



Title	動的信号領域を用いたSOMによるネットワークの可視化
Author(s)	岩田, 泰士; Yasushi, Iwata; 鈴木, 育男 他
Citation	情報処理学会論文誌, 50(11), 2752-2759
Issue Date	2009-11-15
Doc URL	https://hdl.handle.net/2115/50521
Rights	ここに掲載した著作物の利用に関する注意 本著作物の著作権は情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。
Type	journal article
File Information	IPJSJ-JNL5011016.pdf



動的信号領域を用いた SOM による ネットワークの可視化

岩田 泰士^{†1} 鈴木 育男^{†1}
山本 雅人^{†1} 古川 正志^{†1}

ネットワーク可視化技術は、ネットワークを大域的に俯瞰することで、その構造的特徴をとらえ、ノード間のリンク関係だけでは見えにくい新たな情報を見つけ出すうえで有用な技術である。本研究ではネットワークの可視化に対して、自己組織化マップ (SOM) の学習機構を利用した可視化方法を適用する。従来の SOM に基づくグラフィケアウト方法である ISOM (Inverted Self-Organizing Map) は、従来の力学的手法に比べ非常に高速であり、隣接ノードどうしが近い位置に配置されるという利点により、ある程度の意味を持つ結果が出力可能である。しかし、一方で可視化結果が信号領域で歪められる現象が起こる問題点を持つ。本研究ではこの問題点を解決する方法として DSSOM (Dynamically-Signaling Self-Organizing Map) を提案し、その有用性を検証する。

A Study on Visualization of Complex Networks Using DSSOM

YASUSHI IWATA,^{†1} IKUO SUZUKI,^{†1}
MASAHITO YAMAMOTO^{†1} and MASASHI FURUKAWA^{†1}

Visualizing complex networks is a useful technology to acquire new knowledge since it allows us to overlook a whole network characteristic at a glance. In this paper, Self-Organizing Map (SOM) is adopted for visualizing the network. A conventional visualization method named Inverted-SOM (ISOM) brings us a great benefit for visualization, but it has disadvantages to distort the network unnaturally. We improve ISOM to get rid of this disadvantage and propose a new visualization method named DSSOM (Dynamically-Signaling Self-Organizing Map). Numerical experiments prove that DSSOM overcomes network distortion in visualization and that it is free from a scale size of visualization.

1. 序 論

データ関係をノードとエッジといった単純な要素で表すネットワーク構造は多くの分野で共通に利用されている。電力網や航空網など実世界に存在する構造をはじめ、Web のようなリンクで結ばれた仮想的なデータの関係まで、あらゆる場面でネットワーク構造を見ることができる。こうした大規模ネットワークは近年複雑ネットワークといわれる。従来の複雑ネットワーク研究では、現実ネットワークがスモールワールドやスケールフリーなどの特定の構造へ自己組織的に導かれる傾向にあることが報告されてきた。しかし、現実のネットワークには特徴量だけでは表すことのできない構造的特徴が隠されていることも事実である。複雑ネットワークの可視化はネットワーク特徴量だけでは把握できない構造的特徴を実際に見ることができる手段の 1 つとなっている。また、可視化は人間の高い認知能力を利用し新たな知的発見を導く重要な技術として期待されている。

近年では、大規模データを扱う機会が増加し、可視化対象となるネットワーク構造もまた大規模化している。そのため、本研究では大規模ネットワークの可視化に対して高速な可視化を実現する自己組織化マップの学習機構を取り入れたグラフィケアウト方法を提案する。ネットワーク可視化研究では主に力学的手法が取り扱われることが多く、自己組織化マップを利用した可視化に対する検証や改良が十分ではない。そのため、本研究ではその有用性を検証し、問題点を改善することを目的とする。

本論文の構成を以下に示す。2 章では関連研究を述べ、従来の可視化研究から大規模ネットワークを対象とした近年の可視化研究までを述べる。また、自己組織化マップを応用した可視化技術についても取り上げる。3 章では可視化問題の提起と本研究の目的について述べる。4 章では従来手法の問題点と提案手法について述べる。5 章では数値計算実験の結果を述べ、6 章でこれらのまとめを述べる。

2. 関連研究

ネットワーク可視化は可視化空間内にノードを配置するグラフィケアウト問題としてこれまでに多くの研究がされてきた。グラフィケアウト問題の主旨は人間が見たときにグラフの構造を把握しやすいようにノードを可視化空間上に自動配置することにある。一般的に

^{†1} 北海道大学大学院情報科学研究科
Graduate School of Information Science and Technology, Hokkaido University

グラフ構造を描画する際に使われる審美的基準としては、交差するエッジ数の最小化、エッジの長さの均一化、ノード分布の均一化などがあげられている¹⁾。しかし、審美的基準の中には互いに背反する制約も存在し、すべての基準を満たすことは実質的に不可能である。したがって、可視化対象となるグラフの特性や可視化目的にあわせて適当な制約を選択し、適用することが必要となる。また、これら単純な制約に対しても、最適な配置を導き出すことは、ほとんどの場合 NP 困難であることが知られている。そのため、グラフィックアウト問題にはヒューリスティック手法を利用するのが一般的になっている。

