

持続的ホモロジーに導かれた力方向グラフのレイアウト

アシュリー・スー、ムスタファ・ハジ、ベイ・ワン、カルロス・シャイ
デガー、ポール

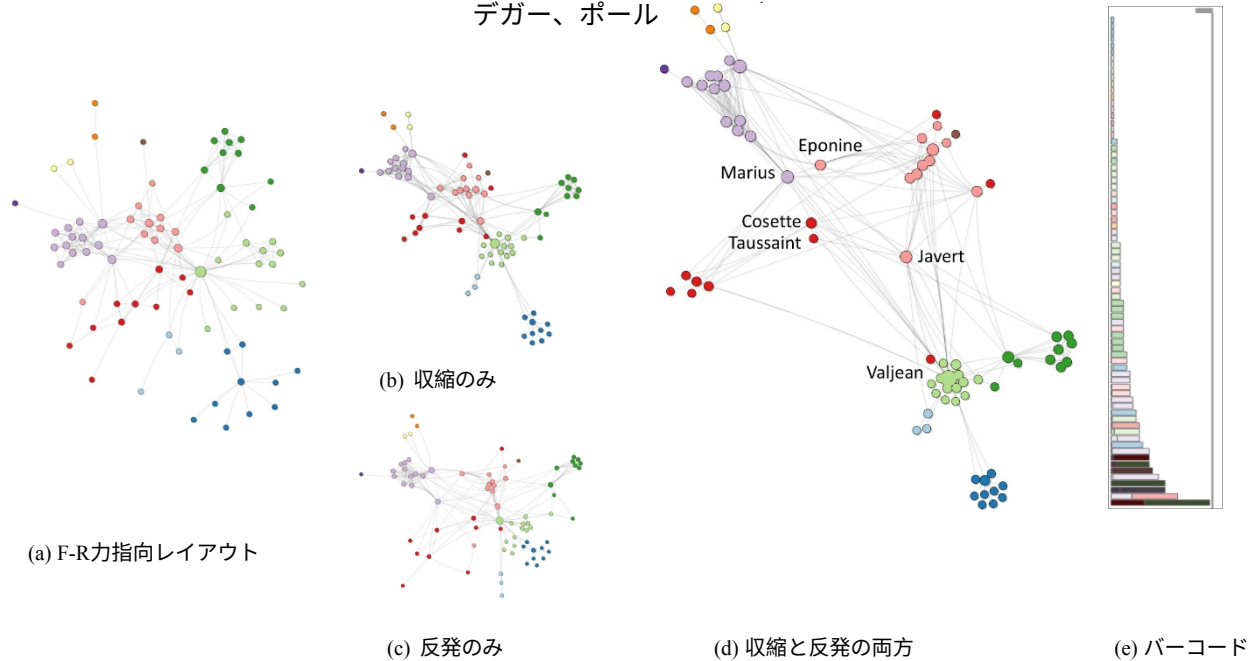


図1. (a) Le's Miserablesのグラフは、Fruchterman-Reingold (F-R) 力指向レイアウト[31]を使って描かれている。我々のアプローチでは、(e)永続性バーコードを使用して、強制指向レイアウトと相互作用するための2つのメカニズムを提供する。(b)最初のメカニズムは、重要性の低い、あるいは永続性の低い特徴に関連するグラフのノードを縮小する。(c) 2番目のメカニズムは、ユーザーが選択した特徴を用いてグラフを分割し、異なる分割のノードを互いに反発させる。(d)このアプローチを組み合わせると、永続的な相同性を利用して、ユーザーが選択したグラフの側面を強調するようにレイアウトをインタラクティブに制御することができる。

抽象-グラフはエンティティ間の関係を符号化するためによく使われるが、その抽象性のために分析が難しい。ノードリンクダイアグラムはグラフを描くのによく使われ、力方向レイアウトは、グラフの大域的な形状を明らかにしようとして局所的な関係を利用するノード配置のための柔軟な方法を提供する。しかし、無関係な構造の乱雑さや重なりは、グラフの視覚化を混乱させる原因となる。本論文では、無向グラフの持続的な相同性の特徴を、力指向レイアウトの対話的操作のための派生情報として活用する。まず、重み付き無向グラフと重みなし無向グラフの両方から、0次元の持続的相同性特徴を効率的に抽出する方法について述べる。次に、力方向グラフレイアウトを操作するために使用される対話的な永続性バーコードを紹介する。特に、ユーザは、持続的ホモロジー特徴によって生成される収縮力と反発力を追加・削除し、最終的にレイアウトを最も改善する持続的ホモロジー特徴の集合を選択する。最後に、様々な合成データセットと実データセットを用いて、本アプローチの有用性を実証する。

索引用語-グラフ描画、力指向レイアウト、トポロジカルデータ解析、持続的ホモロジー。

1 はじめに

グラフは、個人やオブジェクト間の複雑な関係を表現するためのユビキタスなものであり、社会的相互作用、エネルギーグリッド、コンピュータネットワーク、脳の連結性などをモデル化

するためによく使用される。グラフの抽象性の高さは、視覚化において大きな柔軟性を提供する。しかし、密度が高く、直径の小さい部分グラフは、「毛玉」のような分かりにくい可視化につながる。優れたグラフの視覚化は、構

造を迅速かつ明確に示し、データのさらなる調査をサポートするものでなければならない。

- アシュリー・スーは南フロリダ大学およびタフツ大学所属。Eメール: ashley.suh@tufts.edu。
- Mustafa Hajij はオハイオ州立大学。電子メール: hajij.1@osu.edu。
- Bei Wang はユタ大学。電子メール: beiwang@sci.utah.edu。
- カルロス・シャイデガーはアリゾナ大学。電子メール: cscheid@email.arizona.edu。
- ポール・ローゼンは南フロリダ大学。電子メール: prosen@usf.edu。

原稿受領 *xx xxx.201x* 年。201*x*. 出版日 *xx xxx.201x*; 最新版の日付 *xx xxx.201x*. 本論文の別刷り入手に関する情報は、reprints@ieee.org。デジタルオブジェクト識別子: *xx.xxxxxx/TVCG.201x.xxxxxx*

ノード・リンク図の重要な要素は、ディスプレイ上に（半）自動的にノードを配置するレイアウト・アルゴリズムである。au-tomatic グラフレイアウトの問題には豊富な文献があり、多くのアプローチは、グラフの対称性、辺の長さ、辺の交差の数などの読み取り能力メトリック[77]を最適化することによってグラフの埋め込みを見つけることに焦点を当てている。重要な進歩は、ノードランク、グラフ距離、近似クラスティングなどの派生情報を使用することで、グラフレイアウトを改善することに気づいたことである[23, 34, 68]。しかし、これらの技術の多くは、グラフ・レイアウトを対話的に操作する能力を欠いているか、そのような対話を効果的に行うために必要なレイアウトの時間的一貫性を欠いている。

インタラクティブ性をサポートするグラフ・レイアウトを考えると、おそらく最もポピュラーな（必ずしもベストではないが）方法は、力指向またはバネ質量レイアウト[36]である。これらのシステムは、グラフの全体的な形状を明らかにするために、局所的な関係に依存している。その結果、特定のグラフ、特に疎なグラフにおいて、トポロジカルな構造を示す手法となる。しかしこの手法では、無関係な、あるいは離れたトポロジカル構造が重なったり、パスが交差したりすることが多く、区別が難しくなる。この問題に対処するために、ユーザーとのインタラクションが用意されている。残念なことに、インタラクションはほとんどの場合

これは、個々のノードをクリック&ドラッグすることによって行うが、大きなグラフでは効果がなく、グラフレイアウトに適応される力によって制約を受ける。本論文では、無向グラフの可視化のために、*持続的相同性* (PH) [27,35]を非依存的情報として活用することで、力によるグラフレイアウトのインタラクティブな操作に取り組む。PHは最近、グラフのロバストな記述子であることが示されており[41, 78]、このアプリケーションに理想的ないくつかの重要な性質を持っている。第一に、PHの計算では、パラメータを選択することなく、グラフから0次元のホモロジー群の形でPHの特徴を抽出する。第二に、PHの特徴は定量化することができ、その重要性に応じてランク付けすることができる。第三に、PH特徴量は小さな変形に対して不変であり、ノイズやデータの小さな変化に影響されない（例えば、重みの小さいエッジを削除してもグラフは大きく変化しない）[41]。最後に、すべてのPH特徴の集合は、*持続性バーコード*[35]を用いて表現できるグラフの圧縮された記述を生成する。

ノードを直接操作する。

簡単に説明すると、我々のアプローチは以下のように機能する。全てのノード間の距離を誘導することにより、無向グラフをメトリック空間に埋め込む。我々は、メトリック空間構造[28]のPH特徴（すなわち、0次元ホモロジー特徴）を抽出し、それらの有意性（すなわち、永続性）を用いてソートする。Fruchterman-Reingold force-directed layout [31]からスタートし、ユーザが選択可能な2つの方法でPH特徴を採用する。第一に、選択されたPHフィーチャーは、そのフィーチャーを作ったノード間に強い吸引力を発生させ、それらを収縮させることができる（図1(b)参照）。第2に、選択されたPH特徴を用いてグラフを2つの部分集合に分割し、互いに反発させることができる（図1(c)参照）。ユーザーは、関心のあるグラフ要素を強調するために、収縮力または反発力を好きなだけ用いることができる（図1(d)参照）。

貢献我々は、0次元のPH特徴を用いて、力指向のレイアウトを制御することの有用性を実証する。要約すると、1)重み付きグラフと重みなしグラフの両方からPH特徴を抽出することについて議論する。2)PH特徴に由来する新しい力をレイアウトに導入する。3)PH特徴を用いてレイアウトをインタラクティブに操作できる、パーシスタンスバーコードに基づくインタラクティブなインターフェースを提供する。

2 過去の作品

グラフの可視化。グラフの可視化は、von Landesbergerらの調査[88]が示すように、幅広い分野である。ノードリンク図[36]

