

# 地図投影法を応用したグラフデータ可視化手法

伊藤 貴之

お茶の水女子大学 理学部情報科学科 E-mail: itot@computer.org

グラフデータ可視化手法は、金融・交通・通信・社会組織・科学・計算機システム・インターネットなど、非常に幅広い分野において実用事例が報告されている、きわめて有用な技術である。近年の情報科学技術で適用されるグラフデータは、ノードを連結するアークの密度が非常に高いことが多く、その全てを一画面に表示するのは無意味である場合が多い。そこでユーザーの注視部分や検索対象部分に焦点を絞り、その局所を理解できるようにグラフデータを可視化する手法が有用であり、既に多くの手法が発表されている。

本報告では、地図投影法を応用したグラフデータ可視化手法を提案する。提案手法では、まず既存のグラフデータ画面配置手法を応用して、楕円領域にノード群を配置する。続いてその楕円領域に、地図投影法の一つである正積図法を適用して、各々のノードに経度と緯度を与える。注視部分や検索対象部分が指定されたとき、提案手法は正距方位図法を適用することで、指定された部分を中央とするノード再配置を高速に実現する。

## A Graph Data Visualization Technique Applying Map Projection Methods

Takayuki ITOH

Department of Information Sciences, Faculty of Science, Ochanomizu University

Graph data visualization is a very vital and useful research topic because it is very useful for analysis and arrangement for various data such as financial, traffic, communicational, social, scientific, computer system, and internet data. Recent real graph data applied in vital information technology areas often contain very messy arcs, and therefore it is sometime nonsense to visualize every connection of arcs of graphs in one display. Many recent graph data visualized techniques concentrate to focus on local interesting or query regions. Such techniques are useful for such messy graph data.

This paper proposes a graph data visualization technique that applies multiple map projection methods. The technique first places nodes of a graph onto an ellipsoidal area applying existing graph data layout techniques. It then calculates longitudes and latitudes to the nodes by applying equiarea projection to the ellipsoidal area. When an interest area is manually specified, the proposed technique quickly provides the new viewpoint for the graph data by applying equidistant projection.

### 1. はじめに

グラフデータの可視化技術は、非常に幅広い分野のデータ分析およびデータ整理に有用である。最近では、

- ウェブサイトのリンク構造の表示。
- 金融・通信・交通・社会組織などの各種ネットワークの表示。
- 化学・生物などのサンプリングデータの傾向

理解のために、データを関連性で連結したグラフの表示。

- テキストデータや画像データの整理のために、データを関連性で連結したグラフの表示。
  - 例えば並列計算機のプロセスなどのように、関連あるモジュールの集合で構成されるシステムの振舞いを表したグラフの表示。
- などの分野での実用例が報告されている。また、グラ

グラフデータ可視化のための基礎技術については、近年になって有用な解説書やサーベイ論文が出版されている [Bat99] [Her00]。

グラフデータ可視化技術で最も活発に議論されている問題は、グラフデータを構成するノードに効果的に画面配置する技術である。文献[Bat99]では、代表的なグラフデータ画面配置手法として、

- アーク間の交差を減らすために、平面グラフ化やノードの位置交換を試みる手法
- 可読性を向上するために、アークの直交化を図る手法
- ノード間の適正距離、アーク間の適正長、を保つために力学モデルを適用する手法
- 有向グラフの方向性を保持する手法

などを紹介している。これらの手法は原則として、グラフデータを構成する全てのノードおよびアークを、最適に近いレイアウトで一画面に表示することを目的としている。しかしこれらの手法の多くは、ノード数  $n$  に対して計算量が  $O(n)$  を超えるため、大規模なグラフデータにおいて大きな計算時間を要する。

また、近年の情報科学技術で適用されるグラフデータには、これらのグラフデータ描画手法が通用しないほど、アークの密度が高いものも多く見られる。例えば、ウェブサイトのリンク構造を表現したグラフの場合、1ノードに接続されるアークの平均本数が10本を超える事例も多い。このようなグラフデータを既存のグラフデータ描画手法で可視化しても、膨大な本数のアークを全部表示しては可読性が失われる。このようなグラフデータの可視化においては、グラフデータを構成するアークを全て表示するのではなく、ユーザーの注視部分に含まれるアークだけを表示し、局所的な接続関係の可視化に専念するのが現実的である、と考えられる。

