

時間変化を考慮したネットワーク可視化方法 Network Visualization Method deal with Time Change

川口 光[†] 山本 雅人[†] 古川 正志[‡]
Ko Kawaguchi Masahito Yamamoto Masashi Furukawa

1. はじめに

近年、ITの普及とコンピューターの性能向上により、各企業は自社や関連会社のデータを膨大に蓄積するようになった。現在、こういったデータからネットワークを生成し可視化することが盛んに行われている。従来のネットワーク可視化方法は単一のデータからネットワークを可視化するという方法が一般的である。しかし、各企業でデータの蓄積は今後も続いていくであろうことから、時間変化する連続的なデータを扱うネットワーク可視化方法が必要とされる。本研究では時間変化する連続的なデータを扱うネットワーク可視化方法の提案と検証を行う。

ネットワークの可視化方法には何を目的として可視化するかにより異なる多くの可視化方法が存在するが、代表的な可視化方法として力学的方法の一種であるFR法[1]があげられる。FR法ではノードをランダムに選択し、ノードにかかる力を計算し、その力によりノード座標を更新する、この繰り返しによりノードごとにかかる力を積分して得られるポテンシャルエネルギーは下がる。FR法はポテンシャルエネルギー最小化を目的とした可視化方法である。ここで次のような状況を考えてみる。2つの異なるネットワークデータが存在する。このデータ1とデータ2でノードのIDが一致した場合は同じものを示すノードである。2つのネットワークデータはノードの接続関係はほとんど同じだが若干異なる（つまり、ネットワークを俯瞰したグローバルな可視化結果はあまり変わらないであろうが、ネットワークの特定の部分に注目したローカルな可視化結果は変化しているであろうデータ）。この2つのネットワークデータに対してFR法で可視化を行うとデータ1、データ2それぞれで得られるノード座標配置は同一IDでも異なる座標となる。したがって、データ1の可視化結果とデータ2の可視化結果を並べたとして、データ1であるノードに注目しても、データ2でそのノードがどこにあるか予測できないため、違いの比較がしづらい。

時間変化を考慮したネットワーク可視化方法を考える上で、今回想定するネットワークデータは、ノードを会社、リンクがノード同士の取引関係を示すこととする。こういったネットワークデータは業界ごとに異なる規模となるが、直接ネットワーク全体を可視化するには規模が大きすぎることが多い。その場合はいくつかのノードをまとめて1ノードに可視化する、ネットワークを階層化し部分的に可視化するなど、1度に可視化するノード数を減らしてから可視化することが多い。また、想定したネットワークデータについては、研究所や会社などの実際にネットワーク的分析を行いたい人々はある特定のノードに注目し、その周りでのノード同士の接続関係を調べ、他のノードとの違いを比較したいというニーズが存在する。そこで本研究において、ネットワークは可視化を行う以前に階層化や特定ノードに注目などで、同時に可視化を行う必要のあるノード数が小規模となっていることを前提とする。

2. 提案方法

以下に提案方法について説明する。

1. 扱うネットワークデータセットからデータごとにノードとエッジからなるグラフ G_0, G_1, \dots, G_n を作成する
2. グラフ G_i を準備する
3. G_i のノードの初期座標をランダムに配置する
4. G_i に対して FR 法を用いてノード座標配置を一定時求める
5. 求めたノード座標配置にしたがい G_i を描画する
6. 次に可視化する対象のグラフ G_{i+1} を準備する
7. G_i と G_{i+1} で次の条件を満たすノードをリスト L に入れる

(ア) 次数が減ったノード

(イ) 次数は変わらないが接続ノードが変わったノード

(ウ) 次数が増えたノードで下記の基準を満たすもの

8. G_{i+1} のノードの初期座標配置を G_i の描画した座標配置にする
9. G_{i+1} のノードに対して FR 法を用いてノード座標配置を求める。ただし、 L からランダムにノードを選択する
10. 求めたノード座標配置にしたがい G_{i+1} を描画する
11. $i = i+1$ として6に戻る

7. (ウ) の基準は

$C_j(G_i, G_{i+1}):G_i$ でのノード j の1近傍ノードかつ G_{i+1} でもノード j の1近傍ノードであるものの個数

$k_i(G_{i+1}):G_{i+1}$ でのノード j の次数

$$R_j = \frac{C_j(G_i, G_{i+1})}{k_j(G_{i+1})}$$

としたとき、 R_j の値が設定した値を下回るノードである。つまり、 G_{i+1} となり次数が増えたノードで、 G_i では1近傍になかった新しい接続先が増えたノードは L に入りやすくなる。たとえ、 G_{i+1} となり次数が増えたノードであっても、 G_{i+1} での1近傍の接続先ノードが G_i でのものと大差なければ L には入らない。

3. 実験

3.1 実験目的

提案方法を用いて小規模のネットワークデータセットを連続的に可視化し、振る舞いを調べる。ノードの接続関係が変わらない部分はノード座標があまり変化してほしくないなど、人の直感的な欲求に合致した振る舞いになっているかを確認する。

3.2 実験条件

ネットワークデータセットを作るため、現実世界に近いネットワークが生成できる CNN モデルを用いる。ノード数 10、エッジ生成確率 0.3 でネットワークを生成し、ランダムにエッジを切り替え、ネットワークデータセットを作

成する。エッジの切り替えは孤立ノードができない、自己ループができないように行う。また、今回の実験ではネットワークのノード数、エッジ数は変化しない。ノードの接続関係のみが変化する。

