

類似度に基づき成長する複雑ネットワーク生成モデルの提案

日大生産工(院) ○寺山 敬佑

日大生産工 山内ゆかり

1 まえがき

近年注目されている複雑ネットワークには、現実のネットワークの特徴を表現した代表的なモデルとして、スモールワールドネットワークとスケールフリーネットワークがある。

スモールワールドネットワークの生成は、WattsとStrogatzらによってレギュラグラフからの再結合により構成する手法(WSモデル)が提案されている[1]。また、スケールフリーネットワークの生成は、BarabasiとAlbertらにより、新規ノードの結合に制約を持たせながらネットワークを成長させる手法(BAモデル)が提案されている[2]。

BAモデルでは現実のネットワークで現れる次数分布がべき則となるスケールフリー性が見られるが、もう一つの現実のネットワークの特徴である、近くのノードと結合されやすいというクラスター性を満たしていない。このことから、ネットワークを成長させてクラスター性の高いスケールフリーネットワークを生成する手法が盛んに研究されてきた。代表的なモデルに頂点コピーモデル[3][4][5][6]、適応度モデル[7][8]、頂点非活性化モデル[9]などが知られている。

また現実のネットワークはスモールワールド性を示すことが知られている。脳は細胞分裂により成長させながらネットワークを構成し、学習により動的に構造を変化させ、シナプス結合の物理的構造や機能のネットワークでスモールワールド性やスケールフリー性を持つことが知られている[10][11]。しかし、WattsとStrogatzらが提案したWSモデルのような、再結合によるネットワークは現実のネットワークのモデルとしては不自然である。

本研究では、新規ノードの結合に類似度に基づく優先選択を取り入れた制約を与え、成長させながらネットワークを構成し、スモールワールドネットワークを生成する方法を提案する。

提案手法は大脳皮質の視覚野をモデルにしたKohonenによる自己組織化マップのアルゴリズム[12]を基にしている。また現実のネットワークでは、ソーシャルネットワークサービスのようなネットワークに新たに参加する場合、自分と共通点の多い友人などの紹介で参加し、またその友人の知り合いの中で趣味や所属など共通点のある人物と関係を持つような、現実のネットワークが構成される様子を想定している。

提案手法により構成したネットワークと従来の構成法によるネットワークを複雑ネットワークの指標を用いて比較し、スモールワールド性、スケールフリー性、及び次数分布、またハブノードの与える影響について検討する。

2 ネットワークの特徴量

ネットワークの特徴量について説明する。ネットワークの特徴量としては、クラスター係数、平均頂点間距離、次数分布が挙げられる。これ以降の説明では、ネットワークのノード数を N 、平均次数を K とする。

2. 1 クラスター係数

クラスター係数 C とは、自分と繋がっているノード同士が繋がっているかどうかを表す係数である。クラスターとは群れ、集団といった意味である。

クラスター係数は次の式で求められる。

$$C = \sum_{i=1}^N C_i \quad (1)$$

C_i は i 番目のノードと繋がっているノード同士が結合されている枝の本数 m を、 $k_i \times (k_i - 1) / 2$ で割ったものと定義する。 k_i は i 番目のノードが持つ枝の本数(次数)である。 C_i は次の式で求められる。

$$C_i = \frac{m}{\frac{k_i \times (k_i - 1)}{2}} \quad (2)$$

2. 2 平均頂点間距離

平均頂点間距離 L は2点間最短距離 d_{ij} の全体にわたる平均である。自分への距離は0とし、平均には含まない。以上のことより平均頂点間距離は次の式で求められる。

$$L = \frac{1}{\frac{N \times (N - 1)}{2}} \sum_{i=1}^N \sum_{j=1}^N d_{ij} \quad (3)$$

3 代表的な複雑ネットワーク

複雑ネットワーク研究における代表的なネットワークについて説明を行う。従来のグラフ理論をもとにしたネットワーク(グラフ)は、レギュラグラフ

Proposal for a Growing Complex Network Generated by Similarities

Keisuke TERAYAMA and Yukari YAMAUCHI

フ、ランダムグラフがある。複雑ネットワークの代表的なモデルとして、スモールワールドネットワーク(WSモデル) [1]、スケールフリーネットワーク(BAモデル) [2]、また成長させながらクラスター性の高いスケールフリーネットワークを生成する頂点非活性化モデル [9]を取り上げる。

3. 1 スモールワールドネットワーク(WSモデル)

スモールワールドネットワークの生成は、WattsとStrogatzらによってレギュラーグラフからの再結合により構成する手法(WSモデル)が提案されている [1]。

3. 2 スケールフリーネットワーク(BAモデル)