ここで、グラフデータの全体像と、そのグラフデータ中のノード群の相対的な位置関係を提示しながら、しかもユーザーの注視部分にあたる局所の可視化に重点をおく、というような可視化手法の実現について考える。ユーザーの注視部分が明確であるとき、多くの場合において、ユーザーの注視部分が画面の中央に配置されることが好ましい。この要件を満たすために本報告では、

1. 既存のグラフデータ画面配置手法を一度だけ実行して、基本的なグラフデータ配置を決定する。
2. 1.で確定したグラフデータ配置によるノード間の相対的な隣接関係は崩さず、しかも高速な処理でグラフデータ中の任意の局所を中心とする

画面再配置を実現する。

という考え方によるグラフデータ可視化手法を提案する。本報告で提案する手法は、正距方位図法を含む複数の地図投影法の組み合わせにより、グラフデータ中の注視部分を中央に集めた可視化結果を高速に提示するものである。提案手法は以下の手順により、グラフデータ中の特定部分を中心としたノード配置を高速に実現する(図1参照)。

1. 既存のグラフ画面配置手法を応用して、楕円領域にノード群を配置する。
2. 楕円領域内の位置から経度・緯度を算出することで、各ノードを球面に配置する。
3. 球面上の注視部分を指定した際に、正距方位図法を適用して、その注視部分を中心としてノード群を再表示する。
4. 注視部分に含まれるノードおよびアークで構成される部分グラフを表示する。

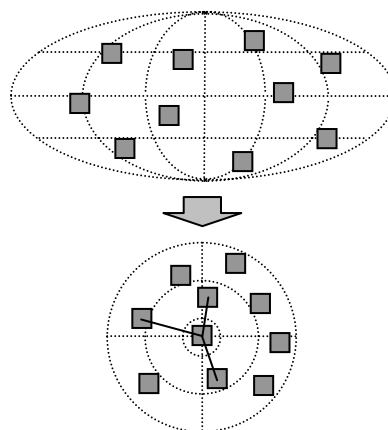


図 1. 提案手法の概要。(上)楕円領域上のノード配置、および経度・緯度の算出。(下)正距方位図法を適用した再表示。

本報告で提案する手法は、航空路図にヒントを得た手法である。航空会社が提示する航空路図には、正距方位図法を用いて、特定の空港を中心とした航空経路を提示するものが見られる。この表現は、中心点からの方位と距離が正しいという正距方位図法の性質により、航空経路を短く表示できるという利点がある。航空経路に限らない一般的なグラフデータの可視化においても、アークの長さが短いほうが好ましい場合が多いことから、正距方位図法を応用した可視化手法は有用であると考えられる。

提案手法の特徴は以下の通りである。

- ユーザーによる注視部分の指定を変更するたびに、既存のグラフ画面配置手法を再適用せずに、単なる座標変換でグラフデータを

再描画するため、高速に新しい可視化結果を提示できるとともに、初期配置結果におけるノード間の相対的な近隣関係を持続する。

- 多くのデータにおいて、表示対象となるアークが短く表示され、可読性が高まる。
- 非幾何的な操作で適切な視点を提供できることから、CG や GUI に習熟していないユーザーや、対話的操作を好まないユーザーに向いている。

## 2. 地図投影法を応用したグラフデータ可視化手法

本章では、以下の記法により入力グラフデータ  $G$  を記述するものとする。

- グラフデータ  $G = \{N, A\}$  は、ノード群  $N$  とアーク群  $A$  で構成される。
- $N$  は  $n$  個のノード  $N = \{n_1, \dots, n_n\}$  で構成される。
- $A$  は  $m$  本のアーク  $A = \{a_1, \dots, a_m\}$  で構成される。

また、 $G$  の局所を指定する方法として、本章では以下の3種類を想定する。

### [方法 A: 1 個のノードを指定する方法]

ノード  $n_i$  および位相的距離  $d$  を入力する。提案手法は、 $n_i$  から  $d$  本以下のアークを経由して到達できるノード群、および経由したアーク群で構成される部分グラフ  $G_A$  を、局所として扱う(図 2(左)参照)。

