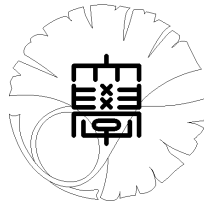


数理科学実践研究レター 2023-2 March 30, 2023

複雑ネットワーク上の情報拡散

by

林 晃平、矢野 良輔



UNIVERSITY OF TOKYO
GRADUATE SCHOOL OF MATHEMATICAL SCIENCES
KOMABA, TOKYO, JAPAN

複雑ネットワーク上の情報拡散

林晃平¹ (東京大学大学院数理科学研究科)

Kohei Hayashi (Graduate School of Mathematical Sciences, The University of Tokyo)

矢野良輔² (東京海上ディーアール株式会社)

Ryosuke Yano (Tokio Marine dR Co., Ltd.)

概要

本研究では、ネットワーク上の情報拡散を扱う。実社会にみられるネットワーク構造の多くは複雑性を有し、特に遠くのエージェントにも影響を与えうる長距離相互作用を持つことが特徴的である。本稿ではまず一般のネットワークにおける情報拡散現象を扱うため、離散無向グラフ上のSIRモデルを考える。次に、具体的にいくつかのグラフ構造を与えることでより詳細な解析を行う。特に複雑ネットワークについては、長距離相互作用をモデル化し連続極限をとることで、非整数階ラプラシアンで駆動される偏微分方程式が導出されることを示す。

1 導入

人やものの繋がりを可視化したものをネットワークと呼び、数学的には頂点集合と辺集合で構成される「グラフ」を用いてモデル化することができる。(本稿では特に、無向グラフを用いてモデル化される場合を扱う。) ネットワークは実社会における様々な局面でみられ、構成要素間に密接な関係があることで利便性をもたらす一方、ネットワーク上の情報拡散においてはしばしばリスクを伴う。例えば、特に災害時において頻発して問題視されたインターネット上でデマ情報や風評被害の拡散や、昨今の新型コロナウイルス感染症の拡大など、人同士の繋がりが過去に比べていかに密接になっておりそれに伴うリスク管理が重要かを実感する事例は枚挙にいとまがない。このような関係性の緊密化に加え、実社会に見られるネットワークの多くは空間的に一様なモデルや独立同分布に従う確率変数を用いたモデルなどでは説明できないような複雑性を有することが知られている。このような実社会にみられるネットワークの特徴は、以下の性質に集約される。

- 1) 2点間距離が小さい。
- 2) 三角形や四角形などの閉経路(クラスター)が多く存在する。
- 3) 頂点と繋がっている辺の本数を表す次数の分布の裾野が広い。

これら3つの特徴を持つネットワークを「複雑ネットワーク」と呼ぶ。最初の2つの性質は「スモールワールド性」と呼ばれ、例えば人同士のネットワークでは共通の知人を通じて遠くの人も含めた多くの人と繋がることができることを象徴している。一方、ネットワークの複雑性を表す3番目の性質は「スケールフリー性」と呼ばれ、解析のしやすさのために強い対称性などを仮定する古典的な数理モデルによる解析を困難にしている。(例えば各頂点が一定の次数を持つ正則グラフは数学的に扱いやすく、全ての頂点同士が繋がっている完全グラフに近いものを取ればスモールワールド性は満たすがスケールフリー性は満たさない。) このような背景から、上記のような複雑性を反映でき、かつ適宜必要な近似を用いて詳細な解析が可能であるようなモデルが求められる。そこで本稿では、まず一般のグラフに付随する情報拡散モデルを考え、次に適切な近似やスケール操作の下で従来の古典的なモデルでは説明できないような現象を表現することを目指す。以下では、説明のしやすさのために感染症拡大のモデルとして解釈することにする。しかし、本稿で考える数理モデルは上述のような様々な局面に応用でき、その用途は必ずしも感染症拡大に限定されない。

