

ネットワーク可視化のための拡大描画に適したエッジバンドリング手法

秋山 桂一[†] 藤田 秀之^{††} 大森 匡^{††} 新谷 隆彦^{††}

[†] 電気通信大学 情報理工学域 〒182-8585 東京都調布市調布ヶ丘 1-5-1

^{††} 電気通信大学 大学院情報理工学研究科 〒182-8585 東京都調布市調布ヶ丘 1-5-1

E-mail: †akiyama@hol.is.uec.ac.jp, ††{fujita,omori,shintani}@is.uec.ac.jp

あらまし Web ページのリンク関係や SNS のフォロー関係など、さまざまなデータがネットワーク構造をなしている。このようなデータを活用する為に、可視化することは重要であるが、昨今ではデータが大規模化しており、大量のエッジが重なり合って描画されるエッジクラッタなどの問題が発生する。これを解決するために似た傾向をもつエッジを力学的に束ねるエッジバンドリングという技術が盛んに研究されているが、拡大描画に適したバンドリング手法は研究されていない。本研究では無向ネットワークを対象に、拡大描画に適したバンドリング手法を提案する。可視化結果の定量的評価により提案手法の有効性を確認した。

キーワード 情報可視化, エッジバンドリング, ネットワーク構造

1 背景

Web ページのリンク, 被リンク関係, SNS のフォロー, フォロワー関係, 場所と交通量の関係のようなネットワーク構造を持つデータは日常の至る所に存在しており, これらを可視化することは重要で, さまざまな研究がなされてきた。昨今ではデータが大規模になってきており, 単純にノードとエッジを描画するだけでは, 大量のエッジが重なり合って描画され, ネットワーク構造が把握できなくなるエッジクラッタという問題が発生してしまう。この問題を解決するため, エッジを束ねて描画するエッジバンドリングという技術が盛んに研究されているが, データの大規模化に伴って需要が増してきている拡大縮小といった操作を考慮した研究は見当たらない。また, ネットワークの可視化のみならず, 可視化結果の全体像を拡大したときに, 併せて周辺の情報提示するオフスクリーン可視化という技術は重要であり, 研究が行われているが, エッジバンドリングを考慮したオフスクリーン可視化の研究は見当たらない。そこで本研究は, 拡大描画に適したネットワーク可視化のためのエッジバンドリングとそれに適したオフスクリーン可視化を検討する。

2 関連研究

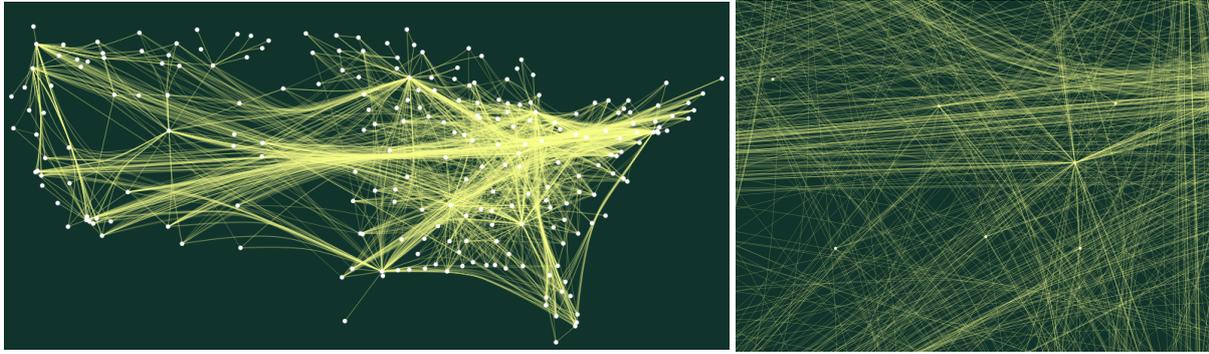
ネットワーク構造を可視化する試みはグラフ理論などの立場から, 個々のノードやエッジを正確に読み取るために, エッジが重ならない, 重なるときは直交する, などのいくつかの条件を満たすノード配置を最適化する問題として研究されてきた [1]。しかし, 昨今ではデータが大規模化しており, 大量のエッジが重なり合って描画され複雑な印象を与えるエッジクラッタという問題を始めとする, 新たな問題が発生しており, 従来のように, 個々を正確に把握するのではなく, ノードの分布やエッジを群として, おおまかに把握することが求められるようになっており, グラフマイニング, ネットワーク科学, 情報可視化な

どの新しい分野からさまざまなアプローチが取られている。

情報可視化の分野では, エッジクラッタを解決する試みとして, 似た傾向をもつエッジを束ねて描画するエッジバンドリングという技術が存在する。エッジバンドリングの厳密な定義は存在しないが, Lhuillier ら [2] によれば, エッジバンドリングは, エッジが重なり合って描画されることを解消するため, またエッジを容易に追従できるようにするため, エッジを両端点との接続関係を保持したまま, 空間的に薄く, また適度に低い曲率で配置することである。これは, 空間ネットワークなどに挙げられるノードの位置が予め決められているネットワーク構造においても適用可能である。

