

# Scale-free ネットワークの研究動向

林 幸雄 (yhayashi@jaist.ac.jp)  
北陸先端科学技術大学院大学

概要： インターネットや WWW などの通信網, 企業間取引や協業の社会システム, 遺伝子や代謝系などの生物システムにおいて広く共通する Scale-free ネットワーク構造に関する研究動向を概説する.

## Research Trends of Scale-free Networks

Yukio Hayashi (yhayashi@jaist.ac.jp)  
Japan Advanced Institute of Science and Technology

Abstract: We briefly review research trends of scale-free networks, whose structure commonly exist in many complex networks, e.g. communication systems of Internet and WWW, social tradings or collaborations, and biological gene-protein or metabolic systems.

### 1 はじめに

知人関係や企業間取引, 電力網やインターネット, さらに生物の食物連鎖や細胞内のエネルギー代謝反応系など, 我々が知っている多くのネットワークはどれもみな複雑で, 一見すると何ら構造的な類似性がないように思われる. しかしながら, 実は**驚くほど共通の性質**を持ち [21], そうした構造を創り出す**普遍的なメカニズム**が存在することがちょうど世紀末頃に発見され [3], 大きな話題となっている [4].

そのネットワーク構造は scale-free(SF) と呼ばれ, 従来の研究で主に扱われた格子などの規則的あるいは一様ランダムな同質構造とは異なり, 次数が小さい大多数のノードと極少数だが非常に多くの次数をもつハブで構成され, 結合次数  $k$  の分布がべき乗則,  $P(k) \sim k^{-\gamma}$ ,  $2 < \gamma < 3$  に従うものである. 本稿では, この SF ネットワークに関する新しい研究の潮流を表 1 に従って概説する.

1950 ~	ランダムグラフ理論
1955	苗字のべき乗分布 [18]
1965	論文引用のべき乗分布 [15]
1969	6 人の知人の実測
1974	弱い紐帯の仮説
1998	小さな世界 [21]
1999	WWW 頁の分離度は 19 クリック
1999	growth & pref. attach.[3]
1999	癌に関わる遺伝子ネットのハブ p53
2000	インターネットのアキレス腱 [1]
2000	一般化 BA モデル [2]
2001	生物系の duplication[19]
2001	fitness による独占 [5]
2001	経済的かつ効率的な構造 [6]
2001	感染しきい値の不在 [14]
2002	階層的成長モデル [16]
2002	Best Business Books 受賞 [4]
2003	局所生成規則 [20]
2003	結合相関 [13]

表 1: SF ネットワーク研究の主な話題

## 2 歴史的経緯

### 2.1 源流

まず, SF ネットワーク研究の源流とも言える, 「べき乗分布」を創り出す基本的な考え方や, 我々の直感から少しずれた性質を示唆した経験則を紹介する.

富の分布や語の出現頻度, 文化伝承などにおける rich-get-richer 現象によって, べき乗則が現れることを最初に示したのは Simon[18] で, ネットワークとして議論したのは Price[15] である. 現在も議論されている SF 構造を生成する種々のモデル [7] の出発点として, **新たなノードとある種の偏ったノードへのリンクの追加**が「べき乗分布」を創り出す鍵であることが以下から伺えるだろう.

まず, Simon のモデルは, 毎時刻  $t$  にある個人が追加され,

1. 確率  $\alpha$  で新しい苗字を,
2. 確率  $1 - \alpha$  でランダムに選んだ既存の人の苗字をもつ,

成長ネットワークに対応する. ここで, ランダムに選択した人と同じ苗字を持つ人の数は, その人の家族数 (次数  $k$ ) に比例するので, 実質的に優先的選択 (pref. attach.) を行っていることに注意しよう.

