

## SF ネットワークモデルの特徴比較

林幸雄,<sup>1</sup> 宮崎敏幸<sup>2</sup>

<sup>1,2</sup> 北陸先端科学技術大学院大学, 知識科学研究科, 〒923-1292 石川県能美郡辰口町旭台 1-1

**概要:** 実世界に存在する多くのネットワークは, それぞれ種類が異なるにも関わらず共通の特徴を持つことが, これまでの物理や計算機科学における研究により明らかにされている. 例えば, インターネットや WWW, 電力網, 電子メールなどが挙げられ, これらのネットワークはベキ乗則に従う次数分布を持ち, スケールフリー(SF)ネットワークと呼ばれている. 本稿では, シミュレーションによる SF ネットワークの構築方法に注目し, これまでに提案されたいくつかの SF ネットワーク生成モデルの特徴を比較する. 特に, しきい値や複写等の局所メカニズムに基いたネットワーク構築方法を用いて, それらがつネットワーク構造の特徴について考察した.

## Characteristic comparison of SF network models

Yukio Hayashi,<sup>1</sup> Toshiyuki Miyazaki<sup>2</sup>

<sup>1,2</sup> School of knowledge science, Japan Advanced Institute of Science and Technology,  
1-1, Asahidai, Tatsunokuchi, Ishikawa 923-1292 Japan

**Abstract:** Previous research in physics and computer science reveals that real-world networks have common characteristics in spite of many different kinds of networks. For example, Internet, the World Wide Web, and power grid, and E-mail, and others, these networks with power-law degree distribution are called Scale-Free(SF) network. In this paper, based on computer simulations, we compare characteristics of SF networks. In particular, we consider several characteristics of these network structures constructed by local rules, such as threshold and duplication.

### 1. はじめに

ネットワーク構造に関するこれまでの研究により, 実世界に存在する多くのネットワークは, それぞれ種類が異なるにも関わらず, 共通の特徴を持つことが明らかにされている[1,2]. 例えば, インターネットや WWW, 電力網, 電子メールの送受信関係などが挙げられる. これらのネットワークは, 結合次数  $k$  がベキ乗則  $P(k) \sim k^{-\gamma}$  に従う分布を持ち, スケールフリー(SF)ネットワークと呼

ばれている.

基本的な SF ネットワーク生成モデルは, Barabasi-Albert(BA)モデル[2]である. BA モデルは, 成長(頂点と辺の追加)と優先的選択(次数に比例する確率で選択した頂点との連結)という単純な二つの規則を用いて, SF ネットワークを生成できる. これは成長過程で古い頂点がハブ(多数の連結を持つ頂点)になる傾向がある.

本稿は, これまでに提案されたいくつかの SF

ネットワーク生成モデルの中でも、BAモデル以外の局所メカニズムに基づく、いくつかの生成方法に注目し、シミュレーションにより生成したネットワーク構造の特徴を比較する。

## 2. ネットワーク構造の尺度

次数分布以外の尺度として、以下に示す平均結合相関: $\langle k_m \rangle$ , Assortativity 係数: $r$ , 平均経路長: $L$ , クラスタリング係数: $C$ , 平均クラスタリング係数: $\langle C \rangle$ を用いる。

### ・平均結合相関

次数  $k$  の隣接頂点の平均次数は、平均結合相関と呼ばれ、次数間の連結性質を捉えられる。平均結合相関の分布が正相関ならば、高い次数の頂点同士、低い次数の頂点同士が連結しており、負相関ならば、低い次数の頂点が高い次数の頂点に連結している。俳優の競演関係や電子メールの送受信関係のような社会的ネットワークの場合は正相関になり、インターネットや WWW のネットワークの場合は負相関になる。この相関関係を定量的な数値で表したものが、Assortativity 係数  $r$  である[3]。この係数の正負でネットワークの正相関、負相関を表せる。 $r$  の計算式は以下に従う。

$$r = \frac{\sum_{jk} jk(e_{jk} - q_j q_k)}{\sigma_q^2} \quad (1)$$

$$\sum_{jk} e_{jk} = 1, \quad \sum_k e_{jk} = q_j, \quad \sum_j e_{jk} = q_k$$

$e_{jk}$  は、次数  $j$  と次数  $k$  の頂点が接続されている確率で、 $\sigma_q$  は  $q_k$  の標準偏差である。

### ・平均経路長

頂点  $i$  の平均経路長  $L_i$  は、任意の頂点  $j$  への最短パス長  $d_{ij}$  の平均値である。ネットワークの平均経路長は、 $L_i$  の平均値である。

