

Holme-Kim モデルで 生成したネットワークの Uncorrelated 性について

2019/1/31

日本大学文理学部情報科学科

谷聖一研究室

太田一輝

目次

1. はじめに
2. 実験方法
3. 実験結果
4. 考察
5. 今後の課題

1.はじめに

目次

1.1 複雑ネットワークとは

1.2 グラフの基礎

1.3 ネットワークの特徴量

1.4 ネットワークの性質

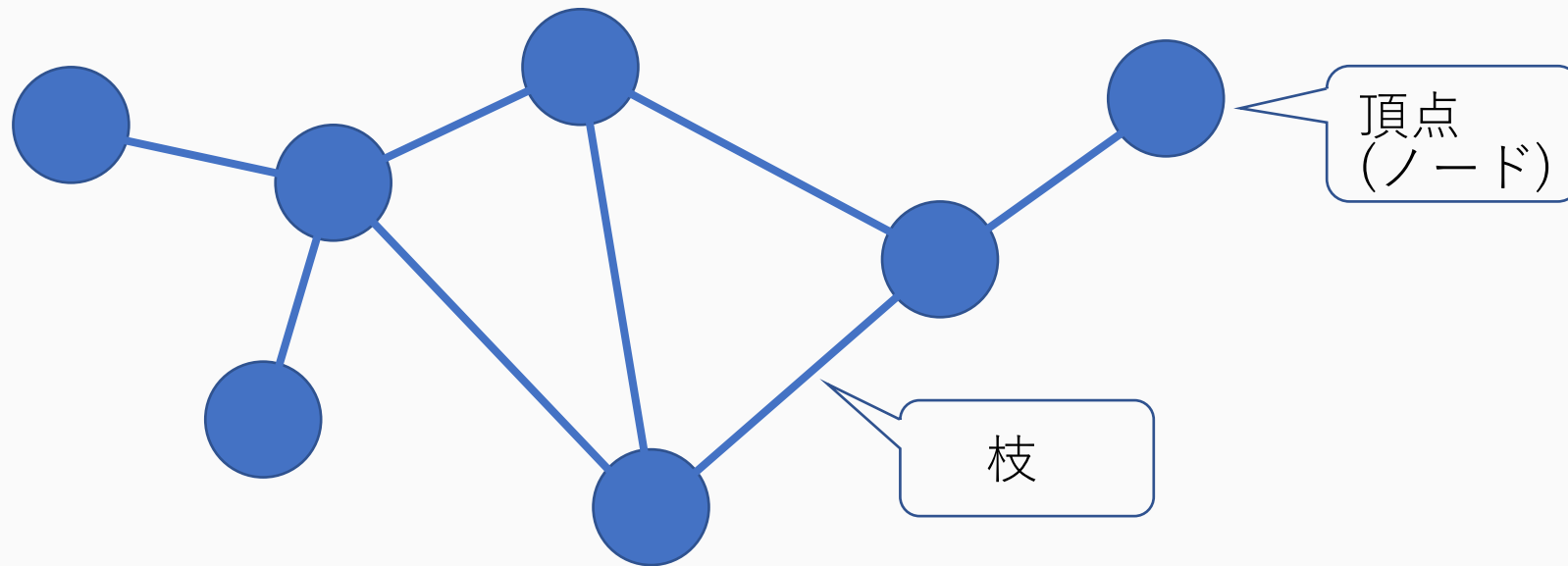
1.5 背景

1.6 目的

1.1 複雑ネットワークとは

ネットワークとは

→頂点と枝から構成され、繋がりや関係を表したものの



1.1 複雑ネットワークとは

- 日常ではインターネットなど，つながり全般

頂点	枝	ネットワーク
人	人間関係	社会ネットワーク
駅	線路	路線網
交差点	道路	道路網
コンピュータ	ケーブル	コンピュータネットワーク

現実世界のネットワーク
→ 巨大で非常に複雑



複雑ネットワーク

1.2 グラフの基礎

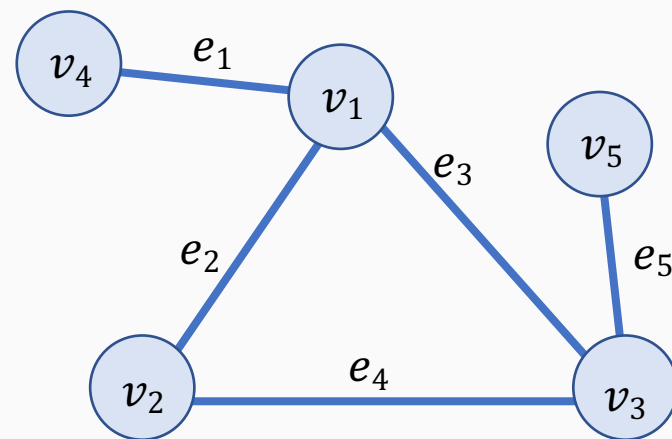
- ネットワークの図：数学ではグラフ
- グラフの定義

$$G = (V, E)$$

- ▶ 頂点の集合 $V = \{v_1, v_2, \dots, v_N\}$
- ▶ 枝の集合 $E = \{e_1, e_2, \dots, e_M\}$

N : 頂点数

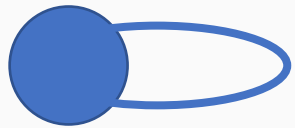
M : 枝数



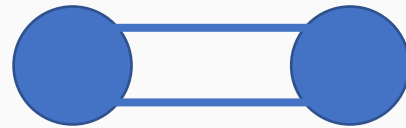
$N = 5, M = 5$

1.2 グラフの基礎

ループ



多重辺



無向グラフ



有向グラフ



本演習では、ループ・多重辺を持たない無向グラフのみを扱う

1.3 ネットワークの特徴量

- ネットワークを特徴づける概念
- ネットワークを解析するための指標となる

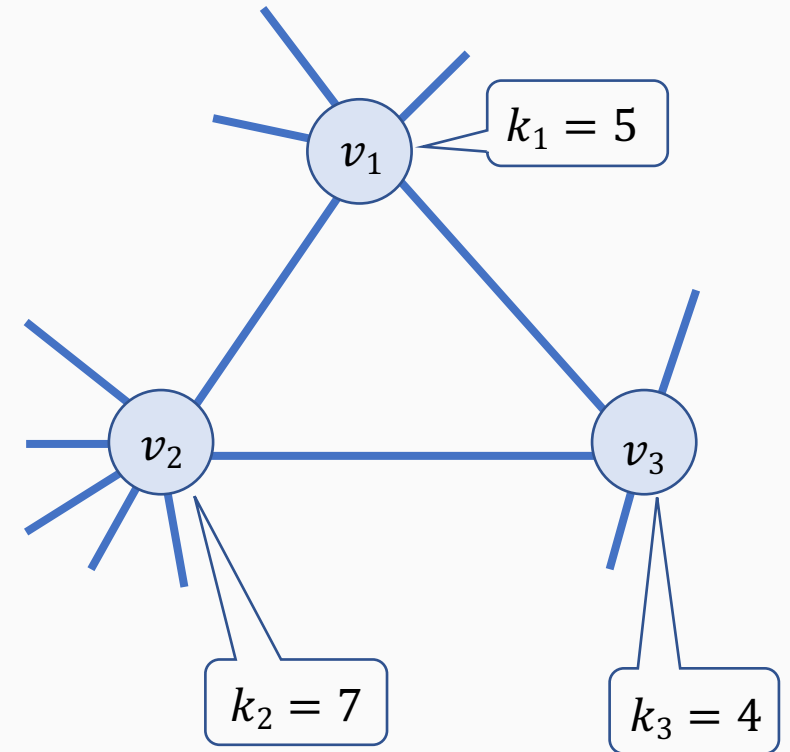
1.3 ネットワークの特徴量

- 頂点 v_i の次数

次数 k_i = 頂点 v_i と接続する枝の数

- 次数分布

次数分布 $p(k)$ = 次数が k の頂点の割合



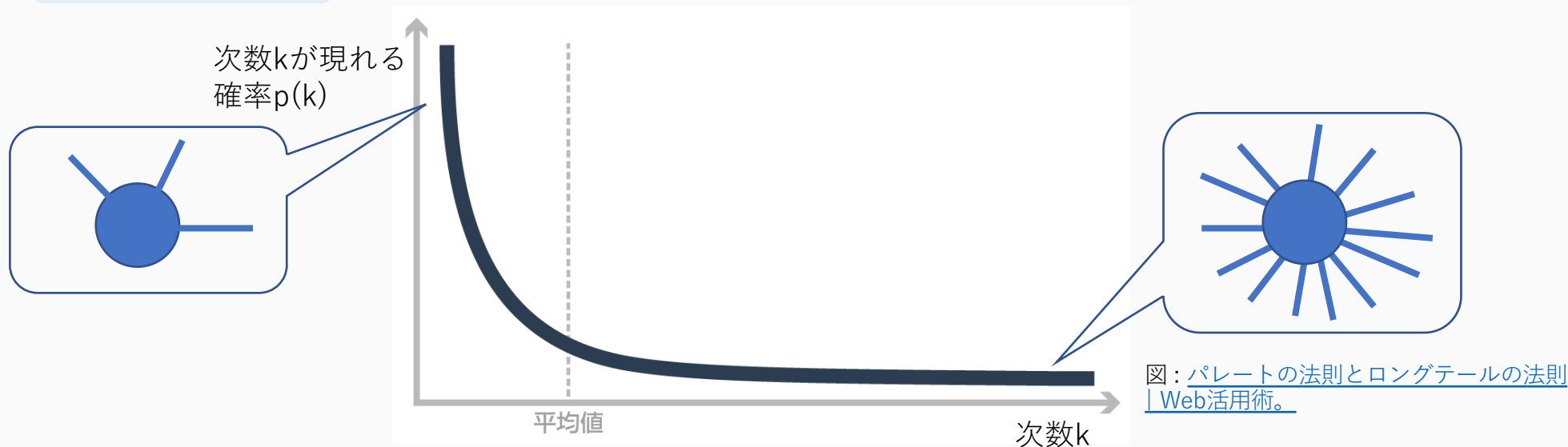
1.3 ネットワークの特徴量

- べき則

次数分布 $p(k)$ が次の式に従うこと

$$p(k) \propto k^{-\gamma}$$

γ : べき指数



1.3 ネットワークの特徴量

- クラスター係数