ヒューリスティック手法の中でも広く知られている方法は力学的解法である。Eades²⁾ はノード間に張られたエッジをバネと仮定し、グラフィックアウト問題を単純な物理モデルに基づいたものに置き換えた。この方法はバネ埋め込み法と呼ばれる。エッジをバネに見立てることで、離れた位置にある隣接ノードは引き寄せられ、近づきすぎた隣接ノードは引き離される効果がもたらされる。この作用が結果的に様々な審美的基準に対応する。しかし、非常に計算量が大きく、出力結果が収束しないという問題もある。Kamada ら³⁾ はノード間のパス長をグラフ的理想距離とし、可視化空間上に配置されたノード間のユークリッド距離がグラフ的理想距離に近くなるようにノード配置を行う方法を提案した。この方法は KK バネモデルと呼ばれ、力学的手法の中でも有名な方法の 1 つである。KK バネモデルでは、隣接ノード以外のノード間にもグラフ的理想距離に基づく自然長を持ったバネを仮定する。そして、系全体のエネルギーを最小化するようにノード配置を最適化する。Fruchterman ら⁴⁾ は Eades のバネモデルにおける引力と斥力の定義を単純なものに改変し、最適化の過程にアニーリングを導入した (FR 法)。力学的手法は最適配置となる保証はないが、一般的なグラフに対しては適度なノード配置を発見するうえで良い方法論である。しかし、力学的手法はいずれも計算コストが高いことが問題である。

近年では、Web の発達などにより可視化対象となるデータ量が増加傾向にある。そのため、大規模ネットワークを実用可能時間で描画する要求が増加し、その研究がさかんに行われている。たとえば、前述の力学的手法では大規模ネットワークに対して計算時間が極端に増大してしまうため、計算の並列化方法や改良アルゴリズムなどが多く提案されている。Adai ら⁵⁾ の提案した LGL (Large Graph Layout) ではネットワーク構造から最小全域木を抽出することで、可視化対象を単純化し高速なノード配置を可能にしている。ただし、これは最小全域木として抽出されなかったエッジはノード配置の段階では無視されることを意味する。そのため、隣接ノードであっても近い位置に配置されとは限らないという問題点もある。また、松林ら⁶⁾ は FR 法に天文力学の方法を応用し、各ノードに固有の更新頻度

を設定することで大規模ネットワークの可視化を高速に実現する方法を提案した。

力学的手法とは異なるアプローチから大規模なネットワーク構造を可視化する目的で考案された方法も存在する。ISOM (inverted self-organizing map) は Meyer^{7),8)} によって提案された可視化方法で、自己組織化マップ (SOM: self organizing map)⁹⁾ を利用し、高速な可視化を実現している。また、同時期に Bonabeau ら¹⁰⁾ も同様のアプローチを試みている。ISOM は高速な可視化を実現する一方で、可視化結果を不自然に歪めてしまう現象が発生する。

現在、ネットワーク可視化に必要とされていることは、大規模なデータを、高速に、意味ある形で描画することである。本研究では、自己組織化マップを利用した従来方法である ISOM を改良し、問題点であった可視化結果が歪められる現象を防ぐ新たな方法として DSSOM (Dynamically-Signaling Self-Organizing Map) を提案する。

3. ネットワーク可視化問題

ネットワーク可視化問題には明確な最適解が存在しない。そのため、可視化問題では目的にあわせて擬似的に最適な解を設定し、最適化問題を解くことで可視化を実現している。

本研究では、以下を可視化目的とする。

- [1] 非隣接ノードを隣接ノードよりも相対的に離れた位置に配置する。
- [2] ISOM の利点である高速さをできる限り維持し、高速に大規模なネットワークを可視化する。
- [3] あらかじめ設定される静的な信号領域により可視化結果が歪められるという ISOM の問題点を改善する。

また、可視化精度については、ネットワーク可視化問題に最適解が存在しない以上、目的にあわせて適切に可視化結果を評価する測度が必要である。そこで、本研究では山田ら¹¹⁾ が提案した接続 F 尺度を測度として導入する。接続 F 尺度は全ノードで、隣接ノードが非隣接ノードより相対的に近い位置に配置された場合に最大値をとる。これは、上記の可視化目的 [1] を数値的に評価する有用な評価方法である。

4. 動的信号を用いた SOM による可視化

4.1 SOM を利用した従来手法 (ISOM)

ISOM は自己組織化マップ (SOM) を利用した可視化方法であり、グラフィックアウトに対する最初のコネクションズアプローチである。SOM は本来、多次元ベクトルからなる入