一方, Price のモデルでは, 引用ネットワークにおいて各時刻に,

1. 新しい論文（常に一定数でなくてもよい）ノードが追加され,
2. 既存の論文へ平均  $m$  本の引用リンクが生成される,

ことが繰り返されるものとする. 但し, 新たな引用を得る確率はそれまでの引用実績  $k$  に比例するもの (pref. attach.) とし, 新しい論文自身への引用も一定値  $k_0$  で加味する. すなわち,  $k + k_0$  に比例するものと考え (Price はこの効果を cumulative advantage と呼んだ). 詳細は省略するが, 最近の理論解析法を適用すれば, これらのモデルから結合次数の「べき乗分布」が導出できる [7].

ところで, ちょうど同じ頃に米国で行われた社会実験があった. それは, 全く見ず知らずの人に手紙を送る際, 知人のみを介して (その人に近そうな知人に次々と) 転送すると平均何人で届くかを測ったもので, 我々は**わずか 6 人で互いに結び付いている**という予想外の結果が得られた. また別の例では, 転職する際に有益な情報を提供してくれた人は, 日頃良く付き合っている人よりも極まれにしか会わない人であることが知られている. つまり, 人的ネットワークでは繋がり弱い人が重要と考えられた (今ではネットワークの中心性 [22] で分析できるが). 但し, それらは経験則として理解されるにすぎなかった.

## 2.2 画期的モデル

こうしたネットワークに潜む性質の科学的根拠が次々と明らかにされる転機が訪れたのは, 比較的最近の出来事 (偶然にも世紀末) であった. それには, 実際のネットワークの結合構造を対象とした膨大な分析や大規模シミュレーションが可能になるまでの年月や, 新しい発見を導くための**統計物理や計算機科学者らによる斬新な解析手法**が必要だった.

'98年当時, 複雑系で著名な SantaFe 研究所にいた Watts らは, 表 2 に示すさまざまな対象において共通に観測された, 高いクラスター係数 (友達の友達はまた友達である頻度が高い) とお互いのノード間は平均的に短い距離という, 規則的グラフと一様ランダムグラフの中間的な性質に着目し, こうした構造を創り出す「小さな世界」モデルを考えた [21]. 但し, 「べき乗分布」を示すものではなかった.

翌年, Barabási らは統計物理の平均場解析を用いて, 「小さな世界」と「べき乗分布」の双方を満たす普遍的な生成規則を提示し [3], 新たなネットワーク研究分野がブレイクした (実際, この後の欧米メディアの反響は絶大: <http://www.nd.edu/~networks/press.htm>). この規則は, ノード追加による growth と選択的結合を行う pref. attach. に従うもので, BA モデルと呼ばれている.

社会関係	知人, 企業間取引, 映画の共演, 論文引用, 性的関係, 言語
インフラ技術	インターネット (ルータおよび AS レベル), WWW, 航空路線, 電力網, 電子メール送受信
生物系	神経回路網, 遺伝子やエネルギー代謝反応, 食物連鎖

表 2: さまざまな対象に見られる SF 構造

## 2.3 最近の研究

SF ネットワークは、これまで数十年間の君臨を続けた従来のランダムグラフ理論では説明できないばかりか、以下のような全く予期しなかった特徴の発見などからも、さらに活発な研究を促す格好の対象となっていた。

**頑健性と脆弱性:** ランダムなノード故障には頑健である（80%の故障でも連結性を保持できる）が、ハブへの集中的な攻撃には脆弱（10%程度で極度に分断される）[1]。この性質は、インターネットの**アキレス腱**として注目を浴びた。

**最適性:** 経済性  $\rho$  と効率  $d$ （リンク総数と距離あるいは伝達時間にそれぞれ相当する）を  $0 < \lambda < 1$  で重み付けた評価関数  $E = \lambda d + (1 - \lambda)\rho$  を最小化するランダムなネットワーク生成で、SF 構造が現れる [6]。すなわち、**経済性と効率の両基準から最適な構造**。

**不可避な感染:** 従来の一様な接触感染（同質構造）を前提とした疫学理論では、感染率をあるしきい値以下に抑えれば感染は拡大しないものと考えられていた [17] が、SF ネットワークにはそうした**ウィルス絶滅のしきい値が存在しない** [14]。