多くのエッジバンドリングの研究ではネットワーク構造データの他にバンドリングを実現するため, 階層的な情報 [3] や構造メッシュ [4] などの補助的な情報を必要とする一方, Holten らが提案した, 力学モデルに基づくエッジバンドリングの端緒である Force-Directed Edge Bundling (以下, FDEB) [5] では, 補助的な情報に頼らずデータのみでバンドリングを実現することができるという特徴があった。この手法はシンプルで拡張性が高く, 今日でもこの手法を元に, エッジの方向や重みを考慮したバンドリング手法 [6], エッジの曖昧さのないエッジバンドリング手法 [7] エッジの属性を考慮したバンドリング手法 [8], など, さまざまな新たな性質を持たせたバンドリング手法が提案されている。しかし, いずれのエッジバンドリングの手段においても情報可視化における重要な要素の一つである, 描画の拡大表示を考慮した研究は見当たらない。

今日のデータは大規模化しており, ビューに表示しきれないことが考えられる。そういった際に, 拡大縮小をはじめとするインタラクティブ操作は効果的である。また, データ可視化において拡大描画した際に周辺の情報提示するオフスクリーン可視化は重要であり, 盛んに研究されている。Dominik ら [10] によると, オフスクリーン可視化の要件は, 拡大した描画対象領域のコンテキストを維持することである。描画対象領域外のオブジェクトの位置や接続関係, 距離を保存するために, 視覚



(a) 既存手法の適用結果

(b) 一部拡大

図1 拡大した可視化結果

的な代替物を描画対象領域内に表示する方法が一般的である。オフスクリーン可視化は、描画対象領域外のオブジェクトを円で提示する技術[11]や、描画対象領域端に描画対象領域外のノードを提示する技術[12]に挙げられるように、地図やグラフの描画などにおいて幅広く研究され使われているが、エッジバンドリングを考慮した研究は見当たらない。

3 Force-Directed Edge Bundling

3.1 手法

FDEBはエッジクラッタやエッジアンビギュイティの解消を目的とし、似た傾向を持つエッジをバンドリングして描画する技術である。FDEBに入力するデータを頂点とエッジの集合 $G = (V, E)$ とし、頂点を $v = (x, y) \in V$ とする。 (x, y) は平面上の座標値である。エッジを $e = \{u, v\} \in E$ とする。 $u, v \in V$ である。エッジを束ねて描画するために、エッジは直線ではなく、曲線で描画する必要がある。エッジを曲線で描画するための制御点のことを分割点と呼び、エッジ P の分割点列を (p_0, p_1, \dots, p_n) と定義する。隣接する分割点同士に互いにばね引力を働かせることで、エッジは伸び縮みが可能になる。エッジ P, Q 間のエッジ親和性を $C(P, Q)$ とする。エッジ親和性 $C(P, Q)$ はエッジの配置や形状の観点から計算される値である。このエッジ親和性 $C(P, Q)$ がユーザの与える閾値 t を越えている場合、エッジ P, Q の分割点ペア p_n, q_n 間に静電気力を作用させバンドリングを実現する。以上を踏まえてFDEBでは以下の流れでバンドリングを実現する。

- step 1 エッジ集合 E に含まれるすべてのエッジペア $\{P, Q\}$ においてエッジ親和性 $C(P, Q)$ を計算する。
- step 2 エッジ集合 E に含まれるすべてのエッジで分割点列をつくる。
- step 3 エッジ親和性 $C(P, Q)$ に基づき力学計算を行う。
- step 4 step2～step3をユーザの与える回数繰り返す。

最終的に E に含まれるエッジの分割点列を出力し、この分割点列を制御点としてバンドリングされた曲線のエッジを描画する。

4 課題と方針

本研究が取り組む課題は以下の二点である。

(1) 入力データに対して既存手法を適用したのち、ある一部分を拡大すると手法の効果が低減して発生するエッジクラッタを解消すること。

(2) 描画対象領域内外の関係を示すため、描画対象領域内外を跨るエッジで画面外のノードの方向や距離を大まかに示すこと。

(1)は既存手法が入力データ全体を可視化するのに適しており、拡大縮小を考慮していないことに起因する。図1(a)はアメリカ合衆国の航空路線網のデータに対し、既存手法であるFDEBを適用した結果である。図1(b)は(a)のインディアナポリス周辺を拡大した結果である。(b)では多くのエッジは、そのエッジの中心付近で大きく曲がるため、画面外でバンドリングされ画面内では直線で描かれてしまったため、エッジクラッタが再び出現している。そこで本研究は既存手法の分割点の配置を変更することで、描画対象領域内でバンドリングの効果が現れるようにし、解決を試みる。

(2)は、拡大した描画対象領域は全体像の一部で、描画対象領域が全体に対してどのような位置付けか把握するために描画対象領域周辺の情報を同時に提示することは、可視化の分野においてオフスクリーン可視化と呼ばれ、一般的な課題である。本研究は画面外のノードの距離や方向をエッジ束でおおまかに示すことで描画対象領域外の情報を提示する、その具体的な方法として、既存手法のエッジ親和性を変更し、配置と形状だけでなく、描画対象領域に対しての画面外のノードのエッジ親和性を考慮することで解決を試みる。

