

感染症流行予測におけるネットワーク構造の影響分析

Influence of Network Structure on Prediction of Epidemic Outbreaks

湯浅 友幸*1

Tomoyuki Yuasa

白山 晋*2

Susumu Shirayama

*1 東京大学大学院工学系研究科

School of Engineering, the University of Tokyo

*2 東京大学人工物工学研究センター

RACE, the University of Tokyo

Although an epidemic spreading through the network of human contacts has been widely studied using computer models, the knowledge about the influence of the network structure on the outbreak is not sufficient. This is because the types of the networks used in simulation are limited. In this paper, we examine the influence of the network structure on prediction of epidemic outbreaks using SIR (Susceptible-Infectious-Removed) model and 12 networks with distinctive statistical properties.

1. はじめに

数理モデルを用いてインフルエンザをはじめとする感染症の流行を予測する試みは古くから広く行われている。これまで微分方程式系を用いた単純なモデルから仮想都市を設計し、シミュレートする複雑なモデルまで数多くのモデルが提案されてきた。しかし、現実性の追求とモデル上の現象の分析のしやすさはトレードオフの関係にある。そのため、これらのモデルの中間に位置し、ある程度現実性を維持したままで分析が容易なネットワーク上の感染症モデルのシミュレーションが行われている。

このモデルでは人間をノード、人間どうしの接触をリンクとして扱うことで人間の接触関係のネットワークを想定し、そのネットワーク上で感染症が伝播するシミュレーションを行う。既存研究においてこのシミュレーションの結果はネットワークの統計的指標の観点から分析されており、平均頂点間距離 L 、クラスター係数 C 、次数分布 $p(k)$ 、次数相関 r [Newman 03] などの様々な指標の影響がこれまで調べられている。例えば、次数 k の分布 $p(k)$ がべき乗分布に従う、いわゆるスケールフリーネットワークにおいては感染率の非常に低い感染症でもネットワーク上に蔓延することが知られている [Moreno 02]。

しかしながら、既存研究においては BA モデル [Barabasi 99] のような限定的なネットワーク上での 1 つの統計的指標のみに着目した数値実験、理論分析しかなされていない。このため、ネットワーク構造の影響をより詳細に分析するには多様なネットワークを用いた実験結果間での比較が必要と考えられる。

本稿では様々なネットワーク上で感染症モデルのひとつである SIR モデルの数値実験を行い、ネットワーク構造と感染症の伝播ダイナミクスの関連性を分析し、予測可能性について検討する。

2. モデル

本稿で用いる SIR モデルおよびネットワークモデルについて述べる。

2.1 SIR モデル

SIR モデルは人口を S (Susceptible: 感受性人口), I (Infectious; 感染人口), R (Removed; 隔離人口) として

連絡先: 湯浅友幸, 東京大学大学院工学系研究科, 〒 277-8568
千葉県柏市柏の葉 5-1-5, yuasa@race.u-tokyo.ac.jp

表現することで、回復に際して患者が恒常的な免疫を獲得できる感染症を表現するモデルである。このモデルは $S(t)$, $I(t)$, $R(t)$ をそれぞれ時刻 t における感受性人口, 感染人口, 隔離人口とすれば以下の微分方程式系で表現される。

$$\frac{dS}{dt} = -\lambda IS \quad (1)$$

$$\frac{dI}{dt} = \lambda IS - \mu I \quad (2)$$

$$\frac{dR}{dt} = \mu I \quad (3)$$

ここで、 λ は感染率を表す比例定数、 μ は治癒率を表す比例定数 (本稿ではすべての実験で $\mu = 1$ としている) である。これをネットワーク上でモデル化すると、ノードに対して S, I, R の 3 状態を付与し、各ノードが周囲のノードの状態を加味して状態遷移を行っていくことになる。各ノードとリンクする状態 I のノード数を $n(I)$ と表す。このとき図 1 のように状態 S のノードは $\lambda n(I)$ の割合で状態 I へ、状態 I のノードは μ の割合で状態 R にその状態を遷移していく。本稿ではこのモデルを連続時間のグラフ上のマルコフ過程としてとらえ、実際にシミュレーションを行う際は離散時間過程へと近似して数値実験を行う。このとき微小時間 dt (1 ステップ, 本稿ではすべての実験で $dt = 0.01$ としている) の進行は次の (a) ~ (c) のようなプロセスによって表される。

- (a) ランダムにノードを一つ選択する。
- (b1) 選択ノードの状態が S であれば、確率 $\lambda n(I) dt$ で選択ノードの状態を I に変更する。
- (b2) 選択ノードの状態が I であれば、確率 μdt で選択ノードの状態を R に変更する。
- (b3) 選択ノードの状態が R であれば、特に処理を行わない。
- (c) (a), (b) をノード数の回数繰り返す。

