

スケルトン抽出による動的グラフレイアウト手法

平野未来[†] 白山晋^{††}

A Graph Layout of Dynamic Networks Using a Skeletal Extraction

MIKU HIRANO[†] SUSUMU SHIRAYAMA^{††}

動的に変化するネットワークの大きな構造から局所的な構造までを直感的に把握するためのネットワーク可視化が模索されている。そのような可視化のためには、コミュニティ構造の表出に加え、構造の変化が小さい場合にはメンタルマップを保持し、変化が大きい場合にはそれに追従するようなグラフレイアウト手法が必要になる。本稿では、我々が提案しているスケルトン抽出法を、動的ネットワークのレイアウトに適した形に拡張し、動的なネットワーク構造を端的に表現できるグラフレイアウト手法を提案する。

1. はじめに

現実世界の様々なネットワークにおいて、スモールワールド性[WS98]、スケールフリー性[BA99]といった性質が見いだされ、ネットワークの性質が社会現象、経済現象、あるいは自然現象に与える影響が注目されている[Mas05]。

そのようなネットワークは、一般的には、ネットワークの統計的指標と注目領域の局所的リンク構造の両者にコミュニティ構造を加えることで、全体像が示されることが多い[BLMCH06]。しかし、それらの3つの構造を組み合わせただけでは、大きな構造から局所的な構造を直感的に把握することは難しい。

このため、統計的指標以外でネットワークの特徴を調べる方法が模索されてきた。その代表がネットワーク可視化である。ネットワーク可視化は、ノードとエッジの配置であるグラフレイアウトと属性値などに従った色付けに分けられる。古くから多くの検討がなされ、適用事例も多く、その有用性も示されている[Sug93][SS06]。しかしながら、ほとんどが、ノードやエッジの構造が変化しない静的ネットワークを対象にしており、動的に構造が変化するものを対象にしたものは少ない。

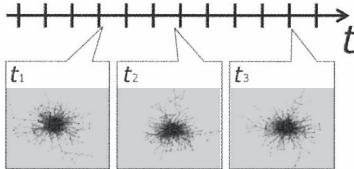


図1 動的ネットワークの可視化例

動的ネットワークとは、ノードの増加、減少、あるいはエッジが増加、減少することにより、ネットワーク構造が動的に変化するネットワークを指す[Bra01]。現実世界のネットワークは動的であることが多い。動的ネットワークは、キーグラフのようにパラメータの変化にともなうエッジの増減によっても現れるが、多くは時間変化の中で生じるものである。本研究では断りの無い限り、“動的”を“時間変化”と同義として扱う。

動的ネットワークに対する可視化は、一般的には構造が時間的に変化する動的ネットワークにおける、各時刻のネットワークをそれぞれ静的ネットワークと見なし、各時刻においてグラフレイアウトを行い、色付けをするという方法になる(図1)。

動的ネットワークに対して、既存の静的ネットワークの可視化を適用するという方法には、静的ネットワークにおける課題に、主として二つの課題が加わる。

一つは、ネットワーク構造がほとんど変化していない場合でも、各時刻の可視化結果が大きく変化することである。これは、各時刻において最適なグラフレイアウトを行うために生じる。この問題は、前後の時刻におけるネットワークとの差分を考慮した可視化によって解決が試みられている。代表的な方法として、Foresighted Layout[DGK01]、Aggregate-Graph Approach [CKNPW03]、WebRelievo [TK05]、Anchored Map に基づく方法[Mis07]などがある。

Foresighted Layout は、時系列的なグラフの和集合であるスーパーグラフを予めレイアウトし、各時間のグラフレイアウトを、スーパーグラフに基づいて決定する手法である。Aggregate-Graph Approach と WebRelievo は、各時刻のグラフをレイヤーとして保存しておき、各レイヤーを力学的手法でレイアウトする際に、時系列上の前後のレイヤーを参考にして、グラフの構造が変化しても同じノードがほぼ同じ位置を保持するよう補正する手法である。

Anchored Map は、時間的に変化しない、あるいは重要なノードの位置を固定し、その他を静的ネットワークに対するグラフレイアウトによって配置するというものである。いずれの方法も、ユーザが動的ネットワークに対する一連の可視化結果をみたときに、変化が認知しやすい(ユーザにとつての認知的負荷が小さい)ようにレイアウトを行うという方針に基づくものである。あるいは、ユーザが生来的に有すると仮定されるグラフレイアウトのイメージ(Mental map) [DGK01]を考慮した可視化手法になっている。