力信号（入力データ）を正方格子，もしくは六方格子からなる二次元トポロジのニューロン上に投影する方法である．SOM による次元削減は入力信号間に定義される距離（ベクトル間距離）が近いものを二次元トポロジ上で近い位置に配置することで行われる．この際，トポロジ上で隣接関係にあるニューロンはその保持するベクトルが距離的に近くなるのが特徴である．

ISOM は，この隣接ニューロンのベクトルが近くなる特徴を利用することで可視化を実現する．具体的には，ニューロンをノードにみたと，ニューロンの保持するベクトルをノードを配置する座標とする．また，ニューロン間の隣接関係は，ノード間の隣接関係であるネットワークトポロジを用いる．たとえば，二次元座標上にノードを配置したい場合には，ノードにみたとされたニューロンに二次元のベクトルを保持させることで，学習後は隣接ノードが近い二次元座標を保持することになる．しかし，このとき入力信号をどのように用意するかが問題となる．この問題に対して Meyer^{7),8)} は，可視化対象となる二次元可視化領域から一様ランダムにベクトルを取得する方法をとった．これは可視化領域を入力信号ベクトルが存在する領域として定義したといえる．つまり，ISOM では可視化領域と入力信号領域は同一のものである．また，この可視化領域は任意の形状を設定することが可能である．

力学的手法と比較したとき，ISOM の最たる利点は結果出力までの時間が高速であることにある．力学的手法では，グラフノードの配置をエネルギー状態が大域的に最小化する最適化問題として扱い，目的関数の度重なる評価を必要とする．一方で，SOM と同様に教師なし学習を行う ISOM では評価値を算出するための処理を必要としない．また，複数のノード座標を同時に更新することも高速な可視化を実現している要因となっている．

4.2 ISOM の問題点

ISOM では，任意の信号領域を設定できることを特徴としている．一方で，この特徴が可視化したネットワークの形状があらかじめ設定した静的な信号領域に合わせて変形してしまう現象を引き起こす（図 1）．この現象は，自己組織化マップそのものの特性に要因がある．自己組織化マップはトポロジを入力信号のベクトル集合に適合させることを目的とする．そのため，トポロジを構成するニューロンの重みベクトルが入力信号すべてを内包するように学習する．この特性は，ノードが信号領域（可視化領域）に均等に充填するような配置傾向をもたらす．

ISOM においてネットワークの形状を無理に歪めることなく可視化を行うためには，最終的なネットワークの形状を予測し信号領域を設定する必要がある．しかし，ネットワークを可視化する際の理想的な形状を予測することは困難な問題である．

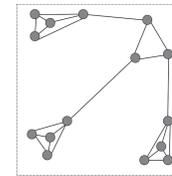


図 1 ISOM の信号領域
Fig. 1 A signal area (ISOM).

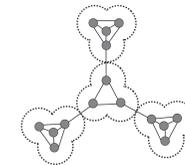


図 2 DSSOM の信号領域
Fig. 2 A signal area (DSSOM).

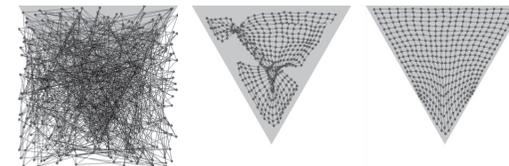


図 3 ISOM の学習過程
Fig. 3 ISOM learning processes.

4.3 動的信号を用いた SOM (DSSOM)

ISOM の問題点を解決するために，あらかじめ信号領域を設定する必要のない新たなアプローチとして DSSOM (Dynamically-Signaling Self-Organizing Map) を提案する．DSSOM の特徴はその信号領域，つまりは可視化領域が動的に変化することにある．具体的には，信号領域を各ノードの周囲に円形に発生させる（図 2）．ここでは，円の半径 r_s を便宜上「信号半径」と呼ぶことにする．この信号領域はノードの座標が更新されるたびにその位置が変化する，そのため学習ステップごとに異なる信号領域が設定される．また，このようにノード周囲 r_s 以内に信号領域を設定することで，入力信号と最も近いノードまでの距離が信号半径 r_s より大きくなることはない．したがって，ネットワークの形状を外側へ向けて無理に押し広げる力を軽減することができる．ISOM における学習過程の様子を図 3 に，DSSOM によるものを図 4 に示す．ネットワークのトポロジは 20×20 の正方形グリッドトポロジを用いている．ここでは，ISOM はあらかじめ三角形の信号領域が設定されているものとする．その結果，正方形のトポロジが三角形に歪められている．一方で，DSSOM では学習の各段階で信号領域が変化している様子が確認できる．また，正方形グリッドのトポロジ特性の影響を受けて最終的な信号領域の形状も正方形となることが分かる．



図 4 DSSOM の学習過程
Fig. 4 DSSOM learning processes.

4.4 可視化アルゴリズム

ネットワークのノード集合 $N = \{i\}$ とそのエッジ集合 (結合関係) $E = \{e_{ij}\}$ を与えたときの DSSOM による可視化アルゴリズムを以下に述べる.