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

上式で、 $N$  は頂点数である。

### ・クラスタリング係数

頂点  $i$  のクラスタリング係数  $C_i$  は、隣接 2 頂点との相関関係の強さ、つまり三角形をクラスタと考え、実際に存在するクラスタ数  $e_i$  と可能なクラスタ数の割合を計算することで、ノードの密集度を計算できる。ネットワークのクラスタリング係数は、 $C_i$  の平均値である。

$$C_i = \frac{2e_i}{k_i(k_i-1)}, \quad C = \frac{1}{N} \sum_{i=1}^N C_i \quad (3)$$

上式で、 $k$  は次数、 $N$  はノード数である。

### ・平均クラスタリング係数

各次数毎の平均クラスタリング係数  $\langle C \rangle$  の分布を考える。これにより次数とクラスタリング係数の相関関係を捉えることができる。

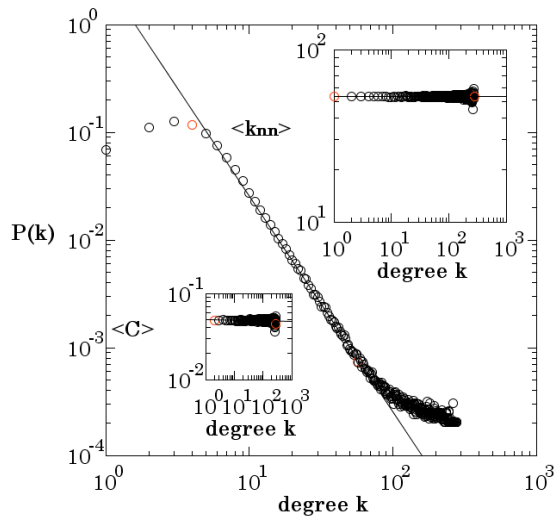
## 3. しきい値モデル[4]

成長原理を導入せずに SF ネットワーク構造を生成できるモデルとして、しきい値モデルを紹介する。頂点生成時点で全頂点に適応度を割り当て、2 頂点間の適応度に依存した確率でリンク構造を生成する。このモデルで生成したネットワークでは、最大適応度を持つ頂点がハブになりやすい。ネットワーク構築の方法は以下に従う。

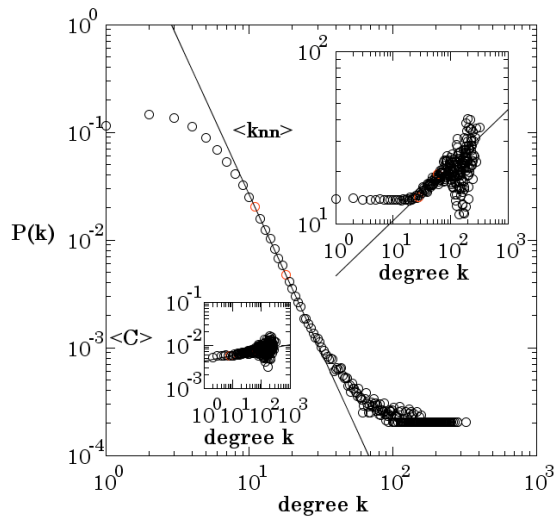
- (1) 最初に  $N$  個の非連結頂点群を生成する。各頂点は生成時に、適応度  $x_i$  が設定されている。適応度は、確率分布  $\rho(x)$  でランダムに割り当てられる実数値であり、重要度やランクを表現している。
- (2) 各頂点ペア  $i, j$  間のリンクは、2 頂点の適応度に依存した確率  $f(x_i, x_j)$  で連結、非連結を決定する。

$$f(x_i, x_j) = (x_i x_j) / x_M^2$$

確率分布  $\rho(x) \sim x^\beta$  で各頂点の適応度を割り当てるとき、リンク生成確率  $f(x_i, x_j) = (x_i x_j) / x_M^2$  を用いる。ここで  $x_M$  は、最大適応度である。この確率でネットワークを生成すると、べき乗則に従う次数分布を持つ SF ネットワークが得られる。



(a)  $\beta = 2$



(b)  $\beta = 3$

図 1.  $\rho(x) \sim x^{-\beta}$ ,  $f(x_i, x_j) = (x_i x_j) / x_M^2$  で生成した次数分布  $P(k)$ , 平均結合相関  $\langle k_{nn} \rangle$ , 平均クラスタリング係数  $\langle C \rangle$ .  $N = 5000$  で行い, 100 回の実験結果の平均値でプロットする.