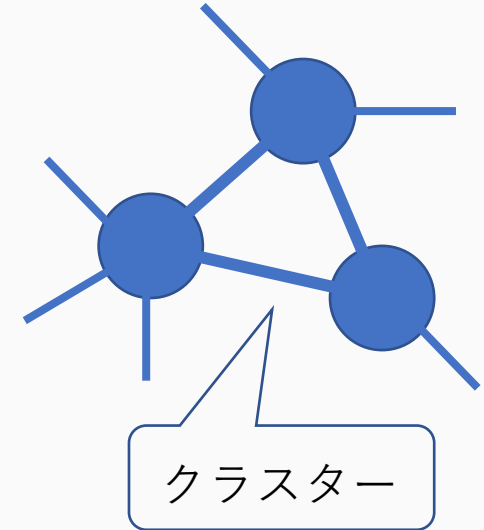
クラスター：ネットワークに含まれる三角形

- ▶ ある頂点 v_i のクラスター係数

$$C_i = \frac{v_i \text{を含む三角形の数}}{k_i(k_i-1)/2}$$

- ▶ ネットワーク全体のクラスター係数

$$C = \frac{1}{N} \sum_{i=1}^N C_i \quad 0 \leq C \leq 1$$



1.3 ネットワークの特徴量

- 平均距離

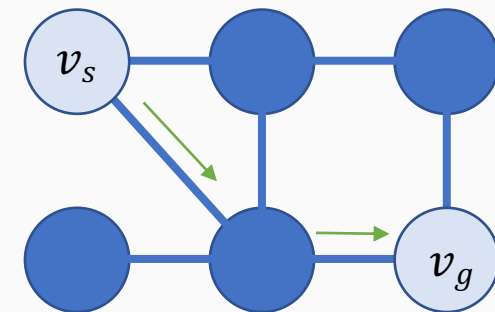
- ▶ v_i と v_j の距離 $d(v_i, v_j)$:

- v_i から v_j に行くために通らなければならない最小の枝数

- ▶ 平均距離 L は $d(v_i, v_j)$ の全ての頂点对に渡る平均

- ▶ 頂点对の選び方 : $N(N - 1)/2$ 通り

$$\text{平均距離 } L = \frac{2}{N(N-1)} \sum_{1 \leq i < j \leq N} d(v_i, v_j)$$



$$d(v_s, v_g) = 2$$

1.4 ネットワークの性質

現実のネットワークは複雑ながらも共通する性質がある

ここでは次の2つを説明

- ▶ スケールフリー性
- ▶ スモールワールド性

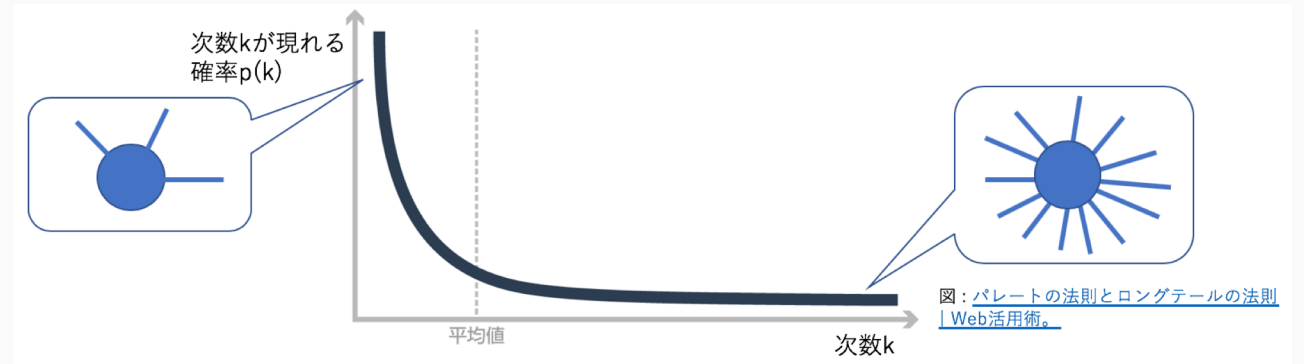
1.4 ネットワークの性質

- スケールフリー性

次数分布がべき則に従う

性質・特徴

- 一部の頂点：多くの枝を持つ
次数が大きい頂点 → ハブ
- 大多数の頂点：わずかな枝を持つ
- ハブが多数存在するため、任意の2点間の距離が小さい



1.4 ネットワークの性質

- スモールワールド性

平均距離 L が小さく，クラスター係数 C が大きいネットワーク

性質・特徴

- ▶ 任意の2点間において，わずかな頂点を中継するだけで結べる
- ▶ 「世間は狭い」と言われる現象

1.5 演習の背景

- 情報をすばやく伝搬したい
→どのように情報を伝搬すれば
全体にすばやく情報共有できるか？

1.5 演習の背景

Uncorrelated 性について

ある伝搬モデルでネットワークに情報を流すとき、ネットワークが Uncorrelated 性を持つ場合、次数の小さい頂点に優先的に流す方法が効率的である

[TTMO]Hiroshi Toyoizumi,Seiichi Tani,Naoto Miyoshi,Yoshio Okamoto,Reverse preferential spread in complex networks. Phys. Rev.E 86, 021103 (2012)より

