

ネットワークの隔たり数とエントロピー

豊田 規人†
Norihito Toyota坂本 智晴†
Tomoharu Sakamoto

1. はじめに

これまで、筆者等は青山等が提案したミルグラム条件[1]が6次の隔たりの鍵となる指標と考え、ネットワークの数学的構造を研究してきた。それを使って、スモールワールドネットワーク[2]とスケールフリーネットワーク[3]において、ミルグラム条件と一般化クラスタリング係数の関係が全く違った振る舞いをする事を示した[4]。この事は、情報の伝達に対する閉路の効果が、両ネットワークでは、全く違う事を意味している。この論文では、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の中間的ネットワークモデルを2種類考察し、その違いを明確化する。さらに次数分布のエントロピーが情報伝達に大きな影響を与えている事を示す。これにより、それらのネットワークの情報伝達における閉路の影響の相違が統一的に理解できる。

2. 2つのネットワークの比較

2つのネットワークを比較する前に、ミルグラム条件とそれに関連する概念について解説する。

- (1) j -string とは、ネットワーク上で、 j ノードからなる紐状の部分である。
- (2) S_j とはネットワーク上の j -string の本数を示す。
- (3) \bar{S}_j とは非縮退[5]-[7]の j -string の本数を示す。
- (4) p 次一般化クラスタリング係数 $C_{(p)}$ [5]-[7]を以下の式で導入する。

$$C_{(p)} = \frac{2p \times \text{number of polygons}}{\text{number of connected } p\text{-plets}} = \frac{2p\Delta_p}{\bar{S}_p} \quad (1)$$

ただし Δ_p はネットワーク内で p 本の枝を持つ多角形構造の数である。従って、 $C_{(p)}$ はネットワーク内の多角形構造の量を表す一つの指標となる。これらの概念および指標は、隣接行列を使い統一的に記述される[5]-[7]。

- (5) ミルグラム条件 M_n

$$M_n \equiv \frac{\bar{S}_n}{N^2} \sim O(1). \quad (2)$$

N はネットワークサイズ (ノード数) である。この条件を満たせば、 n 次の隔たりが成立していると考えられる[1]。この事から n を隔たり数と呼ぶ事にする。尚、 M_n が(2)式を満たす値より大きくなった場合は、より n 次の隔たりが容易になる。

2.1. 2つのネットワークのミルグラム条件

一般化クラスタリング係数の和 X_n を次の式で定義する。

$$X_n \equiv \sum_{p=3}^n C_{(p)}. \quad (3)$$

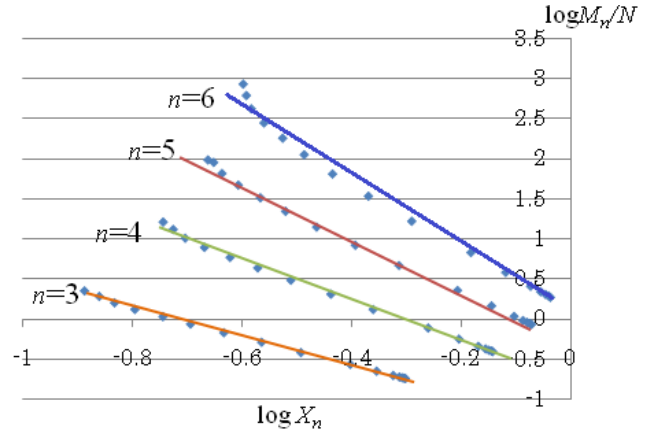


図1 スモールワールドネットワークの X_n と M_n の関係

この X_n とミルグラム条件の左辺 M_n の関係をスモールワールドネットワークで、隔たり数 n ($n=3\sim 6$) ごとに評価したものを図1に示す[8]。ただし、スモールワールドネットワークはニューマン・ワッツモデルで作成した[9]。

