

アリコロニー最適化によるベイジアンネットワークの構造最適化

鈴木 真也† 水野 一徳† 戸谷 太亮‡
 † 拓殖大学工学部情報工学科 ‡ 拓殖大学大学院工学研究科

1 はじめに

ベイジアンネットワーク (Bayesian Network(BN)) とは事象間の因果関係を推定し、ネットワーク構造として表現したものである。BNにおいて適切に因果関係を推定するには、最適なネットワーク構造を見出す必要があるが、BNのノード数が増加すると候補となる構造が指数関数的に増加してしまう。そのため、実用的な実行時間で構造最適化を行えるように、様々なアルゴリズムが提案されている。

近年、蟻コロニー最適化 (Ant Colony Optimization(ACO)) を用いたアルゴリズムが提案されている [1]。ACOは蟻の採餌行動をモデル化したメタヒューリスティクスであり、BNの構造最適化において有効性が示されている。一方、山登り法を用いた先行研究 [2] では、構造最適化を行う前にクラスタリングを行うことで、最適化の精度が上がる事が示されている。そこで、本研究では構造最適化のさらなる精度向上を図るために、クラスタリングを用いたACOを提案する。提案手法ではノードをクラスタリングで分類し、最適解に含まれるようなネットワーク構造を生成しやすくすることで、構造最適化の精度向上を図るものである。ベンチマークデータセットを用いた実験により、提案手法の有効性を確認する。

2 研究分野の概要

2.1 ベイジアンネットワーク (BN)

BNとは、事象をノードとみなし、ノード間の因果関係を有向辺(エッジ)を用いて有向非循環グラフとして表したグラフィカルモデルである。BNを用いることで、事象の起こりやすさをネットワーク構造と条件付き確率表で表すことができる。BNは産業や医学、意思決定支援などの様々な分野で応用されている。

2.2 群平均法

群平均法とは、クラスタ間の距離の計算に各クラスタ内に属する要素の距離の平均を用いるクラスタリングの方法である。群平均法を用いると各クラスタ内の要素数の偏りが小さくなり、距離が近い要素同士が同じクラスタに属しやすくなる事が期待できる。

2.3 蟻コロニー最適化 (ACO)

ACOは、蟻の採餌行動におけるフェロモンコミュニケーションをモデル化したメタヒューリスティクスである。ACOでは解候補の生成とフェロモンの蓄積を繰り返すことで最適化を目指す。具体的には生成された割り当て(解候補)をもとに、解候補の情報をフェロモン蓄積量として記録する。新しい解候補を生成する際に、フェロモン蓄積量をもとに確率的に解候補生成を行

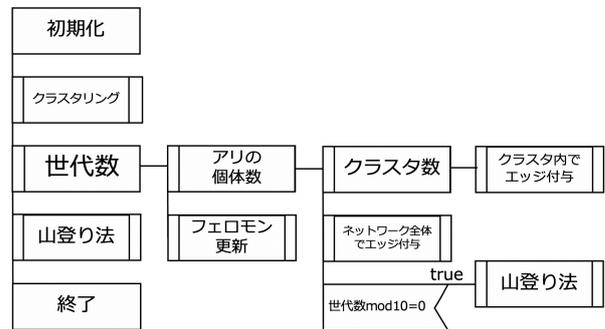


図 1: クラスタリングを用いた ACO-B のアルゴリズム

う。ACOを用いたBNの構造最適化アルゴリズムとしてACO-B[1]がある。ACO-Bはスコア関数を参考にして、ネットワーク構造の最適化を行う。

3 提案手法

3.1 概要

本研究では、ACOを用いたBNの構造最適化アルゴリズムACO-Bを改良する。従来のACO-Bはネットワーク全体でエッジの付与を行うため、最適解になり得ないエッジが付与される可能性がある。そこで本研究では、クラスタリングを行った後、クラスタ内、クラスタ間の2段階でエッジの付与を行う手法を提案する。アルゴリズムの流れを図1に示す。

3.2 クラスタリング

提案手法では相互情報量をもとに群平均法を用いてクラスタリングを行う。相互情報量は同時分布 P_{XY} と周辺分布の積 $P_X P_Y$ をもとに、ノード X とノード Y の2つのノードの独立性を示す尺度である。ノード X とノード Y の距離は相互情報量を用いて以下で表す。

$$Dis(X; Y) = 1 - \sum_{i=1}^{|\mathbf{x}|} \sum_{j=1}^{|\mathbf{y}|} P_{X,Y}(x_i, y_j) \log \frac{P_{X,Y}(x_i, y_j)}{P_X(x_i) P_Y(y_j)} \quad (1)$$

\mathbf{x}, \mathbf{y} はそれぞれノード X , ノード Y がとる状態の集合を表す。 $P_{XY}(x_i, y_j)$ は、ノード X が状態 i をとり、ノード Y が状態 j をとるときの同時分布である。 $P_X(x_i)$ と $P_Y(y_j)$ はそれぞれ、ノード X が状態 i , ノード Y が状態 j をとるときの周辺分布を表す。

3.3 アルゴリズム

3.3.1 K2スコア

提案手法ではK2スコアで解候補を評価する。K2スコア f_{K2} を式(2)に示す。

$$f_{K2}(G : D) = \sum_{i=1}^n f_{K2}(X_i, Pa(X_i)) \quad (2)$$

Learning Bayesian Networks based on Ant Colony Optimization
 †Masaya SUZUKI †Kazunori MIZUNO †Takaaki TOYA
 †Department of Computer Science, Takushoku University
 ‡Graduated School of Engineering, Takushoku University

$$f_{K2}(X_i, Pa(X_i)) = \sum_{j=1}^{q_i} \left(\log \left(\frac{(r_i - 1)!}{(N_{ij} + r_i - 1)} \right) \right) + \sum_{k=1}^{r_i} \log(N_{ijk!}) \quad (3)$$

