

BA モデルおよび Configuration モデルで生成したネットワークの Uncorrelated 性について

谷聖一 研究室 飯田 佳徳 ・ 丸山隆太
Yoshinori Iida, Ryuta Maruyama

概要

有限ネットワーク上の情報伝搬において、ネットワーク全体に情報が行き渡るまでの時間は、情報を持っている頂点が隣接する頂点の中から情報の伝搬先をどのように選択するのかに依存する。各頂点が記憶能力を持たず、ソースノードが単位時間に 1 つの隣接頂点にのみ情報を伝搬するモデルにおいて、ネットワークが Uncorrelated 性を持つならば、次数の逆数に比例する確率分布に従い伝搬先の頂点を乱択する方法が、ネットワーク全体に情報が伝搬されるまでの平均時間が最小になることが証明されている。この定理を適用するにはネットワークが 真に Uncorrelated 性を持つ必要があるが、その判定法は、報告者の知る範囲では確立されていない。そこで本演習では、ソースノードが情報を伝搬させる隣接頂点を選択する際に次数の逆数に比例した確率分布に従い乱択する方法が効率の良いことと、ネットワークが Uncorrelated 性を持つことの関連性を検討するため、Uncorrelated 性を持つネットワークを生成できるとされているヒューリスティックでネットワークを生成し、それらに対して情報伝搬実験を行った。

1 はじめに

1.1 背景

私たちの身の回りには数多くのネットワークが存在している。一般にネットワークというとインターネットの意味で用いられることが多いが、本稿におけるネットワークとは「つながり」全般を指している。例をあげると人間関係、食物網、交通網、脳や WWW などである。現実世界に存在するネットワークは非常に大規模かつ複雑であることから「複雑ネットワーク」とも呼ばれ、それらには共通して見られる普遍的な性質がある。ネットワークの問題はグラフとして抽象化することで、数学的に解析することが可能である。現実世界に多く存在し、数理的に解析が可能ということで、ネットワークの研究は幅広く行われてきた ([1])。

1.2 目的

論文 [4] では次が示された。

定理 1. 各頂点が記憶能力を持たず、ソースノードが単位時間に 1 つの隣接頂点にのみ情報を伝搬するモデルにおいて、ネットワークが Uncorrelated 性を持つなら、ネットワーク全体に情報が伝搬されるまでの平均時間は、情報を伝搬させる隣接頂点を選択する際に次数の逆数に比例した確率分布に従い乱択する方法が最小となる。

BA モデルで生成したネットワーク上でこの定理を適用できるかを検証するために、BA モデルおよび修正 BA モデルでネットワークを生成し情報伝搬実験を行った。すると「次数の逆数に比例した確率分布に従い乱択する方法」が最も効率的なネットワークとそうではないネットワークが存在するという結果が得られた。定理 1 の対偶をとると「各頂点が記憶能力を持たず、ソースノードが単位時間に 1 つの隣接頂点にのみ情報を伝搬するモデルにおいて、ネットワーク全体に情報が伝搬されるまでの平均時間が情報を伝搬させる隣接頂点を選択する際に次数の逆数に比例した確率分布に従い乱択する方法が最小とならないならば、ネットワークが Uncorrelated 性を持たない」となる。このことから「次数の逆数に比例した確率分布に従い乱択する方法」が最も効率的ではないネットワークが存在する原因がネットワークの Uncorrelated 性が弱いことにあるのではないかと考えられる。そのため、[15] では「次数の逆数に比例した確率分布に従い乱択する方法」が最も効率的なネットワークとそうではないネットワークの特徴量を比較し、Uncorrelated 性と関連する特徴量の調査が行われた。しかし、Uncorrelated 性と明確に関連のある特徴量の提案には至らなかった。

そこで、BA モデル以外のスケールフリーネットワーク生成モデルで生成したネットワークを生成し、特徴量の比較を行うことにした。本演習では、修正 Configuration モデル ([9]) を用いてネットワークを生成し、それらに対して情報伝搬実験を行った。

[11, 16] などの先行研究では、BA モデルおよび修正

1.3 構成

本稿は以下のような構成になっている。2 節では、まず本演習を行うにあたっての準備について述べる。3 節では、本演習で用いた伝搬モデルについて述べる。4 節では、Uncorrelated ネットワークについて述べる。5 節では、ネットワーク生成モデルについて述べる。6 節では、実験方法について述べる。7 節では、ネットワークの特徴量について述べる。8 節では、実験結果と考察について述べる。9 節では、今後の課題について述べる。

2 準備

無向グラフ G は頂点集合 V と辺集合 E からなり、

$$G = (V, E)$$

と表す。ただし、辺は頂点の対である。また、頂点から出る枝の本数を次数と呼ぶ。以下に友人関係を例にして紹介をする。

- 頂点：人
- 辺：人と人の間の友人関係
- 次数：一人の人がもつ友人関係の数

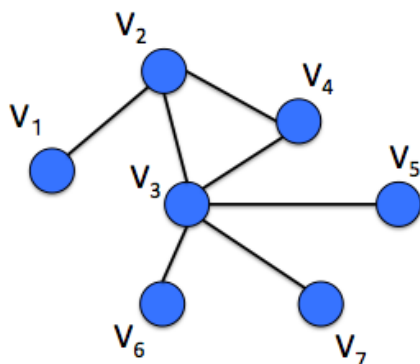


図 1: 友人関係

図 1 のグラフは、

$$V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$$

$$E = \{(v_1, v_2), (v_2, v_3), (v_2, v_4), (v_3, v_4), (v_3, v_5), (v_3, v_6), (v_3, v_7)\}$$

のように表せる。また、2 頂点間に複数の辺があるとき、それらを多重辺と呼び、辺の両端の頂点が同一であるとき、その辺を自己ループと呼ぶ。グラフの任意の 2 頂点間を結ぶ経路が存在するとき、そのグラフは連結であるという。

また、本稿において「グラフ」と「ネットワーク」は同義であり、 V の要素をノードと呼び、 E の要素を枝と呼ぶこともある。

現実の多くの複雑ネットワークは、スケールフリー性を持ったスケールフリーネットワークである。スケールフリー性とは、複雑ネットワークがもつ普遍的な性質の 1 つである。また、ネットワークがスケールフリー性をもつとは、ネットワークの次数分布がべき則であることをいう。

次数分布がべき則であるとは

$$P(k) \propto k^{-\gamma}, (\gamma > 0)$$

を満たしていることである ([1])。($\propto$ は比例を表し、 γ はべき指数と呼ぶ。) つまり、少ない頂点が多数の枝を持ち、他の多くの頂点とつながっており、多数の頂点が少ない枝しか持たないということである。そして、この多数の頂点と接続している頂点のことをハブという。世の中の多くのネットワークはこの性質を有していることが確認されている。またネットワークにまつわる驚きの多くが、この性質から導かれる ([1])。

また、スケールフリーネットワークの次数分布を両対数で描写すると直線となる。

3 本演習で用いた伝搬モデル

本演習では、本演習の発端となった論文 [4] と同様の伝搬モデルを用いて情報伝搬実験を行った。ここでは本演習で用いた伝搬モデルについて説明を行う。

説明を行うにあたり、ソースノードとターゲットノードを以下のように定義する。

- ソースノード：情報を保持し、隣接頂点に情報を発信できる頂点
- ターゲットノード：隣接するソースノードから情報の伝搬先として選択された頂点

3.1 情報伝搬の制約・ルール

情報伝搬のルールは以下の通りである。

- 初期状態でのソースノードは 1 つ
- ソースノードは単位時間 (1 ステップ) に隣接頂点のうちの 1 つをターゲットノードとして選択し情報伝搬
- 情報を受け取った頂点はソースノードとなる (その後もソースノードであり続ける)
- 全ての頂点がソースノードとなると、情報伝搬が完了とする

また、戦略を決めるのに使える情報は次の通りである。

- ソースノードが使える情報は、隣接頂点の次数のみこの使える情報のルールは、以下のような状況で生じる。
 1. 隣接頂点同士は、お互いに情報交換を行わない
 2. 各頂点は、情報伝搬に関して記憶能力を持たない
 3. 各頂点は、隣接頂点の次数は除き、ネットワークの構造はわからない