図1を見ると、 X_n が増加すると M_n が小さくなっている事が分かる。 X_n が増加するという事は、 p 次一般化クラスタリング係数 $C_{(p)}$ が平均的に増加している事を示している。また、 M_n が小さくなっているという事は n 次の隔たりが起こりにくくなっている事を示している。これらの事から、スモールワールドネットワークでは、 $C_{(p)}$ が大きくなると情報が伝わりにくくなる事が分かる。つまりスモールワールドネットワークでは、ネットワーク内の閉路が、情報の伝達を阻害している事が分かる。図2は X_n と M_n の関係をスケールフリーネットワークについて評価したものである。ただし、スケールフリーネットワークはコンフィグレーションモデルを使って作成した[8]。スケールフリーネットワークの次数分布は $p(k) \propto k^{-\gamma}$ で与えられるが、図2ではスケールリング指数 γ は $1.8 \leq \gamma \leq 4$ の範囲で評価している。また、図2のグラフでは、主に左下側のデータ点は γ が大きい場合のものであり、右上側のデータ点は γ が小さい場合のものである。このようにスケールフリーネットワークでは、クラスタリング係数 $C_{(p)}$ の値は、スケールリング指数 γ に依存している。尚、図1では隔たり数ごとにデータ点がストライプ状の構造に分かれたため、4本の直線状のグラフになったが、図2ではすべての隔たり数($n=3\sim 6$)で一本の直線状のグラフとなっている。

図2を見ると横軸の X_n が増加すると M_n が大きくなっている事が分かる。 X_n が増加するという事は、 $C_{(p)}$ が平均して増加している事を示している。また、 M_n が大きくなっているという事は、 n 次の隔たりが満たしやすくなっている事を示している。これらの事から、スケールフリーネットワークでは、クラスタリング係数 $C_{(p)}$ が大きくなると情報が伝わりやすくなる事が分かる。つまりスケールフリーネットワークでは、ネットワーク内の閉路の数の増加が情報伝達を促進していることが分かる。結局、

† 北海道情報大学

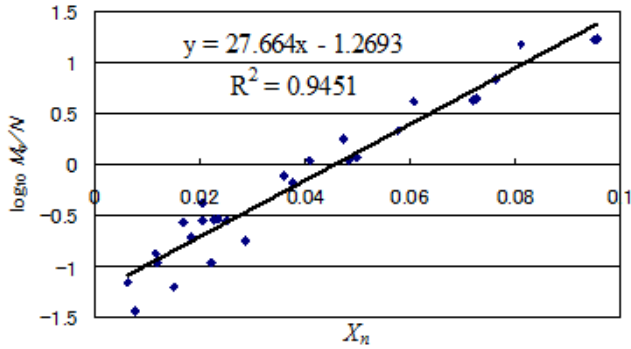


図2 スケールフリーネットワークの X_n と M_n の関係

情報伝播のしやすさについての $C_{(p)}$ 依存性がスモールワールドネットワークと比べ、真逆になっているのである。また、図2では、 γ が小さいほど $C_{(p)}$ が大きくなるが、 M_n も大きくなり、情報が伝播しやすくなる事も分かる。

尚、図1と図2から、両ネットワークにおける X_n と M_n のおおよその定量的関係は以下であると考えられる。

$$M_n \sim B_n(X_n)^{-a}, \quad (4)$$

$$M_n \sim \exp(cX_n). \quad (5)$$

但し、 $B_n, a > 0, c > 0$ は図1と図2のグラフの定量的分析から評価される定数である[5],[10].

3. 中間的ネットワーク

前章で述べた M_n における $C_{(p)}$ 依存性が、スケールフリーネットワークとスモールワールドネットワークで逆であることにに関して、両ネットワークの中間的ネットワークを2つの手法で作成し考察を深めた。ただし、 M_n の代わりに評価の容易な最短平均パス長 L を使用した。それは M_n が L と極めて良い相関が得られているため正当化される[5],[10].

(1) スケールフリーネットワークをスモールワールドネットワーク化したネットワーク

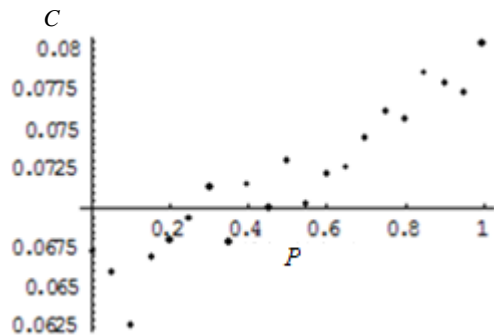
まず1つ目の中間的ネットワークとして、スケールフリーネットワークをスモールワールドネットワーク化したネットワークを作成する。そのために、まずパラバシモデル[2]で、サイズ N 、 $\gamma=3$ のスケールフリーネットワークを作成する。次に、このネットワークにおいて貼り替え率 P の確率で、ハブと連結しているノード間に枝を貼り、代わりにハブから出ている枝を削除する事によってネットワークを作成する。枝を貼る度に枝を削除しているので、最終的に全ノード数は一定になるようになっていく。

ここで頂点 v_i における局所クラスタリング係数 C_i を次の式で定義する[3].