- [1] ノード集合 N の全ノードにランダムに座標ベクトル $\mathbf{x}_i \in X_N$ を与える.
- [2] 全ノードの周囲に信号領域 X_s を定義する.
- [3] 信号領域の範囲からランダムに座標ベクトルを選択し, 入力信号ベクトル $\mathbf{x}_s \in X_s$ とする.
- [4] ノード集合 N から, 入力信号ベクトル \mathbf{x}_s とのユークリッド距離が最も近い座標ベクトル \mathbf{x}_c を持つ勝者ノード c を見つけ出す.

$$c = \arg \min_i |\mathbf{x}_s - \mathbf{x}_i| \quad (1)$$

- [5] 式 (1) により決定された勝者ノードからグラフ距離的に近いノードとして近傍ノード集合 $N_s \in N$ が定義される.
- [6] 座標ベクトルを式 (2) で更新する.
ただし, 近傍ノード集合 N_s 以外のノードの座標は変化させない.
- [7] 指定したループ回数なら終了, そうでなければ [2] に戻る.

ここで, $\arg \min()$ は $|\mathbf{x}_s - \mathbf{x}_i|$ を最小化するノード i を c に与える関数である. 近傍ノード集合 N_s は, 勝者ノード c からの最短経路長 $d(c, i)$ が近傍半径 $\sigma(t)$ 以内に存在するノード群を指す. そのため, 勝者ノードを基点としたダイクストラ法によって逐次的に近傍のノード座標を更新していく必要がある.

ノード座標を入力信号ベクトルに近づける度合いを意味する近傍関数 $h_{ci}(t)$ は式 (3) を採用する.

$$h_{ci}(t) = \alpha(t) \cdot \exp\left(-\frac{d(c, i)^2}{2\sigma^2(t)}\right) \quad (3)$$

近傍関数は, 勝者ノード c からのグラフ的距離が離れるにつれ更新の度合いを減少させる役割を果たす.

DSSOM では, トポロジに可視化対象となるネットワークを利用しているため, ネットワークの特性によっては近傍半径 $\sigma(t)$ が比較的小さな値であっても, 大多数のノードを更新することになる. そのため, 初期の近傍半径 $\sigma(0)$ はネットワークの直径に応じて調整する必要がある. また, $\sigma = 0$ の状態は勝者ノードの座標ベクトルのみを更新することになる, したがって, グラフトポロジとは関係なくノードがランダムに移動してしまう. この状態で学習を継続することは望ましくないため, σ の最小値は 1 とする. $\sigma(t)$ は, $\alpha(t)$ で表される学習率とともに計算ステップ t に応じた減少関数になっている. この $\sigma(t)$ を最大計算ステップ T までに十分減少させることができない場合, ノードが可視化空間内で広がらず混み合ったノード配置になる点に注意されたい.

DSSOM と通常の ISOM のアルゴリズム的な相違点は, ISOM では [2] の処理がなく信号領域 X_s があらかじめ静的に定められていることである.

5. 数値計算実験

5.1 実験条件

DSSOM による可視化を評価するために, 従来方法である ISOM と, 力学的可視化アプローチである KK パネモデルとの比較を行う. 可視化対象とするネットワークデータはネットワーク分野の論文共著ネットワーク (NetScience) と, アメリカ西部の電力網ネットワーク (Power Grid) を利用する. 論文共著ネットワークの元データは非連結ネットワークとなっているため, 本実験では最大連結成分を取り出したサブネットワークを利用する. 論文共著ネットワークの最大連結成分はノード数 367, エッジ数 2,742 であり, KK パネモデルでも比較的短時間で可視化することができる. 一方, 電力網ネットワークはノード数 4,941, エッジ数 6,594 である. 電力網ネットワークはノード数が多いデータではあるが, 地理的ネットワークであるため, 2次元空間に描画しやすく, 可視化した際に目視による評価を行うことも十分に可能である. また, 近傍半径 $\sigma(t)$, 学習率 $\alpha(t)$ の減少関数は論文共著ネットワークで $\sigma(t) = \max(8 \exp(-\frac{t^2}{4 \times 10^6}), 1)$, $\alpha(t) = \max(\exp(-\frac{t^2}{9 \times 10^6}), 0.01)$, 電力網ネットワークで $\sigma(t) = \max(20 \exp(-\frac{t}{5000}), 1)$, $\alpha(t) = \max(0.7 \exp(-\frac{t}{5 \times 10^4}), 0.01)$ とした. これらは予備実験で経験的に得たものである.