表 1. 解析結果( $\rho(x) \sim x^{-\beta}$ )

	平均次数	$L$	$C$	$r$
$\beta = 2$	11.5564	3.2944	0.0451	-0.0036
$\beta = 3$	5.9667	4.6638	0.0051	-0.0065

表 2. ベキ指数( $\rho(x) \sim x^{-\beta}$ )

	次数分布	平均結合相関	平均クラスタリング係数
$\beta = 2$	2.0039	-0.0006	-0.0062
$\beta = 3$	2.9044	0.3219	0.1302

$\beta = 2, 3$  でネットワークを生成した結果, 得られた次数分布  $P(k)$ , 平均結合相関  $\langle k_{nn} \rangle$ , 平均クラスタリング係数  $\langle C \rangle$  を図 1 に示し, 解析結果を表 1, 2 に示す. 表 1 の結果は, 平均経路長は 20 回計算した平均値で, 他の値は 100 回計算した平均値で示す.

次数分布のベキ指数は,  $\beta$  とほぼ同じ値になっている(表 2). また  $\beta$  の増加により, 平均結合相関, 平均クラスタリング係数のベキ指数は, 共に正相関になる傾向がある(表 2).

#### ・しきい値モデル

各頂点の適応度分布がベキ乗則に従わないときを考える. 確率分布として指数関数分布  $\rho(x) \sim e^{-x}$  を用いる. このときリンク生成確率  $f$  は, 次のヘビサイド関数を用いる.

$$f(x_i, x_j) = \theta[x_i + x_j - z(N)] \quad (4)$$

各頂点ペアは, 互いの適応度の合計がしきい値  $z(N)$  より大きい場合のみリンクを張れる.

しきい値モデルでネットワークを生成した結果, 得られた次数分布  $P(k)$ , 平均結合相関  $\langle k_{nn} \rangle$ , 平均クラスタリング係数  $\langle C \rangle$  を図 2 に示し, 解析結果を表 3, 4 に示す. 表 3 の結果は, 平均経路長は 20 回計算した平均値で, 他の値は 100 回計算した平均値で示す.

しきい値モデルで重要なことは, ベキ乗でない適応度分布からも, SF ネットワークを生成できる点である.

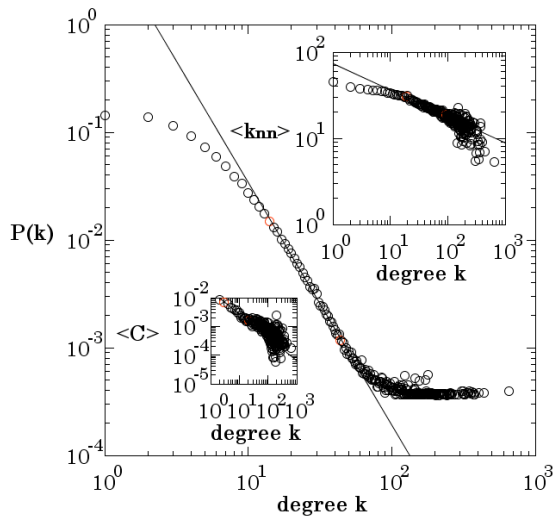
表 3. 解析結果(しきい値モデル)

平均次数	$L$	$C$	$r$
7.5369	1.9917	0.5147	-0.4789

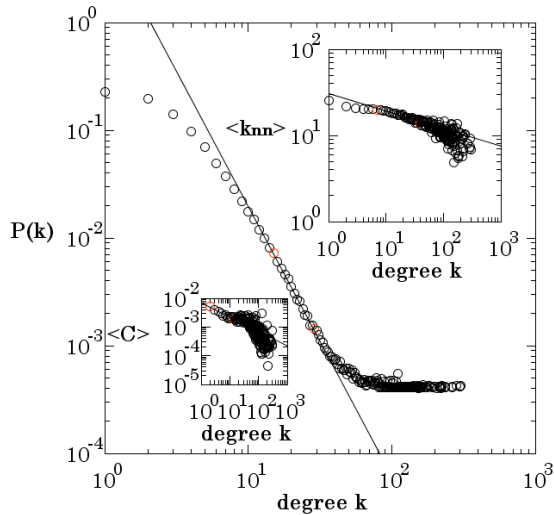
表 4. ベキ指数(しきい値モデル)