これらの手法を利用すれば、ネットワークの構造が変化したとしても、各時刻の可視化結果が大きく変わらないため、ユーザに与える認知的負荷が少ないと考えられる。

動的ネットワークに対する可視化のもう一つの課題は、ネットワーク構造が変化した場合、その変化を可視化結果に反映することが難しいというものである。

Foresighted Layout は時刻毎の統計的指標の変化を加味することは原理的に難しい。Aggregate-Graph Approach と WebRelievo では、補正の際に変化を考慮できるが大きく変化する場合を扱うことは難しい。Anchored Map に基づく方法では、固定されるノードの選択時にネットワークの統計的指標は利用できるが、その変化を十分に考慮しているわけではない。

また、動的ネットワーク可視化を対象とする既存研究では、静的ネットワーク可視化の問題を解決しているわけではない。特に、ネットワークが大規模化すると、可視化結果が煩雑になり、大局的構造、コミュニティ構造、局所構造を捉えることが難しくなる。

[†] 東京大学大学院工学系研究科環境海洋工学専攻
Department of Environmental and Ocean Engineering, School of Engineering, The University of Tokyo
^{††} 東京大学人工物工学センター
Research into Artifacts, Center for Engineering, The University of Tokyo

本研究では、ネットワーク構造の変化が可視化結果に反映されないという点に着目し、構造の変化が小さい場合には、可視化結果も大きく変化せず、構造の変化が大きい場合にその変化を直感的に捉えることができ、大規模ネットワークに対応できる動的可視化手法を提案する。この際に、先行研究において、提案したスケルトン抽出法[HS08]を利用する。

2. 提案手法

変数、指標、記号は著者らの先行研究[HW08]と同じである。ここでは特に本稿で必要になるものを記しておく。

ノードを v_i 。あるいは識別子のみで示す (例えば、ノード i)。ノードの総数を n 、エッジの総数を m とする。ノード v_i の次数を k_i とする。隣接行列を A 、その成分を a_{ij} とし、

$$a_{ij} = \begin{cases} 1 & i \text{ から } j \text{ へのエッジがあるとき} \\ 0 & i \text{ と } j \text{ にエッジがないとき} \end{cases} \quad (1)$$

とする。

グラフ距離において d_{ij} をノード v_i と v_j の間の最短経路長とする。近接中心性 (CC) を、

$$C_c(v_i) = (n-1) / \sum_{j \in V, j \neq i} d_{ij} \quad (2)$$

とする。エッジに対する媒介中心性 (EBC) を次式とする。

$$C_e(e) = \sum_{i \neq j \in V} \frac{\rho_{ij}(e)}{\sigma_{ij}} \quad (3)$$

ノードに対する媒介中心性 (NBC) を次式とする。

$$C_v(v) = \sum_{i \neq v \neq j \in V} \frac{\sigma_{ij}(v)}{\sigma_{ij}} \quad (4)$$

2.1 スケルトン抽出

静的、動的を問わず、グラフィケアウトには、ネットワークの大規模化にともないレイアウト結果が煩雑になり、大局的構造、コミュニティ構造、局所構造を捉えることが難しくなるという欠点がある。これを解決する一つの方法として、スケルトン抽出法が注目されている。

Kim らはエッジの重みとして EBC を与え、クルスカルのアルゴリズム (Kruskal's algorithm) によって最小全域木 (MST) を抽出し、それをスケルトンと呼び、スケルトンによって元のネットワークの概略を示した [KNJ04]。得られる MST は EBC の大きなエッジが削除されたものになる。van Ham らは、この MST がコミュニティ構造を反映したものになることを後述する媒介測度 (bm: betweenness metric) によって示している [HW08]。

Kim ら、van Ham らのスケルトン抽出法はネットワークの直感的な把握に有効な方法であるが、いくつかの問題がある。一つは同じ EBC の値を持つ複数のエッジが存在する場合、MST が一意に定まらないというものである。

先行研究では、任意性が少なく、コミュニティ構造の抽出に適した手法を提案した [HS08]。本稿では、その手法から得られるスケルトンを用いて動的ネットワークに対するレイアウトを行う。

