

複雑ネットワークにおけるコンタクトプロセスとその可視化

本田泰, 中島陽介, 諸岡貴之

室蘭工業大学 情報工学科

概要

本研究では, 様々な複雑ネットワーク上でコンタクトプロセスのシミュレーションを行いネットワークの種類による違いを比較, 可視化した. その結果, スケールフリーネットワークの方がランダムネットワークより低い伝染率で病気が蔓延する事がわかった.

Contact process in complex networks and the visualization

Yasushi Honda, Yosuke Nakajima, Takayuki Morooka

Department of Computer Science and Systems Engineering, Muroran Institute of Technology

Abstract

We simulate a contact process on complex networks and compare differences between these networks, and visualize them. As a result, a disease spreads at lower infection rate for a scale free network than for a random network.

1 はじめに

本研究では, コンピュータウイルスや伝染病の蔓延についてコンピュータシミュレーションを用いて調べる. ランダムネットワーク及びスケールフリーネットワーク (BA モデル) 上でコンタクトプロセス [1] のシミュレーションを行い, ネットワークによってどのような違いがあるか比較した. また, その結果を視覚で捉えるために, 仮想エネルギーモデル [2] を用いて各ネットワークを 3D 空間上で可視化した.

2 複雑ネットワークにおけるコンタクトプロセスのシミュレーション

病気の伝染する確率 (伝染率) を λ , 病気から回復する確率を μ とし, $x_i(t)$ を時刻 t のノードの健康状態とし 0 ならば健康, 1 ならば病気とする. さらに, $m_i(t)$ を時刻 t において x_i が接しているノードのうち病気のノードの数とする.

コンタクトプロセスにおける病気の伝染及び回復

の確率は以下のようにモデル化される.

$$P(x_i(t+1) = 0 | x_i(t) = 0) = (1 - \lambda)^{m_i(t)}$$

$$P(x_i(t+1) = 1 | x_i(t) = 0) = 1 - (1 - \lambda)^{m_i(t)}$$

$$P(x_i(t+1) = 0 | x_i(t) = 1) = \mu$$

$$P(x_i(t+1) = 1 | x_i(t) = 1) = 1 - \mu$$

本研究では $\mu = 0.1$ に固定し, 病気のノードが 1 つだけの状態を初期状態とした. またその場所はランダムに選んだ.

また, ノード数 N , 平均次数 $\langle k \rangle$ については以下の値を持つネットワークを調べた.

- $N = 10000$, $\langle k \rangle = 4, 6, 8, 10$
- $N = 50000$, $\langle k \rangle = 4$
- $N = 100000$, $\langle k \rangle = 4$

実際にシミュレーションで使ったネットワークの次数分布は図 1, 図 2 に示したとおりポアソン分布及びべき分布となる.

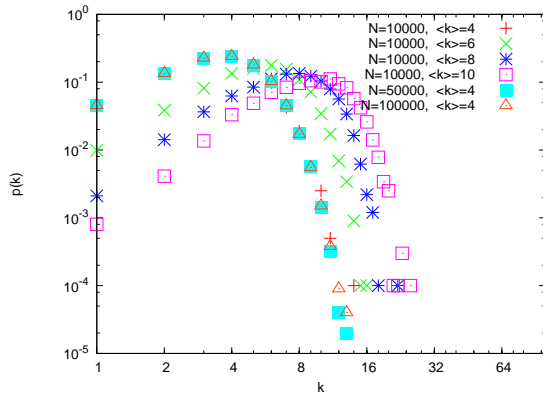


図 1: シミュレーションで使したランダムネットワークの度数分布

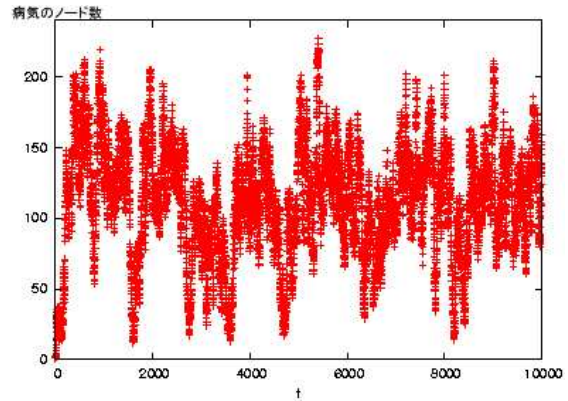


図 3: $N = 10000$, $\langle k \rangle = 4$ の BA モデルで $\lambda = 0.01$ のときの病気のノード数の推移例