は、Gephi[9]、NodeXL[42]、Graphviz[29]などの一般的な可視化システムでグラフを表示するために使用されています。

ノードリンク図をレイアウトするための最初の自動化技法は、Tutteのバリセントリック座標埋め込み[86]であり、その後、線耳プログラミング技法[34]、力指向／マスをプリング埋め込み[31, 47]、グラフメトリック埋め込み[33]、接続性構造の線形代数的特性を利用する技法[11, 53, 55, 56]が続いた。TopoLayout[4]のようなハイブリッドアプローチは、グラフのトポロジーを分析して、最適なグラフ埋め込みタイプを特定する。ストレスメジャーライゼーションを用いた最近の研究では、星、クラスター、円などの特定の特性を強調するレイアウトを可能にする制約を追加する機能が導入されている[89]。多変量ネットワークを可視化するという、より困難な問題は、視覚的分析アプローチによって対処されている[87]。

エッジの乱雑さは、ノードリンク図にとって難しい問題である。高密度のグラフの場合、エッジバンドルはグラフエッジを画面の同じ部分にルーティングすることで乱雑さを軽減することができる[45]。品質に関しては、分割エッジバンドル[83]は高品質な結果を生成するが、階層エッジバンドル[32]は数百万エッジまでスケールする。また、ユーザが選択した関心領域に基づいてエッジの表示を適応させるエッジバンドルリング[91]やフィルタリング[48, 84]の局所化バージョンもある。

グラフのトポロジーやモジュール[25]に基づくモチーフ[24]でノードを置き換えるような比較的保守的な提案から、行列図の変形[21]やグラフ統計の抽象的な表示[50]のようなより積極的な形式まで、全体的な乱雑さを減らすために他の視覚的メタファーが提案されてきた。

大規模なデータセットを表示する場合、グラフの視覚的な限界に疑問を持つのは自然なことである。ピクセルベースの視覚化[51, 52]のような一般的なアプローチは、小さな矩形または表示ピクセル内に大量のデータをエンコードする。空間充填曲線もまた、ピクセルベースのグラフ視覚化を構築するために使用されている[64]。さらに、ネットワークデータに合わせた視覚的ブースト[73]を使用することで、隠された情報をさらに明らかにすることができる。力指向のノードレイアウトのインタラクティブな操作に関する研究には、グラフ内の関心のある領域に焦点を合わせるために使用されるパンやズームなどのインタラクション技術[44, 88]が含まれる。力指向レイアウト[31, 47]のノードと直接対話することに加えて、階層的レイアウトの構築[43]や探索[3, 5]、魚眼レンズ[81, 90]、グラフの対話的洗練化[81, 90]などのアプローチがある。

自動レイアウト [30]、あるいは制約に基づく最適化 [80]。

持続的ホモロジーとグラフPHは、複雑なグラフを研究するための新しいツールである[22, 26, 46, 74, 75]。

7,12]や脳ネットワーク[13,17,58-61,76]に利用されている。PHは最近、可視化コミュニティにおいて、関コミュニティ[78]や時変グラフ[41]を対象としたグラフ解析に使われている。これらの手法の多くは、PHを使ってグラフを要約的に分析、定量化、比較するという類似のアプローチをとっている。対照的に、我々のアプローチは、グラフレイアウトのインタラクティブな操作を可能にするためにPHを使うことに焦点を当てている。

3 グラフの永続的ホモロジー

我々はまず、PHに基づいた我々のアプローチの理論的枠組みを提供する。4.1節では、本論文で用いるPHの限定形が、単一連結階層クラスタリングの特別な場合であることを述べる。4.1節では、本論文で使用するPHの限定形が、単リンク階層クラスタリングの特殊な場合であることを述べる。グラフからPHの特徴を抽出するために、グラフの計量空間表現にPHを適用する[41]。入門的なサーベイについては[27]を参照のこと。

や[28]でPHの正式な扱いについて述べている。

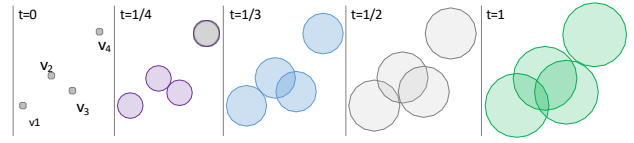
代数的位相幾何学では、グラフの0次元ホモロジー群は、 1 つの空間における計量空間の連結成分を記述する。

$$\begin{bmatrix} 0 & 1/3 & 1/2 & \infty \\ 0 & & 1/4 & \infty \\ & & 0 & 1/1 \\ & & & 0 \end{bmatrix}$$

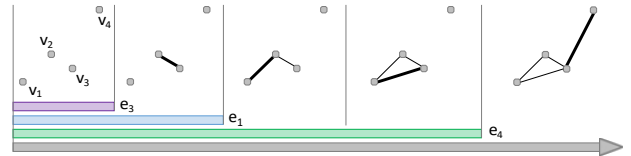
↓

$$\begin{bmatrix} 0 & 1/3 & 1/2 & 3/2 \\ 0 & & 1/4 & 5/4 \\ & & 0 & 1/1 \\ & & & 0 \end{bmatrix}$$

(a) グラフからメトリック空間への埋め込みへの変換

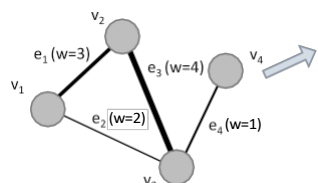


(b) 成長するボールを使ったPH機能の概念的構築



(c) 1-シンプリシーを用いたPH特徴の計算による構築

図2.グラフの0次元PHの抽出例。(a)エッジの重みが w の無向グラフ G が与えられたとき、 $d=1/w$ を用いて重みを距離に変換し、最短経路距離を用いてメトリック空間を完成することにより、メトリック空間表現を得る。(b) 概念上、構成要素 (PH特徴) は、メトリック空間の各点の周りに形成される。構成要素がより大きな構成要素に統合される直径 t を特定するために、ボールが計量空間の点の周りに成長する。(c)2つのボールが交差するときにエッジを追加することによって、 G からフィルトレーションが構築される。つの構成要素が1つに統合されると、永続バーコード (下) の一方の構成要素に関連するバーが終了する。



の解像度を持つ。本論文では、異なる空間解像度における空間の特徴の進化を記述するために、ホモロジーのマルチスケール概念であるパーシステントホモロジー (PH) を用いる。

で定義される正の辺重みを持つ重み付きグラフ $G = (V, E, w)$ が与えられたとする: $E \rightarrow \mathbb{R}$ で定義されるエッジの重みが正である。正¹の逆数を考える。

ここで、各対ノード $x, y \in G$ 間の距離は、それらの間の最短経路である。このメトリックは次のように計算できる。Dijkstraのアルゴリズム[19]を使用している。図2(a)を参照。図2(a)

アルゴリズムの残りは、元のグラフを考慮することなく、メトリック空間に対してのみ動作する。

G のすべてのノードは計量空間の点に対応する。 G の0次元 PH を計算するために、その計量空間表現に簡単な幾何学的構成を適用する。計量空間の各点を中心とする直径 t の球の集合を考える。

t が0から ∞ に増加するにつれて、ボールの結合の構成要素がどのように進化するかを追跡する。 t が増加するにつれて、ボールの連合は階層的に構成要素を形成する。

図2(b)を考えると、 $t=0$ の時点では各点が1つの構成要素であったものが、 t が増加するにつれて、2つの構成要素が合体すると構成要素の数は1つずつ減少する。

イベント²。 $t = 1/4$ で、 v_2 と v_3 を表すボールが接触し、構成要素 $\{v_2, v_3\}$ の併合が起こる。 $t = 1/3$ で、 v_1 と v_2 を表すボールが接触し、集合 $\{v_1\}$ と $\{v_2, v_3\}$ が結合する。 $1/2$ で、 v_1 と v_3 が触れる。しかし、それらはすでに同じ構成要素の一部である $\{v_1, v_2, v_3\}$ 、つまりマージは起こらない。最後に、 $t = 1$ で、 v_3 と v_4 は

v_1, v_2, v_3 と $\{v_4\}$ をマージする。

PHの特徴は、誕生（登場）と死（消滅）に対応している。

ing) のこと。成分の誕生とは、その成分が現れるときの直径である。つまり、2つの成分が合体するとき、一方の成分は他方の成分と合体して消滅する（どちらの成分が消滅するかは次節で議論する）。成分の寿命（すなわち、死滅時間から誕生時間を引いたもの）は、その成分の永続性である。

その始点と終点は、関連するコンポーネントの誕生時刻と死亡時刻に対応し、その幅は存続時間に比例する。図解は図2(c)を参照。

4 持続的ホモロジーの計算

本稿で用いるPHの限定形は、機能的には単結合のクラスタリングと等価である。したがって、グラフのPHは、Kruskalのアルゴリズム[57]を用いてグラフの最小スパニング木(MST)を求めることで計算できる。

4.1 0次元バーコードの高速計算

グラフ G の0次元バーコードを計算するためにクラスカルの

アルゴリズムを使用する際、簡単のために G が連結していると仮定する。しかし、そうでない場合は、グラフの各連結成分は独立に処理される。簡単に言えば、我々のアルゴリズムは、辺の長さ $1/w$ を用いたメトリック空間の埋め込みに基づいて、 $G=(V,E,w)$ のMST T を計算することからなる。

$V = \{v_1, \dots, v_m\}$ を G のノード集合とし、 $E = \{e_1, \dots, e_n\}$ とする。 $1/w$ に関して増加する順にソートされたエッジ・セットである。

このアルゴリズムは、空のスパニング・ツリーを作成することから始まる。次に、構成要素 C_i とバー B_i を、グラフのノードごとに1つずつ作成する。パーシステンス・バーコードの各バーは、実数 (birth, death) の組で表現され、最初には各ノードの birth = 0 をそれ自身のコンポーネントとする。アルゴリズムの第2ステップでは、エッジを1つずつ見ていき、 $1/w_i$ の増加順に並べる。各 e_i について、このエッジのノードが2つの異なるコンポーネントに属するかどうかをチェックする。もしそうであれば、次のようにマージする。

¹ 2つのノード間のエッジ重みの逆数 $1/w(x, y)$ は、それらの間の非類似性を捉える。

² ここでいうトポロジーとは、メトリック空間のホモロジー群のことであり、グラフのトポロジーと混同してはならない。

を構成し、一方の死を $1/w_i$ とする。どの成分が死ぬかは任意であり、結果に影響しない。死亡した成分の持続時間は、その死亡時間から誕生時間を引いたものである。このような成分の出現と消滅は、パーシステンス・バーコードにバー $(0, 1/w_i)$ を生じさせる。このステップは、不連続集合データ構造を用いて効率的に実行できる。