2.2 コミュニティを利用しないスケルトン抽出法

2.2.1 最小全域木の構成

与えられたネットワークに対して、EBC を算出し、それを重みとしてエッジに与える。Kim らの方法と同様に、EBC によって重みづけられたネットワークから最小全域木 (MST) を抽出する。ただし、我々の手法では、プリムのアルゴリズム (Prim's algorithm) によって連結性を保ちながら最小全域木を求める。

2.2.2 ルートノード選択

本稿では、スケルトン抽出に媒介中心性という考えを用いている。この点では、視覚的認知特性を考慮し、周りのノードの最短経路上にルートが存在するように、NBC が最大のノードをルートとすることが考えられる。

一方、表示空間では、物理空間での距離という概念が発生し、次数がリンクの集中度合いを意味するようになる。そして、全体がコンパクトに表示されることが望ましい場合も多い。この場合は、適当なルートから MST を構成し、抽出されたスケルトンに対して近接中心性 CC を計算し、その最大ノードをルートにすることや、最大次数のノードをルートにすることが考えられる。

ルートノードの候補が複数ある存在する場合は、候補の中で他の統計的指標が最も大きなものをルートとする。この処理でルートが一意に選択できない場合、候補をすべて表現し、ユーザが選ぶものとする。

2.2.3 エッジ選択

ネットワークのすべてのエッジがすべて異なる EBC をもつことは稀である。そのため、MST の抽出の際に、エッジ選択の任意性が生じる。

選択の対象となるエッジの端点ノードに対する統計的指標を用いて選択するエッジを決定する。

本稿では、ノードの属性として、 $1/NBC$ 、近接中心性、次数をこの優先順位で用いることにする。なお、次数 1 のノードは NBC の値が 0 になり、必ず全域木に含まれることを付記しておく。

また、これらの値が同じ場合、候補をすべて表現する。あるいは、2パス以上離れたエッジの属性値を用いて一意になるようにエッジを定める。

2.3 コミュニティを利用したスケルトン抽出法

Kim らが原理を述べ、van Ham らがいくつかのネットワークを用いて明らかにしたように、EBC によって重み付けられたネットワークに対するスケルトンは Newman らのコミュニティ構造を表出するものでもある。しかしながら、Girvan と Newman の切断法 [GN02] のように、切断を繰り返しながら EBC を求めるわけではないのでコミュニティ構造が明確に表出される保証はない。そこで、前もってコミュニティを抽出し、その情報を用いてスケルトンを抽出する方法を考察する。

2.3.1 コミュニティグラフの生成

はじめに、Clauset らのアルゴリズムによる Newman コミュニティ抽出法 [CNM04] によってコミュニティを抽出し、コミュニティに番号を振る。番号を α とし、コミュニティを C_α で示す。

次に、ノードの属性に、属するコミュニティ番号を与える (図 2 左では色によって示している)。図 2 右のように、コミュニティをノードとするネットワークを作成する。エッジの重みは、 $e_{\alpha\beta} = (\sum_{i \in \alpha} \sum_{j \in \beta} a_{ij}) / 2m$ とする。

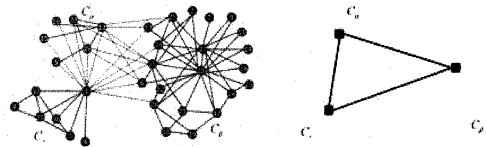


図2 コミュニティ抽出 (左) とコミュニティグラフ (右)

2.3.2 最小全域木の構成

コミュニティグラフに対して最小全域木(MST)を抽出する。この場合、 $e_{\alpha\beta}$ の小さなエッジが選択される。ルートノード選択、エッジ選択は、2.2節と同様の方法で行う。

2.3.3 コミュニティ間のエッジ選択

コミュニティ間に複数のエッジが存在する場合はいずれかのエッジを選択する。

はじめに、コミュニティ α に属するノード i とコミュニティ β に属するノード j の属性値として次式の P_{ij} を与える。

$$P_{ij} = a_{ij} \frac{k_{i\beta} k_{j\alpha}}{k_i k_j} \quad (5)$$

ここで、 $k_{i\beta}$ はノード i のもつエッジの中でコミュニティ β に繋がるエッジの数である。

ノード間にエッジがあり、他方のコミュニティに繋がるエッジの比率が大きいノード同士を選べば P_{ij} の値が大きくなる。本稿では、 P_{ij} が最大となるエッジを選択する。