スケールフリーネットワークの生成は、BarabasiとAlbertらにより、新規ノードの結合に次数の大きいノードと優先時に結合する、という制約を持たせながらネットワークを成長させる手法(BAモデル)が提案されている [2]。

3. 3 頂点非活性化モデル

頂点非活性化モデル [13]は、BAモデルのネットワークの成長と優先選択に加え、ノードの非活性化を考慮する。各ノードは加入時点では活性化状態であり、枝を受け取れるが、ある時点で非活性化状態になり、非活性化状態では永久に枝を受け取れなくなる。

4 提案手法

提案手法のネットワークモデルのアルゴリズムについて説明する。

提案手法のアルゴリズムは、結合するノードの選択に属性値の類似度に応じた確率的優先選択を取り入れる。この属性値は現実のネットワークではノードの座標や、個人の所属、年齢など、そのノードを特徴づける数値を意図している。現実のネットワークではこの属性に何らかの共通点を持つ場合に関係があることが想定されるので、属性値の類似度で優先的に結合することを採用した。

アルゴリズムは以下ようになる。

(1) N個のノードそれぞれにM個の属性値を与える。i番目の属性値は $(A_{i1}, A_{i2}, \dots, A_{im})$ となる。

(2) s個のノードで完全グラフを作成する。この時 $s=K+1$ とする。

(3) ネットワークに新たに追加するノードをxとする。既存のノード数とした時、ノードxの属性値との類似度が最小となるノードwを求め勝者ノードとする。 d_{ix} は以下の式で求める。つまりユークリッド距離を用いる。

$$d_{ix} = \sqrt{\sum_{j=1}^M (A_{ij} - A_{xj})^2} \quad (4)$$

この勝者ノードとxを結合する。

(4) (3)で結合した勝者ノードの隣接ノード (近傍) を対象として類似度に基づく確率選択により $K/2-1$ 個のノードと結合 (学習) する。これは自己組織化マップにおいて勝者ノードの近傍ノードが対象ノードに近づくよう修正される近傍学習を参考にしている。勝者

ノードのi番目の隣接ノードが選択される確率 $P(i)$ は以下の式で求められる。

$$P(i) = \frac{\sqrt{M \times A_{\max}^2 - d_{ix}}}{\sum_{j=1}^{k_w} (\sqrt{M \times A_{\max}^2 - d_{jx}})} \quad (5)$$

ここで $A_{\max}$ は1~N'個目の属性値の最大値、 k_w は勝者ノードの次数である。

(5) (3)(4)の作業をノード数がN個に達成するまで繰り返し行う。

以上が提案手法のアルゴリズムである。従来の成長しながらネットワーク生成を行うBAモデルや頂点非活性化モデルとの違いは、勝者ノードを選択してそのノードの隣接ノードの中で確率選択を行っていること、優先選択の基準に次数ではなく、類似度を用いていることが挙げられる。

5 実験方法

提案手法により生成したネットワークを、従来のネットワーク(レギュラーグラフ、ランダムグラフ、スモールワールドネットワーク(WSモデル)、スケールフリーネットワーク(BAモデル)、頂点非活性化モデル)と比較を行い、スモールワールド性、スケールフリー性を検討する。比較には、複雑ネットワークの特徴を測る指標である、クラスター係数C、平均頂点間距離L、次数分布 $p(k)$ を用いる。

ランダムグラフ、スモールワールドネットワーク(WSモデル)、スケールフリーネットワーク(BAモデル)、頂点非活性化モデル、提案手法は、生成モデルに乱数を使用しているため、CとLの値には100試行の平均値を採用している。WSモデルの張り替え確率PはCが大きくLが小さい値を示す標準的な値である0.1を採用した。また、頂点非活性化モデルの正の定数aは1を採用した。提案手法の属性値の数Mは2を採用した。

6 実験結果

以下に実験結果を示し検討を行う。

6. 1 ネットワークの特徴量による比較及び検討

ネットワークの特徴量である、クラスター係数C、平均頂点間距離Lによる比較を行う。まず表で100試行の数値を示す。

まず、平均次数 $K=12$ と固定し、ノード数Nを変化させ実験を行った。Nが大きくなれば枝の密度が疎となるため、Cが小さくLが大きくなる傾向がある。

表1 N=1024でのCとL

モデル	C	L
レギュラーグラフ	0.6818	43.1261
ランダムグラフ	0.0117	3.0438
WSモデル (P=0.1)	0.5017	4.0507
BAモデル	0.0536	2.8183
頂点非活性化モデル	0.8010	7.0334
提案手法	0.5595	3.2929

表2 N=3600でのCとL

モデル	C	L
レギュラーグラフ	0.6818	150.4585
ランダムグラフ	0.0034	3.5685
WSモデル(P=0.1)	0.5010	4.8787
BAモデル	0.0207	3.1712
頂点非活性化モデル	0.8018	20.2001
提案手法	0.5585	3.8796