データ重み付きグラフ $G = (V, E, w)$

結果最小スパニングツリー T と 0次元バーコード B

```

1 空のスパニングツリー  $T = \{\}$  を作成する。
2 foreach node  $v_i$  do
3   コンポーネント  $C_i = \{v_i\}$  を作成する。
4   出生=0、死亡= $\infty$ の棒グラフを作成する。
5 エンド
6  $E$  の各辺  $e_i = (u, v)$  は次のようにする。
7    $C_u$  と  $C_v$  が異なる成分である場合
8      $C_u$  と  $C_v$  のマージ
9      $B_u$  の死を  $1/w(e_i)$  に設定する。
10     $e_i$  をスパニング・ツリー  $T$  に加える
11 終了
12 エンド

```

スパニングツリーへの辺の追加は、2つの構成要素が合体するイベントと一致する。したがって、MST T のエッジと有限永続性を持つ G の 0次元バーコードの間には1対1の対応がある。

4.2 スパニングツリーとノードの関係

MSTによって符号化された情報をグラフ G に関連付ける。

これは後にグラフィケアウトの修正を定義することになる。MSTを $T(V, E)$ とし、 E は木の辺を表す。 E から辺 $e = (u, v)$ を削除すると、木 T は V_u と V_v の2つの集合に分割される。

各PHフィーチャ（すなわちバーコードのバー）は、以下のように関連付けられる。

図3に示すように、以下の情報である：

- この目的のために、永続性バーコードの各バーを、 $(0, 1/w)$ ではなく、区間 $(0, w)$ として視覚化する。このような視覚化は、長いバーとして重みの高いエッジを強調する。³表記法の乱用により、このようなバーの永続性尺度は w に割り当てられる。
- 死因となる u と v は、コンポーネントのマージを引き起こすエッジのノードである。これらのノードは、グラフのレイアウトをPHの特徴選択を反映するように修正するために使用される。
- ノードの部分集合 V_u と V_v は、MSTからエッジを除去した後の接続ノードの集合を表す。これらの集合は、グラフ・レイアウトを更新する際にも重要になる。
- サブセット比率は、ノードの2つのサブセットに含まれるノード数 $|V_u| : |V_v|$ のノードの数の尺度である。内の中心性の尺度である。

MST。例えば、図4では、MSTの例をサブセット比率で補強している。左右の扇形領域のエッジは比率が低く、1:7 である。中央の2つのエッジは4:4と3:5と、よりバランス

のとれた比率を持ち、MSTの中心であることを示している。サブセットの比率の分布は完全にデータに依存している。

³ この可視化は従来のPHアプローチとは異なるが、我々の設定では重み w の大きいエッジがより重要であると考えられるため、正当化される。

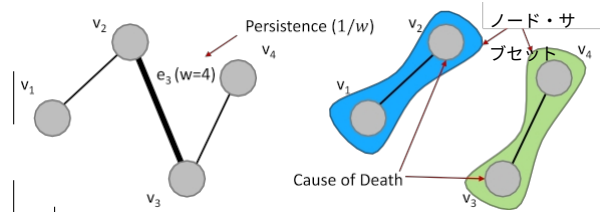


図3 図2(c)のエッジ e_3 （紫色のバー）について、スパニングツリーから抽出された情報の例（左）。右はスパニングツリーから選択されたエッジを削除したときにできるクラスターを示す。

しかし、私たちの例では、バランスのとれた比率よりも低い比率の方がはるかに一般的であることが観察された。

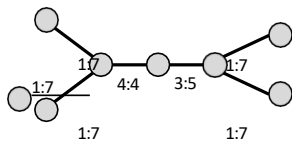
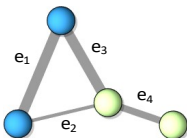


図4.サブセットの比率を調整したMSTの例。周辺部のノードは比率が低く、中心部のノードは比率が均衡している。

4.3 重みなしグラフへの応用

上述の計算では、グラフのエッジに重みが必要である。重みのないグラフの場合、Jaccard index [49]のような類似性尺度が辺の重みの代わりになる。

この手順では、まず各ノードについて、ユーザーが選択可能なホップ数以内の近傍グラフまたはエゴ・グラフ[66]を



収集する。例えば、1ホップのエゴ・グラフは隣接する隣人を含み、2ホップのエゴ・グラフは1ホップの隣人と1ホップの隣人に隣接するノードの両方を含む。理想的なホップ数はデータに依存する。より密度の高いグラフには1ホップ、より疎なグラフには2ホップまたは3ホップを用いる。次に、辺 $e = (v_i, v_j)$ が与えられ、近傍グラフ N_i と

N_j , これらのノード間のJaccard index は $J(v_i, v_j) = \frac{|N_i \cap N_j|}{|N_i \cup N_j|}$ である。である。

辺の重みは $w(e) = J(v_i, v_j)$ となる。このように、Jaccard インデックスは、接続された2つのノードの近傍間の類似性を提供する。最後に、Sect.3.

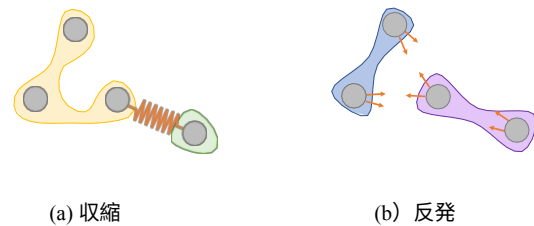
エッジ中心性[37]のような他の尺度は、グラフの異なる側面を強調するために、Jaccard インデックスの代わりに使用することができる。唯一の要件は、その尺度がエッジの重みを提供しなければならないということである。

5 ビジュアル・デザインとインタラクション・デザイン

私たちの設計目標は、グラフのPHをコントロールとして使用し、その力方向レイアウトを操作することである。私たちは、文脈情報を提供し、レイアウトの素早い操作を可能にするために、シンプルな対話型インターフェースを使用する。

5.1 グラフ描画

私たちのデザインは、エッジをオプションでバンドルできるフ



ォース・ダイレクト・レイアウトを採用している。

力指向レイアウト。グラフは最初にFruchterman- Reingold (F-R) 力指向レイアウトで描かれる[31]。このレイアウトは3種類の力から始まります。最初の力は、すべてのノードが互いに反発することを可能にする(図5(c))。2番目の力は、エッジで接続されたノードをバネで引き寄せる力である(図5(d))。最後に、弱い引力がすべてのノードをディスプレイの中央に引き寄せ、レイアウトを実質的に中央に配置する(図5(e))。これらの力のパラメータ(質量、力の強さ、ばねの静止長さ)は手動で調整する必要がある。

(c) 収縮選択

(d) 反発選択

(e) 収縮を加える

(f) 反発を加える

図6. 図5 (a) のグラフに加えられた力の説明図。(a、c、e) : (a,c,e) : レイアウトに加えられた収縮力。(b, d, f) : (e) のレイアウトに反発力が加わる。

エッジバンドル。エッジバンドルは、重なり合うエッジの影響を軽減するためにオプションで使用することができます。レイアウトはアクティブに操作されるため、時間的にコヒーレントなエッジバンドリングの実装が必要です。これを達成するために、我々は力指向エッジバンドル[45]を実装した。この技術は、エッジを細分化し、近接と方向が類似したエッジを引き寄せるために、力指向レイアウトの変形を使用する。図1は、エッジバンドリングを有効にした例である。

ノードの色付け。実験グラフは、カテゴリ・データが付加されており、ノードの色付けに使われている。他のグラフには、ノード度

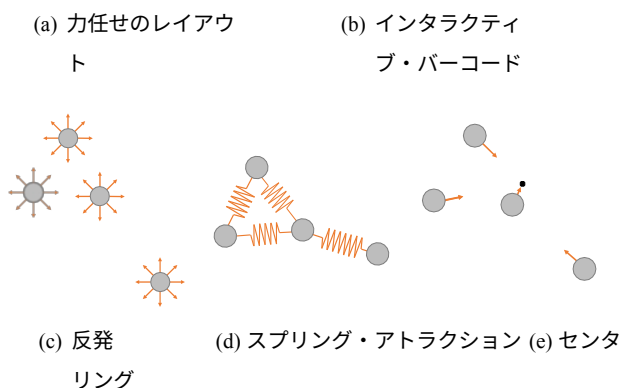


図5.図2のグラフの力指向レイアウト (a) は、(c) 反発力、(d) バネの吸引力、(e) センタリング力を使って構築されている。インタラクティブ・バーコード(b)を使って表示を操作する。

数を使って色付けされたノードがあります。

5.2 永続性バーコード



4.2 節で述べたように、我々は従来とは異なる方法で永続性バーコードを可視化している。4.2 節で述べたように、我々は従来とは異なる方法でパーシステンスバーコードを可視化する。つまり、バーが大きいほど大きいほど、そのPH 特徴が「重要」であることを示す。このバーコードをさらに視覚的な符号化で補強し (図5(b))、相互作用をさらに誘導する。

サブセット比率。各バーは縦線で補強され、サブセット比に基づいて2 つに分割される。例えば、図5(b) の一番下のバーは、図2 と図3 の e_3 を表し、MST (図3 参照) のエッジの両側に2 つのノードが存在するため、50/50 に分割されている。

色。カテゴリデータのグラフでは、PH フィーチャーの死因ノードのカテゴリに基づいて、分割の各辺が色付けされます。

バーの並べ替え。バーは2 つの基準に基づいてソートされる。1 つ目は、低い方から高い方への持続性であり、高い持続性を区別するのに役立つ。

の特徴を、持続性の低い特徴から分離する。2本のバーのパーシステンス値が等しい場合、それらはサブセット比率によってソートされ、50/50の比率を持つバーがMSTの中心であるため、バーコードの下位に表示される。

