

スケールフリーネットワーク上の交通流モデル

本田 泰

室蘭工業大学 情報工学科

概要

たとえば、インターネット上で情報を探する場合、いったいいくつかの Web ドキュメントを訪問すれば自分の目的にぴったり一致した情報にたどりつけるのであろうか? 本研究はこの疑問に対する一つのアプローチである。ある数理モデルに対するシミュレーションを用いる。World Wide Web (WWW) に代表されるような多くのネットワークは、スケールフリー (SF) と呼ばれる性質を持っていることが最近明らかになった。本研究では、SF ネットワーク上で、準 self-avoiding ランダムウォークの方法と総当たりの方法によって経路探索をした場合の比較をおこない、総当たりの方法の方が、準 self-avoiding ランダムウォークを用いるよりも効率がよいことを明らかにした。

1 はじめに

ノードとノードを結ぶリンクの数 (k) が、べき分布 $P(k) \propto k^{-\gamma}$ に従うネットワークは、スケールフリーネットワーク (SF ネットワーク) と呼ばれる [1, 2]。最近、さまざまな分野で SF ネットワークが発見された [3]。たとえば、WWW において、各ドキュメントをノードと見なせばハイパーリンクの構造は SF ネットワークとなっていることが知られている [4]。また、科学者をノードと見なし、共同研究を科学者間のリンクと見なせば、そのネットワークも SF ネットワークとなっているという報告もある [5]。

ネットワークをクラスタリング係数 C と最短パス長 d を用いて分類すると、規則的ネットワーク、ランダムネットワークそしてスモールワールド (SW) ネットワークに分類される。本研究では、この分類において SF ネットワークがどのカテゴリーに属するかを調べる。また、経路情報が存在しない場合の経路探索の方法を二種類 (準 self-avoiding ランダムウォークの方法と総当たりの方法) 提案し、目的ノードに到達するために必要な訪問ノード数を比較した。

2 クラスタリング係数 C と最短パス長 d を用いたネットワークの分類

ネットワーク Ω は頂点 v_i の集合 V とリンク ξ_{ij} の集合 L からなる。

$$\Omega = \{V, L\}, \quad (1)$$

$$V = \{v_i | i = 1, \dots, N\}, \quad (2)$$

$$L = \{\xi_{ij}\}, \quad \xi_{ij} = \{v_i, v_j\}, \quad (3)$$

$$\xi_{ij} = \begin{cases} 1 & (\text{リンクが存在するとき}) \\ 0 & (\text{リンクが存在しないとき}) \end{cases}. \quad (4)$$

クラスタリング係数 C は \mathcal{L}_i, c_v を用いて以下のように定義される。

$$\mathcal{L}_i = \sum_{j=1}^N \xi_{ij} \sum_{k \in \Gamma_i, j < k} \xi_{jk}, \quad (5)$$

$$\Gamma_i = \{v_j | \xi_{ij} = 1\}, \quad (6)$$

$$c_v(i) = \frac{\mathcal{L}_i}{\binom{|\Gamma_i|}{2}}, \quad (7)$$

$$C = \frac{1}{N} \sum_{i=1}^N c_v(i). \quad (8)$$

\mathcal{L}_i はノード i を含む 3 つのノードの組み合わせのうち、お互いにリンクで結びつけられたものの数を意味している。

最短パス長 d は以下の式で定義される。

$$d = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}. \quad (9)$$

ここで、 d_{ij} はノード i とノード j を結びつける経路のうち、経由するノード数が最も少ない経路の長さである。

表 1 に、実際に存在するネットワークのクラスタリング係数と最短パス長の例を示した [6]。いずれも、SW ネットワークに属すると考えられているネットワークである。SW ネットワークは、小さな最短パス長と、大きなクラスタリング係数をもつネットワークである。いっぽう、ランダムネットワークは最短パ

表 1: 実在するネットワークにおけるクラスタリング係数 C と最短パス長 d の例 [6]. N は各ネットワークに含まれるノードの数. $\langle k \rangle$ は各ノードがもつ平均リンク数.

	d	C	N	$\langle k \rangle$
俳優の共演	3.65	0.79	222,226	61
送電線網	18.7	0.08	4,941	2.67
線虫の N.N.	2.65	0.28	282	14