本実験では目視による評価に加え、数値的に評価する測度を導入し結果を比較する。今回の実験では測度として接続 F 尺度を用いる。接続 F 尺度は可視化目的 [1] を評価するうえで適した測度である。また、接続 F 尺度による評価値を時系列に観察することで、可視化目的 [2] における計算速度の評価を行うことが可能である。一方で、目視による評価ではあらかじめ設定された信号領域を用いる ISOM と信号領域が動的である DSSOM の可視化結果の差異を直接的に評価できる。これにより、可視化目的 [3] にあげた信号領域をあらかじめ設定する必要のない可視化が実現されていることを確認する。また、ISOM はあらかじめ設定された信号領域の形状に合わせて結果を歪めることからエッジが不必要に長くなる傾向にあると考えられる。そこで、ISOM と DSSOM の可視化結果の差異をより明確化するためにエッジ長分布についても評価する。

DSSOM は SOM を応用した可視化方法であるため、結果の再現性は保証されない。そのため、実験結果の数値は 5 回の実験を行った平均値とする。

5.2 評価測度

5.2.1 接続 F 尺度

接続 F 尺度は山田ら¹¹⁾によって提案された評価法である。この評価法はネットワークを低次元の可視化空間に埋め込んだ際に、ノードの接続関係を忠実に再現した結果が得られているかを定量的な値 (F 値) として評価できる。F 値の算出には、まず、各ノード i の周囲に半径 r_i の仮想的な超球 $B_i(r_i)$ を想定する。F 値は隣接関係にあるノードがこの球の内側にすべて含まれる状態、かつその他のノードが球の外部に配置されている状態で最大値 1 をとる。ただし、F 値が最大になるような理想的なノード配置が可能となるネットワーク構造は稀である。

各ノード i における F 値は、適合率 $P_i(r_i)$ 、再現率 $R_i(r_i)$ と呼ばれる 2 つの値から定義される。

$$P_i(r_i) = \frac{\#\{j|x_j \in B_i(r_i), a_{ij} = 1, j \neq i\}}{\#\{j|x_j \in B_i(r_i), j \neq i\}} \quad (4)$$

$$R_i(r_i) = \frac{\#\{j|x_j \in B_i(r_i), a_{ij} = 1, j \neq i\}}{\#\{j|a_{ij} = 1, j \neq i\}} \quad (5)$$

ここで、 $\#X$ は集合 X の要素数を意味し、 $[a_{ij}]$ は隣接行列を表す。各式から、 $P_i(r_i)$ は分離球 $B_i(r_i)$ 内のノードにおける隣接ノードの割合を意味し、 $R_i(r_i)$ は隣接ノードのうち分離球 $B_i(r_i)$ 内に含まれるノードの割合を意味している。各ノードにおける F 値はこれらの調和平均であり、以下の式で表される。

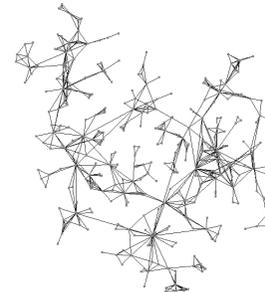


図 5 DSSOM による NetScience の可視化結果
Fig. 5 DSSOM (NetScience).

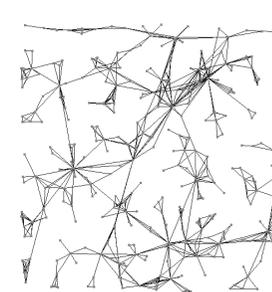


図 6 ISOM による NetScience の可視化結果
Fig. 6 ISOM (NetScience).

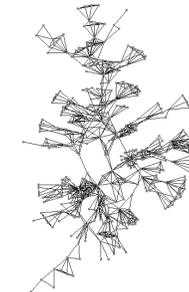


図 7 KK パネモデルによる NetScience の可視化結果
Fig. 7 KK Layout (NetScience).

$$F_i(r_i) = \frac{1}{\alpha/P_i(r_i) + (1-\alpha)/R_i(r_i)} \quad (6)$$

$$F = \frac{1}{N} \sum_{i=1}^N F_i(r_i) \quad (7)$$

本実験では、 $\alpha = 1/2$ を採用する。各ノード i における球の半径は式 (6) が最大となる r_i を各ノードごとに求めることで得られる。式 (6) が最大となる r_i は、ノード i の隣接するノードを含む半径をすべて調べることで見つけることができる。つまり、 r_i の探索には最大でもノード i の次数回の試行を行えば十分である。

また、ネットワーク全体の F 値は、これらの平均として得ることができる。

5.2.2 エッジ長分布

一般的な審美的基準において、ノード間を結ぶエッジの長さは均一なものが良いとされている。そこで、各可視化方法についてそのエッジ長分布を観察し、それらの分散を比較し、エッジ長の均一さを評価する。エッジ長は描画されたネットワークのスケールに依存する。そのため、描画結果からバウンディングボックスを計算し、その対角長 L をもとにエッジ長を正規化する。

5.3 結果・考察

論文共著ネットワーク (NetScience) の可視化結果を図 5, 図 6, 図 7 に示す。ISOM では正方形の信号領域にあわせて可視化結果の形状も歪んでいる (図 6)。一方、DSSOM で

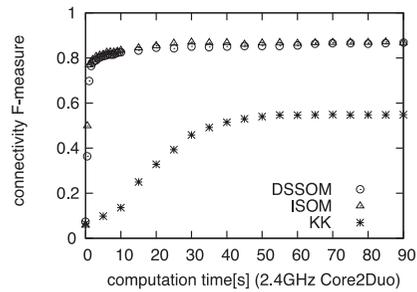


図 8 NetScience の可視化における F 値の時間変化
Fig. 8 F-Measure (NetScience).

