

# 階層型グラフのレイアウト最適化

理学専攻 情報科学コース 2140672 村上 綾菜 (指導教員：伊藤 貴之)

## 1 はじめに

グラフはデータ要素をノードで表現し、それらの関係性をエッジで表現する。近年ではグラフの可視化手法が多様な分野にて使用されている。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が、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。また近年では、深層学習を用いたグラフレイアウト手法が活発に議論されているが、階層型グラフのレイアウト生成には至っていない。

グラフレイアウトの評価も、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、階層型グラフレイアウトに特化した評価手法の研究は非常に少ない。そこで 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 を提唱している。

本研究では、階層型グラフレイアウトに特化した数値評価式 Sprawlter を目的関数とした数理最適化手法を用いることで、階層型グラフレイアウトを最適化する方法を提案する。

## 2 階層型グラフレイアウトとその評価

ノードの関係性をエッジで表現する可視化手法として、グラフレイアウトが活発に議論されており、特に、似た属性の複数のノードがクラスタを形成するグラフを階層型グラフと呼ぶ。Itoh ら [2] は、階層型グラフレイアウト手法 Koala を提案した。Koala では、各ノードの属性に応じてクラスタリングを適用し、同じクラスタに属するノードを近くに配置する。

グラフレイアウトの評価は重要な課題であり、これまでもグラフのトポロジーに着目する指標や、レイアウトの審美性に着目する指標などさまざま提案されている。一方で、階層型グラフに特化したグラフレイアウトの評価指標は少なかった。この課題に対し、Liu ら [1] は、階層型グラフレイアウトの見た目に対する数値評価指標 Sprawlter を提案した。Sprawlter は、空間の浪費を評価する Sprawl と、画面の乱雑度を評価する Clutter の 2 つの指標で構成される。この論文では、評価式の提案と有効性の検証が行われていたが、グラフレイアウトの生成や最適化の目的では使用していない。本研究では、グラフレイアウトの数理最適化の目的関数に評価式 Sprawlter を用いる。

## 3 提案手法

本手法の全体像を図 1 に示す。本手法では、階層型グラフレイアウト描画手法 Koala によるレイアウト描画を反復し、複数の階層型グラフレイアウトを作成したのち、遺伝的アルゴリズムを用いてレイアウトの最適解を探索する。最適化の過程にて、階層型グラフレイアウトに特化した数値評価指標である Sprawlter をレイアウトの評価に採用する。

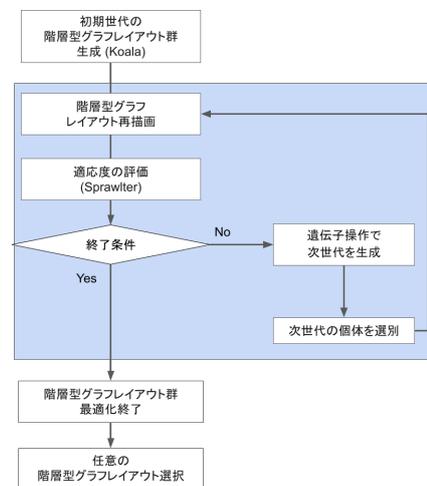


図 1: 本手法の全体像

### 3.1 遺伝的アルゴリズム: NSGA-II

本手法では、レイアウトの最適化において遺伝的アルゴリズムの一種である NSGA-II[3] を採用する。NSGA-II は、非優越ソートと混雑度ソートという 2 種類のソートを用いて次世代に残す遺伝子を選択することが大きな特徴である。非優越ソートは、従来よりも短い時間で前の世代を優越する良い解が得ることを可能にし、混雑度ソートは解の多様性を保持したまま次の世代に良い解を残すことを可能にする。

### 3.2 遺伝子の設計

遺伝的アルゴリズムにおいて、遺伝情報を伝える実態を染色体とよび、その形質を決定するコードを遺伝子と呼ぶ。解の候補は染色体で表現され 1 つの個体が 1 つの解に対応する。本手法では、メタノードの座標値を遺伝子とし、1 つの個体は 1 つのグラフレイアウトに対応する。メタノード  $x$  座標と  $y$  座標のペアを、交互に 1 次元配列に格納することで 1 次元の遺伝子を生成する。

