

エッジの合流を用いた有向グラフ可視化手法

理学専攻 情報科学コース 1640654 十枝 菜穂子 (指導教員: 伊藤 貴之)

1 はじめに

SNS 上の人間関係や遺伝子のネットワーク, 論文の参照関係など, さまざまなデータが有向グラフの形で表現される. これらのデータは情報量が膨大になることも多いため, 全体の構造を短時間で理解するための可視化手法が有用である.

本論文では, 非階層型の有向グラフを可視化する一手法を提案する. 具体的には, ノードがクラスタリングされた有向グラフに対して, なす角が小さい束どうしを合流させるエッジ描画およびそれを効果的にするノード配置を行う.

2 関連研究

グラフ可視化におけるエッジのバンドリング (束化) とは, 近隣するノードを接続する複数のエッジが束として可視化されるように変形または統合する処理を指す. Holten は, 同一クラスタペアに属する 2 ノードを連結するエッジ群をバンドリングすることでクラスタ間の関係性を強調する手法 [1] を提案している. また, 従来のバンドリング手法の多くがエッジをその中央付近で合流させているのに対して, エッジの両端 (特定のノードの近く) でバンドリングすることで曖昧性を排除しながらもエッジの交差を減らす手法 [2] も提案されている. Bach らの手法 [3] では, 共通して接続しているエッジのみを合流させることでエッジの曖昧性を排除している. しかし, [1] は無向グラフを対象としており, [2] や [3] はノードがクラスタリングされていない. 本研究では, ノードがクラスタリングされた有向グラフに対して, クラスタ間の関係を, 誤読の可能性を排除しながら表現するための可視化手法を提案する.

3 提案手法

本章では提案手法について, データ構造および処理手順について説明する.

3.1 データ構造

提案手法では入力データとして有向グラフを想定する. また本研究ではノードの接続性だけでなく任意の数値属性を加味してグラフ配置を実現できるようにするために, 各ノードに対して多次元ベクタとなる特徴量を付与できるものとする.

3.2 ノードのクラスタリング

提案手法では任意の 2 ノード間の距離を算出し, 最長距離法に基づいた階層型クラスタリングを行うことで 2 階層のクラスタを生成する. 2 ノード間の距離は

- 特徴量ベクタに基づくノード間距離
- 隣接ノードの共通性に基づくノード間距離

の 2 種類の距離の一次結合により算出する [4].

3.3 ノード配置

提案手法では, クラスタごとにノードを配置するため, 以下のアルゴリズムにより各ノードの配置場所を

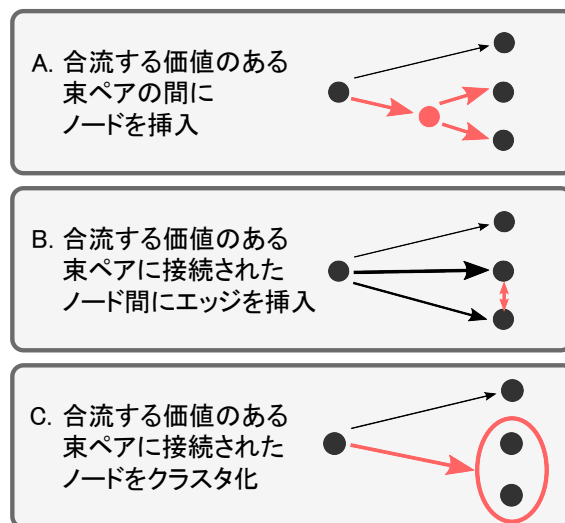


図 1: クラスタ配置のための処理

決定する.

1. クラスタをノードに置き換えたグラフ G を生成する.
2. グラフ G に対し図 1 のいずれかの処理を実行しグラフ G' を生成する.
3. G' にグラフ配置アルゴリズムを適用する.
4. 2 で挿入したノードやエッジを削除する.
5. 各クラスタのノード数からクラスタ半径を計算し, エッジ長が両端のクラスタの半径の合計値にできるだけ近づくようにスムージング処理を適用する.
6. 各クラスタ内にそのクラスタを構成するノードを配置する.

ここで, 2 における「合流する価値のある束ペア」は, 以下の基準により算出した評価値 S から求める.

- エッジの本数
- クラスタの大きさ (ノードの数)

3.4 エッジ描画

提案手法では図 2 の 3 つの条件を満たすエッジ描画を目標とする. 提案手法ではまず条件 C を満たすために合流するエッジ束を選択する. 続いて, 各エッジの条件に合わせて曲線を描画する. 以下, 合流するエッジ束の選択手法, 曲線を描画手法のそれぞれについて述べる.

3.4.1 合流するエッジ束の選択手法

提案手法では, 合流する価値のある束ペアを求める処理 (3.3 参照) で算出した評価値 S をもとに合流するエッジ束を選択する. または, 以下のアルゴリズムによって合流するエッジ束を選択する.

1. 1 つのクラスタを取り出す.
2. 時計回りに 1 つずつエッジ束を参照する.