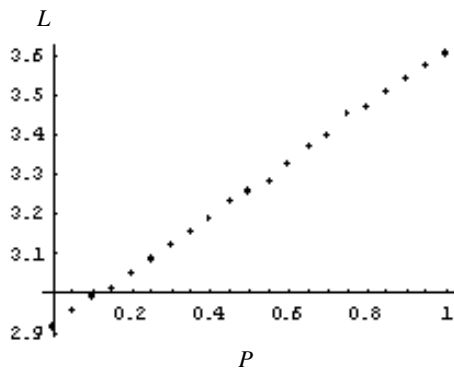
$$C_i = \frac{\text{頂点 } v_i \text{ を頂点に持つ三角形の数}}{(k_i(k_i-1)/2)}. \quad (6)$$

但し k_i は頂点 v_i の次数で、 $k_i \geq 2$ である。また、 C_i の全ノード上での平均を C とする。

このように作成したネットワークで、 C を貼り替え率 P ごとに評価したものが図3の(A)、最短平均パス長 L を貼り替え率 P ごとに評価したものが図3の(B)である。図3を見ると、 P の増加と共に、 C と L が増加する事が分



(A)



(B)

図3 スケールフリーネットワークをスモールワールドネットワーク化したネットワークにおける P と C 、および L の関係

る。つまり閉路が増えると情報が伝達しにくくなっている。この特徴は、元来のスモールワールドネットワークの特徴と一致している。

(2) スモールワールドネットワークをスケールフリーネットワーク化したネットワーク

次に2つ目の中間的ネットワークとしてスモールワールドネットワークをスケールフリーネットワーク化したネットワークを作成する。そのために、まずネットワークサイズ N_0 、次数4の規則的円環格子を作成し、そこに次数4の新たなノードをサイズ N になるまで優先的選択で成長させるパラバシモデル[2]によってネットワークを成長させる。この時、全ノード数 N は一定になる。

このネットワークで C と N_0 の関係を評価したものが図4の(A)、 L と成長前のネットワークサイズ N_0 の関係を評価したものが図4の(B)である。 N_0 が大きくなるにつれ、スケールフリーネットワーク的ではなくなっていくが、その時、図4を見ると、 L も C も共に増加する事が分かる。この特徴は先程と同じく、スモールワールドネットワークの特徴と一致している。

以上の考察から、スケールフリーネットワークで指摘されていた特徴[8]は、このようなモデルでは実現し得ない事が分かる。コンフィグレーションモデルで作ったスケールフリーネットワークの図2のような性質は、スケリング指数 γ の変化によって、 $C_{(p)}$ の変化を実現した事で生じた。一方、この論文のモデルでは、 $\gamma=3$ を持つパラバシモデルに基づいて作成したネットワークであったため、図2のようなスケールフリーネットワークの特徴が実現しなかったと考えられる。

更に、図5では、スモールワールドネットワークをスケ

ールフリーネットワーク化したネットワークにおいて、 M_n と N_0 の関係が示されている。 N_0 が増加すると、 M_n が減少しているの、 L を評価した図 4 の(B)の結果と一致している。

また、図 6 において、 N_0 と個々の一般化クラスタリング係数 $C_{(p)}$ ($p=3\sim 6$) の関係が示されている。 これらを見ると、 $C_{(3)}$ は N_0 と共に増大するが、 p が大きい場合は、一般化クラスタリング係数 $C_{(p)}$ は一旦下がり、その後増大している。 スモールワールドネットワーク及び、折衷的网络において閉路の増大は、最短平均パス長 L の増加を促し、情報の伝播を阻害する役割がある事を指摘して来た。 しかし図 4 (B)における L の増加は N_0 が小さい時は、大きい時と比べ緩やかである。 一方、図 6 のクラスタリング係数 $C_{(3)}$ は常に上昇しているが、 $C_{(4)}$ 、 $C_{(5)}$ 、 $C_{(6)}$ は一旦下がってから増大している。 これらの振る舞いが、 L の増大を阻害していると考えられ、図 4(B)では N_0 が小さい時に緩やかな増大が起こっていると考えられる。

4. 次数分布エントロピー S

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の情報伝達において、閉路の影響が真逆になるという現象をどのように理解したらよいであろうか。 スモールワールドネットワークにおいて、 P を上げる事で、情報が伝わりやすくなった。 P を上げるという事は、ランダムネットワークに近づけるという事である。 スケールフリーネットワークにおいて、 γ を下げる事で、情報が伝わりやすくなった。 γ を下げると、よりランダムネットワークに近づくと推測される。 以上のことから、情報伝達に本質的な事はランダムネスであるかも知れない。 それ故、ランダムネスの一つの指標であると考えられるエン