3.2 ターゲットノード選択方法

3.1 節より、戦略を決めるのにソースノードが使える情報は隣接頂点の次数のみである。そのため、ソースノードが隣接頂点の中からターゲットノードを選択する戦略は以下のようなものが考えられる。

1. 一様分布に従い乱択
2. 次数の大きな頂点を優先的に乱択
3. 次数の小さな頂点を優先的に乱択

本演習では、上記の 3 つの戦略で情報伝搬実験を行った。

また、ソースノード a が、 a の隣接頂点の中から b をターゲットノードとして選択する確率は次の通りである。

$$q(b; a) = \frac{w_b}{\sum_{i=1}^{k_a} w_{a_i}} \quad (1)$$

ここで、 w_i は頂点 i の重み、 k_i は頂点 i の次数、 a_i は頂点 a の隣接頂点の i 番目である。上記の 3 つの戦略のいずれを用いるかによって w_i の値が以下のように変化する。

- 一様分布に従い乱択 :
 $w_i = 1$
- 次数の大きな頂点を優先的に乱択 :
 $w_i = \text{頂点 } i \text{ の次数}$
- 次数の小さな頂点を優先的に乱択 :
 $w_i = \text{頂点 } i \text{ の次数の逆数}$

4 Uncorrelated ネットワーク

Uncorrelated 性とはネットワークにおける様々な研究で用いられる性質の 1 つであるが、論文や文献によって定義が異なる。“correlate” とは「相互に関連がある」

という意味で、“uncorrelated” とは「無相関の、無関係の」という意味である。そこから Uncorrelated ネットワークとは「局所性がなく、どこも同じような構造をしている」ネットワークのことをいう。Uncorrelated 性を持つネットワークは、統計的、平均的に解析できることが多い。

本演習で用いた伝搬モデルにおいて、ソースノード a によって隣接頂点 b がターゲットノードとして選択される確率は 3 節の式 (1) と定まる。論文 [3] では、与えられたネットワーク N が Uncorrelated ネットワークであれば、式 (1) で与えられるのターゲットノードの選択確率が

$$q_b = \frac{k_b w_b}{\sum_{j=1}^n k_j w_j} \quad (2)$$

で近似できることを証明している。式 (2) は全頂点の中から頂点 b がターゲットノードとして選択される確率となっており、ソースノード a の選択に依らない。この近似

$$q(b; a) \approx q_b$$

を用いて定理 1 を証明しているため、この近似が成り立つネットワークでなければ定理 1 の証明は成り立たない。また、「Uncorrelated ネットワークならば、近似 $q(b; a) \approx q_b$ が成り立つ」の待遇「近似 $q(b; a) \approx q_b$ が成り立たないならば、Uncorrelated ネットワークではない」も成り立つことに注意。

5 ネットワーク生成モデル

スケールフリーネットワークを生成するモデルは数多く提唱されており、その 1 つとして BA モデルがあげられる。BA モデルとは 1999 年に Barabási と Albert によって提唱されたネットワーク生成モデルで、2 人の頭文字をとって BA モデルと呼ばれる ([5])。そして、木構造のネットワークを生成するために BA モデルに修正を加えたモデルを修正 BA モデルという。

また、本演習ではより Uncorrelated 性をもつネットワークを生成するため、上記のモデルに加え Configuration モデルと修正 Configuration モデルを用いる。Configuration モデルとは、一般に次数相関が現れにくいモデルで、任意の次数分布を作ることができる ([1])。そして [9] で Configuration モデルに修正を加え、より Uncorrelated 性を持つネットワークを生成することを意図し提案されたヒューリスティックを、本稿では修正 Configuration モデルという。

本節では、BA モデル、修正 BA モデル、Configuration モデルおよび修正 Configuration モデルについて述べる。

5.1 BA モデルのアルゴリズム

BA モデルでネットワークを生成するアルゴリズムは以下のとおりである ([5]) .

1. $m > 1$ 個の頂点からなる完全グラフを置く .
2. 新しい頂点を 1 個追加し, 既に存在している頂点のうちの m 個の頂点に対して, 枝を張る . この時, 新しい枝が張られる確率は, 各頂点のその時点での次数 k_i に比例する .

$$\frac{k_i}{\sum_{j=1}^n k_j} (1 \leq i \leq n)$$

となる (n は頂点数) .

3. 指定の頂点数になるまで Step2 を繰り返す .

新しく追加される頂点の次数は m で固定されている . そのためネットワーク全体の最低次数は m となる . BA モデルの特徴は既に多くの枝を持つ頂点がその後も新たな枝を獲得しやすく, 少ない枝を持つ頂点はその後も新たな枝を獲得しにくいことである ([1]) .

5.2 修正 BA モデルのアルゴリズム

修正 BA モデルでネットワークを生成するアルゴリズムは以下のとおりである ([14]) .

1. 枝を保有しない既存の頂点を 1 個置く .
2. 新しい頂点を 1 個追加し, 既に存在している頂点のうちの 1 つに枝を張る . この時, 新しい枝が張られる確率は, 各頂点のその時点での次数に比例する .
3. 指定の頂点数になるまで Step2 を繰り返す .

修正 BA モデルも BA モデルと同様に, 既に多くの枝を持つ頂点がある後も新たな枝を獲得しやすく, 少ない枝を持つ頂点はその後も新たな枝を獲得しにくい特徴を持つ . そして生成されるネットワークは木構造となる .

5.3 Configuration モデルのアルゴリズム

Configuration モデルでネットワークを生成するアルゴリズムは以下のとおりである ([1]) .

1. べき則の確率分布に従い, 整数列 $(k_1, k_2, \dots, k_n)$ を発生させる (n は全頂点数)
2. n 個の頂点を置き, $v_i (1 \leq i \leq n)$ が k_i 個の枝の片割れを持つようにする .

3. 全ての枝の片割れから 2 つを一様分布に従い乱択し, つなぐ . 2 つの枝の片割れをつなぐと 1 本の枝になる .

4. 枝の片割れがなくなるまで Step 3 を繰り返す .

1 つの頂点から 2 つの枝の片割れが選ばれた場合は自己ループが発生する . また, すでに枝が張られた 2 頂点で 1 つずつ枝の片割れが選ばれた場合は多重辺が発生する . また, 生成の結果ネットワークが非連結となることもある . このように, Configuration モデルは自己ループや多重辺の発生, 非連結となる場合を避けることができない .

Configuration モデルで単純なネットワーク (自己ループ・多重辺を持たないネットワーク) を生成するには, step3 で 2 つの枝の片割れが選ばれた際に, 同一頂点から 2 本の枝が選ばれていないか, すでに 2 頂点間に枝が張られていないかの判定を行う . 判定を行い, 自己ループ・多重辺になるのであれば 2 つの枝の片割れの選択を破棄し, 再度枝の片割れの乱択を行う .

Configuration モデルでは, step1 で発生させた整数列を次数列に持つ単純かつ連結なネットワークが存在しない場合がある . そのため, 効率よくネットワークの生成を行うためには, 与えられた整数列から単純かつ連結なネットワークを生成できるのかを判定する必要がある . 単純かつ連結なネットワークを生成するためには, 与えられた整数列が以下の 4 つを満たしている必要がある . ここで, n は頂点数である .

1. 全ての要素が 1 以上である .
2. 整数列の総和が $2(n-1)$ 以上である .
3. 整数列の総和が偶数である .
4. 整数列がグラフ的である .

2 つ目の「次数の総和が $2(n-1)$ 以上である .」とは, 頂点数 n の木 (連結であることに限らず) よりも枝が多くなければならないということである . 3 つ目の「次数の総和が偶数である .」とは, 枝の片割れは 2 つで 1 つの枝となるので, 枝の片割れの数 = 次数の総和 は偶数でなければならないということである . また, 4 つ目の「整数列がグラフ的である .」とは, その整数列を次数列とする単純グラフが存在するということである . 与えられた整数列がグラフ的であるかの判定をする手法の 1 つに Havel-Hakimi の定理がある ([10]) . 以下で, 与えられた整数列がグラフ的であるかを判定する Havel-Hakimi のアルゴリズムを紹介する .