2.4 評価基準

下記項目で、抽出されたスケルトンの評価を行う。

2.4.1 最小全域木の総数

元のネットワークを重み無し無向グラフとすると、キルヒホッフの行列木定理によって全域木の数が求められる。その値を $\tau_1(G)$ とする。また、Kimらとvan Hamらの最小全域木の数を $\tau_2(G)$ とする。コミュニティを利用しないスケルトン抽出法で得られる最小全域木の数を $\tau_3(G)$ 、コミュニティを利用するスケルトン抽出法で得られる最小全域木の数を $\tau_4(G)$ とし、これらと比較することで任意性が少ないことを示す。

2.4.2 媒介測度

スケルトンにおけるコミュニティの明確さを、van Hamらが提案した媒介測度(b_m)によって数値化する。

両端のノードが異なるコミュニティに属するエッジの数を χ とする。コミュニティの数を c とすると、最小全域木において明確にコミュニティが分離している場合は、 $\chi = n - c$ が成立する[HW08]。 $n = c$ の場合を除き、媒介測度を、

$$b_m = \chi / (n - c) \quad (6)$$

とする。

2.5 グラフレイアウト

動的ネットワークに対して、2.2節、あるいは2.3節の方法によって、Skeleton animation ($S = [s_1, s_2, s_3, \dots, s_n]$)が求められる。ここで、ある s_i はある時刻 i のネットワークに対するスケルトンを示す。 $s_i = (V_i, E_i)$ であり、 V_i と E_i は時刻 t_i のノードとエッジである。

時刻毎に、Radial graph drawingによってスケルトン s_i をレイアウトし表現する。いくつかの試行実験によって、ネットワーク構造の変化が小さい場合、認知的連続性が保持されることがわかっている。

しかし、Radial Graph Layoutではルートノードが変化すると、全体としてはネットワーク構造の変化が小さくてもレイアウトが大きく変わる可能性がある。そのような場合や、ネットワーク構造の変化が大きい場合への対応が必要になる。そこで、Foresighted Layout[DGK01]のスーパーグラフに倣い、 S の和集合 \hat{S} を、

$$\hat{S} = (\hat{V}, \hat{E}), \quad \hat{V} = \cup_{i=1}^n V_i, \quad \hat{E} = \cup_{i=1}^n E_i$$

と定義する。これを、スーパースケルトン(Super skeleton)と名付ける。次に、適当な静的グラフレイアウト手法によってスーパースケルトンに対してレイアウトを行う。その後、 $S = [s_1, s_2, s_3, \dots, s_n]$ を表示する。

また、アンカーマップの考えも重要である。Skeleton animation に対して、 S の積集合 \hat{S} を、 $\hat{S} = (\hat{V}, \hat{E})$ 、 $\hat{V} = \cup_{i=1}^n V_i$ 、 $\hat{E} = \cup_{i=1}^n E_i$ とする。本稿のスケルトンは全域木のため、ノードに関しては元のネットワークと同じである。エッジは、元のネットワークのエッジよりもかなり少ないものになる。そこで、 \hat{E} を固定したレイアウトが考えられる。本稿では、 \hat{E} をアンカーエッジと呼び、アンカーエッジの両端のノードをアンカーノードとする。アンカーエッジとアンカーノードを利用したレイアウトを行う。

また、ノードの積集合 \hat{V} は、時系列の中で共通して存在するノードであるので、この中からルートノードを選択すれば、認知的連続性をより強く保持できるものと考えられる。本稿では、ノードの積集合 \hat{V} の中から2.2.2項で述べた方法によってルートノードを選択する。

ただし、一般には、 \hat{S} は非連結になるのでレイアウトには工夫が必要である。はじめに、連結成分を分離し、各連結成分に関して、その成分に含まれるノードとエッジの時刻歴を調べ、ノードとエッジが存在する時刻の順序付き集合を定義する。例えば、ある連結成分 γ に対する時刻の順序付き集合 T_γ は、

$$T_\gamma = \{i_1, i_1+1, \dots, i_1+n_1, i_2, i_2+1, \dots, i_2+n_2\}$$

のように形成される。次に、 S の部分集合 $S_\gamma = \{s_i | i \in T_\gamma\}$ を求める。 S_γ が占める表示空間の領域を決め、その領域内では、上述した S に対するレイアウト手法を用いることで、ノードとエッジの配置を決める。