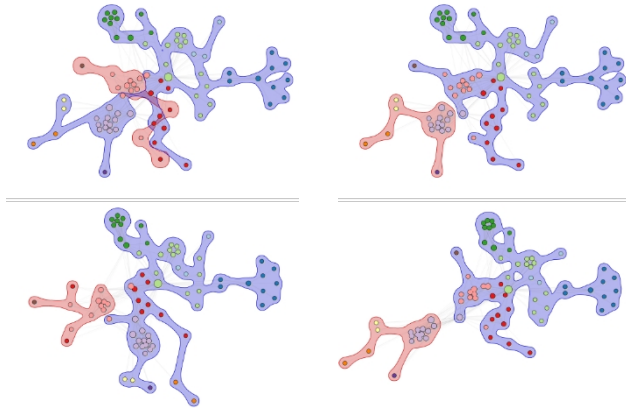
5.3 パーシステンスバーコードを使用したインタラクション

PHフィーチャー縮小。あるシナリオでは、PHを縮小することが望ましい。

レイアウトのノードに割り当てられるスペースを減らし、グラフの他の部分を表示するためのスペースを増やす。我々は、持続性がユーザーが選択した閾値以下である全てのPHフィーチャーの縮小を可能にする持続性簡略化ツールを提供する。この操作は、バーコードの上部にあるフィルターバーをドラッグすることで行う。閾値が左から右にドラッグされるにつれて、PHフィーチャの持続性

そのしきい値を下回ると、色が洗い流され、グラフ・ノードが収縮する。収縮は、イベントの死因ノードの間に強いバネの力を加えることで実現されます。この力のスケールはユーザーが選択できる。図6 (a) はこの力を示し、図6 (e) は収縮の例を示す。

PH特徴の反発。他の状況では、グラフを引き伸ばすことで、



重なりあって追跡が困難なノードやエッジのスペースを確保することができる。バーが個別に選択されると、バーの色が濃くなり、そのPHフィーチャーに関連するノードのサブセット間に強い反発力が加えられる。力のスケールはユーザーが選択できる。これらの2つのグループからのノードの反発は、レイアウトをクラスタリングすることを可能にする。例えば、図6(b)は、図6(d)の一番下のバーが反発のために選択され、サブセットが押し広げられるときの力を示している。言い換えれば、図6(b)では、青色領域のすべての点は、紫色領域のすべての点から追加的な反発を持ち、逆もまた同様である。図6(f)は、この力をグラフに加えた結果を示している。

複数のバーを選択する。力は互いに直接依存しないので、収縮と反発のために複数のバーを選択することができる。新しいバーが選択されるたびに、力は単純にレイアウトに追加されます。

ホバーリングのプレビュー。バーコードのバーの上にマウスを置いたときに、ユーザーがバーの選択の影響をプレビューできるように、バブルセット[15]やケルプダイアグラム[20]などの集合の視覚化を採用することができる。どのノードがどの部分集合に属するかを区別するために、ノードの部分集合にバブル集合を採用する。図7は、データセットにバブル・セットを使用した例を示している。バブル集合は、それぞれPH特徴が反発のために選択される前と後の2つの例を示している。

大きなバーコードの双曲線ズーム。バーの数は、グラフのノ

ード数から1を引いた数に等しい。大きなグラフに対してバーコードを適切に拡大縮小するために、双曲線ズームのスクロールバーがバーの右側に配置される。スクロールバーが移動すると、双曲線ズームのフォーカスが変更され、関連するバーが強調されます。この例は、付属のソフトウェアで見ることができる。

5.4 典型的な使用セッション

グラフがロードされ、標準的な力の強さに調整が加えられた後、ユーザーはグラフのPH機能を探索し始める。始めに、収縮力をゆっくりと閾値を高く（右へ）調整することで探索し、コンパクトさが望ましいレベルに達したときを見つけることができる。

次に、反発力についてPHの特徴を探索。PHは、最も持続性の高い要素を「最も重要」と定義する。従って、我々は通常、サブセット比率の高い持続性の高いバーや、異なるカテゴリー間で分裂するバーを探すことから始める。

(a) Partition mixed in graph before (b) Partition to be separated before (top) and after (bottom) repulsion (top) and after (bottom) repulsion

図7. Le's Miserablesグラフの2つのシナリオを示し、ユーザーが反発を考慮した場合（上）とその結果（下）。

ラベル付きデータの場合。バーの上にカーソルを置くと、まずパーティショニングが興味深いかどうかがユーザーに通知される。興味深いパーティションの例としては、混在しているパーティション（図7(a)を参照）や、ユーザーがグラフの残りの部分から移動させたい、すでに空間的に同位置にあるノードのクラスター（図7(b)を参照）などがあるが、これらに限定されない。ユーザがそのパーティションが面白いと思えば、クリックすることで反発力が有効になる。通常、5-10本のバーを選択することでよいレイアウトが得られることがわかったが、これは決して制限ではない。通常、グラフ・レイアウトを十分に調査し、微調整するのに数分もかからなかった。

表1.結果の定量的分析

データ集合			FR		エス エディ ー		我々のアプローチ				
レ・ミ ス	77	254	1(a)	< 1 ms	n/s	52 ms	1(d)	< 1 ms	< 1 ms	-23%	134%
Bcsstk	110	364	8(a)	3 ms	8(b)	65 ms	8(d)	46 ms	3 ms	46%	201%
6-ary	9,331	9,330	8(a)	49 ms	8(b)	3,514 ms	8(d)	1.97 s	81 ms	63%	1986%
バーベ ル	150	2,501	8(a)	3 ms	8(b)	95 ms	8(e)	29 ms	4 ms	30%	411%
ロブス ター	300	299	8(a)	4 ms	8(b)	116 ms	8(e)	62 ms	4 ms	50%	363%
上院	101	5,048	n/s	4 ms	n/s	110 ms	12(c)	36 ms	5 ms	42%	94%
マドリ ード	70	243	n/s	4 ms	n/s	88 ms	11(b)	46 ms	3 ms	-24%	212%
			FR		モジュラー・ク ラスター		我々のアプローチ				
空港	2,896	15,645	9(a)	15 ms	9(c)	128 ms	9(d)	3.81 s	35 ms	-12%	231%
サイエ ンス	554	2,276	9(a)	7 ms	9(c)	846 ms	9(e)	281 ms	7 ms	-7%	706%
コラボ	379	914	9(a)	7 ms	9(c)	26 ms	9(d)	121 ms	5 ms	14%	544%
カリフ オルニ ア工科 大学	762	16,651	n/s	7 ms	10(d)	110 ms	10(c)	314 ms	14 ms	29%	399%
スミス	2,970	97,133	9(a)	86 ms	9(c)	647 ms	9(e)	2.10 s	38 ms	-239%	426%

n/s: 表示なし

6 結果

PHを使って力方向レイアウトをガイドすることを評価するために、手で調整したレイアウトを使った12個のデータセットを調べた（表1参照）。我々の評価では以下のことを考慮した：

- 最大10,000ノード、最大100,000エッジのグラフを用いて、PHの計算とグラフレイアウトの更新を、インタラクティブな視覚化をサポートするのに十分な速度で行う、我々のアプローチのスケラビリティを検証する。
- 私たちは、私たちのアプローチによって作成されたレイアウトの品質を測定します。
- 我々は、我々のアプローチによって生成されたレイアウトを、一般的な力指向レイアウト手法やクラスタリング手法と比較する。
- 実際のデータを使った3つの簡単なケーススタディを紹介する。

6.1 パフォーマンス

付属のソフトウェアはProcessing [2]を使用して実装されている。すべての計算はCPU上で行われる。力の計算はマルチスレッドで行われる。付属ビデオのライブ映像は、インターフェイスのインタラクティブ性を示すために、3.1Ghz i5プロセッサーを搭載した2017 MacBook Proで制作された。

データがロードされると、0次元のPH特徴が抽出される。MST計算は、メトリック空間上で実行されるKruskalのアルゴリズムのバリエーションである。これは $O(|E|\alpha(|V|))$ 、ここで α は逆Ackermann関数[16]であり、非常に成長が遅い関数である。MSTが見つかった後、ノードの部分集合を決定する(4.2節参照)のに $O(|V|)$ を要する。

F-Rフォース・ダイレクト・レイアウトでは、バーンズ・ハットの近似値を使用する。

mation[8]の反発力と標準的なペアワイズスプリングに対するものである。総コストは反復あり $O(|V|^2 \log |V| + |E|)$ である。さらに、収縮力は単一のペアワイズスプリングです。すべてのPHフィーチャーが収縮のために選択される場合、最悪の場合の総コストは $O(|V|)$ である。反発力はコストがかかる。

力を加えるごとに、バーンズ・ハット・アルゴリズムを追加実行する必要がある。反発力の数は通常少ないので予想される実行時間は $O(|V| \log |V|)$ であり、最悪の場合は $O(|V|^2 \log |V|)$ である。

より大きなデータセットでの可視化のインタラクティブ性を向上させるために、いくつかの重要でない機能は無効化される。例えば、エッジの数が500を超えると、エッジのバンドルが自動的に無効になる。また、ノード数が100を超えると、バブル・セットが無効になる。その代わりに、グラフ・ノードはハロー（後光）で囲まれ、そのハローはノードの属するセットによって色分けされる。

選択されたPH 特徴を強調または非強調するユーザーの意図。このアプローチは、ユーザーによる修正前と修正後のレイアウトの埋め込みグラフノードのPHを比較することで機能する。このアプローチは、ノード間のユークリッド距離を使用して、ソースF-R力指向レイアウト（収縮または反発なし）のノードのPHを計算することから始まります。⁴ 次に、ユーザーが選択したターゲットレイアウトのPHを同様の計算にかかる時間はせいぜい数秒である。

レイアウトでは10ミリ秒以下、テストした大きなグラフでは常に100ミリ秒以下である。

6.2 レイアウト品質

私たちのアプローチの応答性を評価するために、私たちは特に、結果として得られたレイアウトがどの程度反映されているかを定量化するために設計された指標を導入する。

に計算する。

ソースとターゲットが与えられたとき、収縮(C) と反発(R) のためにユーザーによって選択されたバーのセットをそれぞれから抽出する。ソースとターゲットにおけるこれらのバーの持続性は、それぞれ PS と PR である。収縮(EC) と反発(ER) の効果は次のように計算される：

$$EC = \frac{1}{|C|} \sum_{x \in C} \frac{PS(x) - PT(x)}{ピーエスエック}, ER = \frac{1}{|R|} \sum_{x \in R} \frac{PT(x) - PS(x)}{ピーエスエック}$$