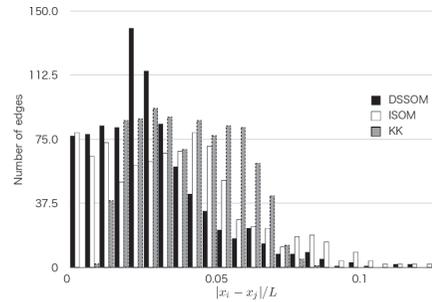


図 9 エッジ長の度数分布 (NetScience)
Fig. 9 Histogram of edge length (NetScience).

表 1 エッジ長の分散
Table 1 Variance of the edge length.

	NetScience	PowerGrid
DSSOM	0.4819×10^{-3}	0.0533×10^{-3}
ISOM	1.8196×10^{-3}	0.0902×10^{-3}
KK	0.2712×10^{-3}	

は信号領域が動的なため ISOM のように特定の形状へ歪められていない (図 5). KK バネモデルは、描画されたネットワークの中心部から外側に向けてノードが放射状に配置される傾向がみられる (図 7). また、DSSOM は KK バネモデルと比べると密につながりあったノード集合が集約されて配置されることが確認できる.

論文共著ネットワークの可視化に際して、F 値の収束までの時間変化をプロットした結果を図 8 に示す. KK と比較すると、DSSOM、ISOM はどちらも F 値の増加が急速であり計算開始後、瞬時に高い値に収束している. これは、どちらの方法も可視化結果までの出力時間が高速であることを示している. DSSOM は ISOM に比べ信号領域が狭いため、各ステップにおいて勝者ノードの移動距離に制限がかかる. しかし、同時に更新される近傍ノードはその制約は受けないため、結果的には ISOM にみられる移動量の大きい座標更新が維持される. そのため、ISOM と比較しても十分な速度を維持されていると考えられる.

エッジ長の分布を調べると、ISOM と DSSOM は KK バネモデルと比較して、短いエッジの割合が高い (図 9). また、分散は KK、DSSOM、ISOM の順に小さい (表 1). つまり、エッジ長の均一性では KK バネモデルが最も良いことになる. ISOM は可視化結果が信号

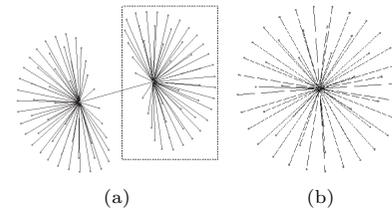


図 10 KK バネモデルの特性
Fig. 10 A characteristic of KK Layout.

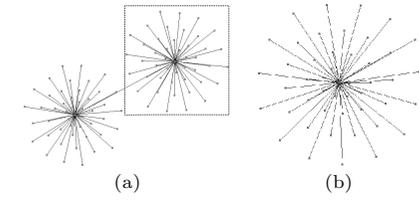


図 11 DSSOM の特性
Fig. 11 A characteristic of DSSOM.

領域全体に引き延ばされているため、エッジ長が他に比べ極端に長いものが存在する. それに比べ、DSSOM では極端に長いエッジは削減され、エッジの分布が全体的に短い方へ偏っている. つまり、DSSOM では KK バネモデルほどエッジの均一性は得られないものの、ISOM と比べるとエッジの均一性は高くなっている.

DSSOM、および ISOM は、F 値の収束値が高い水準を示すのが KK バネモデルと比較した特徴である. これには、近傍半径 $\sigma(t)$ が大きく影響している. 減少関数である $\sigma(t)$ は学習の過程で徐々に減少し、最終的に $\sigma(t') = 1$ となる. そのため、学習の後半では各ノードが一次近傍のノードのみを考慮したうえで配置座標を調整する. 一方で、KK バネモデルでは 2 次近傍は 1 次近傍より 2 倍の距離に配置するなど、可能な限りグラフ距離に即した配置を目指す. そのため、図 10 (a) に示すような構造は、ノードが外側へ向けて放射状に広がる配置がなされる. 一方で、同様の構造を DSSOM で可視化すると図 11 (a) に示すように一次近傍を近い距離に集める効果のみが働いていることが分かる. この特性により、DSSOM は F 値の条件をよく満たすことになる.