次数分布	平均結合相関	平均クラスタリング係数
2.0245	-0.8522	-1.6033





(a)  $\delta = 0.5, \beta = 0.5$



(b)  $\delta = 0.6, \beta = 0.5$

図4. gene duplication model で生成したネットワークの度数分布  $P(k)$ , 平均結合相関  $\langle k_{nn} \rangle$ , 平均クラスタリング係数  $\langle C \rangle$ .  $\delta = 0.5, 0.6, \beta = 0.5, N = 3000$  で行い, 100 回の実験結果の平均値でプロットする.

表5. 解析結果(gene duplication model)

	平均次数	$L$	$C$	$r$
$\delta = 0.5, \beta = 0.5$	8.7633	4.1688	0.0038	-0.1446
$\delta = 0.6, \beta = 0.5$	5.3819	4.9563	0.0023	-0.1002
$\delta = 0.5, \beta = 0.6$	8.8927	4.1570	0.0042	-0.1306
$\delta = 0.6, \beta = 0.6$	5.2216	4.9989	0.0025	-0.0818

表6. ベキ指数(gene duplication model)

	度数分布	平均結合相関	平均クラスタリング係数
$\delta = 0.5, \beta = 0.5$	2.2546	-0.3049	-0.8136
$\delta = 0.6, \beta = 0.5$	2.5093	-0.2005	-0.5041
$\delta = 0.5, \beta = 0.6$	2.2950	-0.2995	-0.8039
$\delta = 0.6, \beta = 0.6$	2.5043	-0.2505	-0.6400

## 5. Coupled duplication-divergence model[6]

4章では gene duplication model を紹介したが, 本章では, 別の生成方法による複製モデルを紹介する. このモデルでは, 複製(Duplication), 発散(Divergence), 自己相互作用(Self-interaction)の三つの操作をある確率に従って行う. gene duplication model との違いは, より局所的に限定したネットワーク内で, 操作を実行している点である. Coupled duplication-divergence model の複製操作を図5に示す.

Coupled duplication-divergence model の成長方法は, 以下のルールに従う.

- (1) 既存頂点群から, ランダムに頂点を選択し, 複製操作を実行する(Duplication).
- (2) 確率  $q_v$  で, 新規頂点とランダムに選択した頂点を連結する(Self-interaction).
- (3) 確率  $1 - q_e$  で, ランダムに選択した頂点を持つリンクと新規に追加した複製リンクの中からランダムに一本選択し, 削除する(Divergence).  $q_e$  の値が減少すると, リンクの損失が増加する.

ネットワークを上の方法で生成した結果, 得られた度数分布  $P(k)$ , 平均結合相関  $\langle k_{nn} \rangle$ , 平均クラスタリング係数  $\langle C \rangle$  をそれぞれ図6, 7, 8に示し, 解析結果を表7, 8に示す. 表7の結果は, 平均経路長は20回計算した平均値で, 他の値は100回計算した平均値で示す.

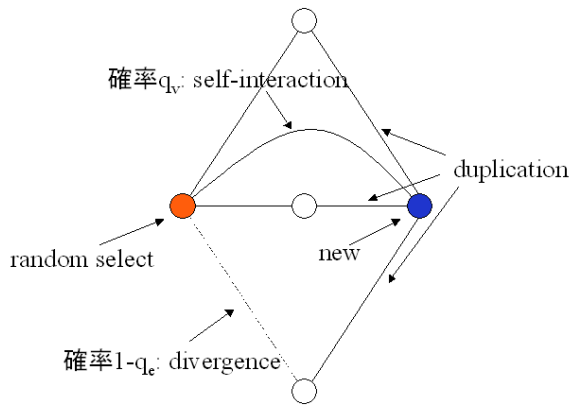


図 5. 複写操作

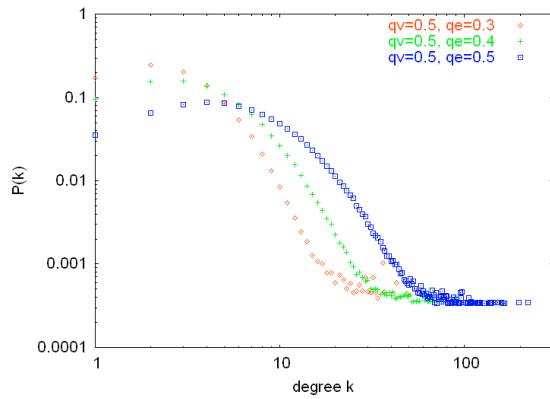


図 6. coupled duplication model で生成した次数分布  $P(k)$ .  $N=3000$ ,  $q_v=0.5$ ,  $q_e=0.3, 0.4, 0.5$  で行い, 100 回の実験結果の平均値でプロットする.