2.2 ネットワークモデル

本稿で用いるネットワークモデルは全て無向、重み無しの単純グラフである。また、ネットワークを構成するノード数 N を $N = 10000$ とし、その次数を k とする。本稿では平均次数 $\langle k \rangle = 8$ とする。

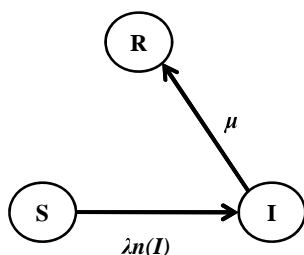


図 1: SIR モデルの概念図

本稿では以下に示す 6 種類のネットワークを用い、パラメータの調整によって 12 種類のネットワークを生成する．表 1 には各ネットワークの平均頂点間距離 L ，クラスター係数 C ，次数相関 r (次数相関の算出は文献 [Newman 03] による) の平均的な値を示している．

2.2.1 8 次格子

2 次元格子状にノードを配置し，各ノードはムーア近傍にあるノードとリンクでつながる．周期境界条件を用いる．以後 8 次格子と呼ぶ．

2.2.2 レギュラーランダムグラフ (RRG)

全てのノードが 8 本のリンクを持つように全ノードをリンクによってランダムにつなぐ．以後 RRG と記す．

2.2.3 WS モデル [Watts 98]

各ノードが周囲 8 近傍のノードとリンクでつながる環状の格子に対して，各リンクについて確率 p でランダムなつながり換えを施す．本稿の実験では p が 0, 0.01, 0.1, 1 の場合を扱い，特に $p = 0.01, 0.1$ の場合は小さな L と大きな C を有するいわゆるスモールワールドネットワークとなる．

2.2.4 KE モデル [Klemm 02]

ネットワークの成長とノードの非活性化というプロセスによって形成されるモデルであり，次数分布 $p(k)$ がべき乗分布 $p(k) \propto k^{-\gamma}$ に従うスケールフリーネットワークである．本稿では BA モデル [Barabasi 99] の生成に用いられる優先的選択によるネットワークの成長を確率 q で施す手法を用いる． q が 0, 0.01, 0.1, 1 の場合を扱い， $q = 0$ の場合には純粋な KE モデル， $q = 0.01, 0.1$ の場合には KE モデルと BA モデルの中間にあたるスモールワールド性を有するスケールフリーネットワーク， $q = 1$ の場合には純粋な BA モデルを生成する．特に q が 0, 0.01, 0.1 の場合には強い負の次数相関を持っている．

2.2.5 HK モデル [Holme 02]

成長と優先的選択による BA モデルの生成手法にクラスター構造の形成というプロセスを導入したモデルであり，スケールフリー性とスモールワールド性を持つネットワークである．0 に近い次数相関を持つ．

2.2.6 CNN モデル [Vázquez 03]

ネットワークの成長と潜在リンクのリンク化というプロセスによって形成されるモデルであり，スケールフリー性とスモールワールド性を持つネットワークである．強い正の次数相関を持つ．

以降の議論においては 8 次格子，RRG，WS モデル ($p = 0, 0.01, 0.1, 1$) の 6 種のネットワークをまとめて非スケールフリーネットワークと呼び，KE モデル ($q = 0, 0.01, 0.1, 1$)，HK モデル，CNN モデルの 6 種のネットワークをまとめてスケールフリーネットワークと呼ぶ．

表 1: 各ネットワークの統計的指標

	L	C	r
8 次格子	33	0.43	なし
RRG	4.8	3.0E-4	なし
WS($p = 0$)	625	0.64	なし
WS($p = 0.01$)	18	0.62	5.9E-4
WS($p = 0.1$)	7.0	0.47	-0.01
WS($p = 1$)	4.7	6.9E-4	-0.06
KE($q = 0$)	157	0.77	-0.12
KE($q = 0.01$)	8.8	0.74	-0.11
KE($q = 0.1$)	5.6	0.54	-0.12
KE($q = 1$)	3.9	6.5E-3	-0.03
HK	4.3	0.45	-0.03
CNN	6.4	0.38	0.16

3. 数値実験と考察

3.1 実験内容

前述の 12 種類のネットワーク上において初期条件として状態 I のノードをランダムに 10 ノード配置し，状態 I のノードが系から消失するまでシミュレーションを行う． λ の値は 0.05 から 1.0 までを 0.05 刻みの 20 パラメータで実験する．また，各実験は 100 回ずつ行っており，提示する結果は各試行の平均値となっている．