1.5 演習の背景

Uncorrelated 性とは

- ネットワークにおける様々な研究でよく用いられる性質の一つ
- uncorrelate → 無相関の
- 局所性がなく，ネットワーク全体で同じような構造

1.6 演習の目的

Uncorrelated 性に関係のある特徴量を見つけない

ネットワークがUncorrelated 性を持つ場合，次数の小さい頂点に優先的に流す方法が効率的である

対偶



次数の小さい頂点に流す方法が効率的でないなら，ネットワークはUncorrelated 性を持たない

1.6 演習の背景

過去の演習

- ▶ 2016年^[1] スケールフリーネットワークにおいて Uncorrelated 性と関連のある特徴量を求めることはできなかった
- ▶ 2018年^[2] スケールフリーではなく、スモールワールドネットワークの WS モデルは Uncorrelated 性を持たない。Uncorrelated 性にクラスター性が関係する可能性

[1]飯田 佳徳, 丸山隆太, 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

[2]新井 天馬, WS モデルおよび修正 Configuration モデルで生成したネットワークの Uncorrelated 性について

1.6 演習の目的

本演習

- ▶ 2016年の演習で調査が行われなかったクラスター性についてスケールフリーネットワークで調べる
- ▶ スケールフリー性を持ち、スモールワールドなネットワーク
→ Holme-Kim モデル

2. 実験方法

目次

2.1 ネットワーク生成モデル

2.1.1 Barabási – Albert モデル

2.1.2 Holme – Kim モデル

2.2 情報伝搬モデル

2.3 ネットワーク生成

2.4 情報伝搬実験

2.1 ネットワーク生成モデル

ネットワーク生成モデル

- ▶ 現実のネットワークを模して人工的にネットワークを作成
- ▶ 数理的に解析が可能な

本演習で扱うモデル

- ▶ Barabási – Albert モデル(BAモデル)
- ▶ Holme – Kim モデル(HKモデル)

2.1.1 Barabási – Albert モデル

Barabási – Albert モデル(BAモデル)

1999年にバラバシとアルバートによって提案された

不規則で乱雑な構造のスケールフリー性を持つネットワーク

- ▶ 成長

頂点がネットワークに新しく加わる

- ▶ 優先的選択

新しく加わる頂点は次数の大きい頂点ほど、高い確率で結びつく

2.1.1 Barabási – Albert モデル

BAモデルの作り方

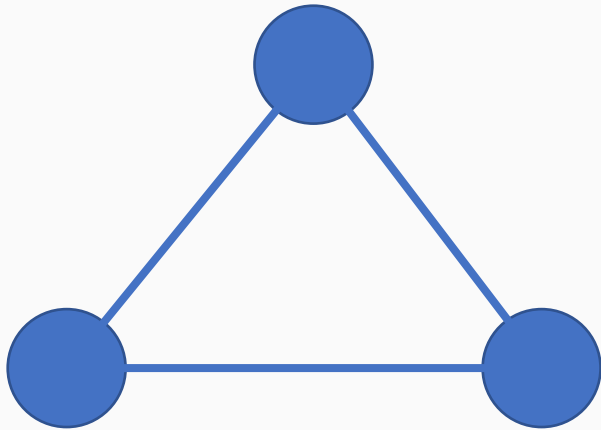
- ① m_0 個の頂点を置き，連結なネットワークとする
- ② $m(\leq m_0)$ 本の枝を持つ頂点を1つずつネットワークに追加
 - ▶ 頂点が N' 個あり，既存の頂点 $v_i(1 \leq i \leq N')$ の次数を k_i とする
 - ▶ 新しい枝のそれぞれが v_i に結びつく確率は