ここで G は生成されたネットワーク構造, D はデータセット, n はノードの個数, $Pa(X_i)$ はノード X_i の親ノードの集合, r_i は X_i のとる状態の数, q_i は $Pa(X_i)$ に属するノードの状態の組み合わせの数, N_{ijk} はデータセット内のサンプルの中で X_i が状態 k の時, $Pa(X_i)$ に属するノードの状態の組み合わせが j であるサンプルの数, N_{ij} は $Pa(X_i)$ に属するノードの状態の組み合わせが j であるサンプルの数 ($\sum_{k=1}^{r_i} N_{ijk}$) を表す。

3.3.2 解候補生成

提案手法では, 解候補をネットワーク構造の隣接行列として表す. クラスタリングによりノードをいくつかのクラスタに分類し, 各クラスタ内でエッジの付与を行った後, 各クラスタ内のエッジ情報を保持したままクラスタ間でエッジの付与を行う. 付与されるエッジの選択は, 以下の式によって行なわれる.

$$\eta_{uv} = f(X_u, Pa(X_u) \cup X_v) - f(X_u, Pa(X_u)) \quad (4)$$

$$u, v = \begin{cases} \operatorname{argmax}_{u, v \in F_G} [\tau_{uv}]^\alpha [\eta_{uv}]^\beta & \text{if } q \leq q_0 \\ U, V & \text{otherwise} \end{cases} \quad (5)$$

U, V は以下の式より確率的にランダムに選ばれた2つのノードを表す.

$$P(U, V) = \begin{cases} \frac{[\tau_{uv}]^\alpha [\eta_{uv}]^\beta}{\sum_{u, v \in F_G} [\tau_{uv}]^\alpha [\eta_{uv}]^\beta} & \text{if } u, v \in F_G \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

U, V は以下の式より確率的にランダムに選ばれた2つのノードを表す. τ_{uv} は子ノード X_u と親ノード X_v のエッジのフェロモン蓄積量, α はフェロモン蓄積量の重み係数, η_{uv} は子ノード X_u と親ノード X_v のエッジのヒューリスティック情報, β はヒューリスティック情報の重み係数, F_G は選択できるエッジの候補 ($\eta_{uv} > 0$ かつ循環グラフにならないエッジ) の集合を表す. 山登り法では各個体が生成した解候補からハミング距離が1のネットワーク構造の中で最も K2 スコアが高いものを選択する.

3.3.3 フェロモン更新

生成された解候補に付与されているエッジに対してフェロモンの蒸発と蓄積を行なう.

4 実験

4.1 実験条件

提案する手法の有効性を確認するために実験を試みる. ここでは従来手法と提案手法の2つの手法について, 次の2点で比較する.

- 全ての試行中で生成されたネットワーク構造の K2 スコアの中で最も高い K2 スコア (最高 K2 スコア)
- 各試行で生成されたネットワーク構造の最も高い K2 スコアの平均 (平均 K2 スコア)

本実験で用いるパラメータは表1, データセットは表2に示す. 表1における G_0 は空グラフ, n はノードの個数を表す. 試行回数は各クラスタ数ごとに20回とした.

表 1: パラメータ設定

項目	値
世代数	100
個体数	10
初期フェロモン量	$1/n \times f_{K2}(G_0 : D) $
フェロモン τ の重み係数 α	1
ヒューリスティック η の重み係数 β	2
フェロモン蒸発率	0.4
エッジ選択で用いる確率 q_0	0.8

表 2: 用いるデータセット

データセット名	ノード数	サンプル数
INSURANCE	27	10000

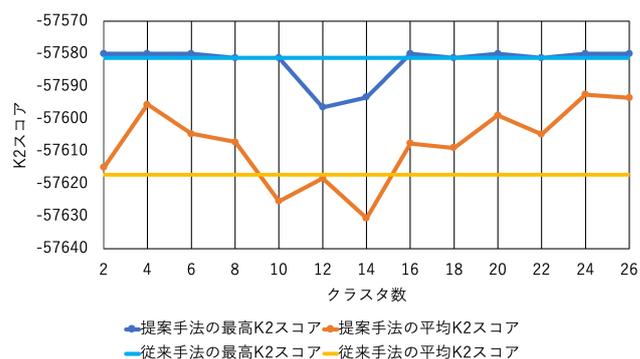


図 2: INSURANCE 20 回の結果

4.2 実験結果

図2はINSURANCEの各クラスタ数ごとの最高 K2 スコアと平均 K2 スコアを表している.

図2よりクラスタリングによる最高 K2 スコアの向上は見られなかったが, 平均 K2 スコアの向上が見られた. この結果より, ACO-B に対してクラスタリングを用いる改良を行うことで, 従来手法よりも比較的 K2 スコアの高いネットワーク構造が見つかる確率が向上した可能性がある.

5 まとめ

本研究では, ACO-B を改良した手法を提案した. 提案手法は, 相互情報量をもとに群平均法でクラスタリングしたノードの情報を用いて最適化を行う. 実験により, 改良による ACO-B の性能の向上を確認した. 今後は他の様々な規模のデータセットを用いて性能を確認する必要がある.

参考文献

- [1] M.C. Luis, M.F. Juan, A.G. Jose, M.P. Jose, Ant colony optimization for learning Bayesian networks, International Journal of Approximate Reasoning 31 (2002)291-311.
- [2] Zhang, Y., J., Liu, Y, Bayesian Network Structure Learning: The Two-Step Clustering-Based Algorithm, Proceedings of the AAAI Conference on Artificial Intelligence 32(1) (2018).