## 4 実行例

### 4.1 使用データ

1998年から2013年までに NERC Biomolecular Analysis Facility (NBAF) で発表された論文に対して、著者をノード、共著関係をエッジとしたグラフを可視化する。

### 4.2 レイアウト最適化結果

最適化の過程として図2に初期世代と最終世代の評価値の比較を示す。横軸は Sprawl の評価値、縦軸は Clutter の評価値であり、ともに評価値が小さいほど評価は高い。それぞれの点が1つの個体、つまり1つの階層型グラフレイアウトを示し、赤い点が初期世代、青い点が最終世代の個体の評価値である。なお、初期世代の番号と該当代の番号に対応関係はない。

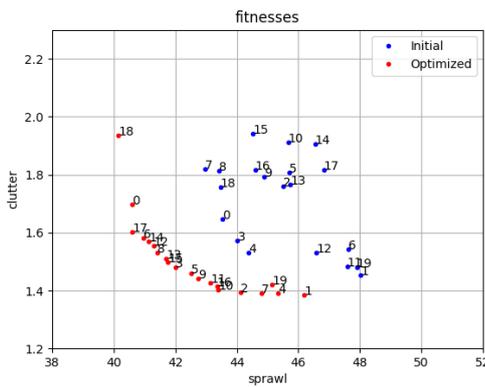


図2: 初期世代と最終世代の評価値の比較

図2における青い点4番に対応するレイアウトの一部を図3に、赤い点11番に対応するレイアウトの一部を図4に示す。レイアウト全体は付録を参照されたい。

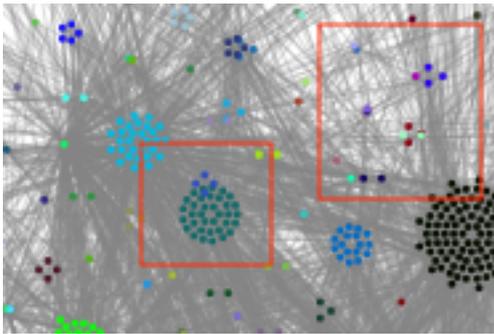


図3: 青い点4番に対応するグラフレイアウトの一部

初期世代(青い点)のレイアウト4番は、図2によると初期世代の中では評価が高い。しかし、図3のオレンジの枠で囲んだ箇所には要素の重なりが散見される。一方で、最適化後のレイアウトの一部である図4を見るとメタノード同士の干渉がほぼなく、異なるメタノードに属するノードが適度に離れて配置されていることがわかる。またそれによりエッジ同士の重なりも減少していることが観察できる。

また、最適化前後における Sprawl および Clutter の評価結果を図5aと図5bにそれぞれ示す。これより、

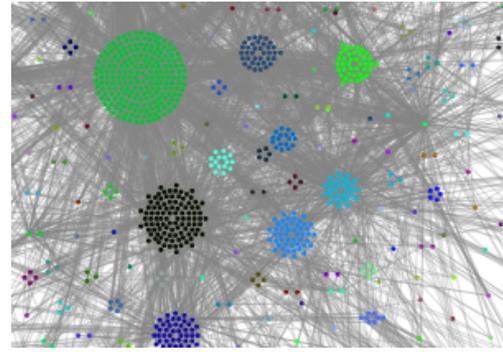


図4: 赤い点11番に対応するグラフレイアウトの一部

最適化前後において全体的に評価が良くなったことが確認できる。一方で、最適化後であっても最大値はほとんど変わらないことから、最適化後も一定数評価の悪いレイアウトが含まれることがわかる。以上のことから、最適化後のレイアウト群の中からユーザーが好みのレイアウトを選択するプロセスが重要になると考えられる。