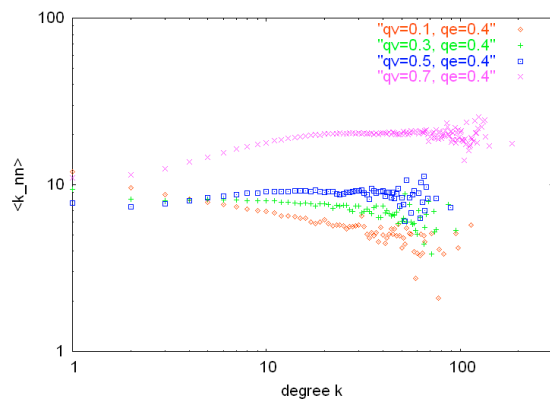


図 7. coupled duplication model で生成した平均結合相関  $\langle k_m \rangle$ .  $N=3000$ ,  $q_v=0.1, 0.3, 0.5, 0.7$ ,  $q_e=0.4$  で行い, 100 回の実験結果の平均値でプロットする.

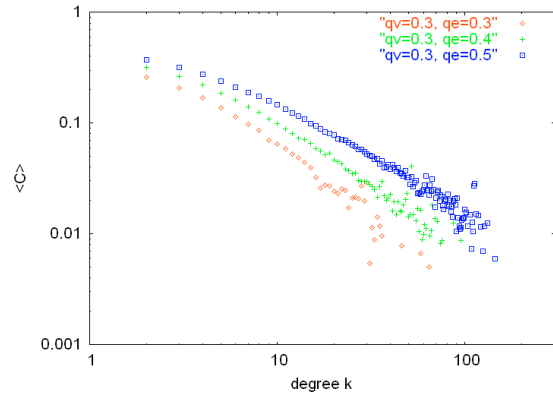


図 8. coupled duplication model で生成した平均クラスタリング係数  $\langle C \rangle$ .  $N=3000$ ,  $q_v=0.3$ ,  $q_e=0.3, 0.4, 0.5$  で行い, 100 回の実験結果の平均値でプロットする.

表 7. 解析結果(coupled duplication model)

	平均次数	$L$	$C$	$R$
$q_v=0.1$ , $q_e=0.3$	2.7463	11.2201	0.0578	-0.2156
$q_v=0.1$ , $q_e=0.4$	3.9487	8.5536	0.0707	-0.1701
$q_v=0.1$ , $q_e=0.5$	7.0366	5.1959	0.0812	-0.1721
$q_v=0.3$ , $q_e=0.3$	3.0671	12.1785	0.1450	-0.0749
$q_v=0.3$ , $q_e=0.4$	4.5105	7.7910	0.1800	-0.0430
$q_v=0.3$ , $q_e=0.5$	8.0686	5.4336	0.1953	-0.0415
$q_v=0.5$ , $q_e=0.3$	3.4491	12.4917	0.2107	0.0295
$q_v=0.5$ , $q_e=0.4$	5.2521	7.9568	0.2529	0.0533
$q_v=0.5$ , $q_e=0.5$	9.6612	5.5706	0.2612	0.0365
$q_v=0.7$ , $q_e=0.3$	3.9258	11.2955	0.2559	0.1189
$q_v=0.7$ , $q_e=0.4$	6.1327	7.8411	0.3007	0.1163
$q_v=0.7$ , $q_e=0.5$	11.8411	5.0735	0.2945	0.0889

$q_v$ の増加でクラスタリング係数  $C$ が増加し,  $q_e$ の増加で平均経路長  $L$ が減少する傾向がある(表 7). 平均結合相関グラフ, 平均結合相関ベキ指数の結果より,  $q_v$ の増加で平均結合相関 $\langle k_{\text{nn}} \rangle$ が, 負相関から正相関に変化する(図 7, 表 8). 次数分布ベキ指数は,  $q_v = 0.1$  以外のとき,  $q_e$ の増加で減少している(表 8). 平均クラスタリング係数ベキ指数は, すべて-1.0に近い値を取る(表 8).