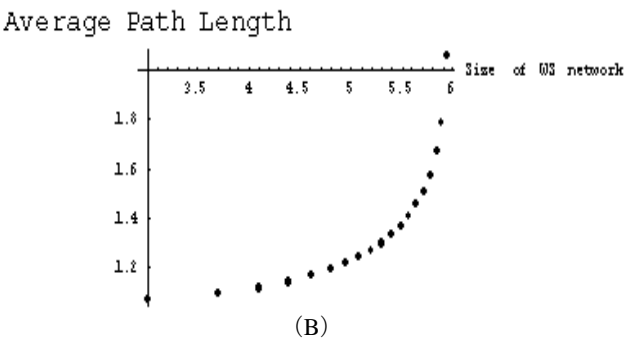
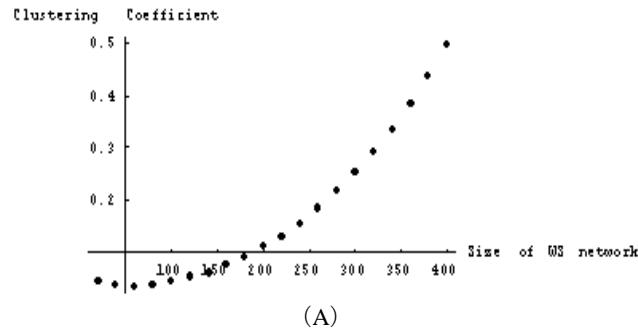


図 4 スモールワールドネットワークをスケールフリーネットワーク化したネットワークにおける N_0 と C 、および L の関係

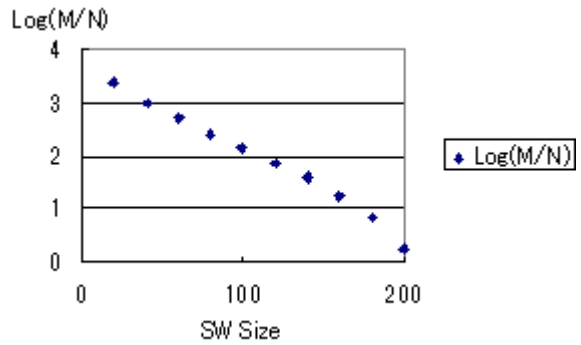


図 5 スモールワールドネットワークをスケールフリーネットワーク化したネットワークにおける N_0 と M_n の関係

トロピー S を導入して考察していく。ここでは Wang 等が導入した次数分布のエントロピーによって S を定義する事にする[11]。

$$S = - \sum_{k=1}^{k_{max}} P(k) \log_2 P(k). \quad (7)$$

ただし $P(k)$ は次数分布である。

スモールワールドネットワークをスケールフリーネットワーク化したネットワークについて、 S と N_0 の関係の評価したものが、図 7 に挙げられている。 これを見ると N_0 が大きくなるにつれて、次数分布エントロピー S が減少していく事が分かる。 さらに、図 4、及び図 5 の結果から N_0 が大きくなると L が大きくなり、 M_n が小さくなるため、情報伝達がしにくくなる事が分かっている。 つまり、 S が下がると情報の伝達がしにくくなっている。

又、スモールワールドネットワークでは、ショートカットを入れる確率を大きくすると、 L が小さくなり、 S は大きくなる。 この事からスモールワールドネットワークでは