5 提案手法

提案手法では前項で述べた問題の解決を目標としており、拡大描画した際の描画対象領域におけるエッジクラッタの解消を目的としエッジバンドリングを行い、描画対象領域内外のノードの関係を提示するため、描画対象領域外のノードまでの距離、描画対象領域外のノードへの接続関係をエッジで表現する。これらを実現するために既存手法のFDEBの分割点の配置とエッジ親和性に着目し、拡大描画に適したそれぞれの計算方法を検討する。

5.1 対象とするエッジ

拡大描画したとき、エッジは以下の三種類に分類できる。

- 1 描画対象領域外に両方の端点をもつエッジ
- 2 描画対象領域内に両方の端点をもつエッジ
- 3 描画対象領域内外のそれぞれに端点をもつエッジ

1のエッジは拡大の興味の対象から外れており、描画対象領域内や内外の関係の可視化の観点の関係で不要と考え、本研究ではバンドリングの対象から除外し、非表示にしている。2のエッジは拡大の興味の対象という観点では重要であるが、描画対象領域にそのすべてが含まれているため、既存研究で対応可能であると考え、本研究ではバンドリングの対象から除外し、非表示にしている。したがって、本研究は3のエッジを対象とし、バンドリングを行い描画する。

5.2 エッジ親和性

既存研究ではエッジ親和性 $C(P, Q)$ はエッジの長さ、エッジの角度、エッジの位置、エッジのビジビリティにより計算されていた。提案手法ではオフスクリーン可視化のため、描画対象領域に対してどのようなエッジであるかを重視する本手法では、二つのエッジの関係からエッジ親和性を計算するよりも、描画対象領域に着目し、描画対象領域からの相対的なエッジの関係に基づいて計算するほうが、エッジのもつノードとノードを結ぶという特徴に加え、描画対象領域内と描画対象領域外との関係を考慮することができると考えたため、以下の二つのエッジ親和性を提案する。

5.2.1 描画対象領域の中心から見た角度によるエッジ親和性

ネットワークを可視化し、拡大した場合、描画対象領域に含まれているノードを片方の端点とするエッジに対する興味は、そのエッジが画面外のどの方向へ向かうのか、ということが考えられる。従来のエッジ同士の角度によるエッジ親和性 C_a では、拡大描画において、描画対象領域外のあるひとつのノード、または描画対象領域外の互いが近くに配置されているノードを端点とするエッジが複数描画されていたとき、描画対象領域内のノードの配置によっては、エッジペアのなす角はおおきくなり、相互作用しないこともあった。単なるエッジの類似度という観点では確かに相互作用しない方がいい場合も考えられるが、拡大描画においては、画面外のどの方向に向かうのか、ということにより興味がある場合も考えられる。したがって、以下のように計算し、描画中心から描画対象領域外のノードのなす角をエッジ親和性 (図 2) に用いることを提案する。

$$C_{ac}(P, Q) = \cos(\angle P_0 C Q_0) \quad (1)$$

ここで、 P_0 、 Q_0 、 C はそれぞれ、エッジ P の描画対象領域外の端点、エッジ Q の描画対象領域外の端点、描画対象領域の中心点である。

5.2.2 描画対象領域の中心からの距離によるエッジ親和性
前項と同様に、拡大描画での興味として、エッジが画面外のどれ程遠距離のノードを端点にもつか、という点が挙げられる。

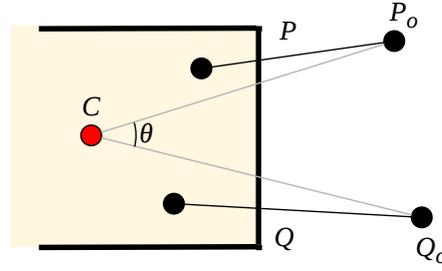


図 2 描画対象領域の中心から見た角度によるエッジ親和性

従来のエッジ同士の長さによるエッジ親和性 C_s では、特に描画対象領域近傍のノードにおいて、描画対象領域内のノードの配置によっては、相互作用しないことも考えられた。したがって、以下のように計算し、描画中心からの描画対象領域外のノードまでの距離をエッジ親和性 (図 3) に用いることを提案する。

$$C_{dc}(P, Q) = \frac{2}{\frac{d_{ave}}{\min(d_p, d_q)} + \frac{\max(d_p, d_q)}{d_{ave}}} \quad (2)$$

ここで、 $d_p = \|C - P_0\|$ 、 $d_q = \|C - Q_0\|$ である。

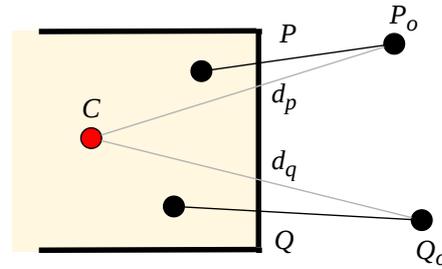


図 3 描画対象領域の中心からの距離によるエッジ親和性

5.3 分割点の配置

既存手法の FDEB では拡大した際、描画対象領域内でエッジが直線のように描かれ、バンドリングの効果が現れていない、という問題があった。この原因として、FDEB では力学計算を適用する分割点がエッジに等間隔に配置されており、等間隔に配置された場合、エッジの中央付近でバンドリングの効果が強く現れ、拡大の倍率によっては、エッジの根本付近しか描画できないためバンドリングの効果が低減してしまう、ということが考えられる。したがって、この問題を解決するために本手法ではエッジの分割点の配置を変更し、バンドリングの効果を描画対象領域内でも認められるようにし、改善を試みる。具体的には、多くの分割点が描画対象領域内に収まるように配置し、バンドリングの効果が強く現れる部分をエッジの中央付近から描画対象領域に寄せる。本手法では以下の 4 つの分割点の配置方法を提案する。