### [方法 B: 2 個のノードを指定する方法]

2 個のノード  $n_i$  および  $n_j$  を入力する。提案手法は、 $n_i$  から最低本数のアークを経由して  $n_j$  に到達するまでに、途中通過するノード群およびアーク群で構成される部分グラフ  $G_B$  を、局所として扱う(図 2(右)参照)。

### [方法 C: 多数のノードを指定する方法]

$k$  個のノードを入力する。提案手法は、 $k$  個のノードの幾何的な重心を中心として、 $k$  個のノードを全て含む部分グラフ  $G_C$  を、局所として扱う。

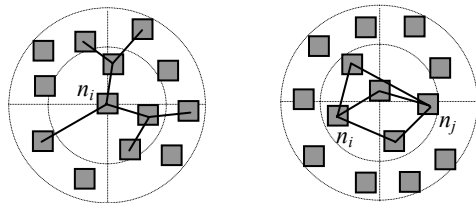


図 2. (左)1 個のノードを指定する方法。(右)2 個のノードを指定する方法。

グラフデータの局所可視化では、CG や GUI の技術を駆使した探索的な対話的操作だけでなく、例えば明示的に指定したノードの周囲を可視化したい、あるいは検索条件に合致するノード群だけを可視化し

たい、というような非幾何的な操作が要求される。上記 A~C のような方法は、このような非幾何的な入力手段で注視部分の指定を実現するものであり、CG や GUI に習熟していないユーザーに向いている方法であると考えられる。

## 2.1 楕円領域上のノード配置

提案手法においても、楕円領域上の初期配置の品質は重要である。特に、

- ノード同士が適度な距離を保ち、領域全体にわたって均一な密度で配置されること
  - アークが適度な長さを保ち、交差を減らすこと
- などの条件は重要であると考えられる。そこで提案手法でも、これらの条件を満たすことを目標とした既存のグラフデータ画面配置手法を応用して、楕円領域にノードを配置する。

楕円領域上にノードを配置する方法は、何通りか考えられる。バネや分子間力などの力学モデルを適用したグラフデータ画面配置手法[Doi04]を適用する場合には、以下の2種類の実装方法が考えられる。

**[方法 1]** あらかじめ楕円領域を設定し、力学モデルの反復計算の課程でノードが楕円領域から外に出してしまうときには、楕円領域の境界線上に強制的にノードを残しながら、力学計算を続行する。文献[Doi04]に示す手法のように、ノードをインクリメンタルに画面配置する手法の場合には、楕円領域も順次拡大する必要がある。

**[方法 2]** 楕円領域を設定せずにグラフデータを画面配置し、配置領域の中心点が原点になるようにノードを平行移動する。続いて配置結果の長軸を算出し、その長軸が画面の水平方向に平行になるように、原点を中心にしてノードを回転移動する。最後に、これらのノードを包括する楕円領域を定義する。

## 2.2 経度・緯度の算出

画面領域の水平方向および垂直方向に楕円領域があり、楕円の中心点が画面領域の原点にあり、楕円の長軸および短軸の長さが  $l_a$ 、 $l_b$  であるとする。このとき提案手法では式(1)によって、座標値  $(x, y)$  を有するノードの経度  $\alpha$  と緯度  $\beta$  を算出する(図 3(左)参照)。

$$\alpha = \frac{\pi x}{l_a}, \beta = \frac{\pi y}{2l_b} \quad \dots(1)$$

続いて提案手法は、経度  $\alpha$  および緯度  $\beta$  を有するノードを、半径 1 の球面に投影したときの直交座標系での座標値  $(x', y', z')$  を、式(2)により算出

する(図3(右)参照)。

$$\begin{aligned} x' &= \cos \alpha \cos \beta \\ y' &= \sin \alpha \cos \beta \quad \dots(2) \\ z' &= \sin \beta \end{aligned}$$