3.2 実験結果及び考察

実験結果の特徴量として本稿では終状態における状態 R のノード数 R_∞ および終状態に至るまでの時間 T を計測し，比較，考察を行う．

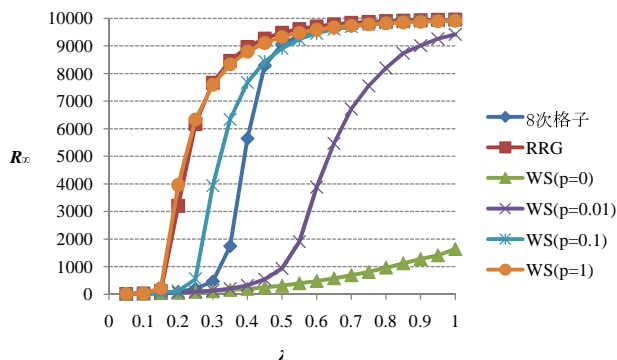
3.2.1 非スケールフリーネットワークの場合

図 2, 3 はそれぞれ非スケールフリーネットワークにおける感染率 λ と総感染ノード数 R_∞ ，感染収束までの時間 T の関係を示している．

はじめに図 2 を見ると，WS モデル ($p = 0$) の場合を除いた 5 種の結果ではそれぞれ λ が一定の値を超えるまではほとんど感染が広がらないが，ある一定値を境に R_∞ が急激に増加するという類似した挙動を示すことがわかる．Moreno らの既存研究 [Moreno 02] では SIR モデルにおいて感染拡大に必要な伝染率の閾値 λ_c が存在するとされており，今回得られた結果はこれに合致するものと言える．しかし，Moreno らが λ_c と感染の拡大傾向を次数分布のみに基づいて分析した一方で，今回我々の得た結果は感染症の伝播に対するネットワークの他の統計的指標の影響を示唆するものとなっている．

この結果を平均頂点間距離 L の観点から評価すると，まず R_∞ の急増が見られない WS モデル ($p = 0$) は表 1 のように他に比べて非常に L が大きいことがわかる．また， L の特に小さな RRG，WS モデル ($p = 1$) においては小さな λ において R_∞ の急増が見られており，感染が拡大しやすいことがわかる．これらの結果より L の小さなネットワークほど感染が拡大しやすいといえる．この現象は L の小さなネットワークの方がより多くのノードを感染するために経由すべきノード数が少なくなり，感染が途中で途絶える可能性がより低くなるためと考えられる．

しかし，クラスター係数 C の観点から今回の結果を評価すると， R_∞ の傾向が L のみに依存するとはいえない．図 2 に

図 2: 非スケールフリーネットワークにおける R_{∞}

において 8 次格子と WS モデル ($p = 0.01$) の結果を比較すると、8 次格子の方が感染がより拡大しやすいことがわかる。一方、表 1 においてこれらのネットワークの統計的指標を見ると、 L に関しては 8 次格子が大きく、 C に関しては WS モデル ($p = 0.01$) の方が大きい。これらのことから R_{∞} は L の減少により増加するのみではなく、同時に C の増大によって減少するといえる。

一般にネットワークにおいては L と C の間には正の相関があるが、 R_{∞} の挙動に対してこの相関の中で両指標がどの程度ずつ影響を与えているかを解明するのは今後の課題である。

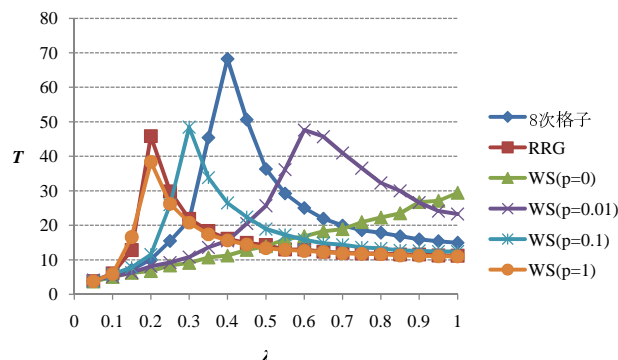
次に実験結果を図 3 に示される感染収束までの時間 T に基づいて考察する。 λ の増加に伴う T の傾向を見ると WS モデル ($p = 0$) の場合を除いてはある一定の λ でピークを迎えるまで T が上昇し、その後減少する傾向を示している。これらの T がピークを示す λ はすべて図 2 における感染がほとんど生じない領域から爆発的な感染がおこる領域への過渡期に合致する。 λ が小さい間は感染が広がらず短期間に収束、大きい間は感染が高速に大規模に広がって収束、中間の領域では時間をかけて大規模に広がって収束しているものと考えられる。また、WS モデル ($p = 0$) では大きな λ においてもネットワーク全体へは感染が拡大しないが、この場合は λ の増加とそれに伴う R_{∞} の増加にしたがって T も増加していく。これは感染の規模の拡大とともにその収束までの時間も長くなるためと考えられる。