結果は表1の最後の2列で、ソース・レイアウト（表中のF-Rレイアウト）とターゲット・レイアウト（我々のアプローチ）を比較しています。直感的に言えば、この尺度は、あるレイアウトの特徴がターゲット・レイアウトにおいて平均してどれだけ縮小されたか、あるいは反発されたかを定量化するものである

⁴ なお、この計算ではグラフの連結性は考慮していない。

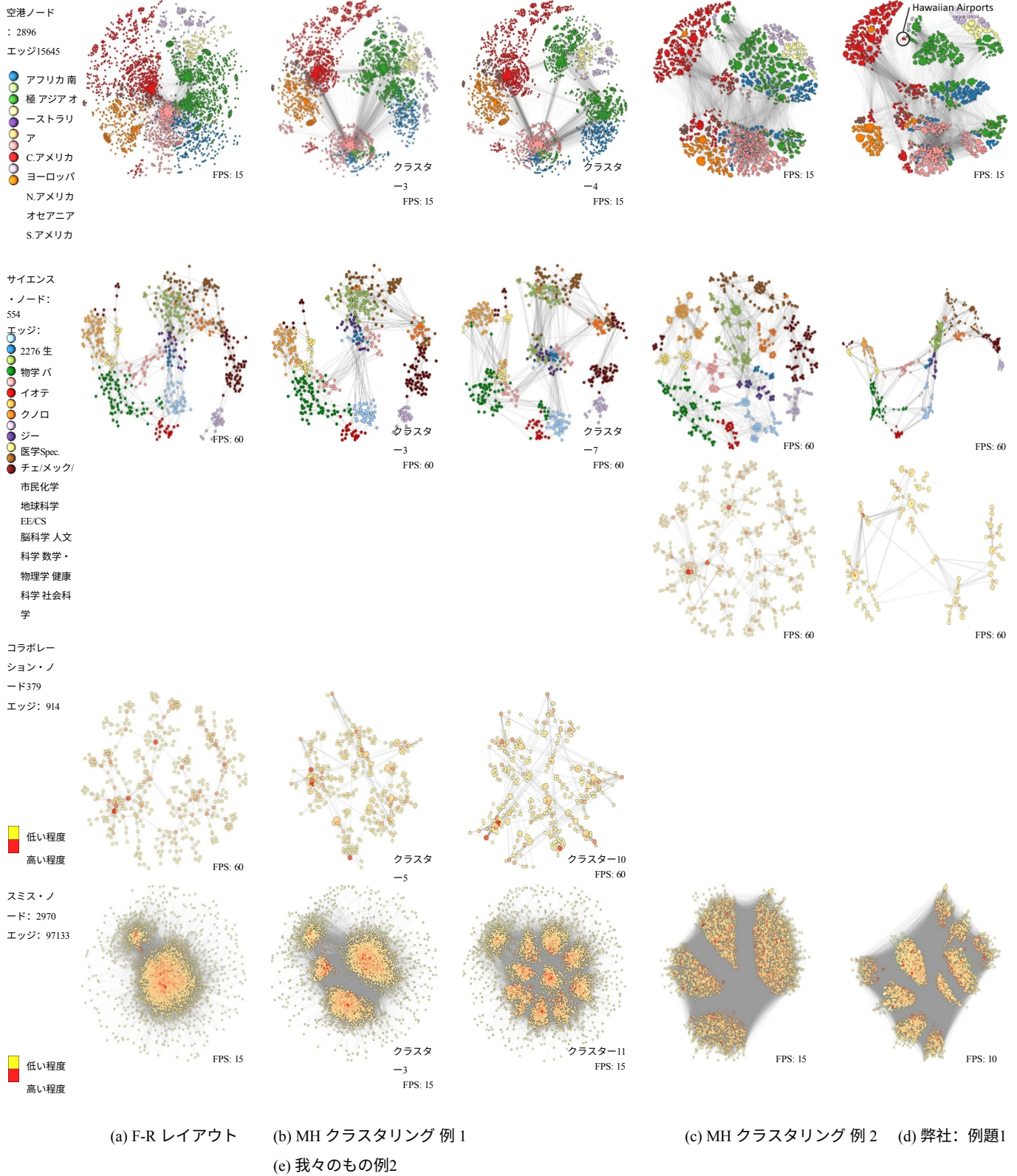


図9.一連のグラフ例に対する我々のアプローチの説明: (a)手作業で調整したFruchterman-Reingold力方向レイアウト、(b-c)手動でクラスタ番号を選択したモジュール性階層クラスタリングの2つの例、(d-e)我々のアプローチを用いた2つの例。

ソース・レイアウトに対する相対値で、マイナスは望ましく ない。ゼロは影響がないことを意味し、プラスは望ましく、

大きければ大きいほどよい。

ほとんどの場合、私たちのアプローチは、適用された収縮と反発による実質的なプラスの影響を示し、ユーザーの意図がレイアウトに反映されたことを示している。特筆すべき例外は、収縮効果がマイナスであるデータセットがあることである。現実には、反発効果がレイアウトの一部で収縮効果を上回ったため、平均的な効果がマイナスになっただけである。

6.3 他の手法との比較

6.3.1 一般的な力指向レイアウト手法との比較

図8では、F-R力方向レイアウト、Neato [72]のレイアウトと比較するために、4つの合成重みなしグラフで我々の方法をテストしている。

Graphviz[29]と我々のアプローチ。BcsstkはUF Sparse Matrix Collection [18]から、6-ary、Barbell、LobsterはNetworkX [40]を使って生成した。Bcsstkは、110個の

ノードと364本のエッジを持つ。6-aryは、9331個のノードと9330個のエッジを含む深さ5のバランス木である。バーベルは、それぞれ50ノードの2つの完全な部分グラフと50ノードのブリッジを結ぶ単純なグラフで、合計150ノードと2501エッジを持つ。Lobsterは、葉を取り除くとイモムシグラフになるという性質を持つ木である[38]。すべてのレイアウト方法において、各グラフには以下の方法で重みが適用される。

で説明されている。4.3.で説明したように、近傍サイズは以下の通りである：Bcsstk：Bcsstk：1ホップ、6-ary：2ホップ、バーベル：1ホップ、ロブスター：ロブスター：1ホップ。図8は、収縮のみ、反発のみ、両方の組み合わせを用いた我々のアプローチの3つの例を示している。

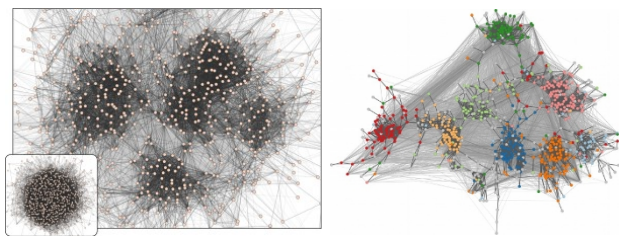
6.3.2 階層的クラスタリングとの比較

次に、モジュール性に基づく階層的クラスタリングと我々のアプローチを比較する。モジュール性を含むグラフクラスタリングのサーベイについては、[82]を参照のこと。最適バージョンはNP困難[14]であるため、我々はGraphviz[29]で利用可能なクラスタ[67, 69]を形成するために貪欲なアプローチを用いる。このアルゴリズムは、グラフの各ノードをそれ自身のクラスタに初期化することから始まる。その2つのクラスタをマージすることで、モジュラリティが最大に増加する。が組み合わされる。加重モジュラリティは[62]のように計算される：

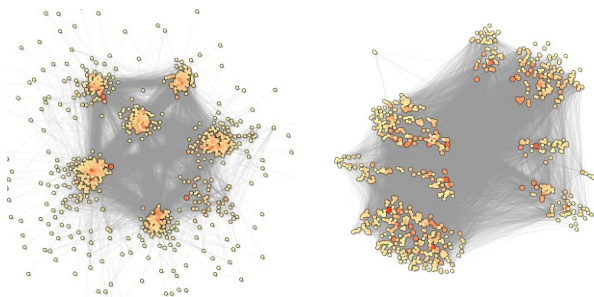
$$Q_w = \frac{1}{2w_s} \sum_{i,j} w_{ij} - \frac{1}{2w_s} \sum_{i,j} \delta_{c_i, c_j} w_{ij}$$

ここで、 $2w_s = \sum_{i,j} w_{ij}$ 、 $w_i = \sum_j w_{ij}$ 、クロネッカー・デルタ δ_{c_i, c_j} は、 i と j の両方が同じコミュニティにいれば1、そうでなければ0である。このプロセスは、1つのクラスタが残るまで繰り返される。

グラフのレイアウトにクラスタリングを反映させるために、異なるクラスタのノード間のバネの静止長を長くしている。例を図9に示す。表1にクラスタリングの計算時間を示す。



(a) [71]の適応的洗練 (b) [70]の適応的洗練



(c) モジュール性・階層的クラスタリング (d) 我々のアプローチ

図10.762個のノードと16,651個のエッジを含むCaltechデータセットを、(a,b)2つの適応的精密化手法、(c)モジュール性階層クラスタリング、(d)我々の手法で比較した。

2011年国際空港 (Airport) (図9の1行目) は、Openflights.org から取得したもので、各ノードは大陸地域別にラベル付けされた国際空港であり、辺は2つの空港を結ぶルート数によって重み付けされている。このデータセットから得られた最大の連結成分は、2,896個のノードと15,645個のエッジを含む。収縮力と反発力の組み合わせを用いると、図9の列4と5は、西部、東部、中部の空港間の分裂を明らかにする。注目すべきクラスターは、北米とアジアの分岐点に形成され、ハワイの空港がいくつか含まれている。図9の2列目と3列目の結果は、同様のクラスター形成の洞察を示しているが、ハワイの空港は直接見えない。

UCSD科学マップ (Science) (図9の2行目) [10]は、科学の554のサブディシプリンをノードとして表し、分野横断的共著を2,276のエッジとして表したマップである。我々の手法をこの特定のデータセットに適用すると、F-Rレイアウトの全体的な形状を保持し、類似の分野を共有するコミュニティのクラスタリングが追加されたグラフが得られる。一方、クラスタリングされたグラフ (列2と3) は、クラスタリング構造を強調するが、文脈 (ground truth labels) を失う。