表1-2より、提案手法のCは、ノード数を変えてもWSモデルとほぼ同じ値を示している。この値は頂点非活性化モデルの値より小さいが、ランダムグラフ、BAモデルの値より遥かに大きく、十分大きいCを実現していると言える。また提案手法のLはWSモデルと比較すると小さい値を示している。ランダムグラフ、BAモデルの値よりは大きい、十分小さいLを実現していると言える。

次に $K=\sqrt{N}$ とし、ネットワーク規模を変えて実験を行った。

表3 N=1024,K=32でのCとL

モデル	C	L
レギュラーグラフ	0.7258	16.4848
ランダムグラフ	0.0312	2.3248
WSモデル(P=0.1)	0.5337	2.7596
BAモデル	0.1035	2.2730
頂点非活性化モデル	0.8172	2.8949
提案手法	0.3918	2.4060

表4 N=3600,K=60でのCとL

モデル	C	L
レギュラーグラフ	0.7373	30.4918
ランダムグラフ	0.0167	2.3441
WSモデル(P=0.1)	0.5401	2.7813
BAモデル	0.0640	2.2763
頂点非活性化モデル	0.8244	3.0641
提案手法	0.2757	2.4059

提案手法のCは、Kが大きくなるにつれ、Cが小さくなる傾向が見られるが、ランダムグラフやBAモデルよりは十分大きな値を示している。また提案手法のLは、WSモデルと比較すると小さい値を示している。ランダムグラフ、BAモデルの値よりは大きい、十分小さいLを実現していると言える。

以上のことから、提案手法により生成したネットワークは大きいCと小さいLの特徴を示していると言える。

6. 2 次数分布による比較及び検討

ランダムグラフ、スモールワールドネットワーク(WSモデルP=0.1)、スケールフリーネットワーク(BAモデル)、頂点非活性化モデル、提案手法の次数分布の比較を行う。N=3600、K=12とし、横軸を次数、縦軸を出現頻度とした対数グラフを図1に示す。

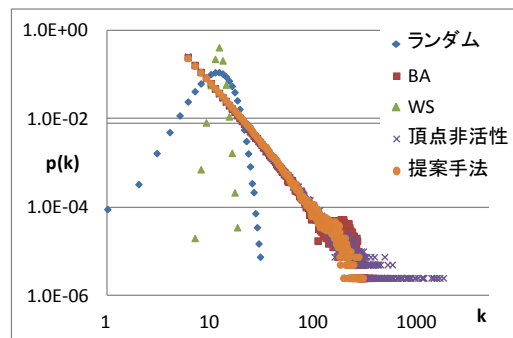


図1 ネットワークの次数分布

図1に従来の生成法によるネットワークと提案手法の次数分布を示す。提案手法では勝者ノードの隣接ノードを類似度により優先選択を行っているため、BAモデルや頂点非活性化モデルと同様にべき則の次数分布が得られている。

6. 3 スケールフリー性に基づく検討

6.2で述べたように、提案手法はべき則を示した。これは、スケールフリー性とも呼ばれる。スケールフリー性を持つネットワークは次数の高いハブノードに結合が密集しているため、ハブノードが攻撃を受けた際、CやLに影響を受けることが知られている[14]。

スケールフリー性を持つBAモデル、頂点非活性化モデル、提案手法で生成したネットワークのハブノードを破壊した場合のCとLに見られる影響について比較を行う。まず、次数の高いノードを5個壊した時の比較を行った。この破壊とは、該当ノード及びそのノードから出ている枝をネットワークから除く作業を意味する。比較結果が表5,6である。

この時、ネットワークから除いた枝数にばらつきが見られた為、ネットワーク内の枝数を10%減少させるまで次数の高いノードを破壊した時の比較も行った。比較結果が表7,8である。表5,6,7,8のデータは100試行の平均値である。

表5 N=1024 K=12 壊したノード数5の時のC

モデル名	C	破壊後のC	増減比
BA	0.054	0.034	0.632
頂点非活性化	0.801	0.779	0.973
提案手法	0.560	0.560	1.002

表6 N=1024 K=12 壊したノード数5の時のL

モデル名	L	破壊後のL	増減比
BA	2.818	2.951	1.047
頂点非活性化	7.033	13.737	1.953
提案手法	3.293	3.674	1.116

表7 N=1024 K=12 ネットワーク内の枝数10%減少の時のC

モデル名	C	破壊後のC	増減比
BA	0.054	0.021	0.391
頂点非活性化	0.801	0.768	0.959
提案手法	0.560	0.550	0.983