本稿の構成は以下の通りである。まず第2章でモデルの定義を行う。ここでは有限状態、離散時間Markov連鎖から出発し、適切な仮定の下である常微分方程式が導出されることをまず示す。以降の章では一般のネットワーク及びこの常微分方程式系を最初に与え、決定論的な枠組みで感染症拡大の時間発展を考えることにする。続く第3章では、具体例なネットワークに対する個別の解析結果を

¹kohei@ms.u-tokyo.ac.jp

²ryosuke.yano@tokiorisk.co.jp

述べる。まず対称性の高い有限ネットワークの場合の結果を述べ、より複雑な場合として実社会においてしばしば観測される super-spreader のモデル化及びその解析を行う。ここで、super-spreader は感染症に罹患していて、かつ多数の他エージェントと接触するようなエージェントである。次に、第4章では複雑ネットワークの特徴の一つである長距離相互作用があるような場合を考える。特に、エージェント数を無限大にする極限操作により、非整数階ラプラシアンで駆動される偏微分方程式が導出されることを示す。最後に、第5章で本稿の結論を述べる。

2 ネットワーク上の情報拡散モデル

以下、 V を頂点集合、 E を辺集合とする無向グラフ $G = (V, E)$ で与えられるネットワーク上の感染症拡大モデルを考える。頂点集合の元をエージェントと呼ぶことにし、各エージェントは S , I , R のいずれかの状態をとるものとする。ここで、 S (Susceptible) は免疫を持たず感受性を持つ状態、 I (Infected) は感染性を持つ状態、 R (Recovered) は感染症から回復し免疫を獲得した状態を表す。このとき、系全体の状態を表す空間 (状態空間) を $\mathcal{X} = \{S, I, R\}^{|V|}$ とし、 \mathcal{X} の各要素を $\eta = \{\eta_j : j \in V\}$ のように書く。例えば、 $\eta_j = I$ であれば、エージェント $j \in V$ が感染症に罹患している状態を表している。また、 β, γ を正定数、 $a = (a_{i,j})_{i,j \in V}$ をエージェント間の相互作用を表す非負値行列とし、次の規則に従い時間発展する \mathcal{X} 上の離散時間 Markov 連鎖 $\{\eta(t) : t = 0, 1, \dots\}$ を考える：

- 時刻 t において状態 S であるエージェント j が状態 I である別のエージェント j' と隣接しているとき (即ち $\{j, j'\} \in E$ であるとき)、時刻 $t+1$ においてエージェント j は確率 $\beta a_{j,j'}$ で状態 I に遷移し、確率 $1 - \beta a_{j,j'}$ で状態 S のまま維持される。
- 時刻 t において状態 I であるエージェント j は、時刻 $t+1$ において確率 γ で状態 R に遷移し、確率 $1 - \gamma$ で状態 I のまま維持される。

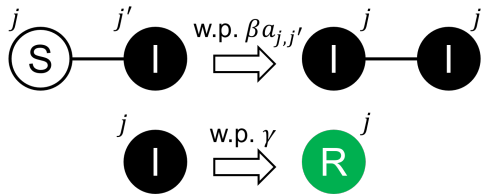


図 1: ネットワーク上の SIR モデルの時間発展則

定義より、時刻 $t+1$ においてエージェント j が状態 S である確率を次のように再帰的に計算することができる。

$$\begin{aligned} \mathbb{P}(\eta_j(t+1) = S) &= \sum_{j' \in V} \mathbb{P}(\eta_j(t) = S, \eta_{j'}(t) \in \{S, R\}) \\ &+ \sum_{j' \in V, \{j, j'\} \in E} (1 - \beta a_{j,j'}) \mathbb{P}(\eta_j(t) = S, \eta_{j'}(t) = I) \\ &+ \sum_{j' \in V, \{j, j'\} \notin E} \mathbb{P}(\eta_j(t) = S, \eta_{j'}(t) = I). \end{aligned}$$