3. 実験結果と考察

3.1 最小全域木の総数

カラテクラブネットワークに対して、各手法によって抽出したスケルトンの総数を求めた。その総数は、van Hamらの手法においては809個以上、コミュニティを利用する場合、しない場合の両提案手法においては2個となった。van Hamらのスケルトン抽出法と比較して99.7%の削減をすることができた。

3.2 媒介測度

各ネットワークから各手法によって抽出したスケルトンの媒介測度 b_m を測定した。van Hamらの手法とコミュニティを利用しないスケルトン抽出法の b_m は0.74であり、コミュニティを利用したスケルトン抽出法では理想値である1.0を得た。

コミュニティ構造を顕在化するという観点では、コミュニティを利用するスケルトン抽出法は理想的な手法であるといえる。

3.3 グラフレイアウト

図3に、動的ネットワークから時刻毎に抽出したスケルトンをレイアウトした結果を示す。図3上段は、ばねモデルによるレイアウト[Pjk]、中段はコミュニティを利用しないスケルトン抽出法、下段はコミュニティを利用するスケルトン抽出法によって抽出したスケルトンである。三手法共に左図が1008ステップ目、中図が1009ステップ目、右図が1010ステップ目のレイアウトである。1008ステップ目から1009ステップ目はネットワークの構造がほとんど変化せず、1009ステップ目から1010ステップ目までは構造が変化した。構造の変化に関わらず、ばねモデルによるレイアウト結果は大きく変化した。一方両提案手法共に、1008ステップと1009ステップのレイアウト結果はほとんど変化せず、1009ステップと1010ステップでは多少の変化が見られた。結果として、安定性や敏感性に優れたグラ

フレイアウトが可能であることを示せた。

また、図4はノード1000程度の静的ネットワークのレイアウト結果である。コミュニティを利用するスケルトン抽出法(右図)の方がコミュニティを利用しないスケルトン抽出法(左図)よりも、コミュニティ構造が明確に抽出されていることが分かる。

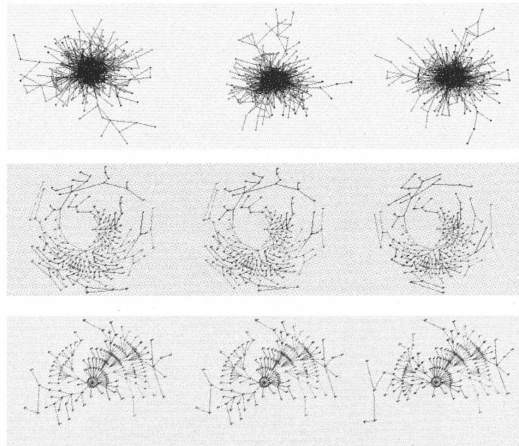


図3 動的ネットワークのレイアウト結果

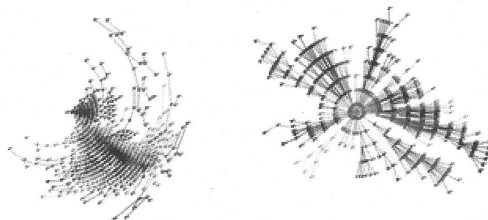


図4 静的ネットワークのレイアウト結果

4. まとめ

本研究では、ネットワークの大きな構造から局所的な構造をつなぐコミュニティ構造の顕在化が可能であり、かつ動的变化を適切に反映できるスケルトン抽出法を構築し、その方法を動的ネットワークに対するグラフレイアウトに適用した。得られた知見は以下のものである。

はじめに、動的ネットワークの可視化に対する要件として以下の三点を指摘した。

- (i) 表示空間におけるノードやエッジの位置関係や距離をネットワークの大きな構造から局所的な構造までを直感的に把握できるグラフレイアウト
- (ii) ネットワーク構造の変化が小さい場合は、認知的負荷が大きくなりやすいグラフレイアウト (安定性)
- (iii) ネットワーク構造が変化する場合には、その変化が可視化結果に反映されるグラフレイアウト (敏感性)