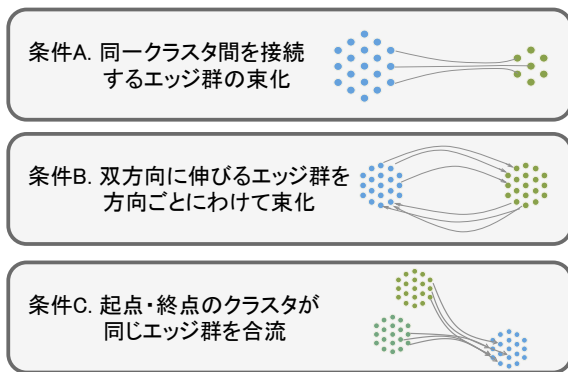


図 2: エッジ描画の条件

3. 左隣のエッジ束が同じ向きの場合、合流するかどうかの評価値を算出し、ユーザの入力した閾値以上ならば合流すると判定する。
4. 全ての束およびクラスタに対して同じ処理を適用する。

ここで評価値は以下の 2 つの評価方法から算出する。

エッジ束の角度

2 本のエッジ束がなす角度が閾値よりも大きい場合は、これらが合流されないように評価値を 0 とする。

合流できる部分の長さ

2 本のエッジ束の角度の二等分線に両クラスタから垂線を伸ばし、その交点と基準となるクラスタとの距離のうち短い方の長さを評価値とする。

3.4.2 エッジの曲線描画手法

本手法では全てのエッジを以下の 3 パターンに分類し、それぞれの曲線を Bezier 曲線または Spline 曲線によって描画する。現在の我々の実装では青から赤のグラデーションで曲線を描画することによってエッジの向きを表現している。

- 他の束と合流するエッジ
- 他の束と合流しないが双方向に接続されたエッジ
- 上記 2 つのどちらでもないエッジ

4 適用事例

適用事例として、1990 から 94 年、2000 から 2010 年までに ACM SIGGRAPH で発表された論文に対し

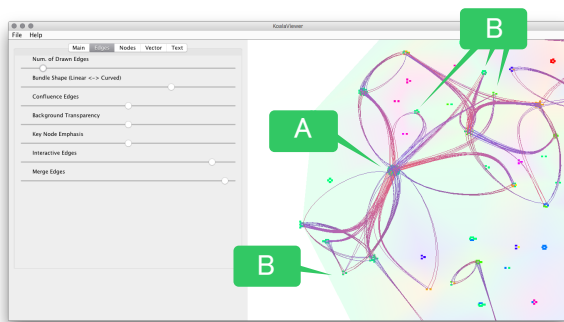


図 3: 可視化結果

て、論文をノード、参照関係をエッジとしたグラフを可視化した事例 [5] を示す。このグラフにおいてノード (論文) は 1072 個、エッジ (参照関係) は 5498 本である。また各ノードに付与される特徴量として、10 の研究分野についてそれぞれの程度当てはまるかを各論文の内容によって分類した 10 次元の特徴量を用いる。また、この 10 次元の特徴量の中で最も値の高い次元ごとに各ノードの色付けを行っている。

図 3 に可視化結果の一例を示す。クラスタ A の論文が様々なクラスタから参照されており、逆にクラスタ A の論文が参照している論文が小さなクラスタ群 B に集中していることがわかる。このことから、この分野について読む場合はクラスタ B の論文、クラスタ A の論文、その他のクラスタの論文と順番に読むことで基礎的な研究から順番に学習ができると考えられる。

5 まとめと今後の課題

本論文では有向グラフ可視化の一手法として、複数のエッジ束を合流させる有向グラフの可視化手法を提案した。本手法は、誤読の可能性を排除しながらクラスタ間の関係性を概観することを可能とし、重要なクラスタの発見を容易にしたと考えられる。今後の課題として、以下の点に取り組みたい。

- エッジ描画における曲線の制御点の最適な位置の自動算出
- ユーザが指定したクラスタやノードに重みを置いたノード配置

参考文献

- [1] D. Holten Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data, *IEEE Trans. on Visualization and Computer Graphics*, 12(5), 741–748, 2006.
- [2] S.-J. Luo, C.-L. Liu, B.-Y. Chen, K.-L. Ma, Ambiguity-Free Edge-Bundling for Interactive Graph Visualization, *IEEE Trans. on Visualization and Computer Graphics*, 18(5), 810–821, 2011.
- [3] B. Bach, N. H. Riche, C. Hurter, K. Marriott, T. Dwyer, Towards Unambiguous Edge Bundling: Investigating Confluent Drawings for Network Visualization, *IEEE Trans. on Visualization and Computer Graphics*, 2016.
- [4] T. Itoh, K. Klein, Key-node-Separated Graph Clustering and Layout for Human Relationship Graph Visualization, *IEEE Computer Graphics and Applications*, 35(6), 30–40, 2015.
- [5] R. Nakazawa, T. Itoh, T. Saito, A Visualization of Research Papers Based on the Topics and Citation Network, *18th International Conference on Information Visualisation (IV2015)*, 283–289, 2015.