以下、エージェント j が時刻 t において状態 S , I , R である確率をそれぞれ $S_j(t)$, $I_j(t)$, $R_j(t)$ とおく。このとき、独立性 $\mathbb{P}(\eta_j = S, \eta_{j'} = I) = \mathbb{P}(\eta_j = S)\mathbb{P}(\eta_{j'} = I)$ が成り立つと仮定すると、上式から $S_j(t)$ に関する漸化式を得ることができる。同様に $I_j(t)$ 及び $R_j(t)$ に対する漸化式を導出し、更に時間間隔を無限小にする極限を取る。このとき、パラメータ β, γ に対しても時間刻みに関するスケールリングを行った後、ネットワーク (V, E) 上の情報拡散モデルとして次の常微分方程式系を得る

ことができる。

$$\begin{cases} \frac{dS_j}{dt} = -\beta S_j \sum_{j' \in V, \{j, j'\} \in E} a_{j, j'} I_{j'}, \\ \frac{dI_j}{dt} = \beta S_j \sum_{j' \in V, \{j, j'\} \in E} a_{j, j'} I_{j'} - \gamma I_j, \\ \frac{dR_j}{dt} = \gamma I_j. \end{cases} \quad (1)$$

以下では、ネットワーク $G = (V, E)$ 及びそれに付随する決定論的な方程式系 (1) を最初に与え、 G 上の感染症拡大のダイナミクスを解析することにする。

3 いくつかの例

ここではグラフ $G = (V, E)$ を具体的に与え、方程式系 (1) で記述される時間発展則により感染症がどのように拡がっていくかをみる。簡単のため、以下 $\gamma = 0$ とおき、初期状態（従って各時刻）において $\mathbb{P}(\eta_j \in \{S, I\}) = 1$ が成り立つ場合を考える。更に、定義より各 $j \in V$ に対し $S_j + I_j = 1$ が成り立つので、方程式系 (1) は $\{I_j : j \in V\}$ のみに関する次の方程式に帰着される。

$$\frac{dI_j}{dt} = \beta(1 - I_j) \sum_{j' \in V} a_{j, j'} I_{j'}. \quad (2)$$

この章では $a = (a_{i, j})_{i, j \in V}$ は $a_{i, i} = 0$ を満たす（自己ループを許さない）グラフ G の隣接行列とする。ここではまず、グラフ G が空間的に均質であり古典的な扱いが可能な場合を考え、全体の感染拡大率

$$\bar{I}(t) = \frac{1}{|V|} \sum_{j \in V} I_j(t)$$

を計算する。また、以下ではエージェントの総数 $|V|$ が十分大きい状況を主に扱うことにする。

例 3.1 (疎グラフ). エージェント同士の関係が希薄である場合（即ち $|E| \approx 0$ が成り立つ場合）のグラフを疎グラフと呼ぶ。このとき、エージェントの人数が十分多ければ、 $d\bar{I}/dt \approx 0$ が成り立ち、全体の感染拡大率 $\bar{I}(t)$ は殆ど時間変化しない。

例 3.2 (完全グラフ). 全てのエージェント同士が繋がっている完全グラフでは、 $d\bar{I}/dt = \beta|V|\bar{I}(1 - \bar{I})$ が近似的に成り立つ。このとき、任意の時刻 $t > 0$ に対し

$$\bar{I}(t) = \frac{\bar{I}_0}{\bar{I}_0 + (1 - \bar{I}_0)e^{-\beta|V|t}} \rightarrow 1, \quad |V| \rightarrow \infty.$$