5.3.1 等間隔配置

一つ目の分割点の配置方法はエッジに等間隔に配置する方法である (図 4(1))。既存手法の FDEB もエッジに等間隔に分割点を配置しているが、本手法と異なる点がある。それは分割

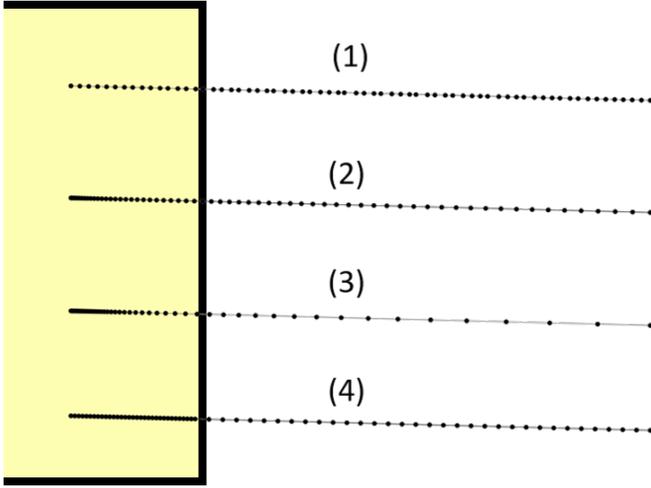


図4 分割点の配置方法 (1) 等間隔配置 (2) 二次関数的配置 (3) 指数的配置 (4) 分割点の割合による配置)

点列の方向を揃えていないことである。本研究が対象としているネットワークは有向ネットワークではない。しかし、FDEBでは計算量を抑えるため、相互作用するエッジの同じ添字番号を持つ分割点同士でのみ静電気力を計算するという特徴があった。すなわち、無向ネットワークでもエッジの分割点には順番が存在し、したがって方向が存在していた。この場合、分割点列の中央付近はあまり影響を受けないが、ノード付近に存在する分割点は、分割点列の方向によって強く引き付けあうときとそうでないときがあった。本研究が提案するこの等間隔配置では、拡大した際の描画対象領域の中心に近いノードから分割点を等間隔に配置してゆき、分割点列の方向を描画対象領域内から描画対象領域外という方向に限定することで、静電気力を適用した際の振舞いを一定にする。なお、以降の3つの分割点の配置方法も同様に分割点列の方向を揃えるものとする。

5.3.2 二次関数的配置

二つ目の分割点の配置方法はエッジに二次関数的に配置する方法である(図4(2))。多くの分割点を描画対象領域内に収めるため、描画対象領域内に存在するノードから分割点を配置してゆき、ノードからの距離を二次関数的に増加させる。エッジの長さを l 、分割点の個数を n としたとき、分割点 p_i と分割点 p_{i+1} の間隔 I_{i-i+1} は以下の様に計算される。

$$I_{i-i+1} = \{(i+1)^2 - i^2\} \times \frac{l}{(n+1)^2} \quad (3)$$

5.3.3 指数的配置

三つ目の分割点の配置方法はエッジに指数的に配置する方法である(図4(3))。二次関数的配置と同様に、多くの分割点を描画対象領域内に収めるため、描画対象領域内に存在するノードから分割点を配置してゆき、ノードからの距離を指数的に増加させる。エッジの長さを l 、分割点の個数を n 、指数の底を b としたとき、分割点 p_i と分割点 p_{i+1} の間隔 I_{i-i+1} は以下の様に計算される。

$$I_{i-i+1} = (b^{i+1} - b^i) \times \frac{l}{b^{n+1}} \quad (4)$$

5.3.4 分割点の割合による配置

四つ目の分割点の配置方法は、あらかじめ描画対象領域に含まれる分割点数の割合を決めておき、その値に従い、等間隔に配置する方法である(図4(4))。既存手法や二次関数的配置、指数的配置とは異なり、明示的に描画対象領域に含まれる分割点の個数を制御する。描画対象領域内のノードを出発したエッジが描画対象領域の端と交わる点までの長さを l_i 、そこから描画対象領域外のノードまでの長さを l_o 、分割点の個数を n 、割合を r としたとき、分割点 p_i と分割点 p_{i+1} の間隔 I_{i-i+1} は以下の様に計算される。

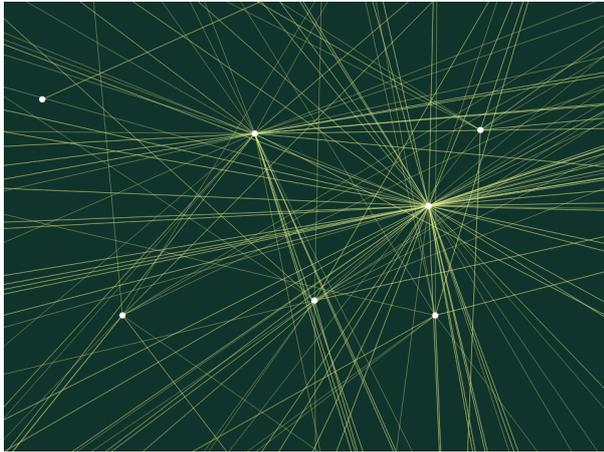
$$I_{i-i+1} = \frac{l_i}{(n+1) \times r} \quad (p_i \text{が描画対象領域に含まれる場合}) \quad (5)$$

$$I_{i-i+1} = \frac{l_o}{(n+1) \times (1-r)} \quad (p_i \text{が描画対象領域に含まれない場合}) \quad (6)$$