— Havel-Hakimi のアルゴリズム —

1. 与えられた整数列を降順に並べ替える .
2. 先頭の要素 k_1 を消去し, それに続く k_1 個の要素から 1 を引く
3. Step 1, Step 2 を全ての要素が 0 になるか, 負の要素が出てくるまで繰り返す .
Step 3 で全ての要素が 0 になれば, その整数列はグラフ的である . 途中で負の要素が発生すれば, その整数列はグラフ的ではない .

5.4 修正 Configuration モデル

Configuration モデルで生成されるネットワークは次数相関がほとんどないが, 単純かつ連結なネットワークを生成しようとするとき次数相関が発生し Uncorrelated 性が弱くなる . そこで, 自己ループや多重辺を持たないネットワークを生成する Configuration モデルに修正を加え, より Uncorrelated 性を持つネットワークを生成するための手法が提案されている ([9]) . ここではその提案手法について述べる .

— 修正 Configuration モデル —

Configuration モデルの Step 1 で整数列を発生させる際に全ての要素を, $2 \leq k \leq n^{1/2}$ に制限 (カットオフ) する (n は全頂点数)

この手法は, Uncorrelated 性を持つネットワークを生成できるとされているヒューリスティックであり, この手法により生成されたネットワークが真に Uncorrelated 性を満たすという数学的な証明はない .

6 実験方法

1. BA モデルで初期完全グラフの頂点数を 1 個, 2 個, 4 個, 6 個, 8 個, 10 個として全頂点数 10000 個のネットワークを各 100 個生成する . また, Configuration モデル (自己ループ・多重辺を許す), Configuration モデル (自己ループ・多重辺を許さない), 修正 Configuration モデルそれぞれで, ベキ指数を 2.5, 2.75, 3.0 として全頂点数 10000 個のネットワークを各 100 個生成する . 生成するネットワークは合計で 1500 個となる .
2. 各ネットワークに対して次の特徴量を求める .
 - 半径
 - 直径
 - 平均距離

- 次数分布
- 中心性
- 次数相関 (ピアソンの積率相関係数)

3. 各ネットワークに対して情報伝搬実験を行い, 次数の小さい頂点を優先的に乱択する方法の伝搬効率を測る .

なお, ネットワーク生成, 特徴量の抽出, 情報伝搬実験を行うプログラムには C++ 言語を使用し, グラフ出力には gnuplot を使用した . また実験プログラムの管理にはシェルスクリプトを用いた .

6.1 情報伝搬シミュレーション

本演習では次の情報伝搬シミュレーションを行った .

入力: ネットワーク N , シミュレーション番号 i

処理: 本演習で用いた伝搬モデルで, ランダムシードを i とし, 初期情報源を乱択し, すべてのノードが情報源となるまで情報伝搬をシミュレーションし, 情報源数の時間変化を記録

ただし, ターゲットノードの選択は乱択である . 公平を期すために, 全ての実験において, 同じシミュレーション番号の実験では同じランダムシードを用いた .

6.2 RPC 伝搬効率

4 節で紹介した 3 つのターゲット選択方法で, 生成した全てのネットワークに対して 100 回の情報伝搬実験を行い, 全頂点が情報源となるまでに要したステップ数を比較する . 「次数の小さな頂点を優先的に乱択する方法 (Reverse Preferential Control, 以下 RPC とする .)」の総ステップ数が最も小さかった割合を百分率で表し RPC 伝搬効率とする .

7 ネットワークの特徴量

本節では, 計測を行った特徴量について述べる . $G = (V, E)$ において, 頂点 v から各頂点への距離の最大値を G の頂点 v の eccentricity という . この eccentricity の最大値を G の直径, 最小値を G の半径という . また eccentricity が最小である頂点を中心という .

平均距離は任意の 2 頂点間の距離の平均である . また, 次数分布は各次数が全頂点に占める割合である .

7.1 中心性の計測

中心性は，ネットワークにおける各頂点の重要性を評価したり，比較したりするための指標である．ここでは，以下の二つの中心性について説明を行う．

7.1.1 近接中心性

近接中心性は，他の頂点との距離が近いほど中心性が高いとする指標である．各頂点の近接中心性は

$$C_c(v) = \frac{N-1}{\sum_{t \in V} d_G(v, t)}$$

として定義される．ここで， $d_G(v, t)$ は v, t 間の最短距離である．

7.1.2 媒介中心性

媒介中心性は，頂点がネットワーク上の流れにどの程度寄与しているのかを表すものである．各頂点の媒介中心性は

$$C_B(v) = \sum_{\substack{s, v, t \in V; \\ s \neq v \neq t}} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

として定義される．ここで， σ_{st} は s, t 間の最短経路数， $\sigma_{st}(v)$ は v を通る s, t 間の最短経路数である．

媒介中心性を定義に従い求めた場合の時間計算量は， N を頂点数とし $O(N^3)$ であるが，本演習では $O(N^2 + NM)$ で媒介中心性を求めることができる Brandes が提案したアルゴリズム [6] を用いて計測を行った．本演習で扱うスケールフリーネットワークは枝数が頂点数の定数倍程度であるため，Brandes のアルゴリズムを用いると本演習においては， $O(N^2)$ の計算量で媒介中心性を求めることができる．しかし，より大規模なネットワークでは Brandes のアルゴリズムを用いても膨大な時間が必要となる．そこで，より高速に媒介中心性を求めるための近似手法が多く提案されている ([7, 8]) ．

7.2 ピアソンの積率相関係数

ピアソンの積率相関係数は 1 つのネットワークに対して与えられ，以下の式で求めることができる．

$$r = \frac{Cov(K, K')}{\sigma_K \sigma_{K'}} = \frac{E[KK'] - \langle K \rangle \langle K' \rangle}{\sqrt{\langle K^2 \rangle - \langle K \rangle^2} \sqrt{\langle K'^2 \rangle - \langle K' \rangle^2}}$$

K と K' は，枝の両端点の残り次数 ($K = k - 1$) である．また， $Cov(K, K')$ は K と K' の共分散を， σ_K は K の標準偏差を， $E[KK']$ は K と K' の積の期待値を，

$\langle K' \rangle$ は K の平均値をそれぞれ表している．ピアソンの積率相関係数は $-1 \leq r \leq 1$ の値を取り，次数の相関が強ければ 1，-1 に，次数の相関が弱ければ 0 に近い値を取る．

8 実験結果と考察

本演習では様々な特徴量を求めたが，BA モデルと Configuration モデルで，ピアソンの積率相関係数以外に明確に差のある特徴量を見つけることはできなかった．ここでは，実験結果の一部として，次数分布，RPC 伝搬効率，ピアソンの積率相関係数についてのみ紹介する．

8.1 次数分布

図 2 は修正 BA モデルで生成したネットワークの次数分布，図 2～7 は BA モデルで生成したネットワークの次数分布，図 8～10 は Configuration モデル (自己ループ・多重辺を許さない) で生成したネットワークの次数分布，図 11～13 は Configuration モデル (自己ループ・多重辺を許す) で生成したネットワークの次数分布，図 14～16 は修正 Configuration モデルで生成したネットワークの次数分布である．いずれも次数分布を両対数で描写している．

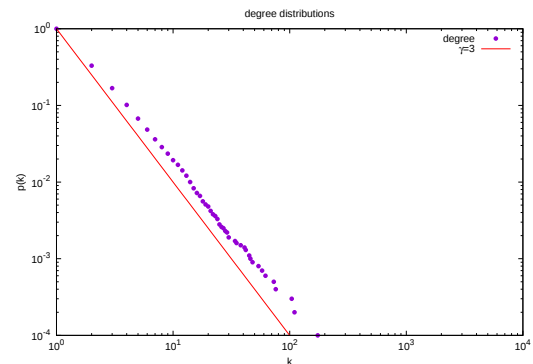


図 2: BA モデル 追加する枝 1 本 (修正 BA)