が成り立ち、 $\bar{I}_0 > 0$ であれば瞬時に感染爆発することがわかる。

例 3.3 (正則グラフ). 各頂点の次数が d （ただし d は偶数）の拡張格子では、平均場近似 $I_i \approx \bar{I}$ の下、 $d\bar{I}/dt = \beta d(1 - \bar{I})\bar{I}$ となり、時刻 t での全体感染率は次のように計算できる。

$$\bar{I}(t) = \frac{\bar{I}_0}{\bar{I}_0 + (1 - \bar{I}_0)e^{-\beta dt}}.$$

即ち、正則グラフの場合には平均的にはロジスティックな挙動が観測される。

例 3.4 (Super-spreader の影響). 感染症を持っており、かつ多数の他エージェントと接触を行うようなエージェントのことを super-spreader と呼ぶ。ここでは $\mathbb{P}(\eta_j = I) = 1$ であるような super-spreader を黒丸で表し、それ以外のエージェントを白丸で表すことにし、図3のような2つの状況を考えることにする。図3の左の場合 (**Case 1**) では、 $d+1$ 人の super-spreader が存在するがそれ以外のエージェント同士は接触しない場合を考えている。一方、図3の右の場合 (**Case 2**) では、super-spreader の人数は1人であるが他エージェントの繋がり方が次数 d の正則グラフの場合を想定している。ど

ちらの場合も super-spreader 以外のエージェントの次数は $d + 1$ である．平均場近似 $I_i \approx \bar{I}$ を用いると，それぞれの場合に対し次のように全体の感染拡大率 \bar{I} を計算することができる．

Case 1 : $(d/dt)\bar{I} = \beta(d + 1)(1 - \bar{I})$ より，

$$\bar{I}(t) = 1 - (1 - \bar{I}_0)e^{-\beta(d+1)t}.$$

Case 2 : $(d/dt)\bar{I} = \beta(1 + d\bar{I})(1 - \bar{I})$ より，

$$\bar{I}(t) = 1 - \frac{d + 1}{(d + 1) + d\bar{I}_0(1 - e^{-\beta(d+1)t})}(1 - \bar{I}_0)e^{-\beta(d+1)t}.$$

従って， $\bar{I}_{\text{Case1}}(t) \leq \bar{I}_{\text{Case2}}(t)$ が任意の時刻 t で成り立ち，**Case 2** の方がより速く情報が拡散することがわかる．即ち，このモデルでは少数の super-spreader の影響よりも大多数の他エージェントによる拡散効果の方が大きいということになり，感染症拡大を抑制するにはエージェント全体に対し接触を減らすよう働きかけていくことが肝要であることを示唆している．

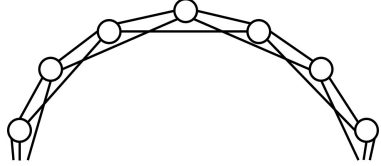


図 2: 正則グラフの例 (次数 $d = 4$) .

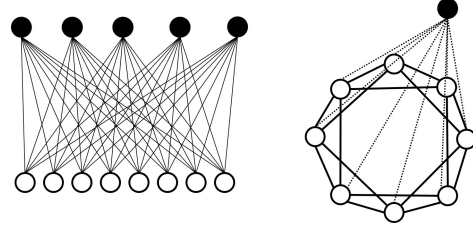


図 3: Super-spreader の影響 ($d = 4$ の場合) .

4 複雑ネットワークにおける異常拡散

ここでは，本稿の主題である複雑ネットワーク上の感染症拡大について考える．第 1 章で指摘したように，エージェント同士が複雑かつ密接に繋がっている複雑ネットワークでは，感染症が短期間に爆発的に拡大する危機的状況が生じやすいと考えられる．実際，[2] では，次のように次数分布の裾野が広いというスケールフリー性が感染爆発を引き起こすことが指摘されている．このことを説明するため，グラフ $G = (V, E)$ として複雑ネットワークを想定し，各エージェント $j \in V$ の次数 d_j に対しスケールフリー性 $\mathbb{P}(d_j = k) \sim k^{-\alpha}$ が (次数が大きい領域で近似的に) 成り立つと仮定する．ここで， $\alpha > 0$ はある定数であり，実社会にみられる多くのネットワークでは α は 2 から 3 の間の値をとることが指摘されている [1]．このとき，方程式 (2) において，互いに隣接するエージェントの感染確率が近い，即ち $\{j, j'\} \in E$ のとき $I_j \approx I_{j'}$ が成り立つと仮定すると，方程式 (2) は明示的に解くことができ，