3.3 実験結果

以下に特徴的な変化を示す。

図1のネットワークデータにおいてノード7はノード0と接続している。図2のネットワークデータにおいてノード7の接続先ノード0との接続を切断し、ノード8を新たな接続先として追加した。図2以降のネットワークデータまでのノード座標更新はノード7とノード0のみである。図2はネットワークデータが切り替わった直後のノード座標配置を示したもので、各ノードは図1での座標となっている。

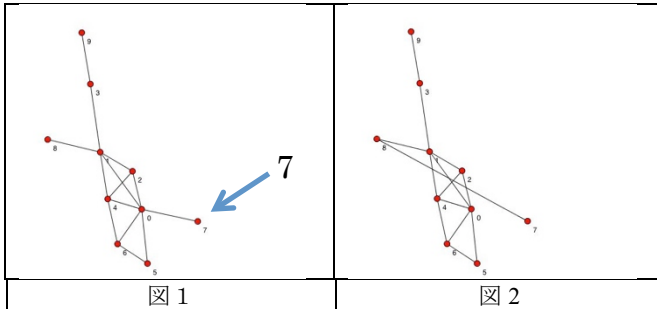


図2で用いたネットワークデータの最終的なノード座標配置は図3となる。カメラ位置が変わっているため、図2と比べて、ノード座標が更新されていないノード群の配置が変わった印象を受けるかもしれないが、それらノード群の座標は図2と同じである。実際動画を見てみるとノード7は徐々にノード8の近くへと座標が変わっていき、ノード0は少しだけ座標が変わり、それ以外のノードは座標が変わらないため、ノードの接続関係が変わらない部分はノード座標があまり変化してほしくないなど、人の直感的な欲求に合致した振る舞いであることが確認できる。

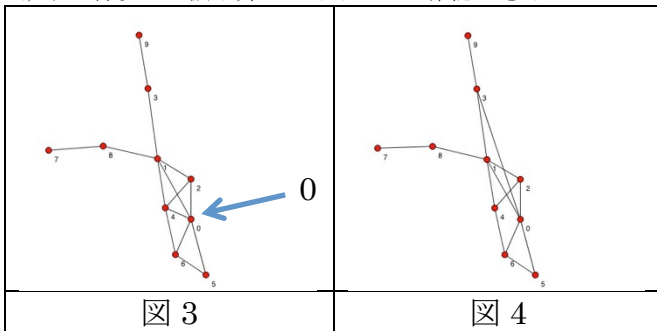
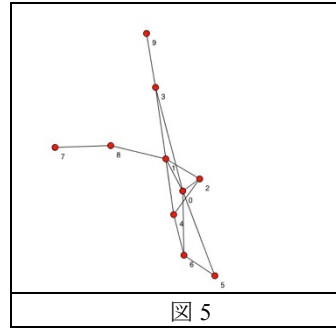


図3のネットワークデータにおいてノード0は左側にあるノード4と接続している。図4のネットワークデータにおいてノード0の接続先ノード4との接続を切断し、ノード3を新たな接続先として追加した。図3以降のネットワークデータまでのノード座標更新はノード0とノード4のみである。図4はネットワークデータが切り替わった直後のノード座標配置を示したもので、各ノードは図3での座標となっている。

図4で用いたネットワークデータの最終的なノード座標配置は図5となる。



動画で振る舞いを確認すると、ノード0は徐々にノード3の近くへと座標が変わっていき、ノード4は少しだけ座標が変わり、それ以外のノードは座標が変わらないため、ノードの接続関係が変わらない部分はノード座標があまり変化してほしくないなど、人の直感的な欲求にある程度合致している。しかし、ノード0とノード4は直接接続をしていないため、その2つのノードだけ考えるならば、もっと離れた配置に変化していったほうが、直接の接続関係がなくなったことがより強調される。

3.4 考察

図1から図2へのネットワークデータの変化は、提案方法で十分に人の直感に合致するよう表現できている。しかし、図3から図4のネットワークデータの変化は、提案方法ではまだ十分に人の直感に合致するような表現ができていない。図1から図2の変化は末端ノードの接続関係の変化、図3から図4の変化はハブ的なノードの接続関係の変化といえるが、ネットワーク規模を大きくしていった場合でも、今回の実験結果と同様にハブ的なノードの接続関係の変化を十分に表現できないかを調べる必要がある。

4. おわりに

本来想定していたネットワークデータを扱うにはノード数変化、エッジ数変化を考慮する必要がある。また、最終的な可視化を行うときは元のネットワークの局所的に可視化するのであることから、局所的にポテンシャルエネルギーを最小化するLEMやノード周りに動的に信号領域を設定しSOMによりノード座標配置を求めるDSSOMなど他の可視化方法を利用することも考慮する予定である。さらに、ノードにハブノードとの経路数やネットワーク特徴量をタグ付けし、その情報をノード座標の計算に利用することで、重要なノードの接続関係の変化とあまり重要でないノードの接続関係の変化の違いをより直感的に表現できるようになるであろう。これらの改良と合わせて、より大規模なネットワークデータへの適用、ネットワークモデルでの振る舞いの違い(ネットワークの特性による振る舞いの違い)を調査する予定である。

謝辞

本研究を行うにあたって、ネットワークの可視化により分析を行おうとしている方々の実際のニーズなど、貴重なご意見をくださった信州大学の松村先生に感謝いたします。

参考文献

- [1] T.Frucherman, E.Reingold, "Graph Drawing by Forcedirected Placement", Software Practice and Experience, Vol.21, No.11 (1991).