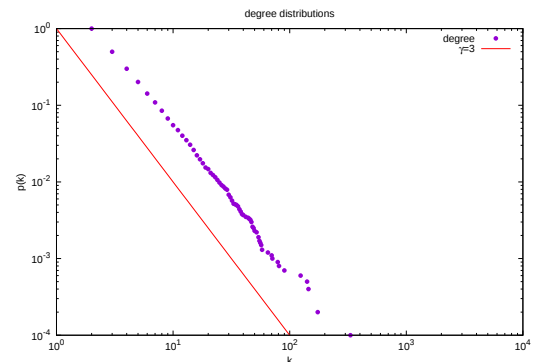


図 3: BA モデル 追加する枝 2 本

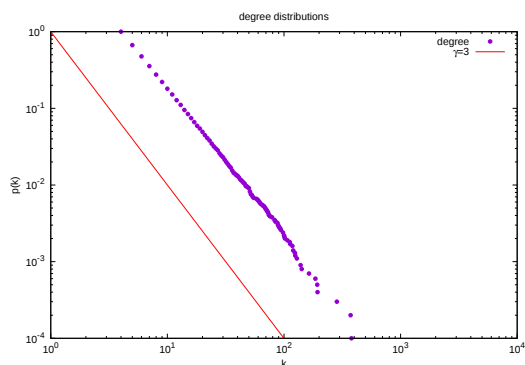


図 4: BA モデル 追加する枝 4 本

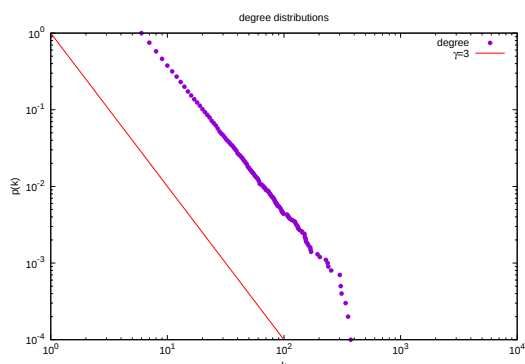


図 5: BA モデル 追加する枝 6 本

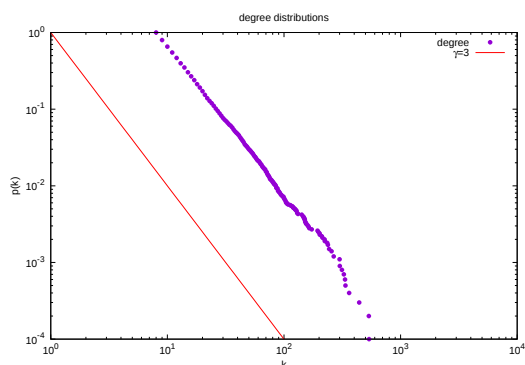


図 6: BA モデル 追加する枝 8 本

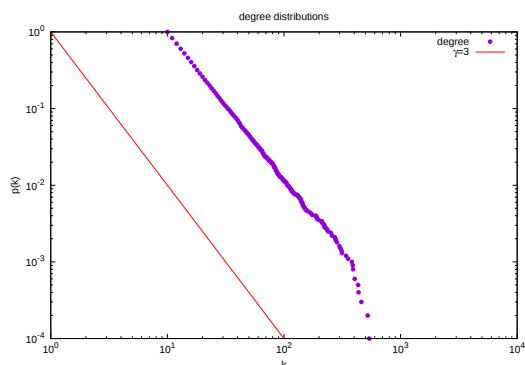


図 7: BA モデル 追加する枝 10 本

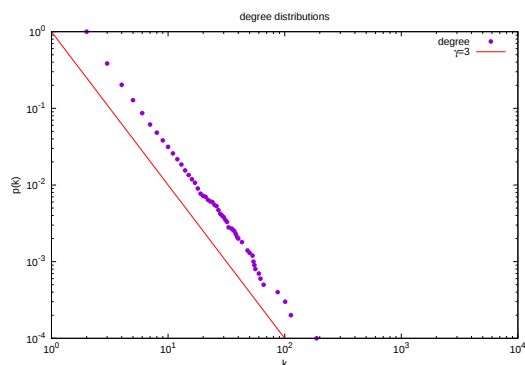


図 8: Configuration モデル (自己ループ・多重辺を許さない) べき指数 3.0

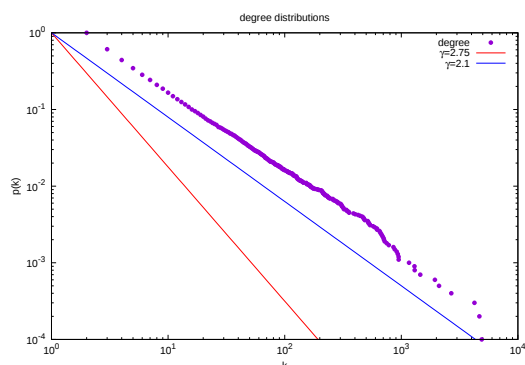


図 9: Configuration モデル (自己ループ・多重辺を許さない) べき指数 2.75

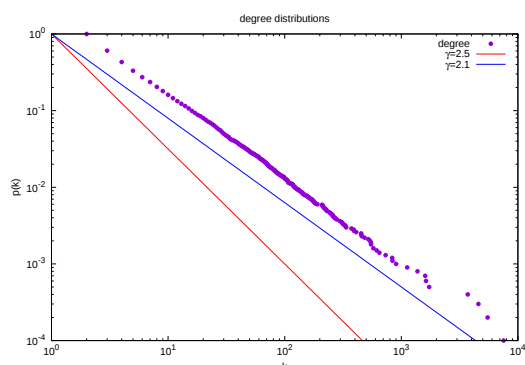


図 10: Configuration モデル (自己ループ・多重辺を許さない) べき指数 2.5

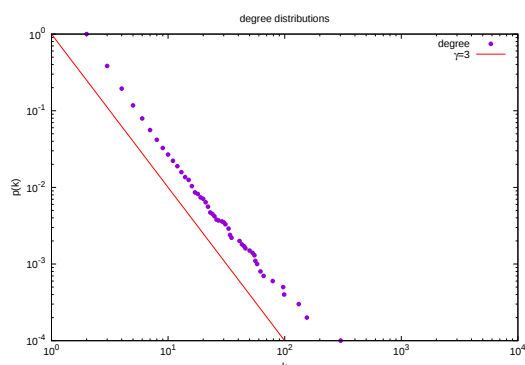


図 11: Configuration モデル (自己ループ・多重辺を許す) べき指数 3.0

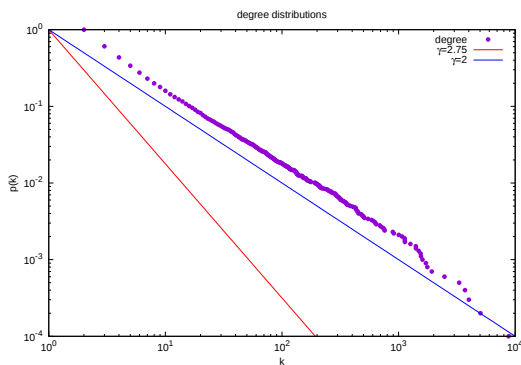


図 12: Configuration モデル (自己ループ・多重辺を許す) べき指数 2.75

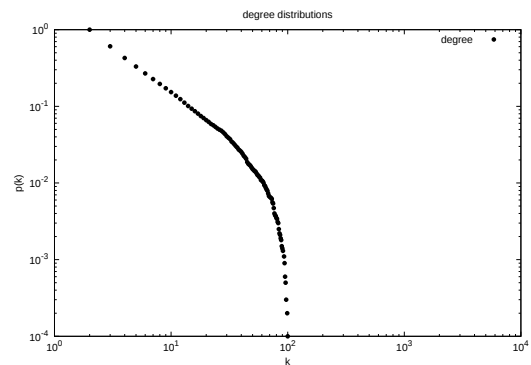


図 16: 修正 Configuration モデル べき指数 2.5

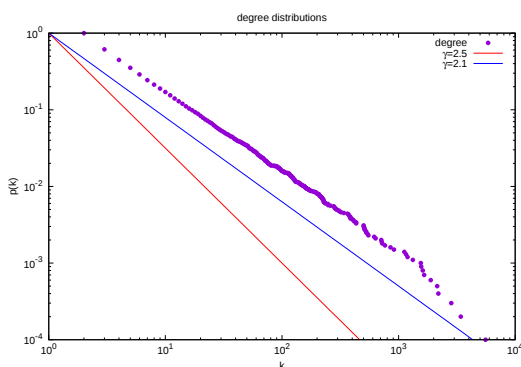


図 13: Configuration モデル (自己ループ・多重辺を許す) べき指数 2.5

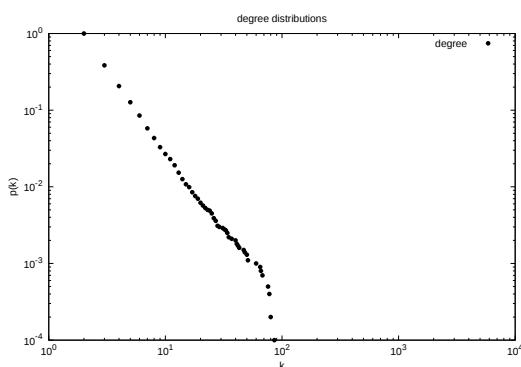


図 14: 修正 Configuration モデル べき指数 3.0

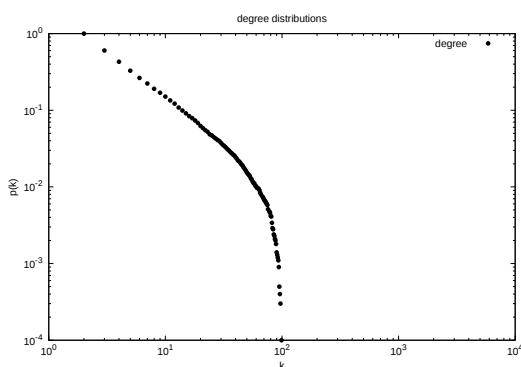


図 15: 修正 Configuration モデル べき指数 2.75

修正 BA モデル，BA モデル，Configuration モデル，修正 Configuration モデルで生成したネットワークはいずれも概ねべき則に従っており，スケールフリー性を十分満たしているとみなせる．

修正 BA モデル，BA モデルで生成したネットワークは，べき指数が 3 程度のべき則に従っていることがわかる．また，べき指数を 2.5, 2.75 として Configuration モデルで生成したネットワークは，いずれもべき指数が 2 程度のべき則になっている．べき指数を 3.0 として修正 Configuration モデルで生成したネットワークは，大きな次数のカットオフが十分行われていないこともわかる．

8.2 RPC 伝搬効率

以下で，情報伝搬実験の結果の一部を紹介する．

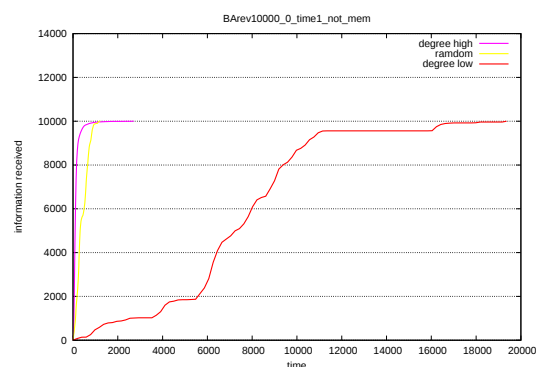


図 17: BA モデル 追加する枝 1 本 (修正 BA)