$$\Pi(k_i) = \frac{k_i}{\sum_{j=1}^{N'} k_j}$$

- ③ 頂点数 N になるまで ② を繰り返す

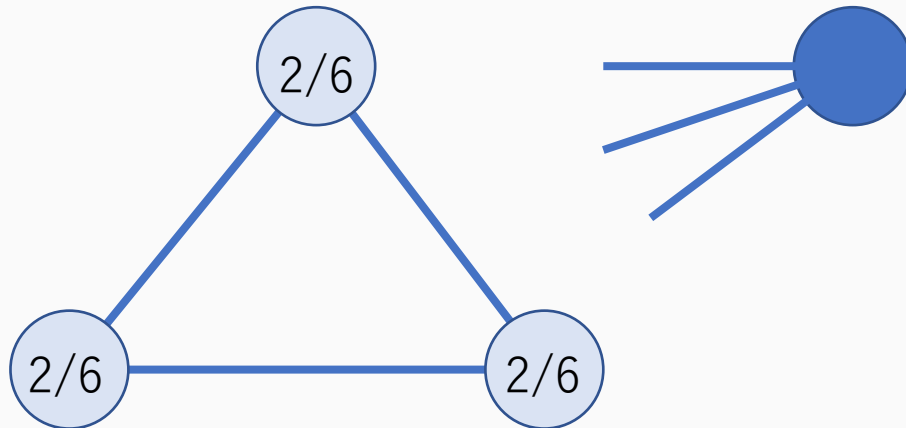
2.1.1 Barabási – Albert モデル

- $t = 0$ (t : 単位時間)
- $m_0 = 3$ の時
- 完全グラフを作成



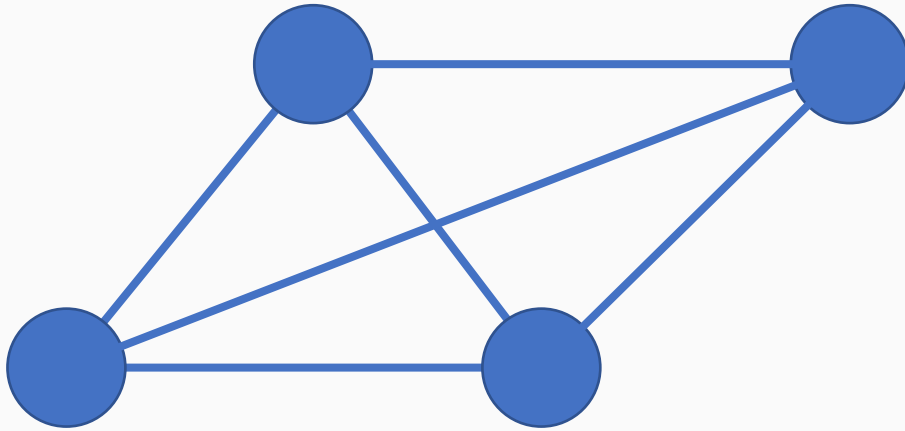
2.1.1 Barabási – Albert モデル

- $t = 1$
- 新しい頂点を追加
- $m=3$ とする



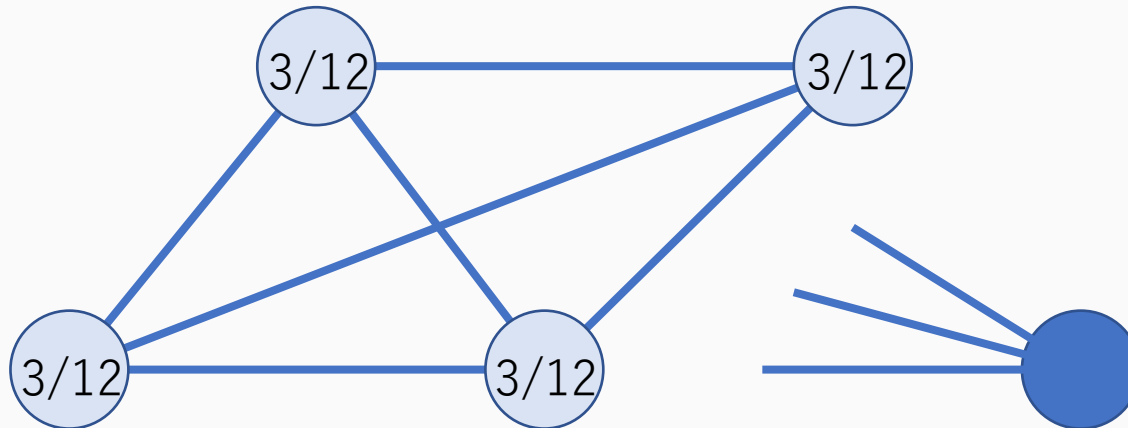
2.1.1 Barabási – Albert モデル

- $t = 1$
- 新しい頂点を追加
- $m=3$ とする



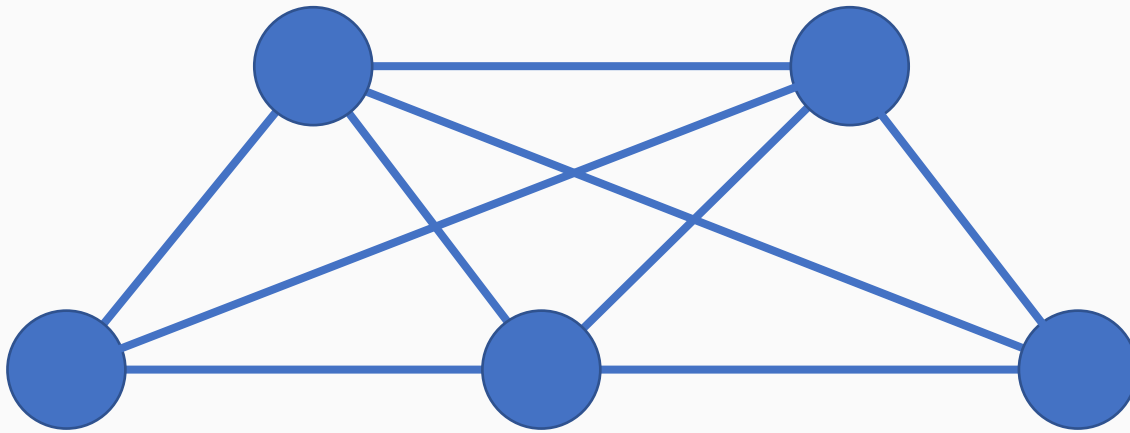
2.1.1 Barabási – Albert モデル

- $t = 2$
- 新しい頂点を追加
- $m=3$ とする



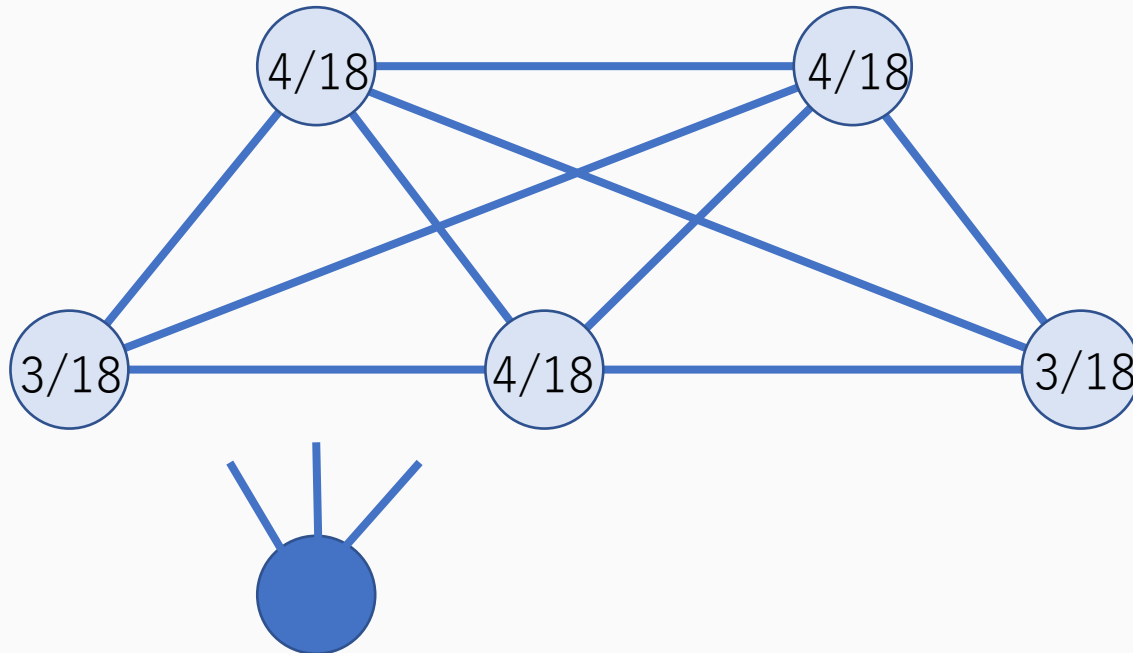
2.1.1 Barabási – Albert モデル

- $t = 2$
- 新しい頂点を追加
- $m=3$ とする



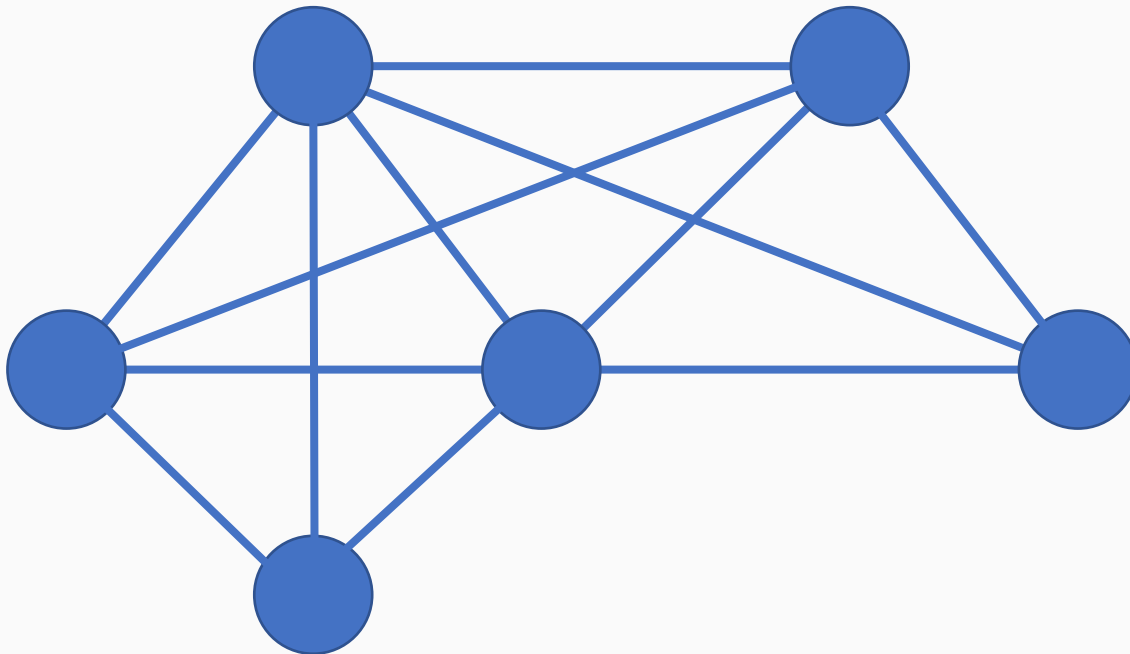