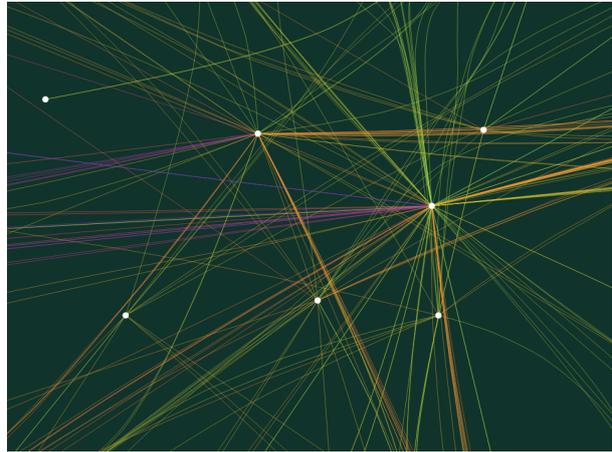
6 適用結果

アメリカ合衆国の航空路線データを可視化し、インディアナポリス周辺を拡大した結果を図5に示す。(0)は、バンドリングを実行していないときの様子、(1)は既存手法を適用したときの結果である。なお、描画対象領域内に両方の端点を持つエッジ、描画対象領域外に両方の端点を持つエッジは非表示にした。(2)は、描画対象領域内外にそれぞれ端点をもつエッジに対して、提案手法である分割点の配置を二次関数的配置にし、エッジ親和性を既存手法の配置によるエッジ親和性と提案手法の2つのエッジ親和性にした結果、また、(3)分割点の配置を分割点の割合による配置にし、同様のエッジ親和性にしたときの結果である。

(1) 既存手法では、描画対象領域内ではバンドリングの効果が現れておらず、描画対象領域内から描画対象領域外へ全方位にエッジが描画され、どの方向への航路が多いのか分かり難く、描画対象領域と描画対象領域外の関係を把握し難い。提案手法の(2)では、描画対象領域内でバンドリングの効果が現れたことにより、描画対象領域外縁部のエッジクラッタは多少解消された。提案手法の(3)では提案手法(2)よりもよりエッジがまとまっているが、一方で、中心部でエッジが過度にバンドリングされてしまい、エッジの追従性が悪くなっている。提案手法の(4)では既存手法(1)に比べるとバンドリングされているが、エッジのほとんどが直線で描かれたままである。提案手法の(5)では、(3)の中心部での過度のバンドリングが解消され、エッジクラッタの解消に成功している。また提案手法(2)(3)(5)では描画対象領域内でバンドリングの効果が現れたことで、描画対象領域外縁部でエッジ群が明確に認識できるようになり、インディアナポリス周辺の空港からは6,7方向の航空路が存在すること、東西方向に遠路の航路が、南北方向に短距離の航空路が多く存在すること、南東方向の航空路が少ないこと、などがよりわかりやすくなり、描画対象領域と描画対象領域外の関係