$$\frac{I_j(t)}{1 - I_j(t)} = \frac{I_j(0)}{1 - I_j(0)} e^{\beta d_j t}$$

が各 $j \in V$ に対し成り立つ．ところが，上述のように近似的に Pareto 分布に従う次数分布はモーメントを持たないため，上式で次数についての平均は， $\inf_{j \in V} I_j(0) > 0$ であれば発散する．換言すれば，複雑ネットワークの場合には特異的な速度で感染症が拡大するため，その時間発展をみるには，近接相互作用の場合と比べてより遅い時間スケールで観測する必要があると考えられる．このことをより明確にするため，ここでは空間的な距離を導入したのちに時間と空間に対するスケール操作を行うことで位置変数について連続な方程式 (偏微分方程式) を導出することを考える．そのため，以下では方程式 (2) に拡散項を加えた以下のモデルを扱う．

$$\frac{dI_j}{dt} = \nu \sum_{j' \in V, \{j, j'\} \in E} c_{j, j'} (I_{j'} - I_j) + \beta(1 - I_j) \sum_{j' \in V, \{j, j'\} \in E} a_{j, j'} I_{j'}. \quad (3)$$

ここで、 $\nu > 0$ は定数、 $c = c_{j,j'}$ はエージェント j 及び j' 間のコンダクタンスであり、感染症の多いところから少ないところへ向かう流れの生じやすさをモデル化している。また以下では、規格化条件 $\sum_{j'} c_{j,j'} = \sum_{j'} a_{j,j'} = 1$ が任意の $j \in V$ について成り立つと仮定する。

4.1 最近接相互作用の場合

ここでは簡単のため、1次元正方格子を考える。このとき、 $|j - j'| = 1$ であるような $j, j' \in \mathbb{Z}$ に対し $c_{j,j'} = a_{j,j'} = 1/2$ とすると、方程式 (3) は次のように書ける。

$$\frac{dI_j}{dt} = \nu(I_{j+1} + I_{j-1} - 2I_j) + \beta(1 - I_j)(I_{j+1} + I_{j-1}).$$

このとき、空間変数 $j \in \mathbb{Z}$ に関するスケール操作を行い偏微分方程式を導出することを考える。そのために、 $n > 0$ をスケールパラメータとして $\nu = \nu_n$ 、 $\beta = \beta_n$ とおき、 $[0, \infty) \times \mathbb{R}$ 上の関数 $I^n = I^n(t, x)$ を次のように定める。

$$I^n(t, x) = \sum_{j \in \mathbb{Z}} I^n(t, j/n) \mathbf{1}_{[j/n, (j+1)/n)}(x).$$

このとき、拡散項及び反応項が残るようにするためには、 $\nu_n = O(n^2)$ かつ $\beta_n = O(1)$ である必要があることがわかる。即ち、 $\nu_n = \nu n^2$ 及び $\beta_n = \beta$ とおくと、 $n \rightarrow \infty$ の極限により、 $I = I(t, x)$ に関する次の1次元反応拡散方程式を導出することができる。

$$\partial_t I = \nu \Delta I + \beta(1 - I)I.$$

ただし、 $\Delta = \partial^2 / \partial x^2$ は1次元ラプラシアンである。

4.2 長距離相互作用がある場合

次に、長距離相互作用がある場合の感染症拡大について検討する。以下、 $V = \mathbb{Z}$ とし、相互作用がある定数 $s \in (0, 2)$ を用いて $c_{j,j'} = a_{j,j'} = (2\zeta(s))^{-1} |j - j'|^{-s} \mathbf{1}_{\{j' \neq j\}}$ で与えられると仮定する。ここで、 $\zeta(s) = \sum_{k \in \mathbb{N}} k^{-s}$ はゼータ関数であり、係数は規格化条件を満たすように定めた。このとき、方程式 (3) の反応項は次のように書ける。