2.1.1 Barabási – Albert モデル

- $t = 3$
- 新しい頂点を追加
- $m=3$ とする



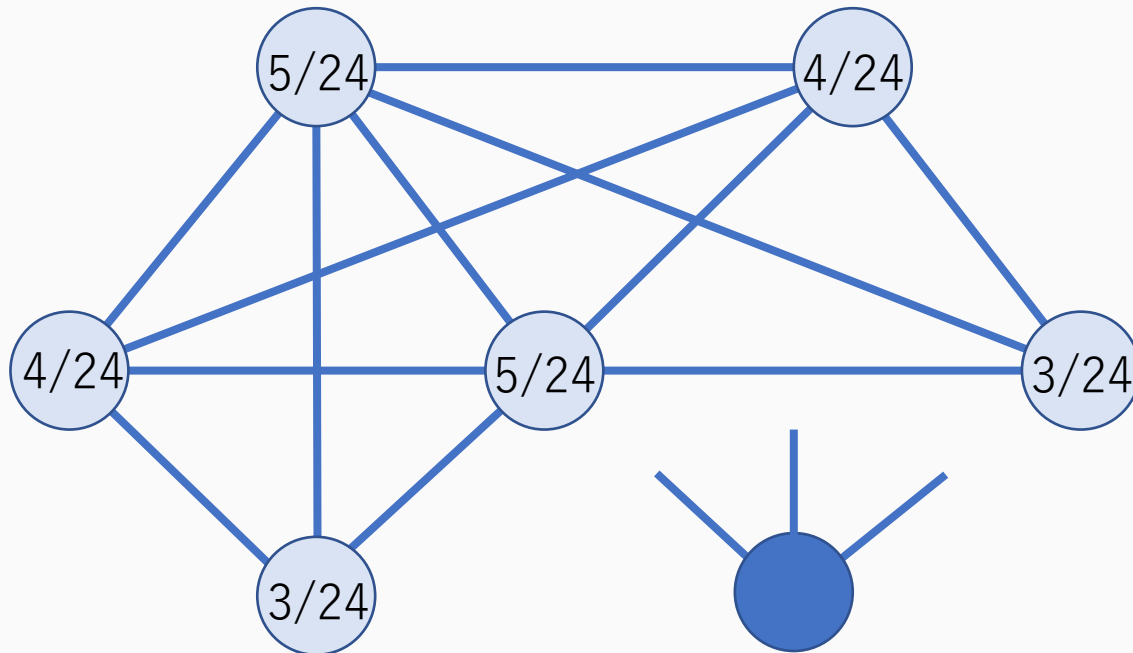
2.1.1 Barabási – Albert モデル

- $t = 3$
- 新しい頂点を追加
- $m=3$ とする



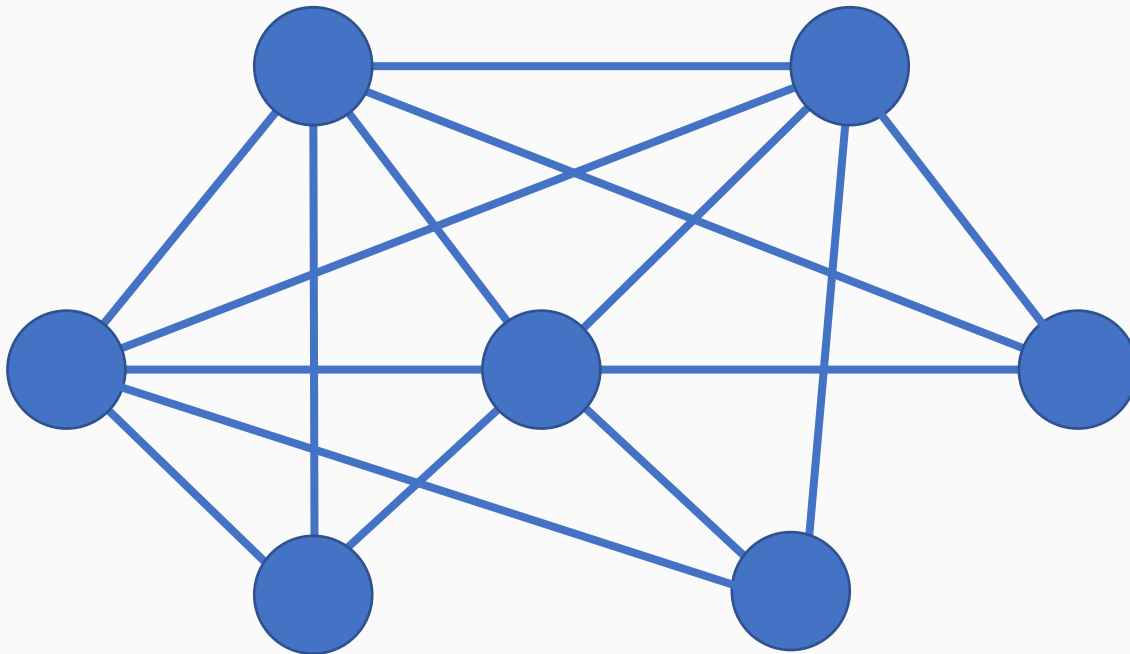
2.1.1 Barabási – Albert モデル

- $t = 4$
- 新しい頂点を追加
- $m=3$ とする



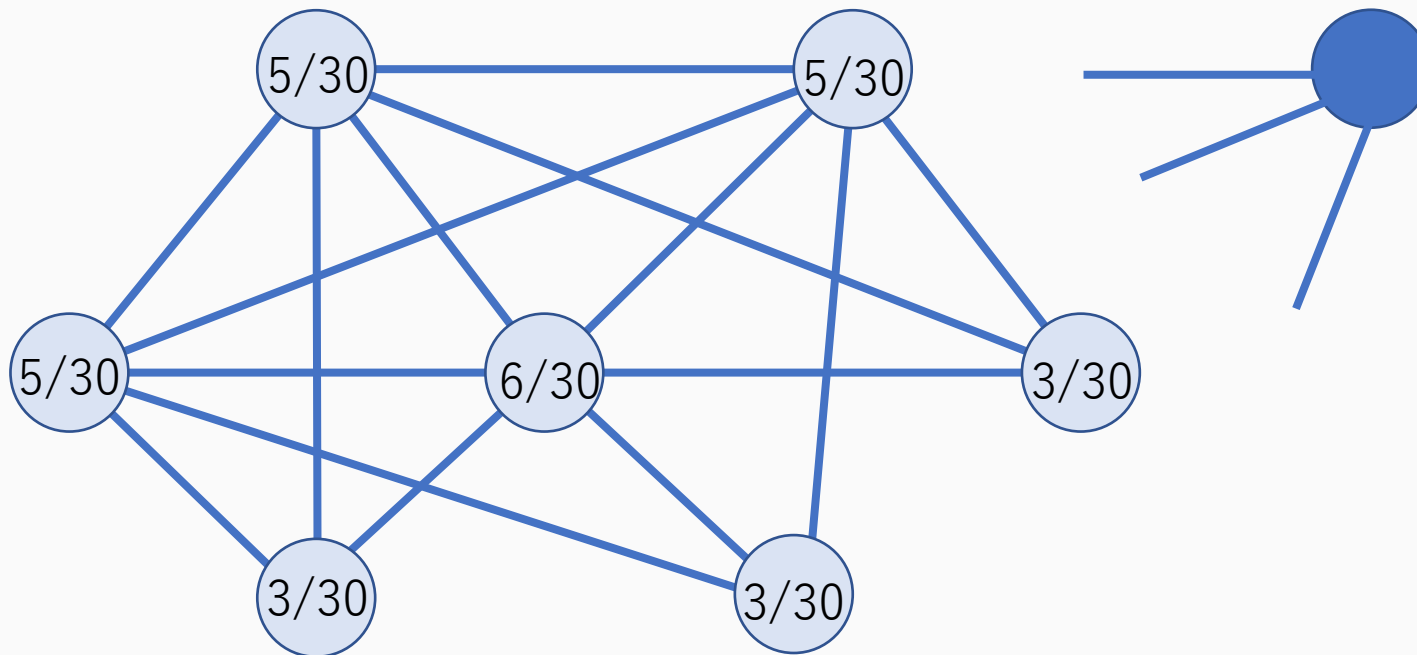
2.1.1 Barabási – Albert モデル

- $t = 4$
- 新しい頂点を追加
- $m=3$ とする



2.1.1 Barabási – Albert モデル

- $t = 5$
- 新しい頂点を追加
- $m=3$ とする



2.1.2 Holme – Kim モデル

Holme – Kim モデル(HKモデル)

ホームとキムによるBAモデルの拡張

- BAモデルのクラスター係数は小さい
→クラスター係数を高めた
- スモールワールドかつスケールフリーなネットワーク

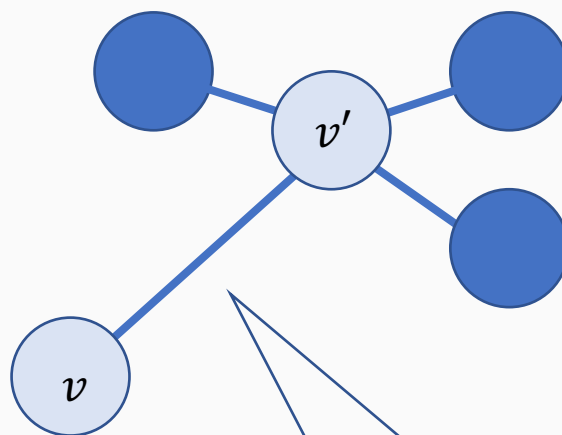
