

深層学習を用いた階層型グラフィックレイアウト生成手法の検討

村上 綾菜[†] 伊藤 貴之[†]お茶の水女子大学[†]

1. はじめに

近年、個々のデータをノードで表現し、それらの関係性をエッジで表現する可視化手法であるグラフが多様な分野において使用されている。グラフはノードのレイアウトによって、可読性の高さが非常に大きく左右される。特に、複数のノードの集合(メタノード)で構成される階層型グラフは、大規模なグラフの概要を効果的に可視化する手法として知られているが、接続関係の複雑化、メタ情報考慮の必要性の観点から可読性の高いレイアウトの生成が非常に難しい。グラフのレイアウト生成には、力学モデルや次元削減手法を適用した手法など、従来から多くの手法が研究されてきた。しかし、従来の手法には各手法の強い特性があり、必ずしも自在なレイアウトが実現できるとは限らなかった。それに対して近年では、深層学習を用いたグラフィックレイアウト手法が活発に議論されている。一例として、学習モデルに LSTM (Long Short Term Memory) を用いてグラフィックレイアウトを生成する手法が発表されているが、この研究では、階層型グラフのレイアウト生成には至っていない。

グラフのレイアウト評価も、長らく議論される課題であり、数値評価と主観評価の両面にわたって多くの研究が発表されている。しかしその大半は非階層型グラフを対象としたものであり、階層型グラフの可視化結果に特化した評価手法の研究は非常に少なかった。それに対して Liu ら¹⁾は、階層型グラフに特化した新しい数値評価基準である *Sprawlter* を提唱する。

本報告では、階層型グラフのレイアウト改善に着目し、LSTM を用いた階層型グラフィックレイアウト生成モデルを提案する。具体的には、数値評価の高いグラフのみを学習データとした階層型グラフィックレイアウト生成モデルを用いて、レイアウトが不十分な階層型グラフのレイアウト改善に取り組む。ここでの学習データは、同一の接続関係を持つレイアウトが異なる階層型グラフ群とする。

学習モデルにより改善された階層型グラフィックレイアウトに対し、再び数値評価式を適用することで評価を行う。また、それらが実際に見た目にも良好な結果であるかについても議論する。

2. 関連研究

2.1 深層学習を用いたグラフィック生成

グラフは、画像データと異なり、隣接ノードの位置、エッジの長さやグラフ全体の大きさが不規則であるため、従来機械学習で扱うことが難しいとされてきた。しかし、You ら²⁾はグラフの学習に特化した RNN をベースとする学習モデルを提案し、従来のグラフを扱う学習モデルと比較し、スケーラビリティと精度の両方を向上させた。また、Wang ら³⁾は、You らの研究を発展させ、LSTM を学習モデルに起用し、複数のタイプのグラフのレイアウト生成を実現した。

2.2 階層型グラフィックレイアウトの評価

グラフ可視化に関する数値評価手法において、階層型グラフに特化した数値評価手法の研究は非常に少ない。その中でも Liu ら¹⁾は、階層型グラフィック可視化のための数値評価を「空間の浪費 (Sprawl)」と「配置の乱雑度 (Clutter)」の2つの観点から提唱し、2つのペナルティの積で可視化結果の品質を評価している。

本研究では、この数値評価手法を適用することで、訓練データとなる階層型グラフ群を評価し、さらに本手法によるレイアウト生成結果についても同様に評価する。

3. 実験

3.1 実験概要

本研究では、訓練データとして、複数の階層型グラフィックを用い、LSTM を用いたグラフィックレイアウト生成モデルを構築する。そこに、任意の階層型グラフィックを入力として与えることで、レイアウトを改善した階層型グラフィックを生成し、Liu らの数値評価式に基づいて生成結果を評価する。

3.2 使用データの生成

本研究において、我々は異なるレイアウト結果を有する階層型グラフィックデータ群を生成し、それらを深層学習において使用した。階層型グラ

A Hierarchical Graph Layout Generation Method

Ayana Murakami, Takayuki Itoh
[†] Ochanomizu University

フデータ群の生成には、村上らの手法⁴⁾を採用している。生成した階層型グラフデータ群は、同じ接続情報をもつが、それぞれレイアウトが異なる。また、ノード数及びエッジ数は全て共通である。また、これらデータ群を訓練データ・バリデーションデータ・テストデータの3つにランダムに分類し、使用した。