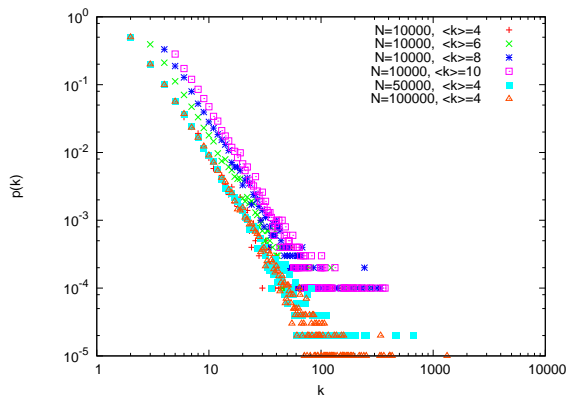


図 2: シミュレーションで使した BA モデルの度数分布

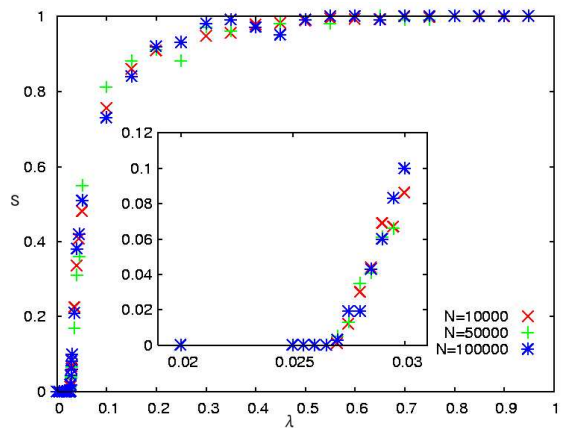


図 4: $\langle k \rangle = 4$ のランダムネットワークの生存確率

3 生存確率

図 3 は病気のノードが残る場合の病気のノード数の推移例であり, $\lambda = 0.01$ の場合百個ほど病気のノードが残っている.

病気のノードが一つでも残っている確率を生存確率 S という. 図 4, 図 5 ではノード数 $N = 10000, 50000, 100000$ のそれぞれの場合について 100 回試行した場合の生存確率 S の値を伝染率 λ の関数として示した. 図 4 はランダムネットワーク, 図 5 はスケールフリーネットワーク (BA モデル) に関する結果である. 平均次数は両者とも 4 である. $\mu = 0.1$ の場合 λ の値に臨界値が存在することが明らかになった. それぞれのネットワークサイズに対する S の λ 依存性はほぼ重なっており, $N = 10000 \sim 100000$ 程度で S のサイズ依存性は十分小さいと考えられる.

次に図 6, 図 7 に異なる平均次数に対する結果を比較した. 予測されたとおり, 平均次数が増加することによって λ の臨界値は小さくなる. 一方, ランダムネットワークと BA モデルを比較するとそれぞれの平均次数の場合において BA モデルの方が λ の低い臨界値をもつ.

4 グラフレイアウト方法

複雑ネットワークの可視化方法について説明する. 構造が複雑なネットワークは, グラフを分かりやすく描画するために計算機によって自動的にレイアウトする必要がある. 一般にグラフレイアウトはいくつかのグラフの美しさや理解しやすさを示す審美的基準に沿って行われる. 審美的基準の例としては, 隣接ノードが近くに配置されている, エッジの長さが均