2.1.2 Holme – Kim モデル

BAモデルにおいて各時点で1つの頂点を追加するとき

- ① 新しい頂点 v を加える
- ② 1本の新しい枝の行き先 v' を優先的選択で決める
- ③ 残りの $m - 1$ 本の新しい枝の行き先を
 - ▶ 確率 p で、 v' の v 以外の隣接点から1つを等確率で選ぶ
 - ▶ 確率 $1 - p$ で、通常の優先的選択に従って1つの頂点を選ぶ

2.1.2 Holme – Kim モデル

新しい頂点 v を加える

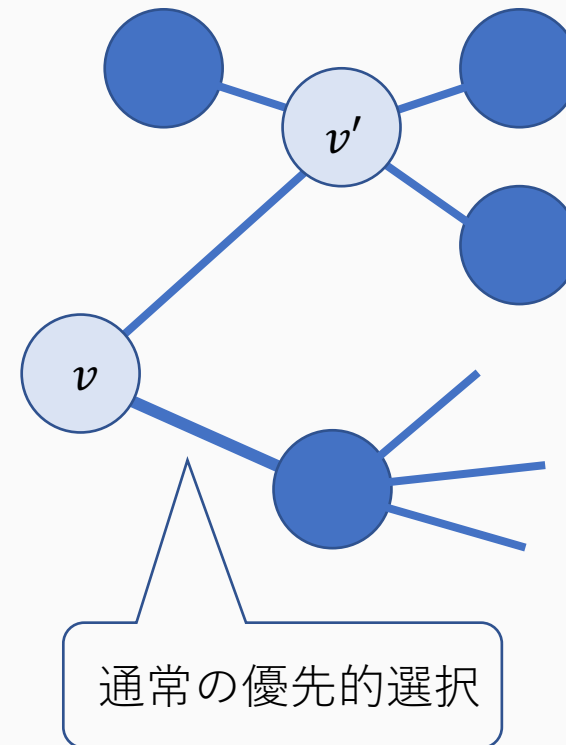
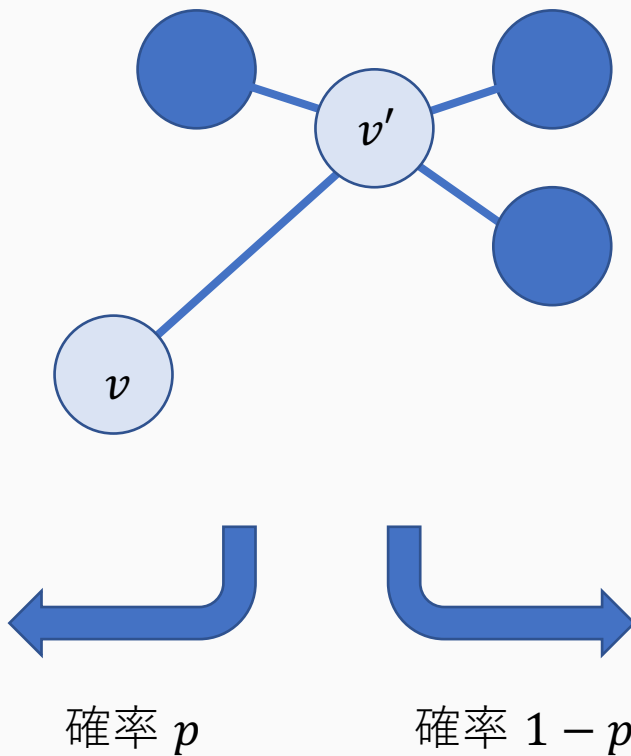
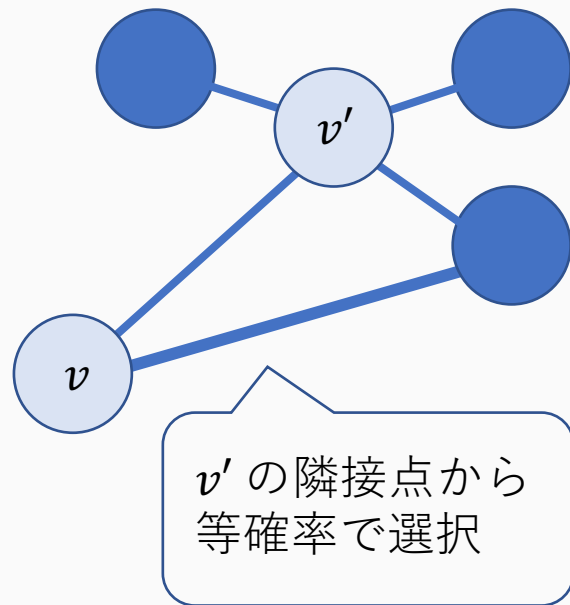


優先的選択：
次数の高い頂点を優先的に
選択する

v の枝1本目を
優先的選択で v' へ

2.1.2 Holme – Kim モデル

新しい頂点 v を加える



2.2 情報伝搬モデル

情報伝搬のルール

- ① 初期状態でのソースノードは1つ
 - ② ソースノードは隣接頂点のうちの1つをターゲットノードとして選択
 - ③ ターゲットノードはソースノードになる
 - ④ 全ての頂点がソースノードとなるまで、②~③を繰り返す
- ▶ ソースノード：情報を保持している頂点
 - ▶ ターゲットノード：ソースノードから伝搬先に選ばれた頂点