(i)に関しては、Kimら、van Hamらが、スケルトンと呼ぶ最小全域木を利用したグラフレイアウト手法を提案している。しかしながら、抽出されるスケルトンには大きな任意性があり、コミュニティ構造を十分に表出できない、また、(ii)と(iii)を考慮しておらず、動的ネットワークのグラフレイアウトには適さないことがわかった。

そこで、Kimら、van Hamらのスケルトン抽出法を改良

し、(i)から(iii)の要件を満たす以下の二つのスケルトン抽出手法を提案した。

- コミュニティを利用しないスケルトン抽出法
- コミュニティを利用するスケルトン抽出法

提案手法によりスケルトンを抽出した結果、両者ともに、スケルトンの任意性が大幅に減少することがわかった。また、グラフレイアウトの結果からは、コミュニティを利用する方法では、コミュニティ構造を十分に表出できることがわかった。

動的ネットワークに対しては、両者ともに、van Hamらの方法よりも安定性や敏感性に優れたグラフレイアウトが可能であることを示せた。また、コミュニティを利用しないスケルトン抽出法の方が、安定性や敏感性の点では有効であることがわかった。

参考文献

- [BA99] A. Barabási and R. Albert: Emergence of scaling in random networks, *Science*, Vol.286, No.5439, pp.509-512, 1999.
- [BLMCH06] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D. Hwang: Complex networks: Structure and dynamics, *Physics Reports*, Vol.424, No.4-5, pp.175 - 308, 2006.
- [Bra01] J. Branke: Dynamic Graph Drawing, In *Graph Drawing - Models and Algorithms*, pp.228-246, Springer-Verlag, 2001.
- [CKNPW03] C. Collberg, S. Kobourov, J. Nagra, J. Pitts and K. Wampler: A System for Graph-Based Visualization of the Evolution of Software, *Proceedings of the 2003 ACM Symposium on Software Visualization*, pp.77-86, 2003.
- [CNM04] A.Clauset, M.E.J. Newman, and C. Moore: Finding community structure in very large networks, *Physical Review E*, Vol.70, No.6, 66111, 2004.
- [DGK01] S. Diehl, C. Görg and A. Kerren: Preserving the Mental Map using Foresighted Layout, *Proceedings of Joint Eurographics - IEEE TCVG Symposium on Visualization*, pp.175-184, 2001.
- [GN02] M. Girvan and M. Newman: Community structure in social and biological networks, *Proceedings of the National Academy of Sciences*, Vol.99, No.12, pp.7821, 2002.
- [HS08] 平野未来, 白山晋: 静的ネットワークに対する新しいスケルトン抽出法, *情報処理学会研究報告(M P S)*, Vol.2008, No.126, pp.97-100, 2008.
- [HW08] F. van Ham and M. Wattenberg: Centrality Based Visualization of Small World Graphs, In *Proceedings of Eurographics*, Vol.27, No.3, pp.975-982, 2008.
- [KNJ04] D. Kim, J. Noh and H. Jeong: Scale-free trees: The skeletons of complex networks, *Physical Review E*, Vol.70, No. 4,46126, 2004.
- [Mas05] 増田直紀, 今野紀雄: *複雑ネットワークの科学*, 産業図書, 2005.
- [Mis07] K. Misue: Anchored Maps: Visualization Techniques for Drawing Bipartite Graphs, *Lecture Notes in Computer Science*, Vol.4551, pp.106, 2007.
- [Pjk] <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>
- [SS06] H.-J. Schulz and H. Schumann: Visualizing Graphs - A Generalized View, *Proc. of Information Visualization*, pp.166-173, 2006.
- [Sug93] 杉山公造: *グラフ自動描画法とその応用*, 計測自動制御学会学術図書, コロナ社, 1993.
- [TK05] M. Toyoda and M. Kitsuregawa: A System for Visualizing and Analyzing the Evolution of the Web with a Time Series of Graphs, *Proceedings of the 16th ACM conference on Hypertext and hypermedia*, pp.151-160, 2005.
- [WS98] D. Watts and S. Strogats: Collective dynamics of "small world" networks, *Nature*, Vol.393, pp.440-442, 1998.