その他にも、SF ネットワーク研究に関連した話題としては、癌の原因遺伝子よりも遺伝子ネットワークの重要性を認識した発病の解明や治療薬の開発をめざす新しい潮流を促したことや、Bose-Einstein 凝縮と企業の独占が同様なメカニズムで起ることを示唆した fitness モデル [5]、図 1 に示すような生物系における局所的な生成規則 [19][20]、階層構造を持つモデル [16]、社会系と技術・生物系とで異なる結合相関 [13] の発見などが挙げられる。

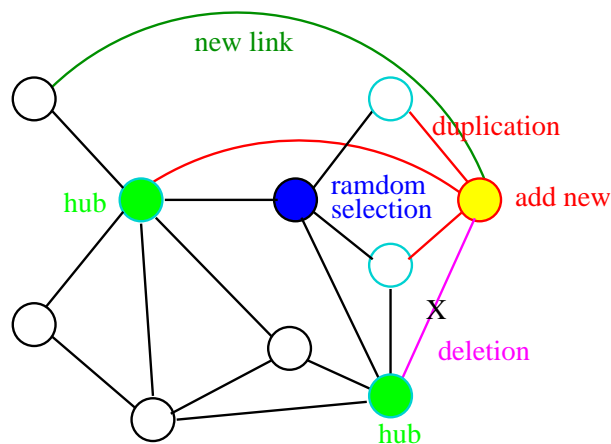


図 1: 生物系の生成規則：ランダムに選ばれたノードに隣接するハブにはリンク複写のチャンス大!

### 3 Pref. Attach. + Rewire + New Node

前章までにSFネットワーク研究の主な流れを概説した。中でも、社会的関係から技術インフラや生物系に至るまで対象や構成要素によらず、「べき乗分布」を創出する普遍的なメカニズムが存在することが重要なものと考えられる。本章では、多くのSFネットワークモデル [7] の基本となる理論解析手法を紹介する。

BAモデル [3] では、常に新ノードから既存ノードにリンクを追加していくが、論文引用以外のWWWなどの現実のネットワークは必ずしもそうではなく、既存のノード間のリンクの追加や削除も行われている。またBAモデルでは、べき指数も3に限定されるが、実測値は主に  $2 < \gamma < 3$  である [7]。そこで、リンクのrewire率などによって、べき指数が調整できる一般化BAモデル [2] が提案された。以下に、BAモデル [3] と同様な平均場近似を用いて、べき指数を導出する。

$m_0$  個の孤立ノードを初期状態として、各時刻  $t_i$  に以下の3つのどれかを確率的に選択して、ネットワークが生成していくものとする。

1. 確率  $p$  で  $m$  本の新しいリンクを生成する ( $m \leq m_0$ )。その際、一方の端のノードをランダムに選び、もう一方の端のノード  $j$  を pref. attach. に従った  $\mathcal{P}(k_j) \stackrel{\text{def}}{=} \frac{k_j+1}{\sum_l (k_l+1)}$  に比例して選ぶことを、 $m$  回くりかえす。
2. 確率  $q$  で  $m$  本のrewireを行う。まず、ランダムに選んだノード  $j$  に結合した  $j - j'$  間のリンクを除去し、確率  $\mathcal{P}(k_{j''})$  で選んだノード  $j''$  とのリンクに置き換える。
3. 確率  $1 - p - q$  で新ノード  $i$  を1個追加し、その新ノードから既存のノード  $j$  に確率  $\mathcal{P}(k_j)$  に従って  $m$  本リンクする。