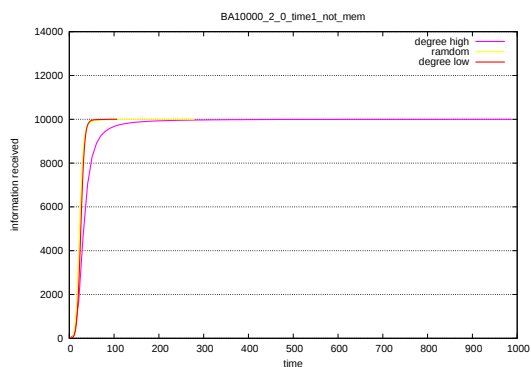


図 18: BA モデル 追加する枝 2 本

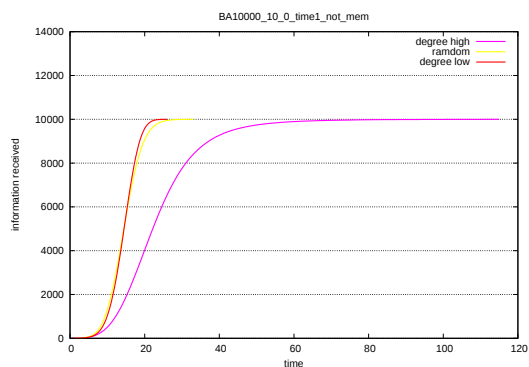


図 22: BA モデル 追加する枝 10 本

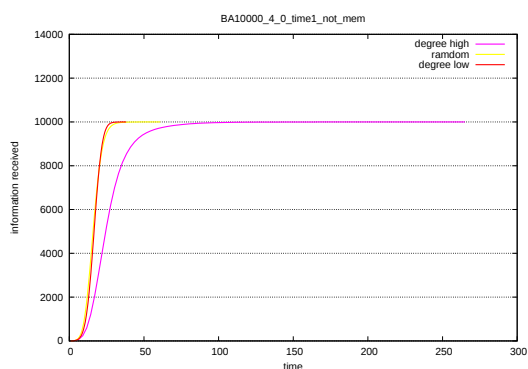


図 19: BA モデル 追加する枝 4 本

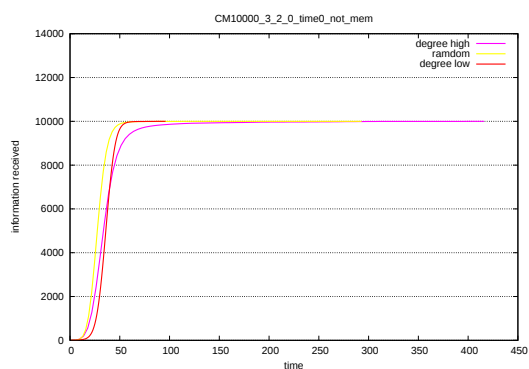


図 23: Configuration モデル (自己ループ・多重辺を許さない) べき指数 3.0

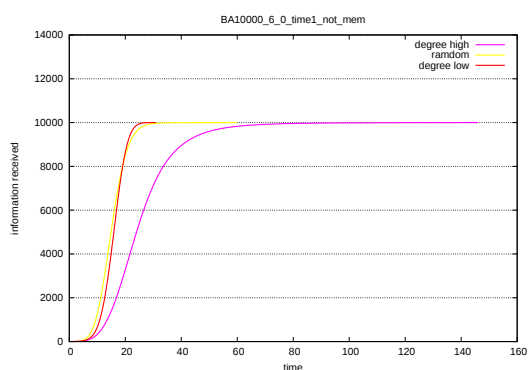


図 20: BA モデル 追加する枝 6 本

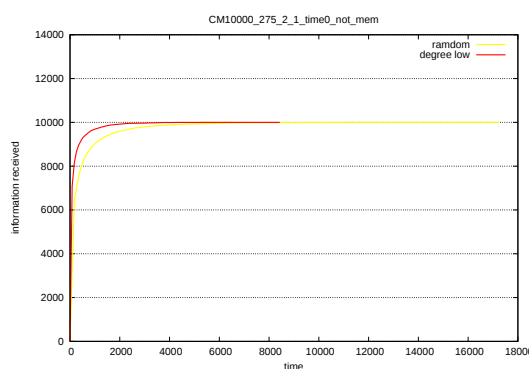


図 24: Configuration モデル (自己ループ・多重辺を許さない) べき指数 2.75

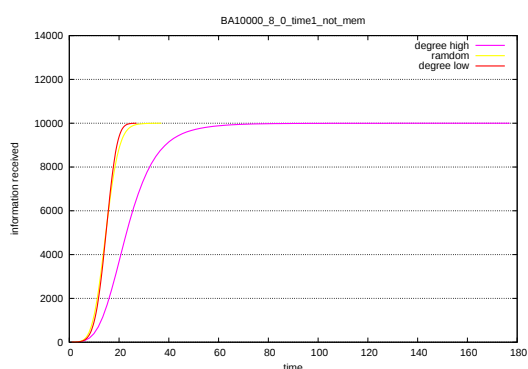


図 21: BA モデル 追加する枝 8 本

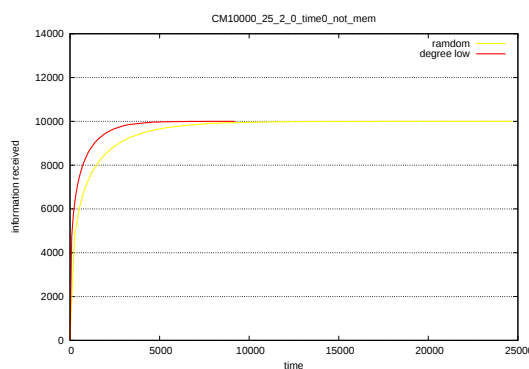


図 25: Configuration モデル (自己ループ・多重辺を許さない) べき指数 2.5

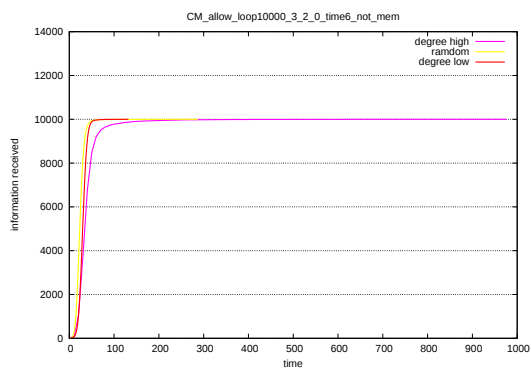


図 26: Configuration モデル (自己ループ・多重辺を許す) べき指数 3.0

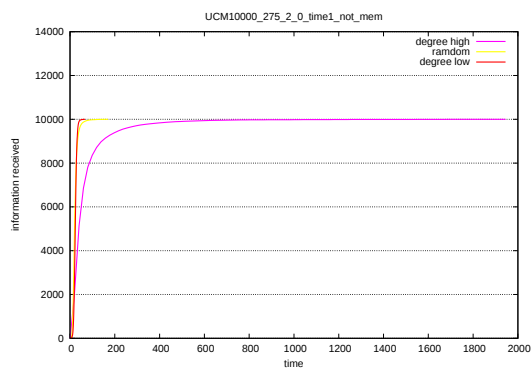


図 30: 修正 Configuration モデル べき指数 2.75

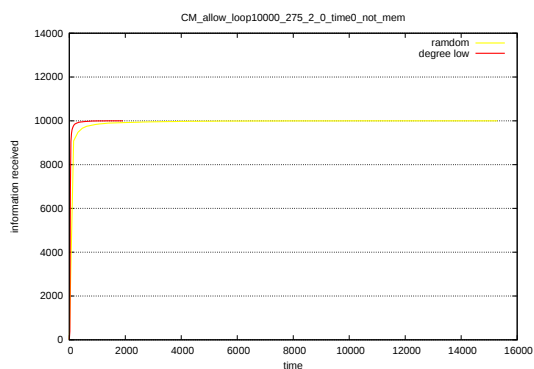


図 27: Configuration モデル (自己ループ・多重辺を許す) べき指数 2.75

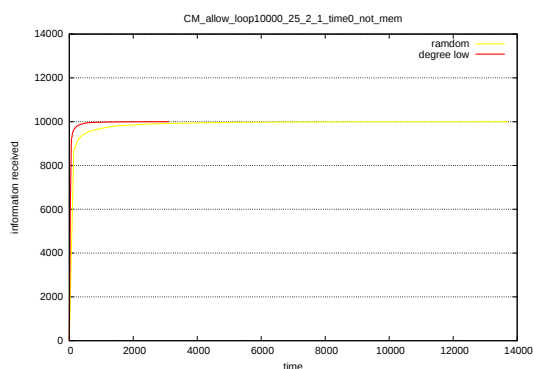


図 28: Configuration モデル (自己ループ・多重辺を許す) べき指数 2.5

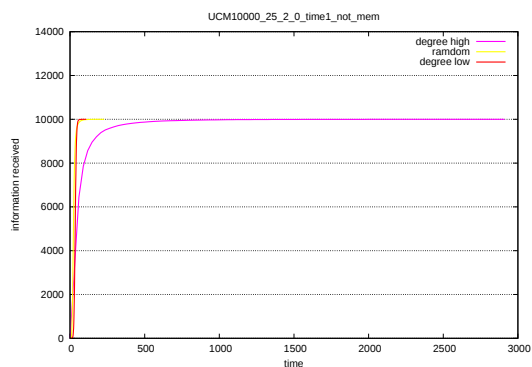


図 31: 修正 Configuration モデル べき指数 2.5

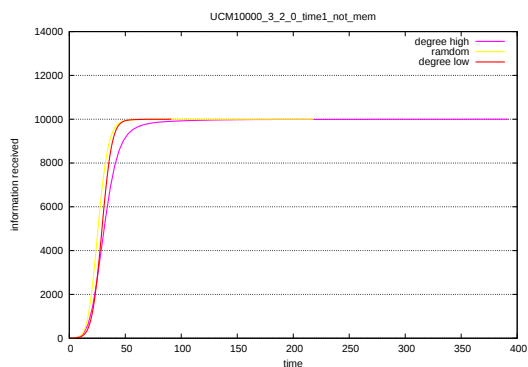