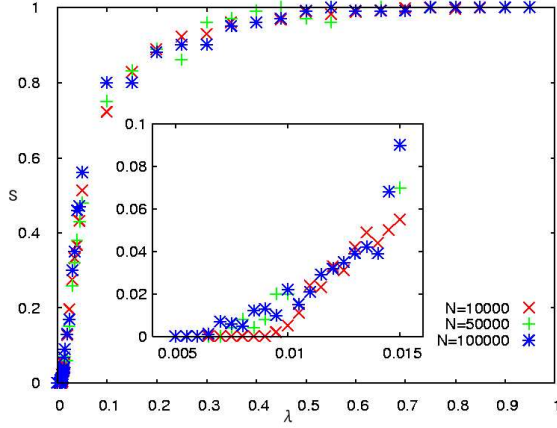


図 5: $\langle k \rangle = 4$ の BA モデルの生存確率

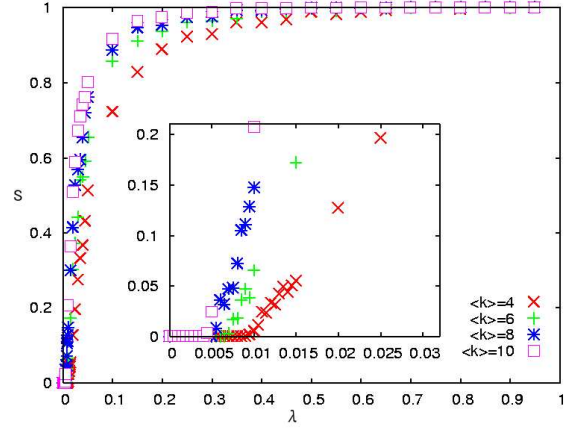


図 7: $N = 10000$ の BA モデルの生存確率

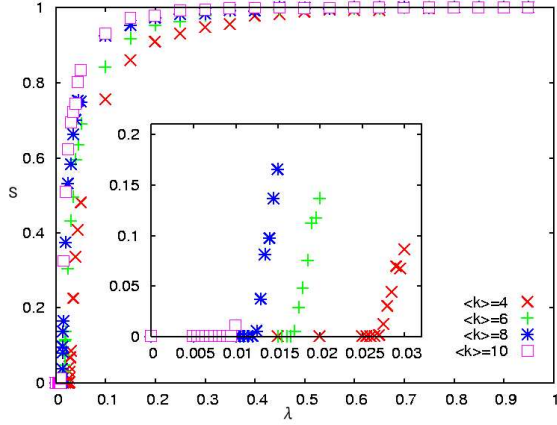


図 6: $N = 10000$ のランダムネットワークの生存確率

一である、ノード、エッジの重なりが少ない、エッジの交差数が少ない等が挙げられる。本研究では審美的基準の表現に用いるモデルとして仮想エネルギーモデルを使用した。この仮想エネルギーモデルのエネルギー関数 E をシミュレーテッドアニーリングの手法を用いて最適解に導く。

4.1 仮想エネルギーモデル

仮想エネルギーを引張圧縮ばねエネルギー E_S 、及びノード間斥力エネルギー E_R の和で表現する。

$$E = \gamma E_S + \delta \alpha E_R \quad (1)$$

γ, δ は重みを表し、 $\gamma + \delta = 1$ となるようにする。各エネルギーの重みを変化させることで審美的基準の優先度を変えることができる。

α はノードをランダムに配置した初期配置での E_S 、

αE_R を等しくするための係数である。具体的には次式を与える事で、どのネットワークにおいても $E_S, \alpha E_R$ の初期値がほぼ等しくなる。

$$\alpha = 170 \times \frac{2|A|}{|V|^2} \quad (2)$$

A は全エッジの集合、 V は全ノードの集合を表す。 $E_S, \alpha E_R$ の初期値を等しくする事で、どちらのエネルギーを重要視するか、重み γ, δ の比率を変えるだけで実現できる。

