフ推定の手法としては、Ordered-based Searchを用い、変数順序配列 $\text{order}[]$ に対して制約を設けることで解空間が削減される。

グラフの推定は、Ordered-based Search^[5]を採用する。この際、複数の変数 X_i を1つのObject O_i に格納する。ABMの文脈では、Objectは個人属性や活動パターン、目的地選択といった計算要素を表現し、変数は事前知識に基づいて近い属性のObjectに格納される。この時、同一のobjectに属する変数を変数順序配列の中で連続させることによって、変数の順序配列 $\text{order}[] = (X_1, \dots, X_i, X_{i+1}, \dots, X_l)$ を決定する問題は、Object順序配列 $\text{order}_L[] = (O_1, \dots, O_l, O_{l+1}, \dots, O_L)$ と、各Object O_i 内での変数順序配列 $\text{order}^l[] = (X_1, \dots, X_i, X_{i+1}, \dots, X_{l_i})$ を決定する問題へと変換することができる。Object内での変数順序配列が変わることによる、Object順序配列 $\text{order}_L[]$ の決定問題への影響を無視すると、グラフの推定問題は、 $\text{order}[]$ の探

索操作から、 $\text{order}_L[]$ の探索と $\text{order}^l[]$ for $l = 1, \dots, L$ の探索を順番に行う操作に変換できる。この際の探索が必要な変数順序配列の総数は、均等に L 個のObjectに変数を格納した場合、 $L!$ から $L! \left(\frac{L}{l}\right)^L$ へと削減される。変数の数と取りうるグラフ構造の数(循環含む)の示した図1から、変数の数が増加するとObjectを導入することによる解空間の削減が、既存のエッジを固定する方法に比べて

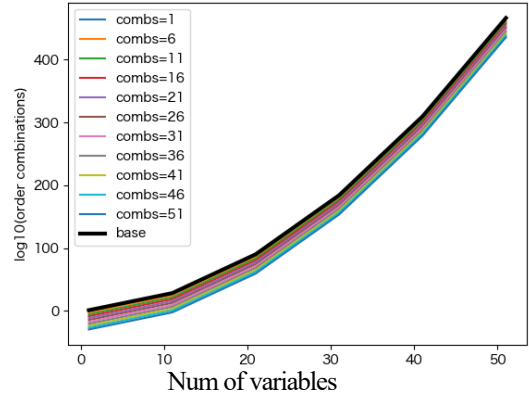


図1-1 エッジの制約による解削減

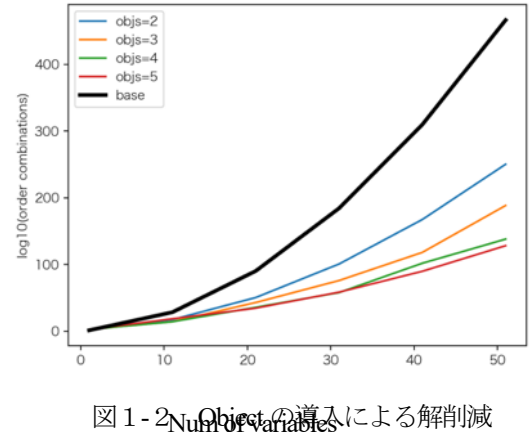


図1-2 Objectの導入による解削減

図1 制約条件ごとの変数の数とグラフ候補数の関係

大きいことが分かる。

3.2 モデル平均化

本節で提案するグラフ構造のモデル平均化手法では深層生成モデルを用いることで、変数ごとに高スコア値を返す部分構造を採用するようにグラフ構造を生成する。グラフ構造を生成する深層生成モデルの研究は、D-VAE(Zhang et al., 2019)^[12]をはじめに多数提案されている。D-VAEはVariational Auto EncoderにDAGグラフ構造を学習させることで低次元の潜在変数へと埋め込み、DAGのサンプリングを可能とした。このDAG生成器を用いて、BNの構造学習を行なっている。