この直交座標系は、地球の北極に相当する点 ( $\beta = 0.5\pi$ を満たす点)の座標値を  $(0,0,1)$  とする座標系である。

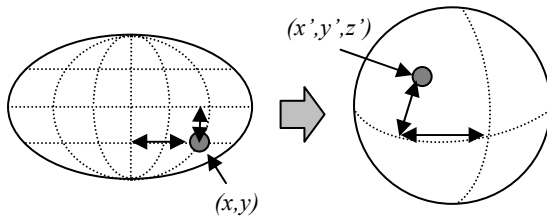


図3. (左)楕円領域上の座標値からの経度・緯度の算出。(右)球面上の座標値の算出。

2.1 節および 2.2 節に示した処理は、グラフデータが与えられた時に一度だけ実行するものである。筆者の実装ではこれらの処理に、数百個のノードを有するグラフデータにて数秒～数十秒程度の計算時間を要するが、2.3 節以降の対話的処理は毎回 10 ミリ秒程度の計算時間で実現できている。

### 2.3 球面上の特定位置を中心とした正距方位図法への変換

ユーザーが注視部分を特定すると、提案手法は注視位置 P を中心とした正距方位図法を適用して、注視部分における部分グラフを表示する。ここで提案手法では、[方法 A]を用いるときは、ノード  $n_i$  の座標値を P とする。[方法 B]または[方法 C]を用いるときは、複数のノードの重心点を正規化した座標値 (= 半径 1 の球に投影した座標値) を P とする。

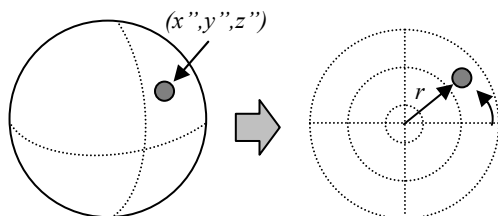


図4. (左)回転変換後の球面上の点。(右)正距方位図法による投影後の円座標系での位置。

まず提案手法は、注視位置 P が座標値  $(0,0,1)$  に写像されるように、各ノードの座標値に回転変換を施す。続いて提案手法は、回転変換後のノードの座標値を

$(x'', y'', z'')$  とした際に、式(3)を用いて、正距方位図法を適用した際の円座標系での位置  $(r, \theta)$  を算出する(図4参照)。

$$r = \frac{0.5\pi - \arcsin z''}{\pi}, \theta = \arctan\left(\frac{y''}{x''}\right) \quad \dots(3)$$

注視位置 P を変更しながら可視化を反復する際に提案手法では、前回の注視位置から今回の注視位置に向けて漸進的に座標変換し、その過程をアニメーション表示することで、可視化結果の急激な変化を和らげる。

### 3. 実行例

提案手法を実装して実行した例を示す。筆者は提案手法を Java SDK 1.4.2 および JOGL 1.1.0 を用いて実装し、IBM ThinkPad T42 (CPU 1.8GHz) および Microsoft Windows XP の上で作動した。本稿では紙面の都合により、小規模なグラフデータによる実行例のみを示し、大規模なグラフデータによる実行例は口頭講演中に示す。

図5は、35ノードで構成される小規模なグラフデータを、楕円領域に配置した例である。図6は、この35ノードの中から各々1ノード  $n_i$  を指定し、 $n_i$  を円座標系の原点として正距方位図法を適用して描画した例である。この例では、 $n_i$  と直接接続されているアーク、またそれらのアークの他端点に接続されたアークを表示している。つまり図6で、表示されているアークに接続されているノードは、 $n_i$  と2本以内のアークを介して接続されているノードであるといえる。

図7は、図5に示した35ノードの中から各々2ノード  $n_i, n_j$  を指定し、 $n_i$  と  $n_j$  の中点を円座標系の原点として正距方位図法を適用して描画した例である。この例では、 $n_i$  および  $n_j$  の両方と直接接続されているノードを抽出し、 $n_i$  または  $n_j$  と抽出されたノード間を接続されたアークを表示している。つまり図7は、 $n_i$  から  $n_j$  に向かって最大2本のアークを介して到達できる経路群を示しているといえる。

図6,7のような可視化結果は、例えば以下のような検索を支援するツールとして有用と考えられる。

- 多数の遺伝子間の相関性を表現したグラフに関して、遺伝子 A から間接的な因果関係をもつ遺伝子群を可視化する。また、直接的な因果関係を持たない遺伝子 A と B が、間接的な因果関係をもつための経路群を可視化する。
- 多数の人物間の面識関係を表現したグラフに関して、人物 A から直接の知人を介して紹介してもらえ人物群を可視化する。また、直接の面