コラボレーション・サイエンス・ネットワーク (コラボレーション) (図9段3行目)

[65]は、379のノード (ネットワーク理論における出版科学者) と914のエッジ (同じ論文に登場する2人の著者のつながり)

を持つ共著ネットワークである。このグラフの重点は、共同研究者のコミュニティを特定することである。大半のバーが縮小のために選択され (4列目と5列目)、小委員会により緊密にまとめられ、全体的なグラフの形がより明確になる。一方、クラスタリング (2列目と3列目) では、解釈が難しい結果が出た。

スミス・カレッジ (Smith College) (図9の4行目) は、Facebook100データセット[85]からのもので、スミス・カレッジの学生の社会的関係を示している。グラフには2,970のノードと97,133のエッジがある。元のグラフには寮、性別などの属性が含まれていたが、このデータを見つけることができなかった。この例では、我々のアプローチのスケーラビリティに重点を置いている (表1参照)。

Caltech (図10) は、California Institute of Technologyのソーシャルリンクのデータセットで、Facebook100データセット[85]からも引用されている。このグラフには762のノードと16,651のエッジがある。この例では、モジュール性階層クラスタリングと適応的精密化技法[70, 71]と比較する。どの手法も似たようなコミュニティを示す。残念ながら、ラベルデータがないため、他の直接比較は難しい。

6.4 ケーススタディ

6.4.1 ル・ミゼラブル

Le's Miserables Co-occurrence ネットワークには77のノードと254のエッジがあり、ノードは登場人物を表し、エッジはVictor Hugoの小説「*Le's Miserables*」[54]の任意の章において2人の登場人物が共有するシーンの数によって重み付けされる。ノードの分類は、小説の登場人物の主なグループ所属に由来し、グループは登場人物に関する知識に基づいて命名される。

図1(a)に、このデータセットのF-R力方向レイアウトを示す。持続性の高いバーの組み合わせを選択し、残りのバーを収縮させると（図1(e)参照）、図1(d)に見られるように、本で取り上げられた主要な登場人物の一部が明らかになる。図1(d)を参照）。レイアウトの対向する2辺には、マリウスとバルジャンを中心とするクラスターがあり、2人の主要な登場人物、マリウスに恋する女性エポニーヌ、バルジャンの主な敵役ジャベール、バルジャンの娘でマリウスの恋人コゼット、そしてコゼットの幼少期を育てた母親的存在トゥーサンがいる。

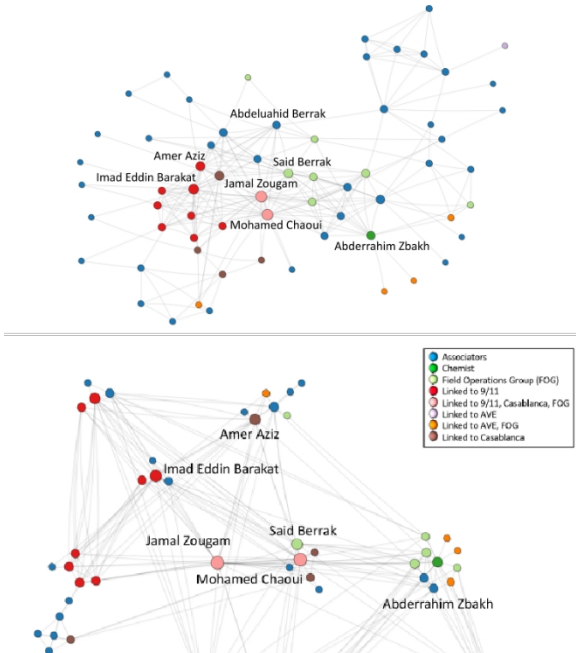
6.4.2 マドリード列車爆破テロ

マドリード列車爆破事件データセットには70のノードと243のエッジが含まれ、ノードは「2004年3月11日にマドリードで発生した通勤列車爆破事件に関与した個人」[79]を表す。各グループは、その人物が以前のテロ行為に関与していたかどうか、そしてその人物がフィールド・オペレーション・グループのメンバーであったかどうかに基づいて識別され、色付けされている。2人の人物が爆弾テロ以前、またはその最中に関係があった場合、リンクが結ばれる。重みは1〜4の間の指数で計算され、次の4つのパラメータがそれぞれペアごとに合計される：信頼-友人関係（接触、親族関係、電話会議でのつながり）、アル・カイダとオサマ・ビン・ラディンとのつながり、訓練キャンプや戦争への共同参加、過去のテロ攻撃（9月11日、カサブランカなど）への共同参加。

(b) 我々のア
フ
ロー
ド

図11.(a)原著論文[79]を再現した従来のレイアウト、(b)ネットワーク内の主要プレーヤーを強調した我々のレイアウトを使った最終的な視覚化。

(a) 従来のレイアウト



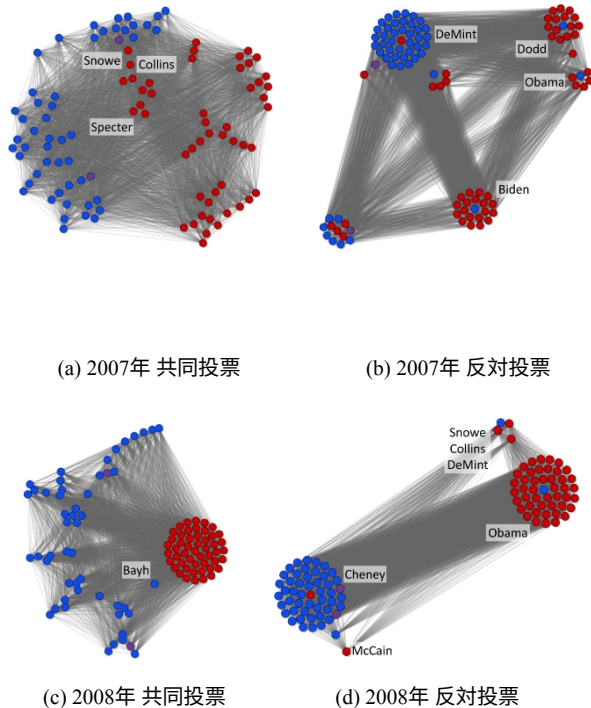


図12.2007年と2008年の米上院における共闘・反共闘グラフ。これらのグラフは、収縮力と反発力の混合を用いて、これらの時間帯における主要政治家の役割を示している。色は民主党: 青、共和党: 赤、無党派層: 紫。

図11(a)は、Rogriguezの論文[79]の元のグラフを再現するために描かれたものである。⁵これは完全なネットワークであり、すべてのプレーヤーとそれぞれのバインディング・タイが含まれている。ロドリゲスはネットワークの中心核を強調するためにこのグラフィックを作成したが、彼がマドリッド爆弾テロに関与した「最も中心的な3人のプレーヤー」と呼ぶ人物を特定するのは難しい。

図11 (b) に示すように、収縮と反発のバーの数を選択すると、グラフは論文で提供された分析と密接に対応する。ジャマール・ズーガム、モハメド・シャウイ、サイード・ベラクの3人が、フィールド・オペレーション・グループのネットワークで最も中心的なプレーヤーであり、3人ともグラフの中心にいる。ズーガムの下には、アブ・デルアヒド・ベラクがあり、彼はフィールド・オペレーション・グループに新しいメンバーをリクルートする役割を担っていたと疑われている。ロドリゲスの分析で強調されたもう一人の中心人物は、"化学者"として知られるアブデラヒム・ズバフである。オリジナルのレイアウトでは、彼のノードは目立たない。しかし、ロドリゲスはズバフを、さまざまなテロリストグループの面識のないメンバーを結

特に、共に投票した。バイは、共和党と最も協調的であった民主党議員である。図12(d)の2008年の反対投票グラフは、この選挙の政治性を浮き彫りにしている。一方では共和党がオバマ大統領に対抗していた。一方、民主党はマケインとチェイニー副大統領に対抗することに力を注いだ。チェイニー副大統領に対する投票は、議会の仕組みによるものであろう。マケインに対する投票は、彼が共和党の候補者であったため、あてはまる。

7 ディスカッション

びつけることによって「ネットワークを強固にした」と評している。他の重要人物としては、アメール・アジズとイマッド・エディン・バラカットがいる。

6.4.3 2007年と2008年の上院共闘と反対投票

米国上院2007年と2008年のデータセットは、101個のノード（100人の上院議員と副大統領）と5,048個のエッジを持つ完全な共同投票と反対投票のグラフである。このデータセットは、GovTrack [1]が提供する投票記録を使用して作成されており、2人の上院議員が一緒に投票する（または互いに反対する）頻度に対応する重みが付けられている。

図12(a)の2007年の共同投票グラフは、民主党は民主党に、共和党は共和党に投票するという通常の党派分断を示している。中道」グループの中で際立っているのは3人である: スノー、コリンズ、スペクターである。スノーとコリンズは党派を超えて投票することで知られている。スペクターは2009年に共和党から民主党に鞍替えした。図12 (b) の2007年反対投票グラフは、各人が誰に反対票を投じる可能性が最も高いかを浮き彫りにしている。このグラフは、反対党の個人を分離した4つのクラスターを示している: デミント、ドッド、バイデン、オバマである。

2008年は大統領選挙の真っ最中で、図12 (c) の共同投票グラフは党派性が支配していたことを示している。共和党

⁵ 元々は何らかの形で力任せのレイアウトが使われていたと思われるが、その情報は文書化されていない。

他の力指向レイアウトアルゴリズムへの移植性。F-R力指向レイアウトを参照レイアウトとして使用しました。私たちの修正は反発力とバネ力を追加するだけなので、sfdp [47]やFM³[39]のような他の力指向レイアウトのバリエーションは、私たちのアプローチをそのアルゴリズムに適応させることができ、私たちが実証したものと同様の効果を見ることができるはずです。

階層的クラスタリングとの関係。0次元のPH特徴を計算することは、階層的クラスターを見つけることと強い関係がある。しかし、主な差別化要因は、PHにおける特徴の扱いである。例えば、PHは重みの小さいエッジを「ノイズ」と見なし、したがってそれらを折りたたむことは理にかなっている。同時に、PHは重みの大きいエッジをシグナルと見なし、したがってそのような特徴を分離することは意味がある。