2.2 情報伝搬モデル

ターゲットノードの選択先を決めるのに使える情報

- ▶ 各頂点が事前に保持している情報は隣接頂点の次数のみ
- ▶ 各頂点はネットワーク全体の構造はわからない

2.2 情報伝搬モデル

ソースノードがターゲットノードを選択する方法

- ▶ 一様分布に従い乱択
(NC; No Control)
- ▶ 次数の大きな頂点を優先的に乱択
(PC; Preferential Control)
- ▶ 次数の小さな頂点を優先的に乱択
(RPC; Reverse Preferential Control)

2.3 ネットワーク生成

モデル	m	m ₀	p	個数
BA	3	3		80
	5	5		80
	7	7		80
HK	3	3	0.3	80
			0.5	80
			0.7	80
	5	5	0.3	80
			0.5	80
			0.7	80
	7	7	0.3	80
			0.5	80
			0.7	80

m_0 : 初期の頂点数
 m : 追加する枝数
 p : 隣接点から選択する確率

2.4 情報伝搬実験

- 1つのネットワークに対し，100回実験
- ランダムのシードは実験番号と同じとする
- 初期ソースノードはランダムで選ばれる
- 1回の実験について NC, PC, RPC で初期ソースノードは統一

3. 実験結果

目次

3.1 RPC 最速割合, 平均ステップ数

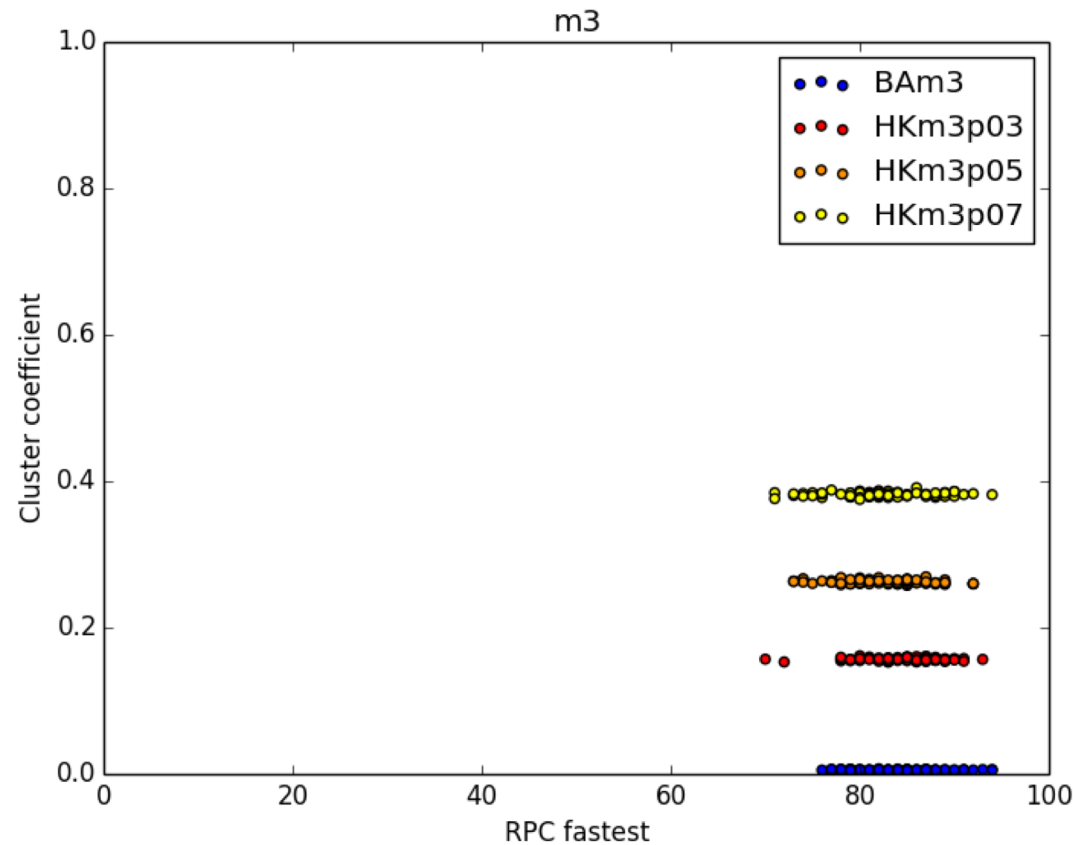
3.2 RPC 最速割合とクラスター係数

3.3 実験結果まとめ

3.1 RPC 最速割合, 平均ステップ数

モデル	m	m ₀	p	NC最速回数	PC最速回数	RPC最速回数	RPC最速割合	RPC平均ステップ数
BA	3	3		1190	0	6754	0.844	74.78
HK			0.3	1175	0	6763	0.845	74.29
			0.5	1314	0	6618	0.827	72.38
			0.7	1365	0	6552	0.819	72.11
BA	5	5		563	0	7322	0.915	39.05
HK			0.3	541	0	7357	0.920	38.63
			0.5	488	0	7415	0.927	38.14
			0.7	445	0	7444	0.930	37.79
BA	7	7		201	0	7733	0.967	31.45
HK			0.3	185	0	7748	0.969	31.32
			0.5	205	0	7729	0.966	31.28
			0.7	168	0	7777	0.972	31.19