図 29: 修正 Configuration モデル べき指数 3.0

情報伝搬実験の結果を生成方法ごとにまとめると以下の表のようになる。なお、以下の表では Configuration モデル (自己ループ・多重辺を許さない) を CM とし、Configuration モデル (自己ループ・多重辺を許す) を CM al (allow loop) とする。また、BA の右隣の整数は BA モデルで追加する枝の数を表し、CM の右隣の少数は Configuration モデルのべき指数を表す。

表 1: 情報伝搬実験の結果

生成方法	RPC 伝搬効率 (%)	RPC 平均 ステップ数	平均枝数
BA 1	0	37252	9999
BA 2	83.19	173.10	19997
BA 4	87.58	49.13	39990
BA 6	94.70	34.45	59979
BA 8	97.19	30.05	79964
BA 10	98.59	28.06	99945
CM 2.5	99.68	7827.89	65164
CM 2.75	99.40	7136.71	64158
CM 3.0	72.72	326.19	15997
CM al 2.5	100.00	2245.34	68382
CM al 2.75	100.00	2156.15	70610
CM al 3.0	73.39	148.55	18868
修正 CM 2.5	99.96	87.39	33018
修正 CM 2.75	99.97	87.15	33011
修正 CM 3.0	73.07	87.72	15736

表 1 を見てわかるように、BA モデルで追加する枝数が増えていくと、RPC 伝搬効率も良くなっていく。また、Configuration モデル (自己ループ・多重辺を許す) の RPC 伝搬効率ももっとも良く、その次に修正 Configuration モデル、Configuration モデル (自己ループ・多重辺を許さない) の順で RPC 伝搬効率が良い。

8.3 RPC 伝搬効率とピアソンの積率相関係数の関係

修正 BA モデルで生成したネットワークでは RPC 伝搬効率が常に 0% という結果を得たため、RPC 伝搬効率はピアソンの積率相関係数に依らないことがわかった。そのため、修正 BA モデルで生成したネットワークについては省略する。

図 32 ~ 45 は RPC 伝搬効率とピアソンの積率相関係数の関係を表し、縦軸にピアソンの積率相関係数、横軸に RPC 伝搬効率をとり、点は実験で使用した 100 個のネットワークを示している。