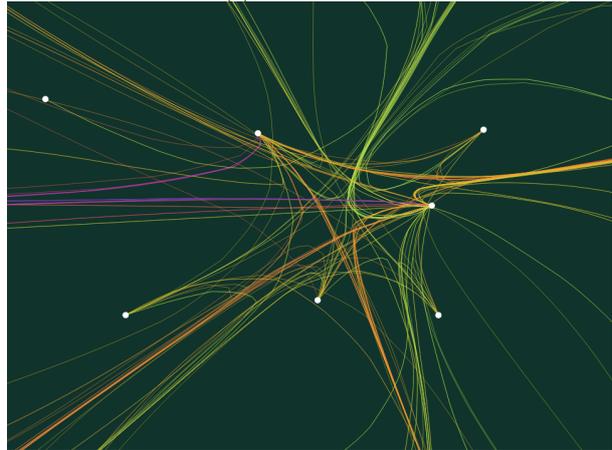
(0) 初期状態



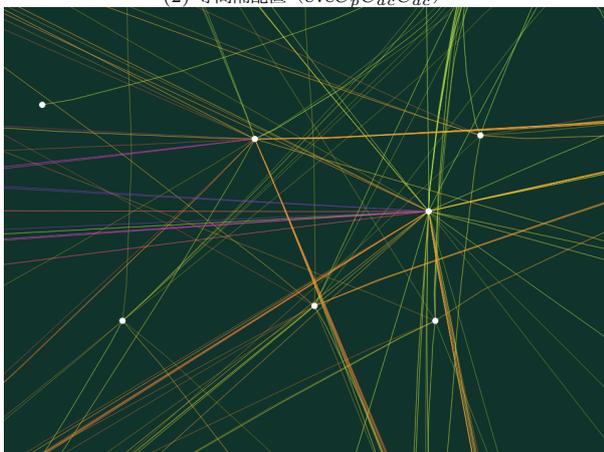
(1) 既存手法 FDEB



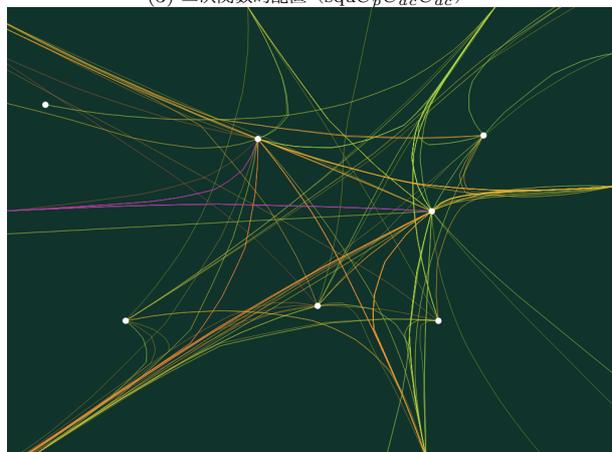
(2) 等間隔配置 (eve $C_pC_{ac}C_{dc}$)



(3) 二次関数的配置 (squ $C_pC_{ac}C_{dc}$)



(4) 指数配置 (exp $C_pC_{ac}C_{dc}$)



(5) 分割点の割合による配置 (bou $C_pC_{ac}C_{dc}$)

図 5 適用結果 (遠距離のエッジを紫色, 短距離のエッジを黄色にレンダリング)

が読み取りやすくなった。

7 評価

エッジバンドリングの定量的な評価は [9] などで研究されている。本研究ではこれらの MELD (Mean Edge Length Difference), MOA (Mean Occupied Area), EDD (Edge Density Distribution) を応用し、エッジバンドリングとしての評価を行う。また、描画対象領域外と描画対象領域内の関係を表しているかという観点から評価を行うため独自の観点から評価を行う。

7.1 構造美学的評価

エッジバンドリングを評価するにあたって、三つの観点が既存研究で挙げられている。一つ目の観点はエッジの長さである。バンドリングしない単純な可視化では、一般的にエッジはノードとノードの間を直線で描画される。これは直線が最短距離であり、エッジの接続関係を理解する上で直感的で、かつ、後述する描画する面積的にも最小化を見込めるからである。無意味な湾曲はエッジの接続関係の把握を困難にし、描画する面積も増加してしまう。したがって、バンドリングにおいても、なる

べく曲率をおおきくしないことが求められる。

二つめの観点は描画する面積の最小化である。エッジの描画されている部分が増大すればするほど、エッジは複雑に重なり合って描画され、理解が困難になる。また、一般的にエッジはノードに比べ、描画に占める面積の割合が大きい。したがってエッジの描画されている面積を小さくすることは、ネットワークの可視化において、理解のしやすさに大きく影響を与えることが予想できる。

三つめの観点はエッジの描画密度である。この観点も同様に、二つ目の観点に関係することであるが、単に描画する面積を小さくするため、関係のないエッジ同士をバンドリングすることは、エッジの接続関係を把握することを困難にしてしまう。関係のあるエッジの束はまとまっていることが好ましいが、関係のないエッジの束とは明確に分離している方が理解がしやすくなる。

7.1.1 平均エッジ長差 (MELD)

一般にエッジバンドリングでは直線のエッジを曲線化して束ねることで視認性の向上を図るが、過剰にエッジを曲線化してしまうとかえってエッジを追従することができなくなってしまい、視認性が低下してしまう恐れがある。したがって、以下により定義される、平均エッジ長差 (MELD) [9] により、バンドリングされたエッジの長さの変化の平均でバンドリングを評価できる。

$$MELD = \frac{1}{n} \sum_{e \in E} |L'(e) - L(e)| \quad (7)$$

ここで n はエッジの本数、 $L(e)$ はバンドリングされる前のエッジ e の長さ、 $L'(e)$ はバンドリングされた後のエッジの長さである。

この指標はバンドリング全体を評価するのに適しており、拡大描画されたバンドリングは考慮されていない。従って式7を次のように変更し、バンドリングを評価する。

$$MELD' = \frac{1}{n} \sum_{e \in E_{inout}} L'_a(e) - L_a(e) \quad (8)$$

ここで E_{inout} は片方の頂点を描画対象領域内に、もう片方の端点を描画対象領域外に持つエッジの集合である。また、 $L_a(e)$ はバンドリングされる前のエッジ e の描画対象領域内の頂点 v からエッジ e が描画対象領域を形作る4辺のいずれかとの交点 v' までの長さであり、 $L'_a(e)$ はバンドリングされたエッジにおいて $L_a(e)$ と同様の操作をして得られる長さである。

この値が大きいほど、バンドリングしていない状態にくらべ、エッジはより強く曲線化されたことを表す。

7.1.2 平均占有量 (MOA)