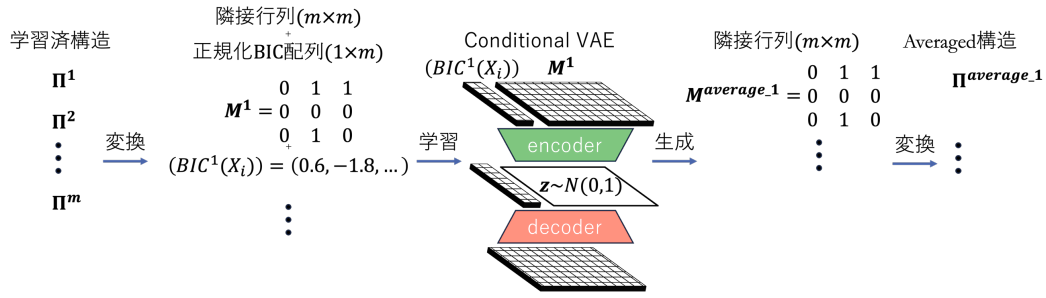


図2 モデル平均化CVAEの構造

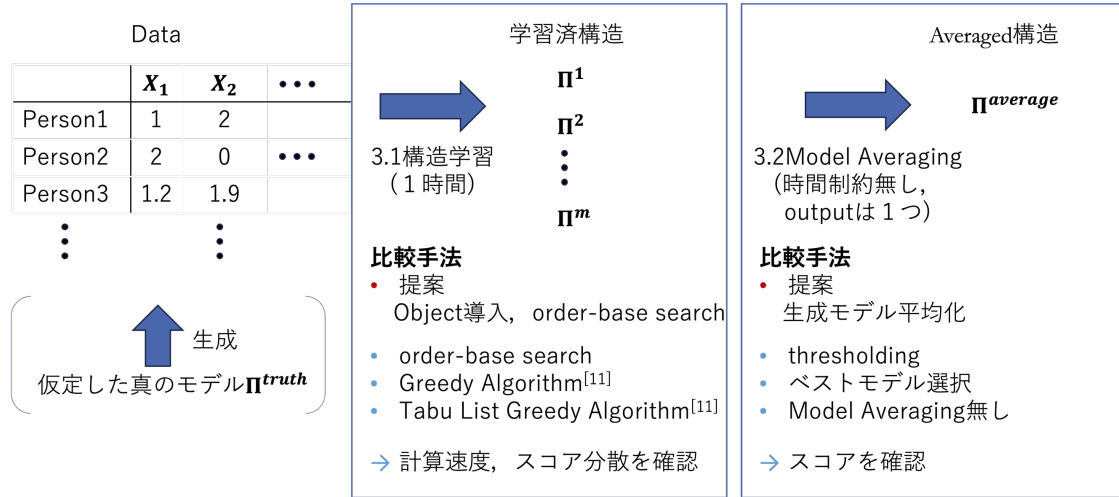


図3 グラフ推定と評価フロー

モデル平均化の問題では、入力は多数の学習で得られたグラフ構造の集合であり、求められる出力は入力集合内のグラフ構造を平均化して得られる高スコアなグラフ構造である。そのため提案手法では、変数ごとの高スコアな親変数集合 Π_i をグラフの部分構造として持つグラフを生成することで高スコアなグラフを生成する、深層生成モデルを用いてモデル平均化を行う。

用いるモデルは、Conditional VAE^[9]である。本研究でのモデル構造を図2に示す。学習時には、学習結果のグラフ構造集合から、グラフ隣接行列 M^k と正規化された各変数のBICスコア配列 $(BIC^k(X_1), \dots, BIC^k(X_l))$ を入力とし、encoderにより潜在変数へと埋め込まれる。その後、潜在変数はBICスコアと共にdecoderによってグラフ隣接行列へと復元される。生成時には、BICスコア配列と分布からサンプルした潜在変数を与えることで、BICスコア配列により条件付けされた分布からグラフ隣接行列 $M^{average}$ が得られる。生成時のBICスコア配列として、全変数に対して高い値を入力することで、各変数に対して高スコアな親変数集合を採用したグラフ構造を生成することが可能である。4章の実験では、全値が2.0(上位98%)のBIC配列で条件付けて構造を生成した。生成されるグラフ構造がDAGであることは保証されていないため、グラフの非循環性を判定し、平均化グラフとする。

4. 双子実験による検証

本章では、提案したグラフ構造推定手法を検証するため、真のモデルを仮定した双子実験を行う。

4.1 実験設定

双子実験では、仮定した真のBNモデルから生成した人工データを観測データとして、提案手法を含む複数のモデル構造推定手法を用いて推定結果のグラフ構造を複数出力する。その上で出力したグラフ構造を複数の指標に基づき評価する。