(半) 自動的な特徴の選択。5.4節で説明した手順を踏まえると、(半) 自動的なヒューリスティックに基づいたアプローチで収縮と反発を初期化することが想像できる。我々は2つの理由から、そのようなアプローチは採用しなかった。第一に、そのようなアプローチによって節約できる相互作用の量はそれほど大きくないだろう。私たちの経験では、収縮と反発のオプションの探索には数分しかかからない。第二に、より重要なことは、相互作用の過程でグラフに関する直感が得られることである。この直感は、最終的なグラフレイアウトを選択する上で重要である。

切断グラフ。切断されたグラフの場合、PHの計算は、切断された成分は決して結合しないことを認識することにより、何の修正も加えることなく機能する。その後、我々のアプローチは、グラフの各連結成分を個別に考慮する。

限界。私たちのアプローチに限界がないわけではない。第一の問題点は、力指向のレイアウトは、その性質上、過度の制約を受けるとのことだ。追加の力を加えることは、この問題を悪化させる可能性があります。このことは、収縮効果に否定的な結果をもたらす多くの例に見られる（6.2節参照）。第二に、レイアウトの選択的エンジニアリングの可能性がある。これは、最終的なレイアウトにおいて、重要なPHの特徴を意図的に無視することにつながる可能性がある。次に、最悪の場合、必要な相互作用の数が非常に多くなる可能性がある。可能な限りの収縮と

反発には $2|N|$ 個の相互作用が必要であり、組み合わせをテストするコストがかかることは言うまでもない。幸いなことに、実際の相互作用過程は

は、高品質な結果を得るためにかなり高速である（5.4節参照）。最後に、我々のアプローチは、入力グラフに対して一意なMSTが存在することを仮定している。そのような仮定が成立しない場合（例えば6.4.2節）、MSTの集合から任意の木が

選択される。様々な制約を用いて最適なMSTを選択することは、今後の研究に値する。

8 結論

われわれは、PHを用いて、力指令レイアウトを対話的に修正するグラフ描画の新しいアプローチを発表した。このアプローチは、PHによって定義されたグラフの特徴を強調するレイアウトを選択するための柔軟なインタフェースを提供する。将来的には、グラフ描画を制御するために、より高次元のPH特徴を使用する可能性に注目したい。例えば、1次元のPH特徴は、メトリック空間内のトンネルを符号化し、グラフ内の類似構造を制約するのに有用であろう。しかし、より高次元のPH特徴では、これらの特徴を生成する点の集合を見つけることは非自明な問題である。

謝辞

貴重なご意見をいただいた査読者に感謝する。本研究の一部は、全米科学財団の助成金IIS-1513616および

DBI-1661375, CRA-W Collaborative Research Experiences for Undergraduates (CREU) program, DARPA CHES FA8750-19-C-0002, and an NVIDIA Academic Hardware Grant.

参考文献

- [1] <https://www.govtrack.us>. アクセス: 2019-03-01.
- [2] <https://processing.org/>. アクセス: 2019-03-01.
- [3] D.D.アルシャンボー、T.マンズナー、D.オーベル。Grouse: 特徴に基づく、操縦可能なグラフ階層探索。In *Proceedings of Eurographics / IEEE VGTC Symposium on Visualization*, vol. 2007, pp. 67-74, 2007.
- [4] D.D.アルシャンボー、T.マンズナー、D.オーベル。Topolayout: トポロジーの特徴による多階層グラフィックレイアウト。*IEEE transactions on visualization and computer graphics*, 13(2), 2007.
- [5] D.D.アルシャンボー、T.マンズナー、D.オーベル。Grouseflocks: グラフ階層空間のステアラブル探索。*IEEE Transactions on Visualization and Computer Graphics*, 14(4):900-913, 2008.
- [6] B.Bach, N. H. Riche, C. Hurter, K. Marriott, and T. Dwyer. Towards unambiguous edge bundling: ネットワーク可視化のためのコンフルエント描画の調査。*IEEE transactions on visualization and computer graphics*, 23(1):541-550, 2017.
- [7] M.Bampasidou and T. Gentimis. *arXiv 電子版*, arXiv:1403.5346, 2014.
- [8] J.Barnes and P. Hut. A hierarchical $O(N \log N)$ force-calculation algorithm. *Nature*, 324(6096):446, 1986.
- [9] M.Bastian, S. Heymann, M. Jacomy. Gephi: ネットワークの探索と操作のためのオープンソースソフトウェア。In *International Conference on Web and Social Media*, pp. 361-362, 2009.
- [10] K.ボーナー、R.クラヴァンス、M.パテック、A.M.ソズ、J.R.ビバースティン、R.P.ライト、V.Larivie`re, and K. W. Boyack. 分類の設計と更新 システム: ucsd map of science. *PLOS One*, 7(7):e39464, 2012.
- [11] U.Brandes and C. Pich. 大規模データの漸進的多面スケーリングのための固有値解法。 *グラフ描画*, pp.Springer, 2007.
- [12] C.C. J. Carstens and K. J. Horadam. コラボレーション ネットワークの永続的ホモロジー. *工学の数学的問題*, 2013, 2013.
- [13] B.キャンディ、C.ライ、V.ソロ。脳活動: Conditional dissimilarity and persistent homology. *IEEE 12th International Symposium on Biomedical Imaging*, pp.1356-1359, 2015.
- [14] S. グラフ可視化のための階層的 クラスターリング. Hierarchical clustering for graph visualization. *arXiv e-prints*, arXiv:1210.5693, 2012.
- [15] C.C.コリンズ、G.ペン、S.カーペンデル。バブルセット: 既存のビジュアライゼーション上のアイソコントロールを使って集合の関係を明らかにする。 *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1009-1016, 2009.
- [16] T.H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *アルゴリズム入門*. MIT press, 2009.
- [17] Y.Dabaghian, F. Me`moli, L. Frank, and G. Carlsson. A topological paradigm for hippocampal spatial map formation using persistent homology. *PLOS Computational Biology*, 8(8):e1002581, 2012.
- [18] T.A.デイビス、Y.フォー。フロリダ大学疎行列コレクション。 *ACM Transactions on Mathematical Software*, 38(1):1, 2011.
- [19] E.W. ダイクストラグラフに関する2つの問題についてのノート. *Numerische Mathematik*, 1:269-271, 1959.
- [20] K.Dinkla, M. J. Van Kreveld, B. Speckmann, and M. A. Westenberg. ケルプ・ダイアグラム: 点集合のメンバーシップの可視化。 *Computer Graphics Forum*, 31(3pt1):875-884, 2012.
- [21] K.Dinkla, M. A. Westenberg, and J. J. van Wijk. 圧縮隣接行列: 遺伝子制御ネットワークを紐解く。 *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2457-2466, 2012.
- [22] I.Donato, G. Petri, M. Scolamiero, L. Rondoni, F. Vaccarino. 高速状態と弱いノードの決定: 持続的ホモロジーによる位相的变化。 *欧州複雑システム会議議事録*, pp.295-301, 2012.
- [23] C.Dunne and B. Shneiderman. 可読性メトリクスを組み込むことによるグラフ描画の可読性の向上: ネットワークアナリストのためのソフトウェアツール。 Technical report, University of Maryland, HCIL Tech Report HCIL-2009- 13, 2009.
- [24] C.Dunne and B. Shneiderman. モチーフの単純化。 *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013.
- [25] T.Dwyer, N. H. Riche, K. Marriott, and C. Mears. 密な有向グラフの可視化のためのエッジ圧縮技術。 *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2596-2605, 2013.

- [26] W.E. J. Lu, and Y. Yao. The landscape of complex networks. *arXiv e-prints*, arXiv:1204.6376, 2012.
- [27] H. Edelsbrunner and J. Harer. Persistent homology-a survey. *Contemporary Mathematics*, 453:257-282, 2008.
- [28] H.H. エーデルスブルナー、J. ハラー 計算トポロジー：入門. American Mathematical Society, Providence, RI, USA, 2010.
- [29] J. Ellson、E. Gansner、L. Koutsofios、S. C. North、G. Woodhull. Graphviz-open source graph drawing tools. In *International Symposium on Graph Drawing*, pp. Springer, 2001.
- [30] M. Fruchlich and M. Werner. 対話的グラフ可視化システム da vinci のデモ. In *International Symposium on Graph Drawing*, pp. 266-269. Springer, 1994.
- [31] T.M. Fruchterman and E. M. Reingold. 力方向配置によるグラフ描画. *Software: Practice and experience*, 21(11):1129-1164, 1991.
- [32] E.R. Gansner, Y. Hu, S. North, and C. Scheidegger. グラフを可視化するための多階層エッジバンドル. *IEEE Pacific Visualization Symposium*, pp. 187-194, 2011.
- [33] E.R. Gansner, Y. Koren, and S. North. Graph drawing by stress majorization. *グラフ描画*, pp. 239-250. Springer, 2005.
- [34] E.R. ガンスナー、E. コウツォフィオス、S.C. ノース、K.-P. Vo. 有向グラフを描くためのテクニック. *IEEE Transactions on Software Engineering*, 19(3):214-230, 1993.
- [35] R. グリストバーコード：データの永続的トポロジー. *Bulletin of the American Mathematical Society*, 45:61-75, 2008.
- [36] H. ギブソン、J. フェイス、P. ヴィッカーズ. 情報可視化のための2次元グラフィックアウト技術のサーベイ. *Information Visualization*, 12:324-357, 07 2012.
- [37] M. Girvan and M. E. Newman. 社会的・二論理的ネットワークにおけるコミュニティ構造. *Proceedings of the national academy of sciences*, 99(12):7821-7826, 2002.
- [38] S.W. ゴロムパズル、パターン、問題、そしてパッキング. Princeton University Press, 1996.
- [39] S. Hachul and M. Jünger. ポテンシャル場に基づくマルチレベルアルゴリズムによる大規模グラフの描画. In *International Symposium on Graph Drawing*, pp. 285-295. Springer, 2004.
- [40] A. Hagberg, P. Swart, and D. S. Chult. Networkx を用いたネットワーク構造、ダイナミクス、機能の探索. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.
- [41] M. Hajij, B. Wang, C. Scheidegger, and P. Rosen. 永続的ホモロジーを用いた時変グラフの構造変化の視覚的検出. In *IEEE Pacific Visualization Symposium*, 2018.
- [42] D. ハンセン、B. シュナイダーマン、M.A. スミス. *NodeXL によるソーシャル・メディア・ネットワークの分析：つながった世界からの洞察*. Morgan Kaufmann, 2010.
- [43] T.R. ヘンリー、S.E. ハドソン. 対話的グラフィックアウト. *ユーザインタフェース・ソフトウェアと技術に関する第4回ACM年次シンポジウム予稿集*, pp. 55-64. ACM, 1991.
- [44] I. Herman, G. Melanc, on, and M. S. Marshall. 情報視覚化におけるグラフの視覚化とナビゲーション：サーベイ. *IEEE Transactions on Visualization and Computer Graphics*, 6(1):24-43, 2000.
- [45] D. Holten and J. J. Van Wijk. グラフ可視化のためのエッジの強制バンドル. *Computer Graphics Forum*, 28(3):983-990, 2009.
- [46] D. Horak, S. Maletic, and M. Rajkovic. 複雑なネットワークの永続的ホモロジー. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(03):P03034, 2009.
- [47] Y. 胡. 効率的で高品質な力方向グラフ描画. *Mathematica Journal*, 10(1):37-71, 2005.
- [48] C. Hurter, A. Telea, O. Ersoy. Moleview: 大規模な要素ベースのプロットのための属性と構造に基づいたセマンティックレンズ. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2600-2609, 2011.
- [49] P. Jaccard. 植物分布の比較調査. *des alpes et des jura. Bull Soc Vaudoise Sci Nat*, 37:547-579, 1901.
- [50] S. Kairam, D. MacLean, M. Savva, and J. Heer. GraphPrism: ネットワーク構造のコンパクトな可視化. *Advanced Visual Interfaces*, 2012.
- [51] D.A. ケイムピクセル指向の視覚化技術の設計：理論と応用. *IEEE Transactions on Visualization and Computer Graphics*, 6(1):59-78, 2000.
- [52] D.A. Keim, J. Schneidewind, and M. Sips. スケーラブルなピクセルベースの視覚データ探索. *Pixelization Paradigm, Lecture Notes in Computer Science*, 4370:12-24, 2007.
- [53] M.M. Khoury, Y. Hu, S. Krishnan, C. Scheidegger. 低ランクストレスメジャライゼーションによる大規模グラフの描画. *Computer Graphics Forum*, 31(3pt1):975-984, 2012.