ス長は小さいが、クラスタリング係数は SW ネットワークよりも小さい。規則的なネットワークの最短パス長は、SW ネットワークやランダムネットワークにくらべて大きい。(図 1 参照)

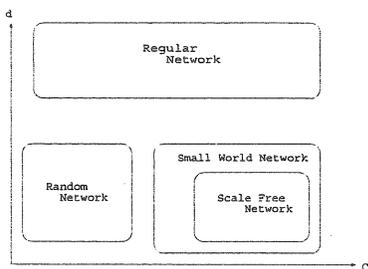


図 1: クラスタリング係数 C と最短パス長 d によるネットワーク分類の概略図。横軸と縦軸のスケールに定量的な意味はない。

3 BA モデルによる SF ネットワーク

SF ネットワークをもたす数理モデルとして、Barabasi と Albert[1] による BA モデルがある。このモデルでは、以下のような過程に従ってネットワークを成長させる。

- ごく少数 (m_0 個) のノードから出発し、ステップが一つ進むごとに新たにノードを一つずつ付け加えて行く。新たに付け加えられるノードは、すでに存在するノードに対して $m (\leq m_0)$ 本のリンクを張る。
- 既存のノードの中で、どのノードに対して新しいリンクを形成するかは、リンクを張ろうとするノードがすでに何本のリンクを持っているかに比例する。すなわち、ノード i に対してリンクが形成される確率 $\Pi(k_i)$ は以下のように与えられる。

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \quad (10)$$

ここで、 k_i はノード i が持つリンクの本数である。

2 番目の条件は Preferential attachment と呼ばれ、リンクをたくさん持つノードはますますリンクを集める傾向にあることを反映している。

この SF ネットワークのもつリンク数分布の一例を図 2 に示した。 $m_0 = m = 2$ の場合、リンク数 k

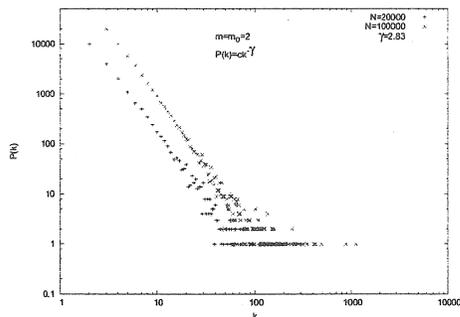


図 2: BA モデルによって構成された SF ネットワークにおけるリンク数 k の分布。

の分布は、べき指数 γ が約 2.8 のべき分布となった。また、クラスタリング係数と最短パス長がノードの数 N および平均リンク数 m にどのように依存するかを図 3 に示した。 C と d の値は N と m の値に依存す

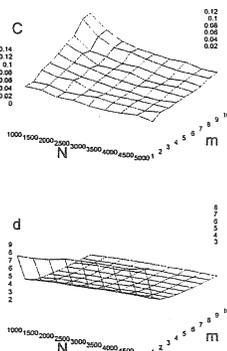


図 3: BA モデルによって構成された SF ネットワークのクラスタリング係数 C と最短パス長 d のネットワークサイズ N とリンク数 m に対する依存性。

るが、SW ネットワークのもつ C, d の値に近い。したがって、BA モデルによる SF ネットワークは SW ネットワークに分類されると考えられる。

図 4 に、SF ネットワークとランダムネットワークのもつ最短パス長のサイズ依存性を示した。両ネッ

トワークとも $N \rightarrow \infty$ の極限において有限の最短パス長をもつと考えられる。また、SF ネットワークの最短パス長はランダムネットワークの最短パス長よりも小さい。