データ設定

人工データを生成する真のモデルは、既往研究^[13]で設定している、34の離散変数からなるBNモデル構造を採用する。個人の社会属性と活動スケジュールの説明変数から成るグラフであり、グラフ構造は既往研究に則り生成し、変数分布や変数間の条件付き確率は無記載のためランダムに与えた。既往研究のデータと同じオーダーである100,000人分のデータを真のモデルから生成し、90%を学習データとしてグラフ構造推定に用い、残り10%のテストデータを用いて指標計算を行なった。

手法設定と実験フロー

図3に実験の手順を示した。はじめに、真のモデルから人工データを生成した後、実験1を行う。実験1では、提案するObjectを用いた解空間削減を行なったグラフ推定手法と、既存のグラフ推定手法を比較する。得られた人工の学習データを用いて、異なる初期値でのグラフ構造推定を1時間の間繰り返し、推定結果グラフの集合を出力する。その後、出力された集合の中のグラフに対して指標を計算し、その平均と分散を算出する。検証に用いる推定手法は提案手法の他、既往研究として記述した Ordered-based Search(Object 導入無し) と Hill-Climbing Search, Hill-Climbing Searchにtabu listを導入した手法、の3つを比較手法として用いる。

実験2では、モデル平均化手法の比較検証を行う。実験1で出力したグラフ集合を入力として、平均化結果の1つのグラフを出力とする操作を10回行い、平均化結果である10個のグラフに対して指標を計算し、その平均と分散を算出する。提案するCVAEを用いた生成モデル手法の他、通常のVAEを用いた手法(D-VAE)と閾値処理手法(2章参照)、学習スコア最大構造選択手法を比較手法とする。また、モデル平均化を行う前のグラフ推定結果との比較も行う。

評価指標

評価指標としては、1回の推定所用時間、テストBICスコア(式(1)), 真モデルのエッジ検出Recall値(式(6)), 変数感度(式(7), 変数"trip1"を目的変数 Q として使用), の4つを採用する。実験1では全ての指標を用い、計算時間の削減を目的としない実験2では推定所要時間以外の指標を

edgeRecall

$$= \sum_i \sum_{X_j \in \Pi_i^{truth}} \delta(X_j \in \Pi_i^{pred}) / \sum_i \sum_{X_j \in \Pi_i^{truth}} 1 \quad (6)$$

$$Sensitivity(Q, F) = H(Q) - H(Q|F)$$

$$= \sum_q \sum_f P(q, f) \log_2 \frac{P(q, f)}{P(q)P(f)} \quad (7)$$

用いて比較する。

4.2 グラフ推定手法検証結果 – 実験1

実験1の結果を、図4に示す。図4-1から4-3では、真のモデルと各推定手法で得られたモデル群について計算した指標を、手法ごとに平均値と上位95%値、下位5%値を示している。図4-4では、各変数が持つ"trip1"への感度値の手法ごとの平均値を示している。

図4-1から、提案したObjectの導入と Ordered-based Searchにより計算所要時間が大幅に減少していることが分かり、解空間の削減に成功している。また図4-2・図4-3から、提案手法の解空間削減によって、推定で得ら

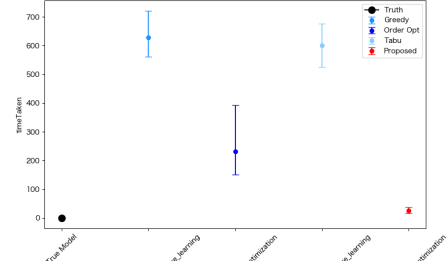


図4-1 計算所用時間

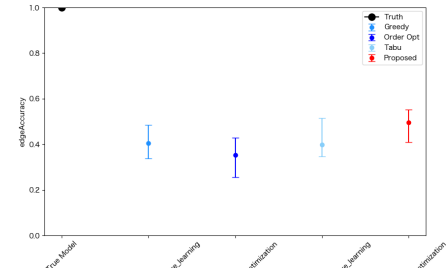


図4-2 エッジ検出 Recall

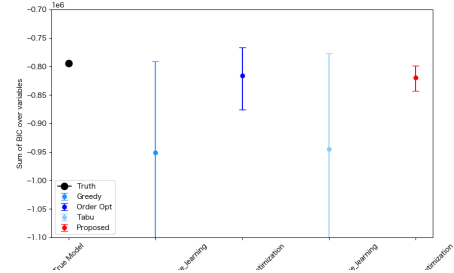


図4-3 テスト BIC スコア

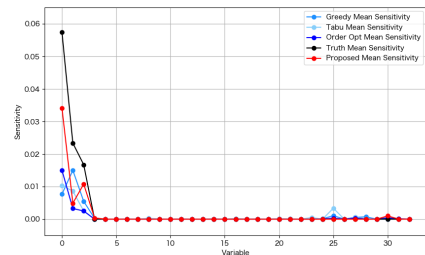


図4-4 変数感度

図4 実験1結果

れるグラフ構造の分散が小さくなっていることが読み取れる。テストBICやエッジ検出率が向上している結果は、解空間削減により貪欲法がより厳密解に近い局所解を見つけやすくなっていると解釈できる。一方で変数感度平均値は、特に感度のある変数において真のモデルから外れており、真のモデルに近い変数感度を実現するために、より精度良くエッジ検出を行う必要があると考えられる。