- [54] D.E. Knuth. *The Stanford GraphBase: a platform for combinatorial computing*, vol.37. Addison-Wesley Reading, 1993.
- [55] Y. コレン. スペクトルグラフ描画について. *計算と組合せ論*, pp.496-508. Springer, 2003.
- [56] Y. Koren, L. Carmel, D. Harel. Ace: 巨大グラフ描画のための高速マルチスケール固有ベクトル計算. *IEEE Symposium on Information Visualization*, pp. 137-144, 2002.
- [57] J.B. クルスカール. グラフの最短スパン部分木と巡回セールスマン問題について. *Proceedings of the American Mathematical Society*, 7(1):48-50, 1956.
- [58] H. Lee, M. K. Chung, H. Kang, B.-N. Kim, D. S. Lee. Kim, and D. S. Lee. Graph filtration and gromov-hausdorff metric using brain networks shape. *International Conference on Medical Image Computing and Computer Assisted Intervention*, pp.302-309, 2011.
- [59] H. Lee, M. K. Chung, H. Kang, B.-N. Kim, D. S. Lee. Kim, and D. S. Lee. 脳ネットワークの識別的持続ホモロジー. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp.841-844, 2011.
- [60] H. Lee, H. Kang, M. K. Chung, B.-N. Kim, and D. S. Lee. デンドログラムから見た永続的脳ネットワークホモロジー. *IEEE Transactions on Medical Imaging*, 31(12):2267-2277, 2012.
- [61] H. Lee, H. Kang, M. K. Chung, B.-N. Kim, and D. S. Lee. ネットワークフィルトレーションによる重み付き脳機能ネットワークモデリング. *NIPS Workshop on Algebraic Topology and Machine Learning*, 2012.
- [62] X. Lou and J. A. Suykens. 加重ネットワークにおける同期によるコミュニティの発見. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 21(4):043116, 2011.
- [63] F. McGee and J. Dingliana. グラフのエッジバンドルがユーザーの理解に与える影響に関する実証的研究. グラフのエッジバンドルがユーザーの理解に与える影響に関する実証的研究. ACM, 2012.
- [64] C. Muehler and K.-L. Ma. 空間充填曲線を用いた高速グラフィックアウト. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1301-1308, 2008.
- [65] M. Newman. ネットワーク科学者間のコラボレーション <http://www-personal.umich.edu/~mejn/centrality/poster.pdf>. Accessed: 2019-03-01.
- [66] M.E. ニューマン. 自己中心のネットワークと波及効果. *Social Networks*, 25(1):83-95, 2003.
- [67] M.E. ニューマン. 複雑なネットワークの構造と機能. *SIAM review*, 45(2):167-256, 2003.
- [68] A. Noack. グラフクラスタリングのエネルギーモデル. *Journal of Graph Algorithms and Applications*, 11(2):453-480, 2007.
- [69] A. Noack and R. Rotta. モジュラリティクラスタリングのためのマルチレベルアルゴリズム. *実験的アルゴリズムに関する国際シンポジウム*, pp.257-268. Springer, 2009.
- [70] A. ノカジ, M. オルトマン, U. ブランドス. 毛玉をほどく. In *Graph Drawing*, pp. Springer, 2014.
- [71] A. Nocaj, M. Ortmann, and U. Brandes. Adaptive disentanglement based on local clustering in small-world network visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(6):1662-1671, 2016.
- [72] S.C. ノース. ノートでグラフを描く. *NEATO ユーザーマニュアル*, 11:1, 2004.
- [73] D.D. Oelke, H. Janetzko, S. Simon, K. Neuhaus, and D. A. Keim. ピクセルベースの視覚化におけるVisual boosting. *Computer Graphics Forum*, 30(3):871-880, 2011.
- [74] G. Petri, M. Scolamiero, I. Donato, and F. Vaccarino. ネットワークとサイクル: 複雑ネットワークへの持続的ホモロジー・アプローチ. *Proceedings of the European Conference on Complex Systems 2012, Springer Proceedings in Complexity*, pp. 93-99, 2013.
- [75] G. Petri, M. Scolamiero, I. Donato, and F. Vaccarino. Topological strata of weighted complex networks. *PLOS One*, 8(6):e66506, 2013.
- [76] V. Pirino, E. Riccomagno, S. Martinoia, and P. Massobrio. A topological study of repetitive coactivation networks in in vitro cortical assemblies. *物理生物学*, 12(1), 2015.
- [77] H.C. Purchase. グラフ描画の美学のための尺度. *Journal of Visual Languages & Computing*, 13(5):501-516, 2002.
- [78] B. Rieck, U. Fugacci, J. Lukasczyk, H. Leitte. 関コミュニティの持続性: 複雑なネットワークのための位相的視覚分析アプローチ. *IEEE Transactions on Visualization and Computer Graphics*, 2017.
- [79] J.A. ロドリゲス. 3月11日のテロネットワーク: その弱さの中に強さがある、2005年. データは <http://moreno.ss.uci.edu/data> .
- [80] K. Ryall, J. Marks, and S. Shieber. グラフを描くための対話的な制約ベースのシステム. 第10回ACMシンポジウム予稿集

ユーザインタフェース・ソフトウェアと技術, pp.97-104.ACM, 1997.

- [81] M.Sarkar and M. H. Brown.Graphical fisheye views of graphs.*SIGCHI Conference of the Human factors in computing systems*, pp.83-91.ACM, 1992.
- [82] S.E. Schaeffer. グラフクラスティング。 *Computer Science Review*, 1(1):27-64, 2007.
- [83] D.Selassie, B. Heller, and J. Heer.方向性ネットワークデータのための分割エッジバンドル。 *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2354-2363, 2011.
- [84] C.C.トミンスキー、J.アベロ、F.ヴァン・ハム、H.シューマン。グラフ可視化のための魚眼ツリービューとレンズ。第10回IEEE 情報可視化国際会議論文集, pp.17-24, 2006.
- [85] A.L. Traud、E. D. Kelsic、P. J. Mucha、M. A. Porter。オンライン大学ソーシャルネットワークにおけるコミュニティ構造と特性の比較。 *SIAM review*, 53(3):526-543, 2011.
- [86] W.T. トウツテグラフの描き方。*Proceedings of the London Mathematical Society*, s3-13(1):743-767, Jan 1963.
- [87] S.Van den Elzen and J. J. Van Wijk.多変量ネットワークの探索と表示: 選択と集約を介した詳細から概要へ。 *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2310- 2319, 2014.
- [88] T.フォン・ランデスベルガー、A.クイパー、T.シュレック、J.コーンルハンマー、J.J.ファン・ヴァイク、J.-D.フェケテ、D.W.フェルナー。Fekete, and D. W. Fellner.大規模グラフの視覚分析: 最先端技術と今後の研究課題。*Computer Graphics Forum*, 30(6):1719-1749, 2011.
- [89] Y.Wang、Y. Wang、Y. Sun、L. Zhu、K. Lu、C.-W.Fu、M. Sedlmair、O.Deussen, and B. Chen.Revisiting stress majorization as a unified framework for interactive constrained graph visualization.*IEEE Transactions on Visualization and Computer Graphics*, 24(1):489-499, 2018.
- [90] Y.Wang, Y. Wang, H. Zhang, Y. Sun, C.-W. Fu, M. Sedlmair, B. Chen, and O. Deussen.Fu、M. Sedlmair、B. Chen、O. Deussen。効率的な大規模グラフ探索のための構造を考慮した魚眼ビュー *IEEE Transaction on Visualization and Computer Graphics*, 25(1):566-575, 2019.
- [91] N.ウォン、S・カーペンデール、S・グリーンバーグ。Edgelens: グラフのエッジの輻輳を管理するための対話的手法。 *IEEE Symposium on Information Visualization 2003 (IEEE Cat. No. 03TH8714)*, pp.IEEE, 2003.