表 8. ベキ指数(coupled duplication model)

	次数分布	平均結合相関	平均クラスタリング係数
$q_v = 0.1, q_e = 0.3$	3.0798	-0.2620	-0.9531
$q_v = 0.1, q_e = 0.4$	3.0085	-0.2049	-1.0485
$q_v = 0.1, q_e = 0.5$	3.0277	-0.2044	-1.0028
$q_v = 0.3, q_e = 0.3$	4.0506	-0.0706	-1.1018
$q_v = 0.3, q_e = 0.4$	3.5074	-0.0116	-1.0418
$q_v = 0.3, q_e = 0.5$	3.0392	0.0388	-0.9028
$q_v = 0.5, q_e = 0.3$	4.4756	0.0818	-1.1082
$q_v = 0.5, q_e = 0.4$	4.0153	0.1361	-1.0778
$q_v = 0.5, q_e = 0.5$	3.4508	0.1695	-0.9516
$q_v = 0.7, q_e = 0.3$	4.7095	0.2012	-1.2098
$q_v = 0.7, q_e = 0.4$	4.0589	0.2516	-1.0058
$q_v = 0.7, q_e = 0.5$	3.5001	0.3015	-0.9496

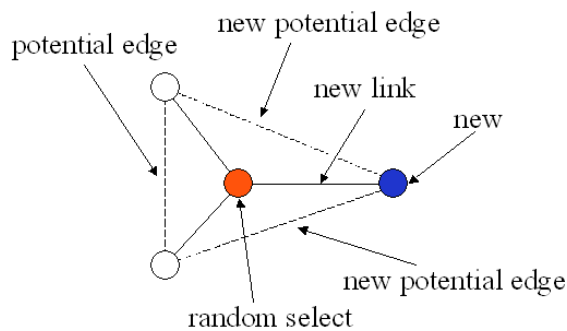


図 9. connecting nearest-neighbor model におけるリンク接続の説明図. 確率  $u$  でポテンシャルエッジ(点線)が実際のリンクになる. 確率  $1-u$  で新規頂点はランダムに選択した頂点と連結する(new link). 同時にポテンシャルエッジも生成される(new potential edge).

## 6. connecting nearest-neighbor model[6]

このモデルでは, 新規頂点がランダムに選択した頂点と連結して成長する. このとき選択した頂点の隣接頂点と新規頂点との間に, 連結する可能性が生じる. つまり単純に言えば「友達の友達はみんな友達」になる. connecting nearest-neighbor model のリンクの生成を図 9 に示す.

connecting nearest-neighbor model の成長方法は, 以下のルールに従う.

- (1) 確率  $1-u$  で新規頂点を追加し, ランダムに選択した頂点と連結する. 選択頂点の隣接頂点と新規頂点との間をポテンシャルエッジとして, 実際のリンクとは別に保存する.
- (2) 確率  $u$  でポテンシャルエッジからランダムに一本選択し, 実際のリンクに変換する.

ネットワークを上の方法で生成した結果, 得られた次数分布  $P(k)$ , 平均結合相関 $\langle k_{\text{nn}} \rangle$ , 平均クラスタリング係数 $\langle C \rangle$ をそれぞれ図 10, 11, 12 に示し, 解析結果を表 9, 10 に示す. 表 9 の結果は, 平均経路長は 20 回計算した平均値で, 他の値は 100 回計算した平均値で示す.

平均結合相関 $\langle k_{\text{nn}} \rangle$ が正相関である SF ネットワークを生成できることが connecting nearest-neighbor model の特徴である(図 11). 平均結合相関のベキ指数は 0.6 に近い値を取る(表 10). 平均クラスタリング係数のベキ指数は確率  $u$  に関わらず-0.6 に近い値を取る(表 10).  $u$ が増加するとクラスタリング係数も増加する(表 9).