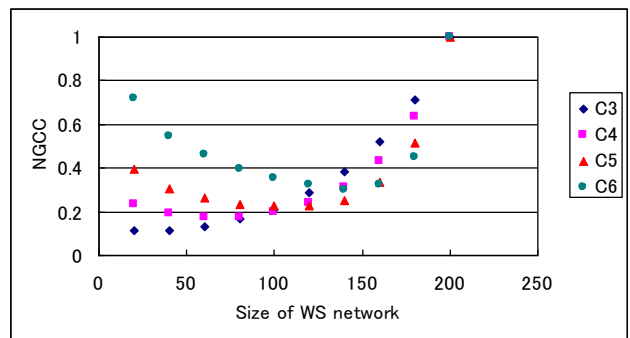


図 6 一般化クラスタリング係数 $C_{(p)}$

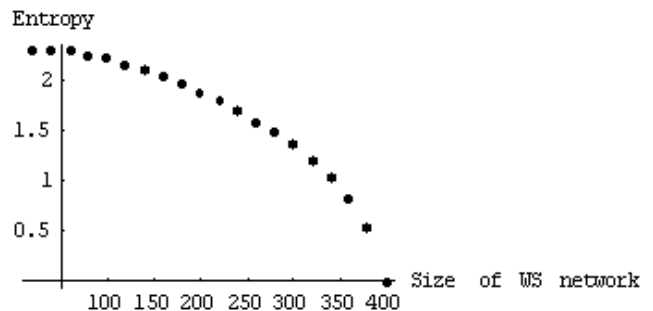


図 7 エントロピー

S が小さくなると L が大きくなり、情報が伝達しにくくなる事が分かる。一方、スケールフリーネットワークでは、明らかにスケーリング指数 γ が大きいと次数分布エントロピー S が減少する。又、その時の情報伝達がしにくくなる事が分かっている。

結局、これらのネットワークでは全て、 S が小さくなると情報の伝達がしにくくなる。従って、情報伝達に、より本質的な量は、ネットワークに内在する閉路の影響ではなく、乱雑さの指標を与えるエントロピー S である事が分かる。

5. まとめ

6 次の隔たりの分析には、ミルグラム条件が重要な役割を演じる。筆者等によって、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において、ミルグラム条件と一般化クラスタリング係数の関係が全く違った振る舞いをする事が示されていた[4]。今回は、この一般化クラスタリング係数のミルグラム条件に与える影響を、より深く調べるため、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をつなぐ折衷的ネットワークモデルを2つ考案し、それらのネットワークについて考察した。その結果として、スケールフリーネットワークをスモールワールド化したネットワークでは、スモールワールドネットワークの特徴が現れた。さらに、スモールワールドネットワークをスケールフリーネットワーク化したネットワークでも、スモールワールドネットワークの特徴が現れた。コンフィグレーションモデルで作成したスケールフリーネットワークの特徴が現れなかったのは、両ネットワークが、 γ を固定したバラバシモデルに基づいたネットワークであったためであると考えられる。又、隔たり数に、より本質的な量は、ネットワークに内在する閉路の数ではなく、次数分布エントロピーである事が示された。このことによって、この論文で議論したネットワークの情報伝達における閉路の影響の相違が統一的に理解できる。

参考文献

- [1]H.Aoyama, "Six degrees of separation; some calculation", (2008); H.Aoyama, Y.Fujiwara, H.Ietomi, Y.Ikeda and W.Soma "EconoPhysics", Kyouritu Shuppan 2008
- [2] A.L.Barabasi and R.Albert, "Emergence of scaling in random networks", Science, 286, 509-512(1999)
- [3]D.J.Watts and S.H.Strogatz, "Collective dynamics of 'small-world' networks", Nature, 393, 440-442(1998)
- [4]N.Toyota and T.Sakamoto, "Six degrees of separation and the generalized clustering coefficients in scale free networks", 42-th lecture meeting of the Society of Instrument and Control Engineers in Hokkaido, 2010.Feb. in Japanese
- [5]N.Toyota, IEICE Technical Report, "String Formalism for p -Clustering Coefficient-Toward Six Degrees of Separation", NLP2009-49(2009) in Japanese.
- [6]N.Toyota, " p -th Clustering coefficients C_p and Adjacent Matrix for Networks: Formulation based on String", arXiv:0912.2807
- [7]N.Toyota and T.Sakamoto, " p -th degrees of separation in the string-adjacent matrix formalism", sixth symposium of network ecology, 2009 Dec. in Japanese
- [8]N.Toyota, "Generalized Clustering Coefficients and Milgram Condition for q -th Degrees of Separation", SICE 2010 Annual Conference in Taiwan, TA16, pp30409-3412, Aug.18-21 2010,

Preprint; arXiv1010.4088

- [9] M.E.J.Newman and D.J.Watts, "Renormalization group analysis of the small-world network model", Phys.Lett.A263,341-346(1999)
- [10]豊田規人, "スモールワールドネットワークにおけるストリング形式による q 次の隔たりと一般化クラスタリング係数の考察", FIT2010 第9回情報科学フォーラム, RA001, pp1-5, 2010
- [11]B.Wang, H.Tang, C.Guo and Z.Xiu, "Entropy optimization of scale-free networks' robustness to random failures", Physica A 363(2006)591-596
- [12] S.Milgram, "The small world problem", Psychology Today 2, 60-67 (1967)
- [13]Norihito Toyota, "p-th Clustering coefficients and q-th degrees of separation based on String-Adjacent Formulation", Preprint arXiv:1002.3431