最適化手法	True Model	Before Averaging	Conditional VAE	VAE (D-VAE)	Threshold	Train Best
Object + Order	1.0	0.50 ± 0.04	0.59 ± 0.01	0.49 ± 0.03	0.40 ± 0.01	0.51 ± 0.04
Order		0.35 ± 0.06	0.39 ± 0.07	0.37 ± 0.07	0.18 ± 0.06	0.41 ± 0.05
Greedy		0.40 ± 0.06	0.31 ± 0.06	0.34 ± 0.10	0.40 ± 0.05	0.41 ± 0.04
Tabu		0.39 ± 0.08	0.31 ± 0.04	0.27 ± 0.07	0.39 ± 0.04	0.37 ± 0.03

表 1-1 エッジ検出 Recall

最適化手法	True Model	Before Averaging	Conditional VAE	VAE (D-VAE)	Threshold	Train Best
Object + Order	-1.20	-1.24 ± 0.004	-1.23 ± 0.0002	-1.24 ± 0.002	-1.25 ± 0.002	-1.24 ± 0.002
Order		-1.24 ± 0.005	-1.24 ± 0.005	-1.24 ± 0.009	-1.28 ± 0.01	-1.23 ± 0.004
Greedy		-1.44 ± 0.06	-1.60 ± 0.54	-7.73 ± 18.22	-1.37 ± 0.09	-1.36 ± 0.04
Tabu		-1.43 ± 0.05	-2.12 ± 2.56	-2.15 ± 1.52	-1.30 ± 0.07	-1.38 ± 0.007

表 1-2 テスト BIC

表 1 実験 2 結果

4.3 モデル平均化手法検証結果 – 実験 2

実験 2 の結果を、表 1・図 5・図 6 に示す。表 1 では、実験 1 で手法ごとに出力したグラフ構造の集合に対して、提案手法と比較手法でモデル平均化を行なった場合のエッジ検出 Recall とテスト BIC 値を示している。エッジ検出 Recall の値は、提案グラフ推定手法で推定したグラフ集合に対して、モデル平均化前から大きく改善しており、他の平均化手法と比べてグラフ構造推定に大きく寄与している。他のグラフ推定手法で出力されたグラフ集合の場合は、1 時間の推定時間において得られたグラフ構造数が少なく、深層生成モデルの学習が進まなかったことが考えられる。またテスト BIC についても、改善が見られる。この改善値は、分散の値を考慮すると大きな改善である。また、モデル平均化前後の一つのグラフ構造について、変数ごとの BIC スコアをプロットしたものが図 5-1 と図 5-2 である。BIC は変数ごとに $[-1, 1]$ に標準化されており、編みかけ部は平均化前グラフの BIC の範囲を示す。提案したモデル平均化手法により、多くの変数に対して BIC が改善されたグラフ構造が生成されていることが読み取れる。また、平均値の改善だけでなく分散の減少を確認できる。このことから BIC での条件付き生成により、より分散の小さい構造出力が行えると解釈できる。

加えて、図 6 では提案のモデル平均化手法を行なった後の変数感度平均を示した。提案グラフ推定手法の出力にモデル平均化をかけた場合の赤線では、図 4-4 と比較して真のモデルに近い変数感度のモデルを構築することができている。エッジ検出精度の向上と、分散が小さいグラフ生成により、平均的な変数感度が真のモデルに近づいたと考えられる。