識を持たない人物 A と B が、面識をもつために仲介しうる人物群を可視化する。

- ニュース記事や学術文献などから抽出される多数のキーワード間の因果関係を表現したグラフに関して、キーワード A から間接的に因果関係を持ちえるキーワード群を可視化する。また直接的な因果関係を持たないキーワード A と B が、間接的な因果関係をもつための経路群を可視化する。

#### 4. 関連手法

提案手法と共通目的をもつグラフデータ可視化手法や木構造データ可視化手法は、既に多く報告されている。以下に、提案手法に関連するいくつかの手法について紹介する。

文献[Lam96]にて提案されている Hyperbolic Tree は、木構造を対話的に表示する有名な手法である。木構造の根ノードが頂上に配置されるように、木構造を双曲面上に配置し、視点操作によって木構造中の局所部分の拡大表示を実現している。木構造に限定した手法であり、一般的なグラフデータには適用されていない。

文献[Mun97]にて提案されている H3 は、木構造を球面上に配置している手法である。提案手法はこの手法に対する拡張と補足であると考えられる。

文献[Sar94]は、魚眼レンズ風の拡大表示によって、グラフデータ中の特定部分を注視させる手法を提案している。遠い位置に配置された複数のノードを同時に注視する目的には向かない。

文献[Hua98][Nor95]は、グラフデータをインクリメンタルに画面配置することで、ユーザーの注視する局所部分から順にグラフデータを提示する手法を提案している。巨大なグラフデータを対象とするときには、提案手法との併用が効果的であると考えられる。

文献[Shi97]にて提案されている「納豆ビュー」は、グラフデータ中の特定ノードを 3 次元的に引き上げることで、特定ノードと他ノードの接続性を注視させる手法である。提案手法は「納豆ビュー」の目的を 2 次元的に実現する手法であるとも考えられる。

文献[Sku01]に提案されている可視化手法は、地理的意味をもたないデータの情報可視化に地図投影法を明示的に適用した、という意味で数少ない関連手法である。しかし、グラフデータの局所注視を目的とした手法は提案していない。

文献[Yee01]は、円座標系に配置されたグラフデータに対して、データ要素間の交差の少ない滑らかな

アニメーションを実現する手法を提案している。提案手法との併用の可能性について、今後検討したい。

文献[Jan03]は、円座標系に配置されたグラフデータに対して、ノードが不均一な面積をもつことを仮定して、それが重ならないように局所拡大表示を実現する手法を提案している。これも提案手法との併用の可能性について、今後検討したい。

#### 5. むすび

本報告では、正距方位図法を含む複数の地図図投影法を用いて、グラフデータの特定部分を注視する可視化結果を高速に提示する手法を提案した。今後の課題として、大規模なグラフデータでの実験、および以下の機能拡張を考えている。

**階層型グラフデータへの拡張:** 大規模なグラフデータの探索的可視化の目的で、グラフデータを階層化することは有意義である。多くの階層型グラフデータ可視化手法[Doi04][Ead96][Sch96]では、複数のノードを括る大きな多角形領域で階層を表現している。しかし提案手法のように、球面に可視化結果を投影する手法の場合、大きな多角形領域の投影結果に歪みが生じることが避けられないため、なんらかの考察が必要である。一方、文献[Azu04]は階層型グラフデータの上位階層を円領域の外側に表示する新しい表現形態を提案している。むしろこの表現形態のほうが、提案手法の拡張に向いている、と考えられる。

**アニメーション手法との併用:** 正距方位図法を用いた放射状の可視化の状態、注視部分を切り替えるような操作を行った場合、グラフデータ全体にわたってノードが移動してしまうので、可読性に影響が生じる恐れがある。これを軽減する一手段として、注視部分を切り替える際に、アニメーションを導入して滑らかにノードを移動させることが考えられる。文献[Yee01]に提案された手法との併用などを含めて検討したい。

#### 参考文献

[Azu04] 我妻他, 対話的 2D/3D 表示技法を用いた階層的因果関係可視化システム, Visual Computing / グラフィクスと CAD 合同シンポジウム 2004.