ネットワークのサイズ (ノード総数) を  $N$  として、上記の1. から3. に従って時刻  $t_i < t$  に追加されたノード  $i$  の次数  $k_i$  の (平均的な) 変化を考えると、

$$\frac{\partial k_i}{\partial t} = pm \frac{1}{N} + pm \frac{k_i + 1}{\sum_l (k_l + 1)} \quad (1)$$

$$-qm \frac{1}{N} + qm \frac{k_i + 1}{\sum_l (k_l + 1)} \quad (2)$$

$$+(1 - p - q)m \frac{k_i + 1}{\sum_l (k_l + 1)}, \quad (3)$$

$$= (p - q)m \frac{1}{N} + m \frac{k_i + 1}{\sum_l (k_l + 1)}, \quad (4)$$

となる。ここで、(1) 右辺第1項は上記1. によるランダム選択で第2項は pref. attach. による変化分、(2) 右辺第1項と2項は2. のrewireによる同様な変化分、(3) 右辺は3. による変化分である。

初期条件  $k_i(t_i) = m$  より (4) を解くと、 $t \geq t_i$  で、

$$k_i(t) = [A(p, q, m) + m + 1] \left( \frac{t}{t_i} \right)^{1/B(p, q, m)} - A(p, q, m) - 1,$$

$$A(p, q, m) \stackrel{\text{def}}{=} (p - q) \left( \frac{2m(1 - q)}{1 - p - q} + 1 \right),$$

$$B(p, q, m) \stackrel{\text{def}}{=} \frac{2m(1-q) + 1 - p - q}{m}.$$

さて、次数  $k$  以下の累積確率  $P[k_i(t) < k] = P[t_i > C(p, q, m)t]$  を考える。

$$C(p, q, m) \stackrel{\text{def}}{=} \left( \frac{m + A(p, q, m) + 1}{k + A(p, q, m) + 1} \right)^{B(p, q, m)}$$

時間的に等間隔な新ノードの追加を伴って、上記の 1. から 3. のどれかが行われることから、 $t_i$  の確率密度は均等な  $1/(m_0 + t)$  となり、

$$P[k_i(t) < k] = 1 - C(p, q, m) \frac{t}{m_0 + t},$$

を得る ( $0 < C(p, q, m) < 1$  で解が存在 [2]) .  $P(k) = \partial P[k_i(t) < k]/\partial k$  より、

$$P(k) = \frac{t}{m_0 + t} D(p, q, m) \times [k + \kappa(p, q, m)]^{-\gamma(p, q, m)} \sim [k + \kappa(p, q, m)]^{-\gamma(p, q, m)}.$$

ここで、 $D(p, q, m) \stackrel{\text{def}}{=} [m + A(p, q, m) + 1]^{B(p, q, m)} \times B(p, q, m) : \text{const.}, t \rightarrow \infty$  で  $t/(m_0 + t) \rightarrow 1$ ,  $\kappa(p, q, m) \stackrel{\text{def}}{=} A(p, q, m) + 1$ ,  $\gamma(p, q, m) \stackrel{\text{def}}{=} B(p, q, m) + 1$ .

ゆえに、パラメータ  $p$  と  $q$  で、べき指数  $2 < \gamma(p, q, m) < \infty$  が調整できる。特に rewire を行わない  $q = 0$  のとき、 $\gamma = 3 + (1 - p)/m$ , さらに新ノードを毎時必ず追加する  $p = 1$  のときは BA モデル [3] のべき指数 3 に一致する。

一般化 BA モデルを例として、べき指数の導出過程を紹介した。これは単なる例である以上に、他の SF ネットワークモデル [7] においても、**マスター方程式を考えて統計的なネットワークの性質を解析する点**は、基本的で共通に適用可能なアプローチであることを強調しておきたい。