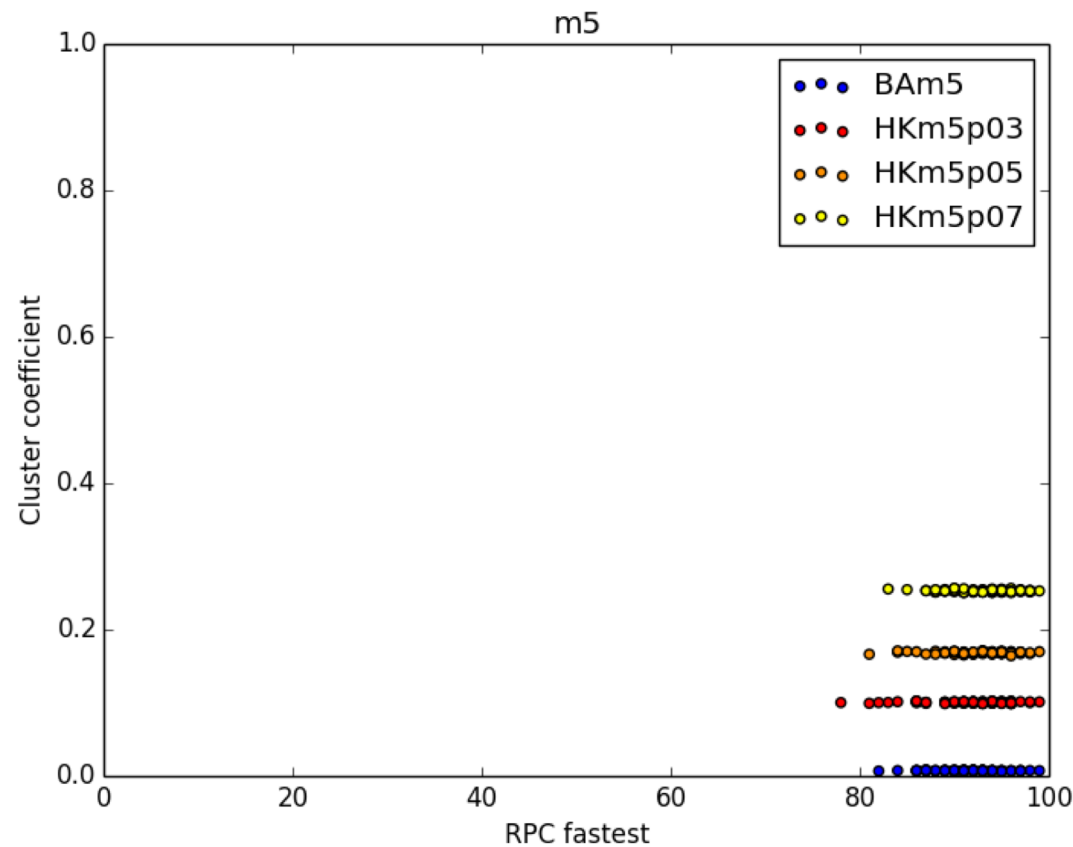
3.2 RPC 最速割合とクラスター係数



RPC 最速割合とクラスター係数の
相関係数：-0.223

クラスター係数と
RPC最速割合の間に
線形の相関関係は見られない

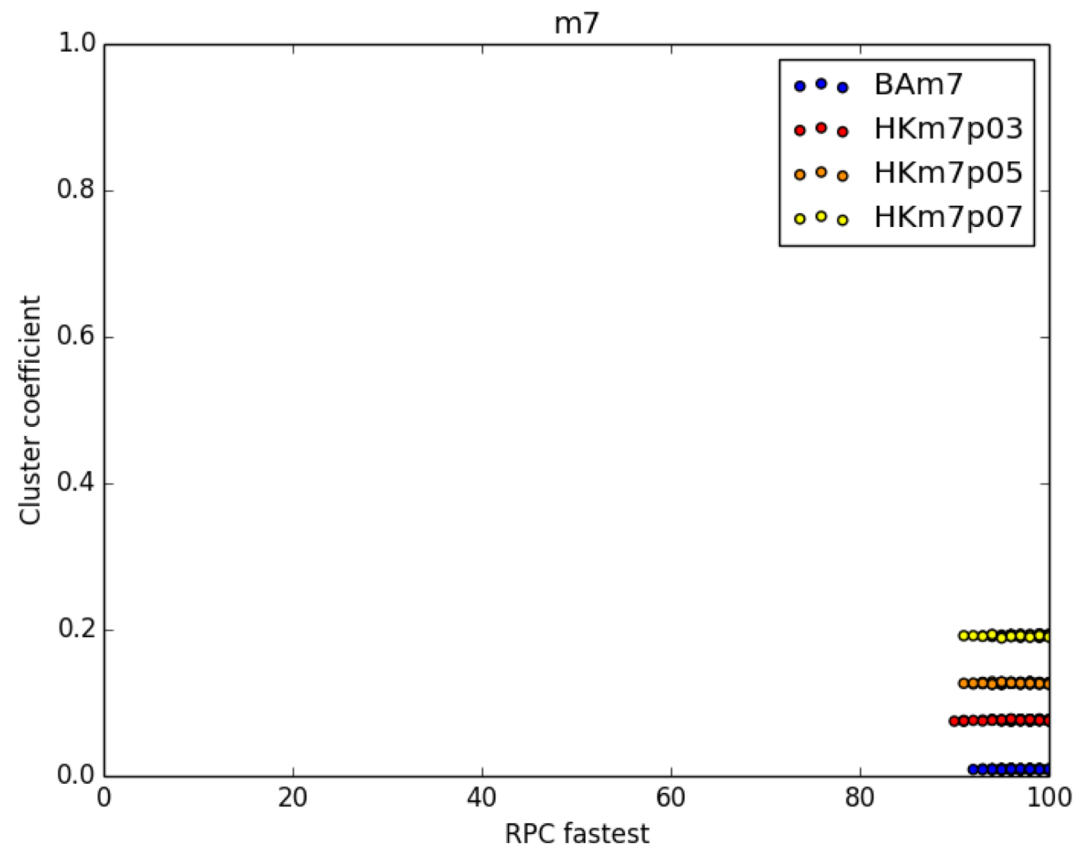
3.2 RPC 最速割合とクラスター係数



RPC 最速割合とクラスター係数の
相関係数：0.164

クラスター係数と
RPC最速割合の間に
線形の相関関係は見られない

3.2 RPC 最速割合とクラスター係数



RPC 最速割合とクラスター係数の
相関係数：0.080

クラスター係数と
RPC最速割合の間に
線形の相関関係は見られない

1.6 演習の目的(再掲)

Uncorrelated 性に関係のある特徴量を見つけない

ネットワークがUncorrelated 性を持つ場合、次数の小さい頂点に優先的に流す方法が効率的である

対偶



次数の小さい頂点に流す方法が効率的でないなら、ネットワークはUncorrelated 性を持たない

4. 考察

RPC最速割合

全てのネットワークモデルで最大

→過去の演習と同様にスケールフリーネットワークでは
RPC最速割合が大きくなると考えられる

→Holme-Kim モデルで生成したネットワークは
Uncorrelated 性を持つ可能性がある

4. 考察

クラスター係数

RPC最速割合とクラスター係数に相関関係は見られなかった
→スケールフリーネットワークにおいてはクラスター係数が
Uncorrelated 性に関係ある特徴量とは言えない

5. 今後の課題

- クラスタ係数に注目した場合
 - スケールフリー性を持たずクラスタ係数が低いネットワークへの伝搬実験（過去の演習を含め実験していない）
- 他の特徴量
 - ▶ 枝数が Uncorrelated 性に関係あるのではないか
 - ▶ 枝数が多いネットワークだと、次数が小さい頂点を優先して情報伝搬する方法が平均伝搬時間が短くなる理由が Uncorrelated 性に関係あるか検証

目次

- | | | |
|----------------|-----------------|----------|
| 1. はじめに | 2. 実験方法 | 3. 実験結果 |
| 1.1 複雑ネットワークとは | 2.1 ネットワーク生成モデル | 4. 考察 |
| 1.2 グラフの基礎 | 2.2 情報伝搬モデル | 5. 今後の課題 |
| 1.3 ネットワークの特徴量 | 2.3 ネットワーク生成 | |
| 1.4 ネットワークの性質 | 2.4 情報伝搬実験 | |
| 1.5 背景 | | |
| 1.6 目的 | | |

情報伝搬先として選択される確率

頂点 b が頂点 a のターゲットノードとして選択される確率 q

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

w_i : 頂点 i の重み

k_i : 頂点 i の次数

a_i : 頂点 a の i 番目の隣接頂点

- ▶ nc : $w_i = \text{一定の値}$
- ▶ pc : $w_i = \text{頂点 } v_i \text{ の次数}$
- ▶ rpc : $w_i = \text{頂点 } v_i \text{ の次数の逆数}$

Uncorrelated ネットワーク

Uncorrelated ネットワークである

次の近似が成立

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

$q(b; a)$:
 b が a のターゲット
ノードとして選択される
確率
(条件付き確率 $q(b|a)$)

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

q_b :
全頂点の中から b が
ターゲットノードとして
選択される確率
(条件付き確率 $q(b|a)$)

$$q_b = \frac{k_b w_b}{\sum_{j=1}^N k_j w_j}$$

平均距離

平均距離が小さいとは

平均距離 L が次に従うこと

$$L \propto \log N$$

クラスター係数

クラスターが大きいとは

同じ次数列でランダムに作られるネットワークより
有意に大きくなること

情報伝搬モデル

伝搬速度を限定

想定される状況

○ウイルス, 口コミ, 噂

×twitter, グループライン