表8 N=1024 K=12ネットワーク内の
枝数10%減少の時のL

モデル名	L	破壊後のL	増減比
BA	2.818	3.123	1.108
頂点非活性化	7.033	18.002	2.560
提案手法	3.293	4.517	1.372

表5,6より、BAモデルはLの変化は少ないが、Cは大きく減少している。頂点非活性化モデルはCの変化は少ないが、Lは大きく増加している。これは、増減の割合を見ても明らかである。提案手法を見てみると、C、L共にあまり変化が見られない。増減の割合も1に近いままである。また表7,8からも、BAモデルはLに、頂点非活性化モデルはCに大きな変化が見られている。提案手法は他の2つのモデルより壊れたノード数が多いが、C、L共に他の2つほど変化は見られていない。

以上のことから、提案手法によるネットワークは、次数の高いノードを破壊しても、BAモデル、頂点非活性化モデルと比較して、CやLの影響が少ないと言える。このことから、同じべき則を持つネットワークであってもハブノードへの依存度が低いのではないかと考えられる。これは提案手法が優先選択を行う際の優先度が次数のような絶対的な値ではなく、現在追加するノードとの類似度という相対的な値で行っていることが関係すると考えられる。

7 まとめ

6章で述べたように、類似度に基づき成長させながらネットワークを構成することで、スモールワールドネットワークの特徴（大きいC、小さなL）と次数分布がべき則に従うネットワークを生成することができた。さらに、6.3より、提案手法が生成したネットワークは次数の高いノードの破壊によっても、CとLともに影響があまり見られなかった。これは現実のネットワークの頑強性と共通しており、提案手法によるネットワークがより現実のモデルを忠実に表現していると考えられる。

今回の提案手法で属性値の数を2としたのは、属性値を2次元の座標(x,y)と見なせることができる為である。このため、提案手法により生成したネットワークが2次元レギュラグラフを基に張り替えを行った、2次元スモールワールドネットワークに近似することが予想される。また、M=3の場合でも提案手法によるネットワークを生成して実験を行ったがM=2の場合と顕著な差は見られなかった。

今後の課題として、提案手法によるネットワークを、連想記憶などの応用に採用して、他のネットワークモデルとの比較及び評価を行うことが挙げられる。

「参考文献」

- [1] Watts, D.J. and Strogatz, S.H.: Collective dynamics of small-world networks, Nature, 393, pp.440-442, (1998)
- [2] Barabasi, A.-L. and Albert, R.: Emergence of scaling in random networks, Science, 286, pp.509-512 (1999)

[3] Kleinberg, M. J., Kumar, R., Raghavan, P., Rajagopalan, S. and Tomkins, A.S.: The web as a graph: Measurements, models, and methods, Lecture Notes in Computer Science, Vol. 1627, pp.1-17(1999)

[4] Krapivsky, L.P. and Redner, S.: Network growth by copying, Physical Review E, Vol. 71, article No. 036118(2005)

[5] Kumar, R., Raghavan, P., Rajagopalan, S., Sivakumar, D., Tomkins, A.S. and Upfal, E.: Stochastic models for the web graph, Proceedings of the 41st Annual IEEE Symposium on the Foundations of Computer Science, pp.57-65(2000)

[6] Sole, R. V., Pastor-Satorras, R., Smith, E. and Kepler, T.B.: A model of large-scale proteome evolution, Advances in Complex Systems, Vol. 5, pp.43-54(2001)

[7] Bianconi, G. and Barabasi, A. -L.: Competition and multiscaling in evolving networks, Europhysics Letter, Vol. 54, pp.436-442(2001)

[8] Bianconi, G. and Barabasi, A.-L.: Bose-Einstein condensation in complex networks, Physical Review Letters, Vol. 86, pp.5632-5635(2001)

[9] Klemm, K. and Eguiluz, V.M.: Highly clustered scale-free networks, Physical, Review E, 65, 036123 (2003)

[10] Bassett, D.S. and Bullmore, E.: Small-World Brain Networks, Neuroscientist, 12(6), pp.512-523 (2006)

[11] T. Suzuki, and T. Ikeguchi: Self-Organizing Small-World Structure of Neural Networks by STDP Learning Rule, 同志社大学理工学研究報告, Vol. 49, No.4, pp.200-204 (2009)

[12] Kohonen, T.: Self-Organizing Maps, Springer (2000)

[13] 増田直紀 今野紀雄, “頂点非活性化モデル, “複雑ネットワーク 基礎から応用まで, pp.117-122, 近代科学社 (2010)

[14] Barabasi, A.-L. and Albert, R.: Statistical mechanics of complex networks, Review of Modern Physics, Vol.74, pp.44-97(2002)