また、WS モデル ($p = 0$) 以外で T のピーク値を比べると、より L の小さなネットワークほどその値が小さいことがわかる。一方、ほとんどすべてのノードが感染する λ が大きな領域で T を比べてみると、より C の小さなネットワークほど T の値が小さいことがわかる。これらのことから感染収束までの時間に関しては L の小さなネットワークほど中規模感染の際の時間が短くなり、 C の小さなネットワークほど大規模感染の際の時間が短くなるといえる。

3.2.2 スケールフリーネットワークの場合

図 4, 5 はそれぞれスケールフリーネットワークにおける感染率 λ と総感染ノード数 R_{∞} 、感染収束までの時間 T の関係を示している。

はじめに図 4 を見ると、KE モデル ($q = 0$)、CNN モデルのケースでは λ の増大に対して R_{∞} が徐々に増加するのに対し、他の 4 ケースでは λ の増大に対して R_{∞} が急増するのがわかる。

図 3: 非スケールフリーネットワークにおける T

前者の KE モデル ($q = 0$)、CNN モデルのケースでは λ が非常に大きな場合にも感染ノードがネットワーク全体には及んでいない。KE モデル ($q = 0$) に関しては非常に大きな L をとっており、非スケールフリーネットワークにおける WS モデル ($p = 0$) の場合と同様にスケールフリーネットワークにおいても長距離の感染症伝播は困難であることを示している。一方、CNN モデルは L が小さいにも関わらず、感染が拡大しない。Moreno らは正の次数相関を持つネットワークでは感染は拡大しにくいと述べている [Moreno 03]。CNN モデルは表 1 のように正の次数相関 r をもつため、また、前項で述べたクラスター係数の効果が存在するために感染が抑制されるものと考えられる。

一方、後者の KE モデル ($q = 0.01$)、KE モデル ($q = 0.1$)、KE モデル ($q = 1$)、HK モデルの場合は L が小さく、 r に関しては負または 0 に近い値をとるが、 λ の増大とともにネットワーク全体に感染が拡大している。KE モデルにおいて q の増加に対する R_{∞} の挙動を見ると、 q の増加に伴い、 L 、 C が減少し、 R_{∞} が増加していることがわかる。 L 、 C の小さなネットワークで R_{∞} が増加する現象は前項と同様のものである。一方、KE モデル ($q = 1$) と HK モデルの結果を比べるとほとんど同じ挙動を示すことがわかる。これら 2 つのネットワークは近い L をもつ一方、 C には大きな差がある。前項の非スケールフリーネットワークの結果では C の増大による感染の抑制効果が顕著な例があったが、スケールフリーネットワークにおいてはこの効果が小さいものと考えられる。

また、特に KE モデル ($q = 1$) と HK モデルに関しては図 2 における非スケールフリーネットワークの場合よりも早い段階で R_{∞} の急増傾向が見られる。これは Moreno らの文献 [Moreno 02] におけるスケールフリーネットワークでは λ_c が 0 に近づくという言及と合致するものといえる。

次に実験結果を図 5 に示される感染収束までの時間 T に基づいて考察する。はじめに各ネットワークにおける T を見ると、図 3 の非スケールフリーネットワークにおける結果よりも全体的に小さな T を示していることがわかる。スケールフリーネットワークにおいては非常に次数の高いハブにあたるノードが存在する。この現象は感染したハブが急激に感染を広め、感染スピードが速くなることでその収束も早まるためと考えられる。

KE モデル ($q = 0$) と CNN モデル以外の 4 ケースに関しては前項の非スケールフリーネットワークと同様に λ の増大とそれに伴う R_{∞} の増大とともに T がピークを迎える。また、

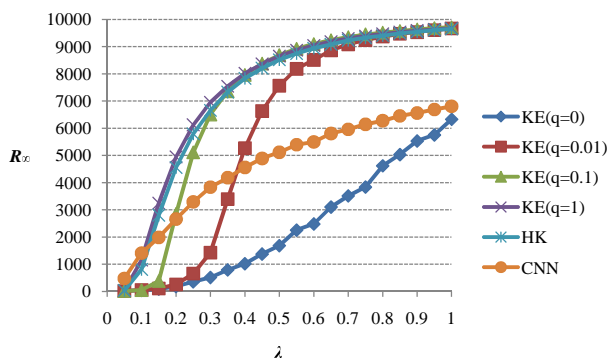


図 4: スケールフリーネットワークにおける R_∞

L, C 双方について差がある KE モデル ($q = 0.01, 0.1$) の間では T の挙動について明確な差がある一方, L が近い値をとり, C について差がある KE モデル ($q = 1$) と HK モデルの間ではその挙動には大きな差が見られない. R_∞ と同様に T についてもスケールフリーネットワークでは L が影響を与える一方で C の影響は小さいものと考えられる.