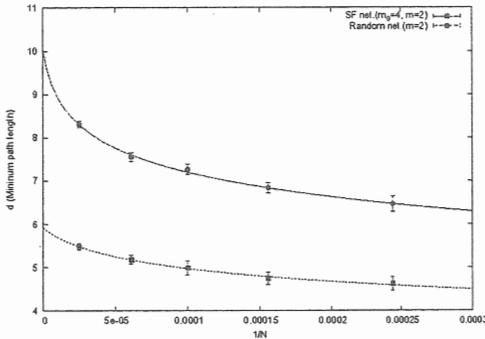


図 4: 最短パス長のネットワークサイズ (N) 依存性。

4 SF ネットワークにおける経路探索

4.1 準 self-avoiding ランダムウォークによる Displacement

目的のノードを探しながらネットワークを渡り歩いていくとき、経路情報が全くない最悪の場合、どのような経路探索が可能だろうか。もっとも単純な経路探索の方法として、目的ノードに到着するまで、ランダムにリンクを選択するが、1度訪れたノードは2度と訪れないように避けながら経路を選択する方法が考えられる。非常に効率は悪いが、時間をかければ必ず目的ノードに到達する可能性がある。これは、ネットワーク上における、Self-avoiding ランダムウォークであると考えられる。

単純な Self-avoiding ランダムウォークをネットワーク上で行くと、目的ノードに到達するまえに袋小路に陥ってしまう場合がほとんどである。これを回避するために、ランダムなリンク選択の試行回数が、そのノードの持つリンク数よりも多くなったら、たとえリンクの先のノードが訪れたことのあるノードであっても、そのリンクをたどる方法をここでは採用する。図 5 にこの経路選択法の概略図を示した。この経路探索法では、パケットが袋小路に陥ってもかみならずそこから抜け出すことが可能である。このような経路探索の方法をここでは仮に「準 self-avoiding ランダムウォーク」と呼ぶことにする。

準 self-avoiding ランダムウォークによって到達した出発点からの displacement $l(t)$ を、ランダムネット

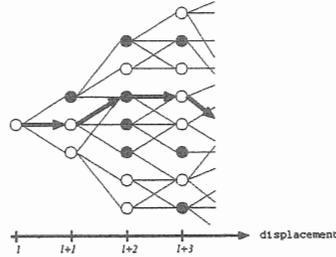


図 5: 準 self-avoiding ランダムウォークの概略。塗りつぶされたノードは、すでに訪れたことのあるノードを表す。Displacement l では3つのリンクを持っている。その中から次に進むべきノードをランダムに選ぶ。訪れたことのないノードが選ばれた場合、そのノードに進む。Displacement $l+1$ では、3つのリンクをもつが、すべてのノードが訪れたことのあるノードである(袋小路)。3ステップ以上このノードにとどまっている場合には、すでに訪れたことのあるノードでも、次に進むことを許す。

ワーク、SF ネットワーク、正方格子について、それぞれ、時間ステップの関数として図 6 に示した。ノード

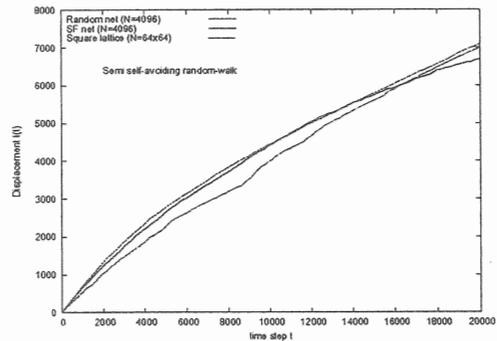


図 6: Displacement の時間発展

の数は、いずれのネットワークも $N = 64 \times 64 = 4096$ である。ノードの数が等しい場合、 $m = 2$ の SF ネットワークの持つリンク総数は、正方格子の持つリンク総数に等しいということに注意してほしい。

ランダムウォークの初期の段階では、パケットの周りは、自身が直前に訪問したノード以外、ほとんど訪れたことのないノードばかりが存在する。したがって、いま、平均リンク数は $\langle k \rangle = 4$ であるから、1時間ステップで次のノードに進む確率は $\frac{3}{4}$ である。これは、図 6 の原点付近では傾きが $\frac{3}{4}$ となることを意味する。

いっぽう、十分に時間がたつと、パケットは全ノ

表 2: Displacement $l(t)$ に対する (11) 式によるフィッティングから得られた τ の値.