文献 11) で述べられている可視化方法 CoPE も、1 次近傍を近くに配置することを目的関数とし、DSSOM と同様に F 値の意味で高い水準を満たしている. しかし、CoPE は KK バネモデルと同様の学習アルゴリズムを用いているため、DSSOM に比べ計算時間がかかる. DSSOM は、1 次近傍ノードのベクトルを近づける SOM の特性を持つ. これが CoPE の可視化特性に近いものをもたらす、結果として CoPE と同様にネットワークの部分構造を切り出して可視化を行う際に良い影響を与える. 具体的には、部分構造を取り出し、再描画すると DSSOM では図 11 (b) のように似た配置を得るが、KK バネモデルでは異なった配置を図 10 (b) のように得る.

電力網ネットワーク (Power Grid) はノード数が比較的多いデータセットである. 大規

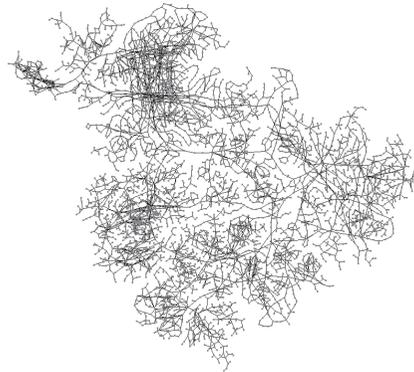


図 12 DSSOM による Power Grid の可視化結果
Fig. 12 DSSOM (Power Grid).

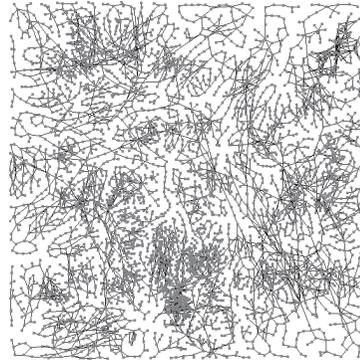


図 13 ISOM による Power Grid の可視化結果
Fig. 13 ISOM (Power Grid).

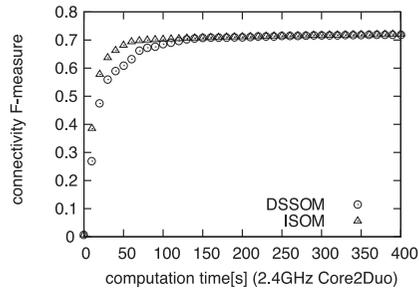


図 14 Power Grid の可視化における F 値の時間変化
Fig. 14 F-Measure (Power Grid).

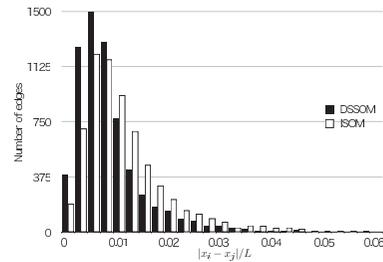


図 15 エッジ長の度数分布 (Power Grid)
Fig. 15 Histogram of edge length (Power Grid).

模データでは、KK バネモデルのような力学的手法では計算時間が極端に増加し、結果出力まで数時間を要する。そのため、ここでは主に DSSOM と ISOM について比較を行う。DSSOM および ISOM による可視化結果を図 12、図 13 に示す。ISOM は正方形の信号領域全体にノードを押し広げている (図 13)。それに対して、DSSOM ではネットワーク全体の概観を確認することができる (図 12)。論文共著ネットワークと同様の結果であるが、ノード数の増加により ISOM の可視化結果ではより構造が把握し難いものになっている。F 値の時間変化からは、どちらの方法も数分で値が収束し、可視化結果が出力されること

が分かる (図 14)。同様のデータを KK バネモデルで可視化すると数時間の計算時間を必要とする。これに比して、DSSOM および ISOM は非常に高速な可視化が実現されている。また、F 値の増加率では、ISOM が DSSOM と比べると早い段階で高い値に収束している。これは、DSSOM では動的信号領域を仮定するため ISOM に比べ余分な計算ステップが必要とされるためである。

エッジ長分布を見ると、DSSOM では ISOM に比べ長いエッジが削減されている (図 15)。また、エッジ長の分散も DSSOM のほうが小さく、ISOM に比べエッジの均一性の面で良い結果を出している (表 1)。しかし、論文共著ネットワークの結果に比べその差は小さい。これは電力網ネットワークの構造的な特性が影響していると考えられる。

6. 結 論

本研究では、ISOM の問題点を改善する新たな方法として DSSOM を提案し、その有用性を検証した。これらをまとめると以下となる。

- [1] DSSOM は従来の ISOM と異なり、動的に変化する信号領域によって、信号領域をあらかじめ設定することなく結果を出力可能である。
- [2] 自己組織化の学習機構を利用するため、従来のバネモデルを主体とした力学的手法に比して結果出力までの計算速度が非常に速い。
- [3] ISOM と DSSOM は F 値の評価基準から“非隣接ノードを隣接ノードよりも相対的に離れた位置に配置する”意味では、高い水準を満たす。
- [4] エッジ長の分散から考えると KK バネモデルと比べ、DSSOM はエッジの均一性に欠けるが、ISOM における極端に長いエッジが削減されている。

一方で、DSSOM では近傍半径、学習率などの各種パラメータが存在し、それらの設定は経験則によるところが大きい。今後の展望として各種パラメータをネットワークの特徴量から推測し、自動設定することが必要である。

参 考 文 献

- 1) Eades, P. and Tamassia, R.: Algorithms for Drawing Graphs: An Annotated Bibliography, Technical Report, Brown University (1988).
- 2) Eades, P.: A Heuristic for Graph Drawing, *Congressus Numerantium*, Vol.42, pp.149-160 (1984).
- 3) Kamada, T. and Kawai, S.: An algorithm for drawing general undirected graphs, *Inf. Process. Lett.*, Vol.31, No.1, pp.7-15 (1989).