[Bat99] Battista G. D., Eades P., Tamassia R., Tollis I. G., Graph Drawing – Algorithms for the Visualization of Graphs, 1999.

[Doi04] 土井, 伊藤, 力学モデルを用いた階層型グラフデータ画面配置手法の改良手法とウェブサイト視覚化への応用, 芸術科学会論文誌, 3, 4, pp. 250-263, 2004.

[Ead96] Eades P., et al., Multilevel Visualization of Clustered Graphs, Graph Drawing '96, pp. 101-112, 1996.

[Her00] Herman I., Melancon G., Marshall M. S., Graph Visualization and Navigation in Information Visualization: A Survey, IEEE Trans. Visualization and Computer Graphics, 6, 1, 24-43, 2000.

[Hua98] Huang, M. L. et al., WebOFDAV – Navigating and Visualizing the Web On-Line with Animated Context Swapping, 7th WWW Conference, pp. 636-638, 1998.

[Jan03] Jankun-Kelly T. J, Ma K., MoireGraphs: Radial Focus+Context Visualization and Interaction for Graphs with Visual Nodes, IEEE Symposium on Information Visualization (InfoVis '2003), 2003.

[Lam96] Lamping J., et al., "The Hyperbolic Browser: A Focus+context Technique for Visualizing Large Hierarchies", Journal of Visual Languages and Computing, 7, 1, pp.33-55, 1996.

[Mun97] Munzner T., H3: laying out large directed graphs in 3D hyperbolic space, IEEE Symposium on Information Visualization (InfoVis '97), 1997.

[Nor95] S. North, Incremental Layout in DynaDAG, Graph Drawing '95, pp. 409-418, 1995.

[Sar94] Sarkar M., et al., Graphical Fisheye Views, Communication of ACM, 37, 12, pp. 73-84, 1994.

[Sch96] Schaffer D., et al., Navigating Hierarchically Clustered Networks through Fisheye and Full-Zoom Methods, ACM Trans. Computer-Human Interaction, 3, 2, pp. 162-188, 1996.

[Shi97] 塩澤他, 「納豆ビュー」の対話的な情報視覚化における位置付け, 情報処理学会論文誌, 38, 11, pp. 2331-2342, 1997.

[Sku01] Skupin, A. Cartographic Considerations for Map-like Interfaces to Digital Libraries. First ACM+IEEE Joint Conference on Digital Libraries (JCDL '01). Workshop on Visual Interfaces to Digital Libraries, 2001.

[Yee01] Yee K., et al., Animated Exploration of Dynamic Graphs with Radial Layout, IEEE Symposium on Information Visualization (InfoVis '2001), 2001.

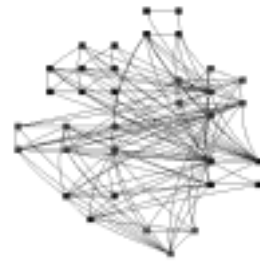


図 5. 35 ノードの楕円領域への配置。

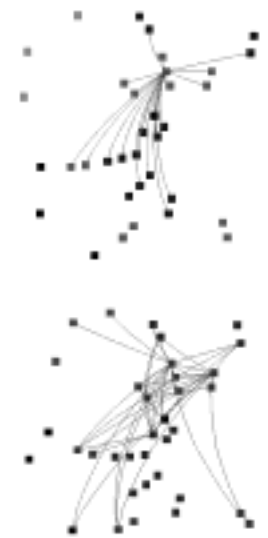


図 6. 1 ノードを中心とした再描画。

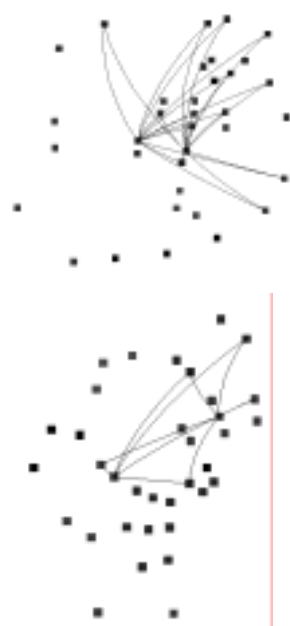


図 7. 2 ノードの中点を中心とした再描画。