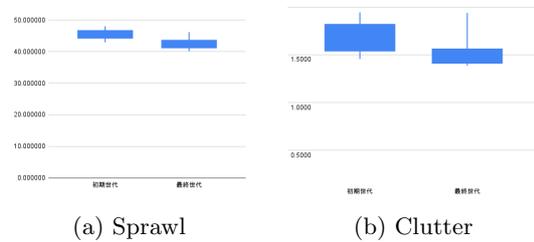


図5: 最適化前後における各評価値の比較

## 5 まとめ

本研究では、遺伝的アルゴリズムと階層型グラフレイアウトの数値評価式 Sprawlter を用いて階層型グラフレイアウトの最適化を行った。これにより、レイアウトを構成するメタノードの干渉が減り、レイアウト全体の評価値も向上したことを確認した。しかし、本実験で用いた2つの目的関数のうち、Sprawlはそれぞれの個体に対する評価値の差異はほぼなかった。今後は目的関数を再設計し、レイアウト最適化の有効性をより明確に視認できるようにしたい。

## 参考文献

- [1] Z. Liu, T. Itoh, J. Q. Dawson, 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] T. Itoh, K. Klein, "Key-node-separated Graph Clustering and Layouts for Human Relationship Graph Visualization.", IEEE Computer Graphics and Applications, vol. 35, no. 6, pp. 30-40, 2015.
- [3] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II." IEEE transactions on evolutionary computation, vol. 6, no.2, pp. 182-197, 2002.

(付録) グラフレイアウトの全体像

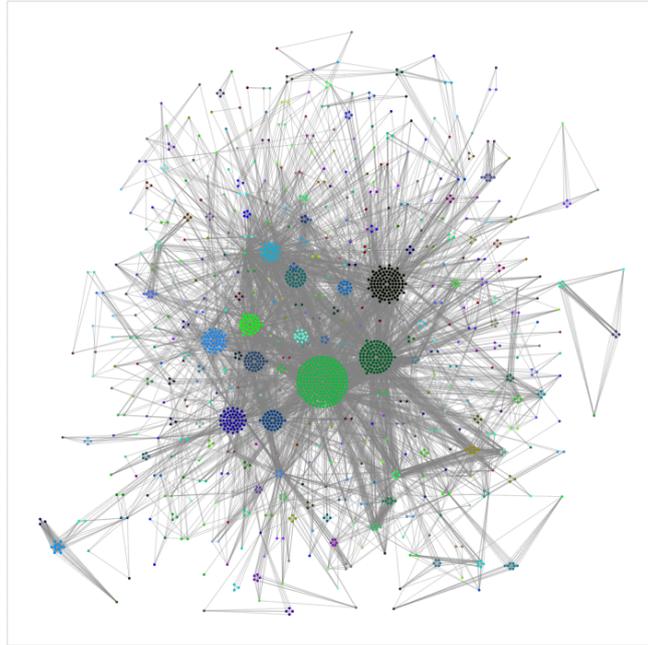


図 6: 図 2 の青い点 4 番に対応する階層型グラフレイアウト

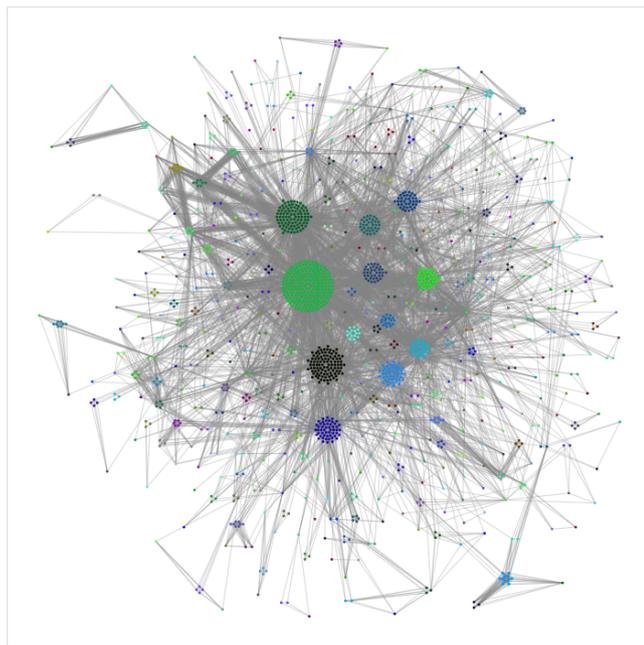


図 7: 図 2 の赤い点 11 番に対応する階層型グラフレイアウト