- 4) Fruchterman, T.M.J. and Reingold, E.M.: Graph drawing by force-directed placement, *Software: Practice and Experience*, Vol.21, No.11, pp.1129–1164 (1991).
- 5) Adai, A.T., Date, S.V., Wieland, S. and Marcotte, E.M.: LGL: Creating a map of protein function with an algorithm for visualizing very large biological networks, *J. Mol. Biol.*, Vol.340, No.1, pp.179–190 (2004).
- 6) 松林達史, 山田武士: 階層的独立固有時間刻み法によるグラフ可視化計算の高速化, 情報処理学会論文誌: 数理モデル化と応用, Vol.48, No.15, pp.126–136 (2007).
- 7) Meyer, B.: Competitive learning of network diagram layout, *Proc. 1998 IEEE Symposium on Visual Languages, 1998*, pp.56–63 (1998).
- 8) Meyer, B.: Self-Organizing Graphs - A Neural Network Perspective of Graph Layout, *Proc. 6th International Symposium on Graph Drawing*, pp.246–262, Springer-Verlag (1998).
- 9) Kohonen, T.: *Self-organizing maps*, Springer-Verlag New York, Inc., Secaucus, NJ, USA (1997).
- 10) Bonabeau, E. and Hénaux, F.: Self-organizing maps for drawing large graphs, *Inf. Process. Lett.*, Vol.67, No.4, pp.177–184 (1998).
- 11) 山田武士, 齊藤和己, 上田修功: クロスエントロピー最小化に基づくネットワークデータの埋め込み, 情報処理学会論文誌, Vol.44, No.9, pp.2401–2408 (2003).

(平成 21 年 1 月 30 日受付)

(平成 21 年 9 月 11 日採録)



岩田 泰士 (正会員)

1986 年生。2008 年北海道大学工学部情報工学科卒業。同年北海道大学大学院情報科学研究科複合情報学専攻に入学。複雑ネットワーク, ネットワーク可視化の研究に従事。電子情報通信学会会員。



鈴木 育男 (正会員)

1973 年生。2004 年北海道大学大学院工学研究科博士後期課程修了。博士 (工学)。同年室蘭工業大学サテライト・ベンチャー・ビジネス・ラボラトリー中核的研究機関研究員。2007 年北海道大学大学院情報科学研究科助手となり, 現在に至る。複雑系, Web マイニングによる感性情報の抽出等に関する研究に従事。日本ロボット学会, 精密工学会, 日本感性工

学会各会員。



山本 雅人 (正会員)

1968 年生。1996 年北海道大学大学院工学研究科システム情報工学専攻博士後期課程修了。同年日本学術振興会特別研究員 (PD)。1997 年北海道大学大学院工学研究科助手。2000 年同大学院工学研究科助教授。同大学院情報科学研究科助教授を経て, 2007 年北海道大学大学院情報科学研究科准教授。この間, 科学技術振興機構さきがけ研究員, デューク大学客員研究員。博士 (工学)。現在は, 進化型計算に基づく仮想ロボット開発, 複雑ネットワークの研究に従事。電子情報通信学会, 人工知能学会, 計測自動制御学会, 日本オペレーションズ・リサーチ学会, 精密工学会等, 日本機械学会会員。

員研究員。博士 (工学)。現在は, 進化型計算に基づく仮想ロボット開発, 複雑ネットワークの研究に従事。電子情報通信学会, 人工知能学会, 計測自動制御学会, 日本オペレーションズ・リサーチ学会, 精密工学会等, 日本機械学会会員。



古川 正志 (正会員)

1948 年生。1973 年北海道大学大学院工学研究科修士課程修了。同年旭川工業高等専門学校電気工学科助手, 同機械工学科助教授, 同制御情報工学科教授を経て 2006 年北海道大学大学院情報科学研究科教授。この間, コーネル大学 NSF 研究員, イーストアングリア大学客員教授。1981 年工学博士 (北海道大学)。自律分散システム, インテリジェント・エンジニアリング, 複雑ネットワーク等の研究に従事。計測自動制御学会, 機械学会 (フェロー), 精密工学会各会員。

リング, 複雑ネットワーク等の研究に従事。計測自動制御学会, 機械学会 (フェロー), 精密工学会各会員。