引張圧縮ばねモデルはエッジを自然長を持つばねとするモデルである。エネルギー関数 E_S は次式となる。

$$E_S = \sum_{e \in A} \{l - d(n_1(e), n_2(e))\}^2 \quad (3)$$

l は自然長、 d はノード間距離、 n はノード、 e はエッジをそれぞれ表す。

ノード間斥力モデルは各ノードを斥力を持った素子とするモデルである。エネルギー関数 E_R は次式となる。

$$E_R = \sum_{n \in V} (1/d(n_1, n_2)) \quad (4)$$

4.2 シミュレーテッドアニーリング

以下の手順で E を最適解へと導く。

- 1) ノード一つを摂動させ、エネルギーを E' とする。
- 2) 以下の確率 $P(E', E, T)$ を元に E' を新たな E として受理するかを判定する。

$$P(E', E, T) = \begin{cases} 1 & (E' - E < 0) \\ \exp(-\frac{E' - E}{T}) & (E' - E \geq 0) \end{cases} \quad (5)$$

T は温度変数である。

3) 1), 2) を繰り返した後, 次式に従い T を下げ, 1) に戻る。

$$T_{k+1} = 0.9 \times T_k \quad (6)$$

5 ネットワークの可視化

ランダムネットワークとスケールフリーネットワーク (BA モデル [1]) を可視化した結果を図 8 に示す。ノード数 100, 平均次数 10, $\gamma=0.5$, $\delta=0.5$ である。

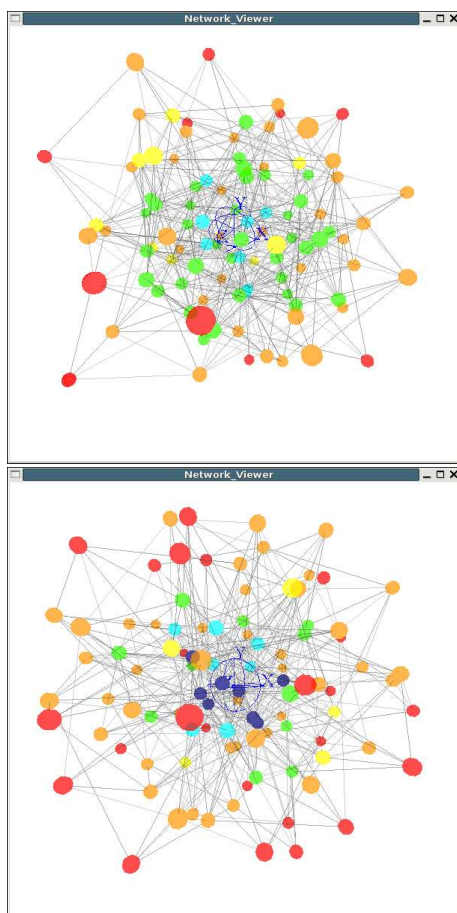


図 8: ランダムネット (上), BA モデル (下) の可視化。21 以上の次数を持つノードは青色で, 16 以上 20 以下のノードは緑色, 11 以上 15 以下のノードは黄色, 6 以上 10 以下のノードは橙色, 5 以下のノードは赤色で表示する。

6 まとめ

BA モデルではハブのまわりで病気が蔓延しやすくランダムネットワークよりも低い伝染率でも病気が蔓延しやすいと思われる。また, ランダムネットワークにおいて $\mu = \langle k \rangle \lambda_c$ という関係が成り立つと

思われるが $\mu = 0.1$ でしかシミュレーションを行っておらず, 今後 μ の値を変えてシミュレーションを行う必要がある。

複雑ネットワークの可視化の結果, 次数の多いノードはネットワークの中心付近に, 次数が少なくなるほど中心から離れるように描画された。BA モデルとランダムネットワークの構造の違いは中心付近のハブの有無によって区別されている。このグラフ上で病気の伝染の様子を表すことで, 複雑ネットワークにおけるコンタクトプロセスをより深く理解できる事が期待される。

参考文献

- [1] 増田直紀, 今野紀雄:「複雑ネットワークの科学」産業図書 (2005)
- [2] 西尾元宏: 板ばねモデルを用いたインタラクティブな曲線グラフ描画手法とその応用に関する研究, 東京大学大学院修士論文 (2003).
- [3] 片川智弘: 複雑ネットワーク上における破壊シミュレーションの可視化, 室蘭工業大学大学院修士論文 (2007).