また, KE モデル ($q = 0$) と CNN モデルを見るとその傾向は全く異なるものとなっている. まず KE モデル ($q = 0$) であるがこの場合は λ とそれに伴う R_∞ の増加に対して T も共に増加の傾向にあることがわかる. 一方, CNN モデルの場合は λ の増加とともに R_∞ が増加するが, これに対して T は変動しないという結果が得られている. KE モデル ($q = 0$) の結果は感染の拡大とともに収束までの時間も長くなるということを示す結果であるが, CNN モデルの結果は伝染率 λ の増大が感染の拡大と同時に感染スピードの増加までもたらしているものと解釈される. CNN モデルにおける結果にはおそらく正の次数相関を持つネットワーク構造が深く関連しているものと推測され, 高次のハブどうしがリンクする構造が感染の拡大スピードを加速するものと考えられる.

4. 結言

本稿では 12 種類の複雑ネットワーク上において SIR モデルのシミュレーションを行い, そのダイナミクスとネットワークのスケールフリー性, 主なネットワークの統計的指標である平均頂点間距離 L , クラスタ係数 C , 次数相関 r との関連性を考察した.

これにより既存研究で得られている知見を再確認するとともに, これらの統計的指標が総感染ノード数 R_∞ , 感染収束までの時間 T に及ぼす影響についての複数の新たな知見を得た.

本稿の範囲では平均頂点間距離 L とクラスタ係数 C が相関の中でいかにネットワーク上の感染症の伝播現象に対して影響を及ぼすかを詳細に検討することができなかった. また, 次数相関の及ぼす効果についてもより詳細な分析が必要と考えられる. さらに多くのネットワーク上でのシミュレーション結果を比較分析することでこれらの解明を行うことが今後の課題である.

参考文献

[Barabasi 99] Barabasi, A. and Albert, R.: Emergence of scaling in random networks, *Science*, Vol. 286, No. 5439,

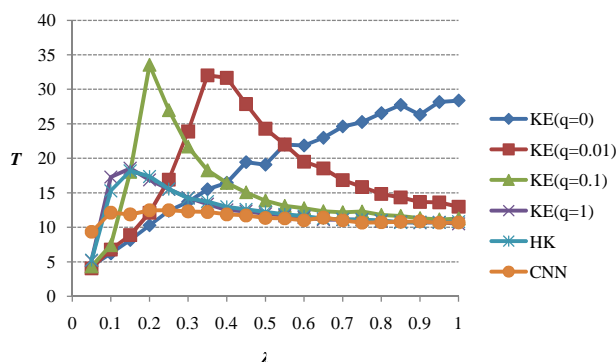


図 5: スケールフリーネットワークにおける T

p. 509 (1999)

[Holme 02] Holme, P. and Kim, B.: Growing scale-free networks with tunable clustering, *Physical Review E*, Vol. 65, No. 2, p. 26107 (2002)

[Klemm 02] Klemm, K. and Eguíluz, V.: Growing scale-free networks with small-world behavior, *Physical Review E*, Vol. 65, No. 5, p. 57102 (2002)

[Moreno 02] Moreno, Y., Pastor-Satorras, R., and Vespignani, A.: Epidemic outbreaks in complex heterogeneous networks, *The European Physical Journal B-Condensed Matter and Complex Systems*, Vol. 26, No. 4, pp. 521–529 (2002)

[Moreno 03] Moreno, Y., Gómez, J., and Pacheco, A.: Epidemic incidence in correlated complex networks, *Physical Review E*, Vol. 68, No. 3, p. 35103 (2003)

[Newman 03] Newman, M.: Mixing patterns in networks, *Physical Review E*, Vol. 67, No. 2, p. 26126 (2003)

[Vázquez 03] Vázquez, A.: Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations, *Physical Review E*, Vol. 67, No. 5, p. 56104 (2003)

[Watts 98] Watts, D. and Strogatz, S.: Collective dynamics of 'small-world' networks, *Nature*, Vol. 393, No. 6684, pp. 440–442 (1998)