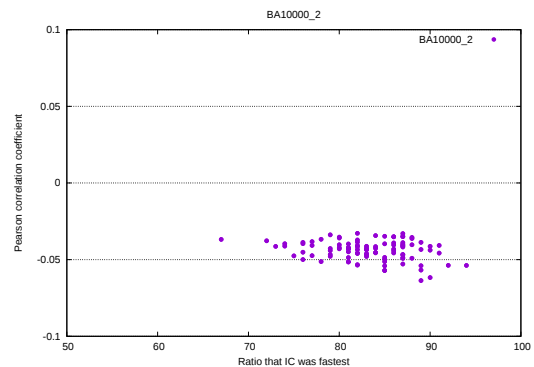


図 32: BA モデル 追加する枝 2 本

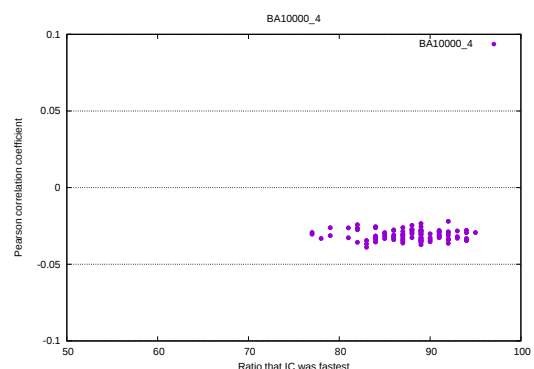


図 33: BA モデル 追加する枝 4 本

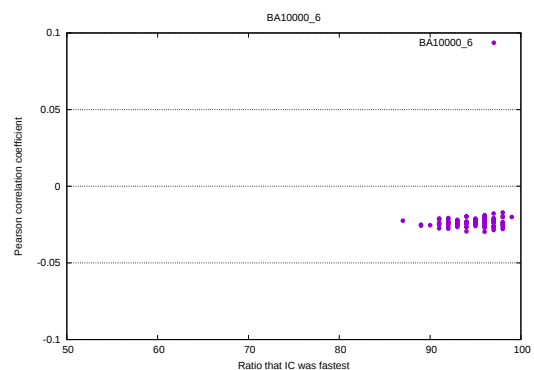


図 34: BA モデル 追加する枝 6 本

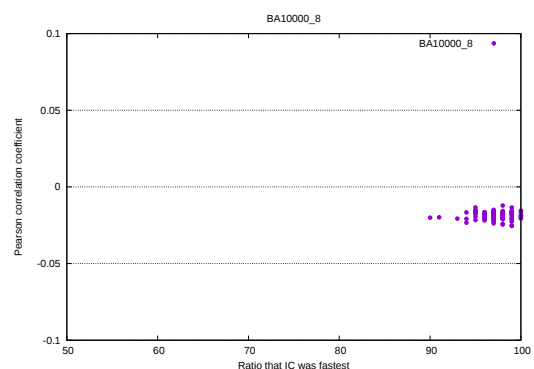


図 35: BA モデル 追加する枝 8 本

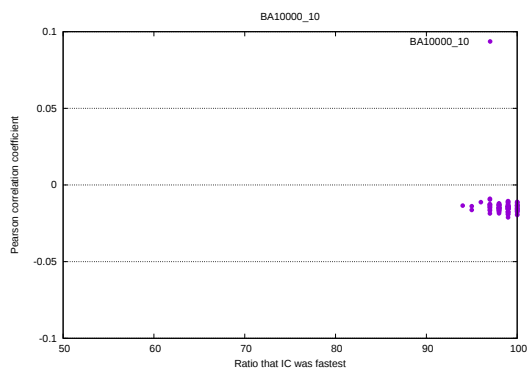


図 36: BA モデル 追加する枝 10 本

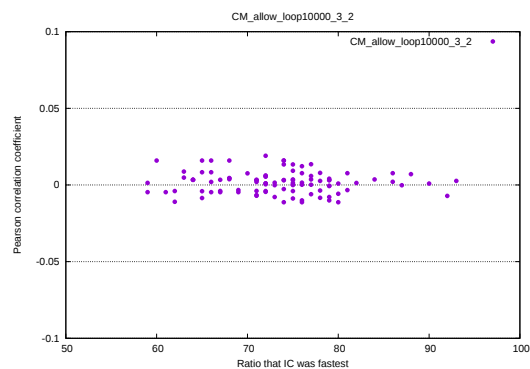


図 40: Configuration モデル (自己ループ・多重辺を許す) べき指数 3.0

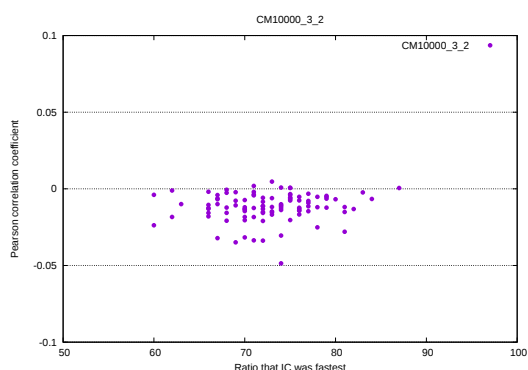


図 37: Configuration モデル (自己ループ・多重辺を許さない) べき指数 3.0

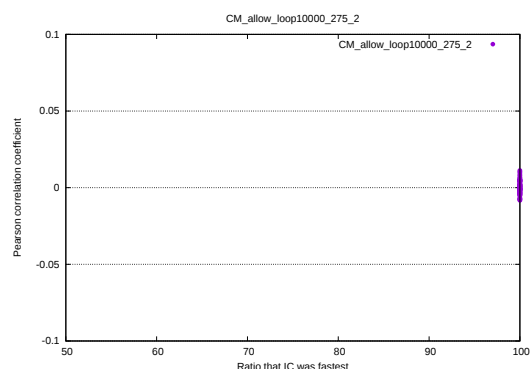


図 41: Configuration モデル (自己ループ・多重辺を許す) べき指数 2.75

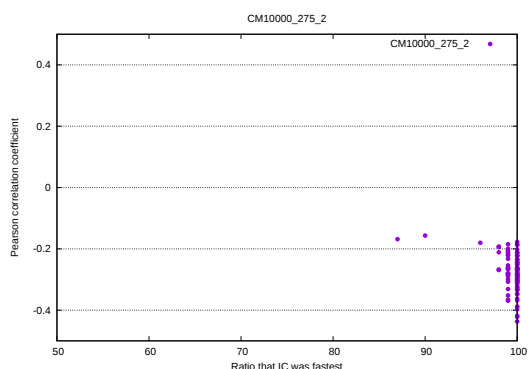


図 38: Configuration モデル (自己ループ・多重辺を許さない) べき指数 2.75

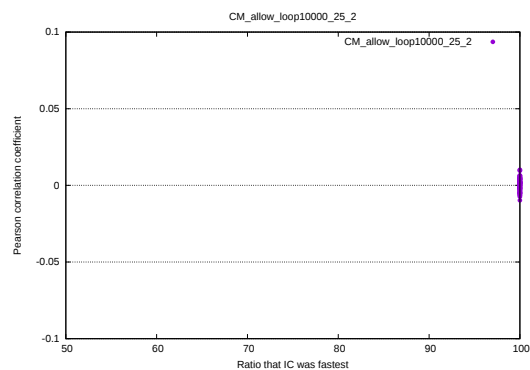


図 42: Configuration モデル (自己ループ・多重辺を許す) べき指数 2.5

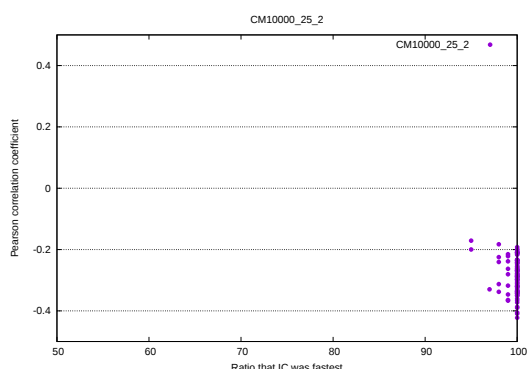


図 39: Configuration モデル (自己ループ・多重辺を許さない) べき指数 2.5

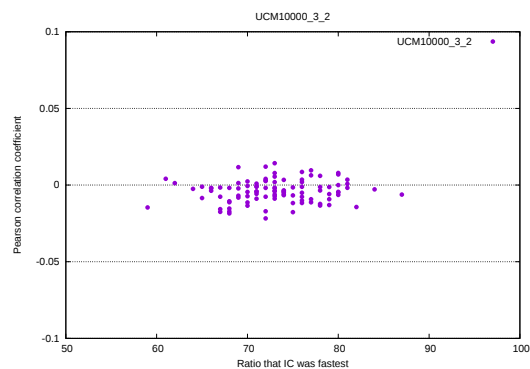


図 43: 修正 Configuration モデル べき指数 3.0

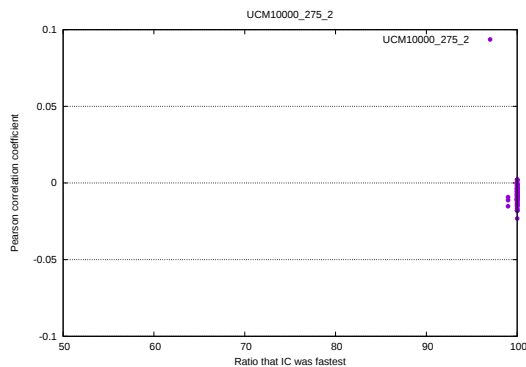


図 44: 修正 Configuration モデル べき指数 2.75

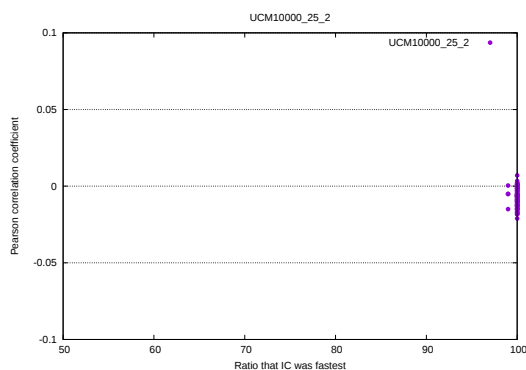


図 45: 修正 Configuration モデル べき指数 2.5

これから以下のことがわかる。

- BA モデルでは、追加する枝数が増えるにつれ RPC 伝搬効率も良く、ピアソンの積率相関係数も 0 に近づいていく
- Configuration モデル (自己ループ・多重辺を許さない) では、べき指数が 2.5, 2.75 の場合は RPC 伝搬効率は良いが、ピアソンの積率相関係数は 0 から離れる (相関が現れている。)
- Configuration モデル (自己ループ・多重辺を許す), 修正 Configuration モデルでは、べき指数が 2.5, 2.75 の場合は RPC 伝搬効率が良く、ピアソンの積率相関係数も 0 に近い。

参考文献

- [1] 増田直紀, 今野紀雄 複雑ネットワーク 基礎から応用まで (近代科学社) 2010/4/26.
- [2] Guido Caldarelli, Michele Catanzaro 著, 高口太郎 訳, 増田直樹 監修 ネットワーク科学 つながりが解き明かす世界のかたち (丸善出版) 2014/4/25
- [3] Hiroshi Toyozumi, Seiichi Tani, Optimal Spread over Finite-Size Statistical Network. Statistics of Networks Workshop 24-25 (2010)

- Configuration モデルでは、べき指数が 3.0 だと、どのモデルもピアソンの積率相関係数は 0 に近いが、RPC 伝搬効率は良いわけではない。

8.4 考察

定理 1 の対偶をとると「各頂点が記憶能力を持たず、ソースノードが単位時間に 1 つの隣接頂点にのみ情報を伝搬するモデルにおいて、ネットワーク全体に情報が伝搬されるまでの平均時間が情報を伝搬させる隣接頂点を選択する際に次数の逆数に比例した確率分布に従い乱択する方法が最小とならないならば、ネットワークが Uncorrelated 性を持たない」となる。本演習の実験では、Configuration モデルにおいて、べき指数が 3.0 で生成したネットワークは、次数の逆数に比例した確率分布に従い乱択する方法が最速とならない場合が多数見られたため、Uncorrelated 性が損なわれている。そのため、べき指数が大きいことが Uncorrelated 性を損なわせている可能性がある。