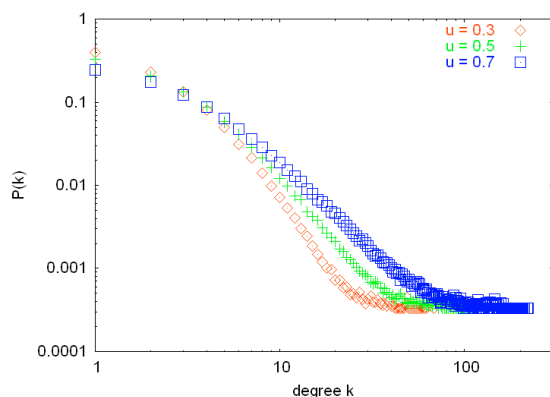


図 10. connecting nearest-neighbor model で生成した次数分布  $P(k)$ .  $N = 3000$ ,  $u = 0.3, 0.5, 0.7$  で行い, 100 回の実験結果の平均値でプロットする.

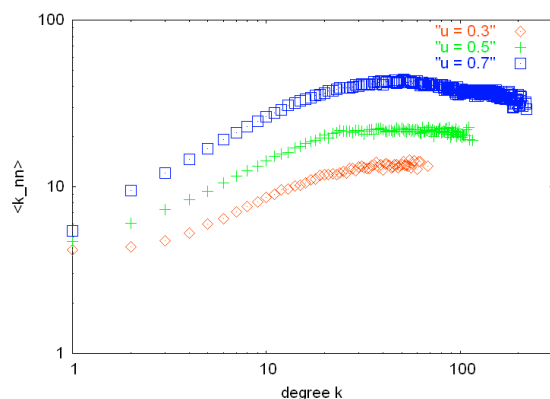


図 11. connecting nearest-neighbor model で生成した平均結合相関  $\langle k_m \rangle$ .  $N = 3000$ ,  $u = 0.3, 0.5, 0.7$  で行い, 100 回の実験結果の平均値でプロットする.

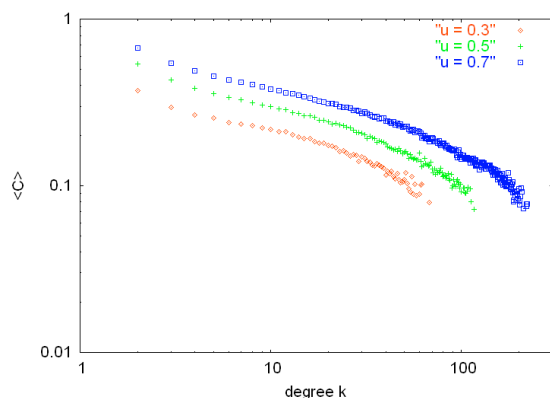


図 12. connecting nearest-neighbor model で生成した平均クラスタリング係数  $\langle C \rangle$ .  $N = 3000$ ,  $u = 0.3, 0.5, 0.7$  で行い, 100 回の実験結果の平均値でプロットする.

表 9. 解析結果(nearest-neighbor model)

	平均次数	$L$	$C$	$R$
$u = 0.3$	2.8606	9.2244	0.1834	0.3190
$u = 0.5$	3.9925	7.4651	0.2789	0.2794
$u = 0.7$	6.6594	5.9951	0.3675	0.1888

表 10. ベキ指数(nearest-neighbor model)

	次数分布	平均結合相関	平均クラスタリング係数
$u = 0.3$	3.3164	0.5042	-0.5990
$u = 0.5$	2.6091	0.5600	-0.6023
$u = 0.7$	2.2001	0.6318	-0.5819

## 7. おわりに

本稿では, 局所メカニズムに注目し, いくつかの SF ネットワーク生成モデルを紹介した. 今回紹介したモデル以外にも, 多くの SF ネットワークモデルが提案されている. 現在, このような SF ネットワークの大規模故障や電子メールウィルスの拡散に対する, トポロジー的な応用研究が多数行われており, その基盤となるモデルの特徴の理解は重要である.

## 参考文献

- [1] S. N. Dorogovtsev, and J. F. F. Mendes, "Evolution of Networks," Oxford University Press, (2003).
- [2] R. Albert, and A. -L. Barabasi, "Statistical Mechanics of Complex Network," arXiv:cond-mat/0106096v1, (2001).
- [3] M. E. J. Newman, Phys, Rev, E **67**, 026123, (2003).
- [4] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Munoz, Phys. Rev. **89**, 258702, (2002).
- [5] S. Bornholdt, and H. G. Schuster, "Handbook of Graphs and Networks," Wiley-VCH, (2003).
- [6] A. Vazquez, Phys, Rev, E **67**, 056104, (2003).