## 4 おわりに

欧米における迅速かつ多大な注目から数年遅れて、やっと日本でも最近話題になってきた SF ネットワーク研究について、その歴史的経緯や普遍的なネットワーク生成メカニズムを中心に概説した。もちろん、一般論としては、欧米に追従する研究の潮流が必ずしも妥当であるわけではないが、**社会システムから通信や生物系も含む将来への技術革新を包括したこの SF ネットワーク研究のスコープの広さ [4]** を考えれば、今後の重要な研究分野となることは間違いないものと思われる。紙面の都合で、著者自身に取り組んでいるウィルスの再流行現象 [9][12] については紹介できなかった。また、ad hoc 通信なども深く関わる連結性（パーコレーション解析）をはじめ、社会ネットワーク分析 [22]、インターネットの渋滞や電力崩壊などを防ぐためのカスケード故障への対策等、SF ネットワークに関連した特筆すべき内容は尽きない。それらについては、他の解説書 [4][8][10][11] や専門書 [7] などを参照されたい。

## 参考文献

- [1] Albert, R., Jeong, H., and Barabási, A.-L., Error and attack tolerance of complex networks, *nature*, **406** (2000), 378-382.
- [2] Albert, R., and Barabási, A.-L., Topology of Evolving Networks: Local Events and Universality, *Physical Review Letters*, **85** (2000), 5234.
- [3] Barabási, A.-L., Albert, R., and Jeong, H., Mean-field theory for scale-free random networks, *Physica A*, **272** (1999), 173-187.
- [4] Barabási, A.-L., *LINKED* -The New Science of Networks, Perseus Publishing, 2002. 邦訳「新ネットワーク思考」NHK出版, 2003.
- [5] Bianconi, G., and Barabási, A.-L., Bose-Einstein Condensation in Complex Networks, *Physical Review Letters*, **86** (2001), 5632.
- [6] Cancho, R.F.i., and Solé, R.V., Optimization in complex networks, *SantaFe Inst. working paper*, 01-11-068, (2001).
- [7] Dorogovtsev, S.N., and Mendes, J.F.F., *Evolution of Networks* -From Biological Nets to the Internet and WWW, WILEY-VCH, 2003.
- [8] 林 幸雄, ネットワーク生態系, *ナレッジサイエンス*, 第5章 知のオデッセイ, 220-223, 杉山 他編, 紀伊国屋書店, 2002.
- [9] 林 幸雄, 箕浦 正人, 松久保 潤, ネットワーク成長によるメール型ウィルスの再流行と重点的なハブの免疫化の効果, *情報処理学会論文誌 数理モデル化と応用*, **44** (2003), 1234-1244.
- [10] 林 幸雄, 成長するネットワークの生態学, *第6回情報論的ワークショップ予稿集*, (2003), 17-24.
- [11] 林 幸雄, 講義紹介 ネットワーク生態学, *計測と制御*, 8月号, 特別企画「創発夏の学校」, **43** (2004), 掲載予定.
- [12] Hayashi, Y., Minoura, M., and Matsukubo, J., Oscillatory epidemic prevalence in growing scale-free networks, *Physical Review E*, **69** (2004), 016112.
- [13] Newman, M.E.J., Mixing patterns in networks, *Physical Review E*, **67** (2003), 026126.
- [14] Pastor-Satorras, R., and Vespignani, A., Epidemic dynamics and epidemic states in complex networks, *Physical Review E*, **63** (2001), 066117.
- [15] Price, D.J.de S., Networks of scientific papers, *Science*, **149** (1965), 510-515.

- [16] Ravaz, E., et al., Hierarchical Organization of Modularity in Metabolic Networks, *Science* **297** (2002), 1551-1555.
- [17] 重定 南奈子, *侵入と伝播の数理生態学*, 東大出版会, 1992.
- [18] Simon, H.A., On a class of skew distribution functions, *Biometrika*, **42** (1955), 425-440.
- [19] Solé, R.V. et al. A model of large-scale proteome evolution, *SantaFe Inst. working paper*, 01-08-041, (2001).
- [20] Vázquez, A., Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations, *Physical Review E*, **67** (2003), 056104.
- [21] Watts, D.J., and Strogatz, S.H., Collective dynamics of small-world networks, *nature*, **393** (1998), 440.
- [22] 安田 雪, *実践ネットワーク分析 -関係を解く理論と技法-*, 新曜社, 2003.