エッジバンドリングでは曲線化されたエッジ同士が重なり合って描画されることでエッジクラッタが解消され視認性が向上する。したがって、以下のように定義された、平均占有量 (MOA) [9] により、エッジを描画するピクセルの変化を比較することでバンドリングを評価できる。

$$MOA = \frac{1}{N} \left| \bigcup_{e \in E} O(e) \right| \quad (9)$$

MOA では area と occupation degree という新たなパラメータを与え計算をおこなう。area は可視化結果を均等なピクセル数で正方形に分割したうちの一つの正方形である。occupation degree は閾値であり、area 内のエッジによって塗られているピクセル数がこの閾値を越えたときに式9の集合 $O(e)$ の要素となる。

この計算方法ではある一つの area に複数本のエッジが描画されていて、一本一本のエッジの描画ピクセルが閾値を超えておらず、一方で複数本のエッジの描画ピクセルの合計が閾値を超えていたとき、集合 $O(e)$ の要素とならないという問題があった。したがって MOA' を以下の様に計算する。

$$MOA' = \frac{1}{N} \sum_{a \in A} O'(a) \quad (10)$$

$$O'(a) = \begin{cases} 1, & \text{if } \sum_{p \in a} o'(p) \geq \text{occupation degree} \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

$$o'(p) = \begin{cases} 1, & \text{if } p \text{ is occupied} \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

ここで A は可視化の対象の拡大描画に含まれるエリアの集合であり、 p はピクセルであり、 N はエリアの総数である。またバンドリングする前の状態で同様に計算した MOA'_{init} からこの値を減算して平均占有量差 (MOAD) を計算し、比較を行う。

$$MOAD = MOA'_{init} - MOA' \quad (13)$$

この値が大きいほど、バンドリングしていない状態と比べ、エッジを描画している area が少なくなったことを表す。逆に、この値が小さいほど、バンドリングしていない状態と比べ、エッジを描画している area が多くなったことを表す。

7.1.3 エッジ密度平均偏差 (EDD)

エッジバンドリングでは似た傾向のエッジが束ねられることで、描画対象領域に均一に配置されていたエッジ郡が複数箇所にまとまり、結果としてエッジが描画されている部分とそうでない部分の粗密が明確になることが期待される。すなわち、描画対象領域内の描画されているピクセルの配置の偏差が大きくなることが期待される。そこで、以下のように定義された、エッジ密度平均偏差 (EDD) [9] により、エッジを描画するピクセルの偏差の比較によりバンドリングを評価できる。

$$EDD = \frac{1}{n} \sum_{a \in A} |p(a) - p_{ave}| \quad (14)$$

ここで $p(a)$ はエリア a に描画されているエッジが占めるピクセルの割合であり、 p_{ave} はその平均である。この値に基づき比較を行い、バンドリングを評価すると、描画されているピクセル総数が多ければ多いほど偏差は大きく計算されるので、比較には向かない。従って、以下のように EDD を p_{ave} で除算し、無次元であるエッジ密度相対平均偏差 (REDD) にする。

条件	MELD'	MOAD	REDDD	pseudoF
(1) 既存手法 (def $C_a C_s C_p C_v$)	0.9389	0.2417	0.0843	449.1663
(2) 等分割配置 (eve $C_p C_{ac} C_{dc}$)	3.6002	0.4058	0.3506	1585.9775
(3) 二次関数的配置 (squ $C_p C_{ac} C_{dc}$)	9.3018	0.4483	0.4516	1168.5115
(4) 指数的配置 (exp $C_p C_{ac} C_{dc}$)	0.2286	0.4117	0.1981	606.2218
(5) 分割点の割合による配置 (bou $C_p C_{ac} C_{dc}$)	4.7937	0.5125	0.5393	9049.6856

表 1 親和性, 分割点の配置を変化させた際の MELD', MOAD, REDDD, pseudoF 値

$$REDD = \frac{1}{p_{ave}} \cdot \frac{1}{n} \sum_{a \in A} |p(a) - p_{ave}| \quad (15)$$

さらにバンドリングを実行する前の状態で計算した $REDD_{init}$ で減算してエッジ密度相対平均偏差差 (REDDD) を計算し, 比較する.

$$REDDD = REDD - REDD_{init} \quad (16)$$

この値が大きいほど, バンドリングしていない状態と比べ, エッジが密に描画されている部分とそうでない部分が明確になったことを表す. 逆にこの値が小さいほど, エッジは描画対象領域に満遍なく描画されたことを表す.

7.2 オフスクリーン可視化の観点の評価

先に述べた構造美学的評価はエッジバンドリング自体の評価指標であり, 本研究の目的である, オフスクリーン可視化という観点を考慮していない. したがって, 本研究ではオフスクリーン可視化の観点から評価をするため, エッジが描画対象領域の端と交わる点に注目する.

7.2.1 概要

本研究では, 拡大描画した際に画面外の情報を画面内に反映するために, 描画対象領域内に存在するノードと描画対象領域外に存在するノードを結ぶエッジに着目し, これらの接続関係を束として認識できるようにすることを目的としている. そこで本研究では, バンドリングされたエッジが描画対象領域の端と交わる点をクラスタリングし, そのクラスタリング結果を評価することで比較を行う.

