

有向グラフ可視化のためのバンドリングとノード配置

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。

本研究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、エッジのバンドリング (束化) に加えて、なす角が小さい束の合流、ノードクラスタ間で双方向に接続されたエッジ群のバンドリングを提案するものである。また、それらを効果的にするノード配置手法も提案する。

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。しかし、[1] は無向グラフを対象としており、[2] 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のクラスタ間の関係を表現するためのバンドリング手法を提案する。なお、本研究では無向グラフ可視化手法 [3] を拡張する形で提案手法を実装する。

3 提案手法

本章では提案手法の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 [3]。

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決定する。現在の我々の実装では、階層をもたないク

ラスタ群に対して以下のアルゴリズム [3] によってノードの位置を算出する。

1. クラスタをノードに置き換えたグラフを生成し、これに対しグラフ配置アルゴリズムを適用する。
2. 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。
3. 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。

階層型のクラスタ群の場合は、まずクラスタを内部に持つ大きいクラスタをノードに置き換えたグラフ配置アルゴリズムを適用する。その後各クラスタの内部のクラスタをノードとしたグラフを生成し上記のアルゴリズムを適用することでノードの位置を算出する。

3.4 エッジ描画

提案手法では図 1 の 3 つの条件を満たすエッジ描画を目標とする。このうち条件 A は従来のバンドリング手法と同じ目標であり、条件 B は有向グラフ描画のために条件 A を拡張したと考えられる。条件 C を満たすために提案手法ではまず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、曲線の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。
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。

ここで評価値の算出は以下の 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

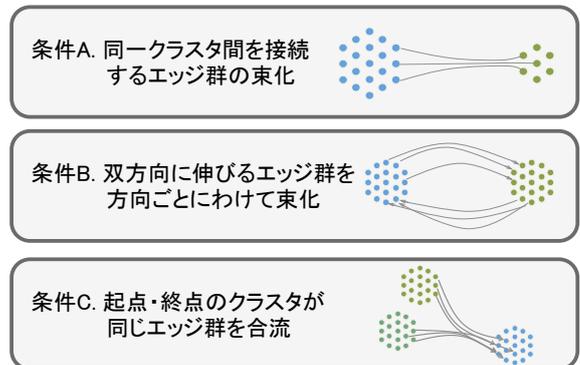


図 1: エッジ描画の条件

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 曲線によって描画する。

他の束と合流するエッジ

2本のエッジの垂直二等分線上に Bezier 曲線の制御点を設定する。

他の束と合流はしないが双方向に伸びているエッジ

2つのクラスタ間の垂直二等分線上に Bezier 曲線の制御点を設定する。

上記2つのどちらでもないエッジ

2ノードを接続する線分と2つのクラスタの中心点を接続する線分から制御点を求める。

4 適用事例

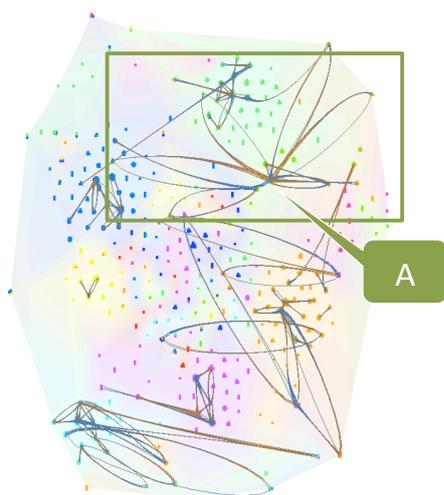


図 2: 提案手法による可視化例

本章では適用事例として、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 [4] 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ど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。

図 2 が階層型のクラスタ群を用いた可視化結果の一例である。図中の各ノードの色は特徴量から求めたその論文の分野を示しており、エッジの色は青が参照する側、橙が参照される側として描画されている。同色のクラスタ間どうして接続されているエッジが多いが、図 3 に示した黄緑色のノード(テキストチャに関する分野)で表されるクラスタ (A) は青のノードで示される

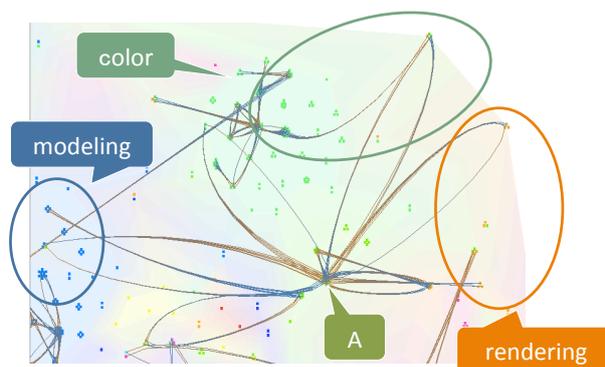


図 3: クラスタ (A) 付近の拡大図

モデリング分野、青緑のノードで示される色彩に関する分野、橙のノードで表されるレンダリング分野とさまざまな分野と接続されている。また、向きごとにエッジが合流していることでクラスタ (A) は多数の論文を参照しているがそれ以上に多数の論文から参照されていることを観察することができ、多分野で応用される基礎的な研究の論文であると考察することができる。

5 まとめと今後の課題

本研究では有向グラフ可視化の一手法として、エッジ束の合流処理などを含むバンドリング手法を提案した。エッジ束の合流という概念の適用により、クラスタ間の関係性の概観がさらに容易になったと考えられる。今後の課題として、起点と終点の両方で合流するエッジ束の曲線のなめらかな描画方法、合流させるエッジ束の選択手法、重要なエッジ束を合流させることに適したクラスタの配置方法などを再検討したい。

6 謝辞

本研究を進めるにあたり、東京農工大学 齋藤隆文教授には論文参照関係のデータを提供していただきました。ここに感謝いたします。

参考文献

- [1] D. Holten. Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 12, No. 5, pp. 741–748, 2006.
- [2] S. J. Luo, C. L. Liu, B. Y. Chen, and K. L. Ma. Ambiguity-free edge-bundling for interactive graph visualization. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 5, pp. 810–821, 2012.
- [3] T. Itoh and K. Klein. Key-node-separated graph clustering and layouts for human relationship graph visualization. *IEEE Computer Graphics and Applications*, Vol. 35, No. 6, pp. 30–40, 2015.
- [4] R. Nakazawa, T. Itoh, and T. Saito. A visualization of research papers based on the topics and citation network. In *19th International Conference on Information Visualization (iV)*, pp. 283–289. IEEE, 2015.