	ランダム	SF	正方格子
τ/N	1.5	2.4	2.2

ドのうち、ほとんどを訪問し終わる。パケットの周りにはすでに訪問したことのあるノードばかりが存在すると考えられる。平均リンク数は $\langle k \rangle = 4$ であるから、4時間ステップの後に次のノードに進むことができる。これは、十分時間がたつと、図 6 で傾きが $\frac{1}{4}$ となることを意味する。

以上の考察から、 $t \ll N$ では $l(t) \sim \frac{3}{4}t$ 、 $t > N$ では $l(t) \sim \frac{1}{4}t + \text{const.}$ となると予想される。したがって、Displacement $l(t)$ として

$$l(t) = ate^{-t/\tau} + (bt + c)(1 - e^{-t/\tau}) \quad (11)$$

と仮定し、シミュレーションによってえられた Displacement に対してフィッティングをおこなった。 τ は、ほぼすべてのノードを訪問するのに要する時間の目安を与えると考えられる。得られた τ の値を表 2 に示した。3種類のネットワークの中で SF ネットワークの最短パス長が最も短いにもかかわらず、 τ の値は最も大きいことに注意してほしい。

4.2 SF ネットワーク上の経路探索に必要な訪問ノード数

経路情報が全く存在しない場合の、もうひとつの経路探索の方法として総当たりの方法が考えられる。これは、出発ノードからリンクをたどって到達できるノードはすべてしらみつぶしに調べて行く方法である。図 7 に、この方法と、準 self-avoiding ランダムウォークの方法を用いて目的ノードに到達するまでに訪問したノード数 v の分布を示した。ネットワークのサイズ $N = 1024$ の場合を調べた。 v が大きいほど、出発ノードから目的ノードに到達する経路を見つけたすのに要した労力が大きかったといえる。総当たりの方法では、ほぼまんべんなく分布しているのに対して、準 self-avoiding ランダムウォークの方法では、 $v \sim N$ に至るまで目的ノードを発見できない場合が若干多いことがわかる。

5 まとめ

BA モデルによって生成した SF ネットワークのクラスタリング係数と最短パス長を調べた。クラスタリング係数は、ノード数と平均リンク数に大きく依

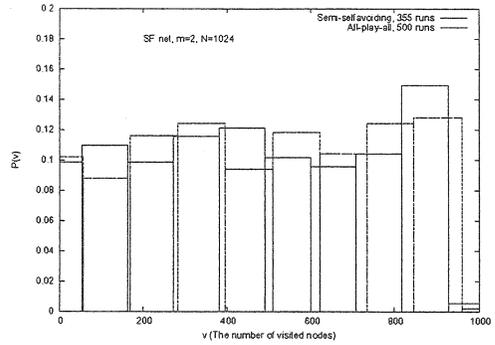


図 7: 準 self-avoiding ランダムウォークの方法と総当たりの方法によって目的ノードを発見するまでに訪問したノード数 v の分布.

存するが、最短パス長は、それらにあまり依存せず $d \sim 5$ 程度である。サイズ無限大のネットワークにおいても $d \sim 6$ と予想される。これらの結果から、このネットワークは SW ネットワークに属すると考えられる。

また、準 self-avoiding ランダムウォークによる displacement をランダムネットワーク、SF ネットワーク、正方格子について調べた結果、ほぼすべてのノードを訪問するのに必要な時間は、SF ネットワークにおいて最も長いことがわかった。さらに、SF ネットワーク上において、準 self-avoiding ランダムウォークと総当たりの方法で、目的ノードを発見するまでに必要な訪問ノード数を調べ、総当たりの方法の方が効率が良いことがわかった。

たとえば、インターネット上で情報を探する場合、次々とハイパーリンクをたどって目的のドキュメントを探して行く (準 self-avoiding ランダムウォークに対応) よりも、関連のありそうなリンク先ドキュメントは、しらみつぶしに探して行く (総当たりの方法に対応) 方が効率が良いそうである。

参考文献

- [1] A.-L. Barabási and R. Albert, *Science* **286** (1999) 509.
- [2] A.-L. Barabási, R. Albert, H. Jeong, *Physica A* **272** (1999) 173.
- [3] R. Albert and A.-L. Barabási, *cont-mat/0106096*.
- [4] R. Albert, H. Jeong, A.-L. Barabási, *Nature* **401** (1999) 130.
- [5] M. E. J. Newman, *cont-mat/0011155*.
- [6] D. J. Watts and S. H. Strogatz, *Nature* **393** (1998) 440.