7.2.2 評価手法

描画対象領域の矩形の頂点の一つを原点 $O = (x_0, y_0)$ とし, その対角線上にある矩形の頂点を $O_d = (x_1, y_1)$ とする. バンドリングされたエッジが描画対象領域の矩形の辺と交わる点を $b = (x, y)$ とし, その座標 b_o を以下のように定める.

$$b_o = \begin{cases} x - x_0 + y - y_0, & \text{if } y - y_0 \leq \frac{y_1 - y_0}{x_1 - x_0} (x - x_0) \\ -(x - x_0 + y - y_0), & \text{if } y - y_0 < \frac{y_1 - y_0}{x_1 - x_0} (x - x_0) \end{cases} \quad (17)$$

このように座標を一次元にすることで描画対象領域の矩形の縦横比を考慮せずに評価が行える. この一次元の座標に対し, DBscan を行う. その際, 点 b_{o1} と b_{o2} ($b_{o1} > b_{o2}$) 間の距離 d を以下のように定義する.

$$d = \min(b_{o1} - b_{o2}, 2 \times (x_1 - x_0 + y_1 - y_0) - b_{o1} + b_{o2}) \quad (18)$$

以上のように計算することで循環を考慮する.

上記を用いたクラスタリング結果に対する pseudo F 値を計算し, 評価を行う. なお, pseudo F 値とは, クラスタリング結果について, クラスタ同士の離散性, クラスタ内での凝縮性を考慮した指標である.

7.3 評価結果と考察

結果を表 1 に示す. 提案手法 (2)(3)(5) では既存手法に比べ MELD' が減少し, エッジの曲率が増加してしまっただが, MOAD, REDDD は共に上昇しており, エッジクラッタが解消したといえる. 提案手法 (4) は既存手法に比べ MELD' が上昇し, エッジの曲率を抑えつつ, なおかつ, MOAD, REDDD は共に上昇しており, エッジクラッタが解消したといえるが, 既存手法 (3)(5) などと比べると REDDD が小さく, エッジが描画対象領域に満遍なく描画されているといえる. 描画対象領域内外の関係という観点から比較すると, pseudoF 値は提案手法 (5) が最も大きく, 描画対象領域端でエッジがまとまり, 方向が明確化しているといえる.

8 おわりに

本研究では拡大描画に適したネットワーク構造の可視化のためのエッジバンドリング手法を, 新たな分割点の配置とエッジ親和性により提案した. また既存のバンドリングの定量的評価手法を拡大描画のために拡張し, オフスクリーン可視化の観点からの定量的評価手法を検討し, 提案手法の可視化結果の評価を行い, 提案手法の有用性を示した. 今後の課題としては, 描画領域内のノードを両端点に持つエッジと, 片方の端点のみが描画領域内のノードであるエッジとの相互作用, エッジの長さの可視化, スムーズなバンドリングの切り替わり等があげられる.

文献

- [1] I. Herman, G. Melancon and M. S. Marshall, "Graph visualization and navigation in information visualization: A survey," IEEE Transactions on Visualization and Computer Graphics, Volume 6, No 1, pp.24-43, 2000
- [2] Lhuillier, A. and Hurter, C. and Telea, A, "State of the Art in Edge and Trail Bundling Techniques," Computer Graphics Forum, Volume 36, Number 3, pp.619-645, 2017
- [3] Danny Holten, "Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data," IEEE Transactions on Visualization and Computer Graphics, Volume 12, No 5, pp.741-748, 2006
- [4] Weiwei Cui, Hong Zhou, Huamin Qu, Pak Chung Wong, and Xiaoming Li, "Geometry-Based Edge Clustering for Graph Visualization," IEEE Transactions on Visualization and Computer Graphics, Volume 14, No 6, pp.1227-1284,

- [5] Danny Holten, Jarke J. van Wijk, “Force Directed Edge Bundling for Graph Visualization,” *Computer Graphics Forum*, Volume 8, Number 3, pp.983-990, 2009
- [6] David Selassie, Brandon Heller, Jeffrey Heer, “Divided Edge Bundling for Directional Network Data,” *IEEE Transactions on Visualization and Computer Graphics*, Volume 17, Number 12, pp.2354-2363, 2011
- [7] Sheng-Jie Luo, Chun-Liang Liu, Bing-Yu Chen, Kwan-Liu Ma, “Ambiguity-Free Edge-Bundling for Interactive Graph Visualization,” *IEEE Transactions on Visualization and Computer Graphics*, Volume 18, Number 5, pp.810-821, 2012
- [8] 山下 敬史, 佐賀 亮介, “多属性グラフへのエッジバンドリングの適用,” *情報科学技術フォーラム講演論文集*, Volume 14, Number 4, pp.43-48, 2015
- [9] Saga, Ryosuke, “Quantitative Evaluation for Edge Bundling Based on Structural Aesthetics,” *Eurographics Conference on Visualization*, 2016
- [10] Jäckle Dominik, Kwon Bum Chul, Keim Daniel, “Off-Screen Visualization Perspectives: Tasks and Challenges,” *Symposium on Visualization in Data Science (VDS) at IEEE VIS*, 2015
- [11] Baudisch Patrick, Rosenholtz Ruth, “Halo: A Technique for Visualizing off-Screen Objects,” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Number 8, pp.481-488
- [12] Frisch Mathias, Dachselt Raimund, “Off-Screen Visualization Techniques for Class Diagrams,” *Proceedings of the 5th International Symposium on Software Visualization*, Number 10, pp.163-172, 2010