3.3 学習モデル

本研究では、LSTMを採用した学習モデルによって階層型グラフのレイアウトを生成する。従来のRNNと比較した大きな違いとしてLSTMは直前の依存関係だけでなく、長期的な依存関係を保持できる。ノード情報だけでなくメタノード情報を保有する階層型グラフに深層学習を適用することを考えたときに、我々はこの階層型グラフの特徴に対して、前述のLSTMの特徴は非常に親和性が高いと考えた。

3.2節で述べた訓練データの階層型グラフからは、各ノードの情報とノード間の接続情報を抽出し深層学習に使用する。

4. 結果と考察

4.1 実験結果

深層学習モデル適用後の階層型グラフ生成結果をグラフ情報および画像データの2つの形で出力した。そして、得られたグラフ情報に対して数値評価式を改めて適用することで新たに生成したレイアウトを評価した。

生成結果のグラフレイアウトの中で、数値評価結果が相対的に高いグラフを図1に示す。

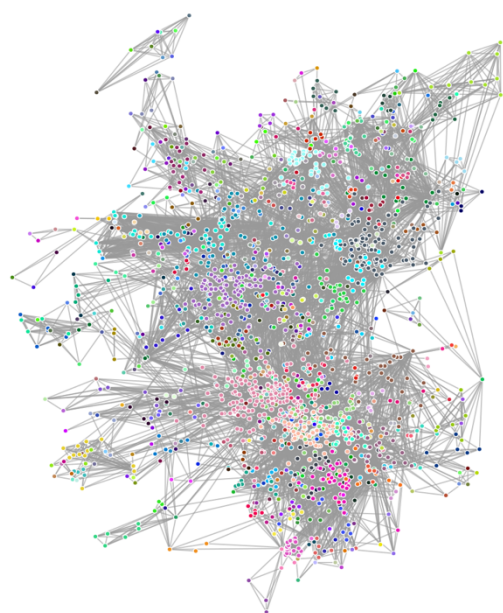


図1. 実験結果の中でレイアウトの評価が高い階層型グラフの一例

4.2 考察

エッジの重なりが多く見られるものの、同じ色を持つノードの色が近くに配置されていることが読み取れる。ノードの色が同じノードは、同じメタノードに属することを示す。特に、グラフ中央付近の紫、ピンクのノードが示すメタノードに着目については、その特徴が顕著に見られる。各ノードの属性情報がグラフレイアウトに反映できていることから、本実験の学習モデルが階層型グラフに有用であると示唆される。

5. まとめと今後の課題

本報告では、深層学習を用いた階層型グラフのレイアウト改善の一手法を提案した。また、学習モデル適用後に得られた新たなレイアウトの階層型グラフを、階層型グラフに特化した数値評価式を用いて評価した。評価値の高いレイアウトにおいて同じ属性を持つノードが近くに配置されたことから学習モデルの有用性を評価した。

今後の課題として、深層学習モデルの改善取り組み、特に損失関数の改善を検討する。本実験では、先行研究にならい深層学習の損失関数にプロクラステス解析を用いているが、形状が複雑な階層型グラフの評価にプロクラステス解析が有用とはいきれない。今後はコスト関数にも階層型グラフに特化した数値評価式 **Sprawlter** を適用することで学習効率をあげ、より評価の高い階層型グラフレイアウトを生成したい。

参考文献

1. Z. Liu, T. Itoh, J. Q. Dawson and T. Munzner, The Sprawlter Graph Readability Metric: Combining Sprawl and Area-aware Clutter, IEEE Transactions on Visualization and Computer Graphics, Vol. 26, No. 6, pp. 2180-2191, 2020.
2. J. You, R. Ying, X. Ren, W. Hamilton and J. Leskovec, GraphRNN: Generating Realistic Graphs with Deep Auto-regressive Models, Proceedings of the 35th International Conference on Machine Learning, pp. 5708-5717, 2018.
3. Y. Wang, Z. Jin, Q. Wang, W. Cui, T. Ma and H. Qu, DeepDrawing: A Deep Learning Approach to Graph Drawing, IEEE Transactions on Visualization and Computer Graphics, Vol. 26, pp. 676-686, 2020.
4. 村上, 伊藤, 深層学習を用いた階層型グラフレイアウトの実験, 可視化情報シンポジウム, 2021.