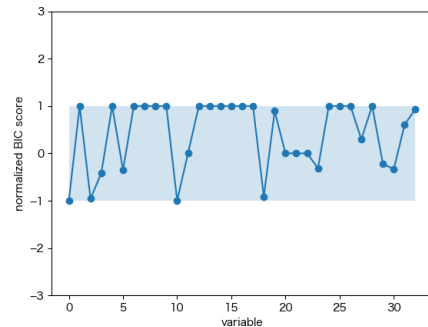


図 5-1 モデル平均化前の変数ごと BIC

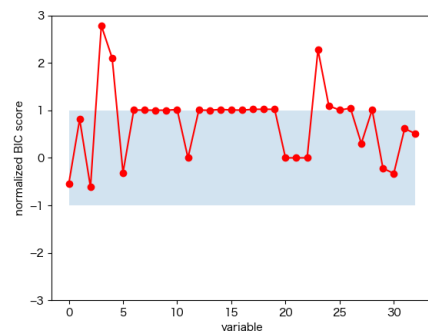


図 5-2 提案モデル平均化後の変数ごと BIC

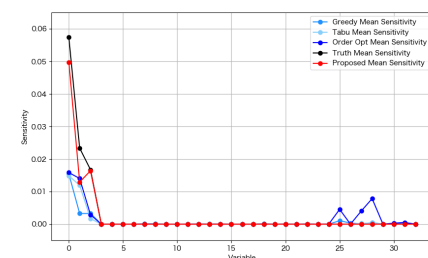


図 6 提案モデル平均化による、変数感度改善

5. 結び

本研究では、Activity-based ModelにおけるBayesian Networksモデルの、グラフ構造推定の速度と安定性を向上させる手法について提案を行なった。Objectの導入により解空間を削減は高速化と初期値依存性の解決に貢献し、スコアで条件付けしたグラフを生成する深層生成モデルを用いたモデル平均化は、グラフ集合から高スコアな部分構造を抽出することで、より真のモデルに近いモデル構造を出力した。これらの手法は、Bayesian Networks分野で用いられるエッジ検出率、BICに加え、Activity-based Modelの設定において重要な変数感度を用いて検証した。

今後の研究では、Bayesian Networksの限界である変数間関係記述の単純さや、個人属性に応じた行動選択構造の変化といったモデルの改善研究に加え、Activity-based Modelとしての活動生成精度や感度の検証といった実証研究が望まれる。

参考

本研究で開発されたObject-Oriented Bayesian Networksとそのグラフ構造推定、Conditional VAEを用いたモデル平均化手法については、以下のurlからプログラムが確認できる。

[\[https://github.com/yosuke-civil-tokyo/master\]](https://github.com/yosuke-civil-tokyo/master)

参考文献

[1] Rasouli, Soora & Timmermans, Harry. (2013). Activity-based models of travel demand: Promises, progress and prospects. *International Journal of Urban Sciences*. 18. 31-60.
[2] Cooper, G. F., & Herskovits, E. (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9(1-3), 309-347.
[3] Bayesian Networks. *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence, UAI 2005*.

[4] Daphne Koller and Avi Pfeffer. (1997). Object-oriented Bayesian networks. In *Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence (UAI'97)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 302-313.
[5] Teyssier, Marc & Koller, Daphne. (2012). Ordering-Based Search: A Simple and Effective Algorithm for Learning
[6] K. Sohn, H. Lee, and X. Yan. *Advances in neural information processing systems*, page 3483-3491. (2015).
[7] Ma, T. (2015). Bayesian Networks for Multimodal Mode Choice Behavior Modelling: A Case Study for the Cross Border Workers of Luxembourg. *Transportation Research Procedia*, 10.
[8] Aurore Sallard, & Miloš Balać. (2023). Travel demand generation using Bayesian Networks: an application to Switzerland. *Procedia Computer Science*. Volume 220. Pages 267-274. ISSN 1877-0509.
[9] Ma, Tai-yu & Chow, Joseph & Xu, Susan. (2016). Causal structure learning for travel mode choice using structural restrictions and model averaging algorithm.
[10] Ma, T.-Y., & Klein, S. (2018). Bayesian networks for constrained location choice modeling using structural restrictions and model averaging. *European Journal of Transport and Infrastructure Research*, 18(1).
[11] Broom, Bradley & Do, Kim-Anh & Subramanian, Devika. (2012). Model averaging strategies for structure learning in Bayesian networks with limited data. *BMC bioinformatics*. 13 Suppl 13. S10.
[12] Zhang, M., Jiang, S., Cui, Z., Garnett, R., & Chen, Y. (2019). D-VAE: A Variational Autoencoder for Directed Acyclic Graphs. *Neural Information Processing Systems*.
[13] Alta de Waal, & Johan W. Joubert. (2022). Explainable Bayesian networks applied to transport vulnerability. *Expert Systems with Applications*. Volume 209. 118348. ISSN 0957-4174.