また、修正 Configuration モデルで生成したネットワークの次数分布を見ると、べき指数が 2.5, 2.75 で生成したネットワークは大きな次数のカットオフが十分行われているが、べき指数が 3.0 を生成したネットワークは大きな次数のカットオフがあまり行われていない。このことが、Uncorrelated 性に影響している可能性がある。

9 今後の課題

今回、多くの特徴量を求めたが、Uncorrelated 性と関連のある特徴量を求めることはできなかった。しかし、本演習では調査を行わなかったクラスター係数など、ネットワークの特徴量はまだ存在する。また、今回は BA モデル、Configuration モデルによりネットワークの生成を行ったが、BA モデル、Configuration モデル以外にもスケールフリーネットワークを生成できるモデルは多数存在する ([1])。従って、様々なスケールフリーネットワーク生成モデルでネットワークの生成を行い、Uncorrelated 性とネットワークの特徴量の関連をより詳細に調べることが今後の課題である。

- [4] Hiroshi Toyoizumi, Seiichi Tani, Naoto Miyoshi, Yoshio Okamoto, Reversr preferential spread in complex networks. Phys. Rev. E 86, 021103 (2012)
- [5] Barabási, A.-L., and Albert, R., "Emergence of scaling in random networks", Science 286, pp509-512 (1999).
- [6] Ulrik Brandes. A Faster Algorithm for Betweenness Centrality, Journal of Mathematical Sociology, 25:163-177 (2001).
- [7] U.Brandes and C.Pich. CentralityEstimationinLargeNetworks. Int. J. Bifurcat. Chaos, 17(07):2303-2318, (2007)
- [8] Y.Yoshida.AlmostLinear-TimeAlgorithmsforAdaptiveBetweenness Centrality using Hypergraph Sketches. In KDD, pp. 1416-1425, (2014)
- [9] Michele Catanzaro, Marián Boguñá, and Romualdo Pastor-Satorras. Generation of uncorrelated random scale-free networks. Phys. Rev. E 71, 027103 (2005)
- [10] Hakimi, S. L., "On realizability of a set of integers as degrees of the vertices of a linear graph. I", Journal of the Society for Industrial and Applied Mathematics 10:496-506, MR 0148049 (1962)
- [11] 小幡 正博 作, 伝播速度限定モデルにおける Scale Free Network 上の情報拡散 Dynamics (http://www.tani.cs.chs.nihon-u.ac.jp/g-2012/kujira1220/obata_resume.pdf)
- [12] 蔡 延安 作, 伝播速度限定モデルにおける Scale Free ネットワーク上の情報拡散について (<http://www.tani.cs.chs.nihon-u.ac.jp/g-2012/choy/choy.pdf>)
- [13] 田中勇歩 作, 修正版 BA モデルで生成したネットワークのスケールフリー性判定計算実験 (http://www.tani.cs.chs.nihon-u.ac.jp/g-2013/ufo/resume/tanaka_resume.pdf)
- [14] 阿部光太郎 大谷舞 作, BA モデルおよび修正 BA モデルで生成したネットワークの Uncorrelated 性について (http://www.tani.cs.chs.nihon-u.ac.jp/g-2014/nw/network_resume.pdf)
- [15] 香取 秀柄 逸見 優聡 作, BA モデルで生成されたネットワークと同じ次数列を持つネットワークの Uncorrelated 性について (http://www.tani.cs.chs.nihon-u.ac.jp/g-2015/n2/n2_resume.pdf)
- [16] 枝國 雄太 作, BA モデルで生成されたネットワークにおける次数と情報伝搬速度の関係について (<http://www.tani.cs.chs.nihon-u.ac.jp/g-2015/n1/resume.pdf>)