$$-\frac{\beta_n n^{-s}}{2\zeta(s)} (1 - I_j) \frac{1}{n} \sum_{j' \in \mathbb{Z}} \frac{I_j - I_{j'}}{\left| \frac{j}{n} - \frac{j'}{n} \right|^{1+s}} + \beta_n (1 - I_j) I_j. \quad (4)$$

ここで、規格化条件 $\sum_{j' \in \mathbb{Z}} a_{j,j'} = 1$ を用いた。最近接相互作用の場合と同様のスケール操作を空間変数について行い連続方程式を導出することを考えると、式 (4) の第1項が非自明な項に収束するためには $\beta_n = O(n^s)$ が必要である。ところが、第2項については $\beta_n = O(1)$ が適切なスケールリングなので、もし $\beta_n = O(n^s)$ と仮定すると反応項は $n \rightarrow \infty$ で発散してしまう。よって、反応項が非自明になるためには、第2項を基準にして $\beta_n = O(1)$ を満たすように取る必要があり、このとき導出できる反応項は s の値に依らず最近接相互作用の場合と同じ形となる。一方、式 (4) の第1項と同じ形である拡散項については $\nu_n = O(n^s)$ が適切なスケールリングとなる。以上より、 $\nu_n = \nu n^s$ 及び $\beta_n = \beta$ とすると、次の偏微分方程式を $n \rightarrow \infty$ の極限により導出することができる。

$$\partial_t I = -\frac{\nu}{2\zeta(s)c_{1,s}} (-\Delta)^{s/2} I + \beta(1 - I)I.$$

ここで、 \mathbb{R}^d 上の非整数階ラプラシアン $(-\Delta)^{s/2}$ を次で定めた： $s \in (0, 2)$ に対し、

$$(-\Delta)^{s/2} f(x) = c_{d,s} \int_{\mathbb{R}^d} \frac{f(x) - f(y)}{|x - y|^{d+s}} dy.$$

ただし、 $c_{d,s} = 2^s \pi^{-d/2} \Gamma((d+s)/2) / |\Gamma(-s/2)|$ は規格化定数であり、Fourier 変換を $\mathcal{F}[f](\xi) = \int_{\mathbb{R}^d} f(x) \exp(-ix \cdot \xi) dx$ とするとき、 $\mathcal{F}[(-\Delta)^{s/2} f](\xi) = |\xi|^s \mathcal{F}[f](\xi)$ が成り立つ。(非整数階ラプラシ

アンを含む偏微分方程式に関する基本的な性質については、文献 [3] を参照されたい。) 上記の場合、拡散項と反応項に対するスケーリングの比は $\nu_n/\beta_n = n^s, s \in (0, 2)$ となるので、最近接相互作用の場合と比較して拡散項に対してより遅いスケール操作を行っている。即ち、長距離相互作用の存在により拡散項に対し通常のラプラシアンの場合とは異なるスケール操作が要求される異常拡散がみられる。

5 結論

本稿ではグラフ上の SIR モデルを導入し、正方格子などの古典的な場合に対する解析を行ったのち、super-spreader の影響と複雑ネットワーク上の情報拡散について検討した。Super-spreader の影響については、spreader が多く存在するが他エージェントの繋がりが希薄である場合と、spreader が少ないが大多数のエージェントが正方格子として繋がっている場合を比較し、検討したモデルでは後者の方がより感染症が拡大することを示した。一方、複雑ネットワークに対して導入部で述べたような特徴を総合的に検討するのは困難であるため、本稿では長距離相互作用の存在に着目し、正方格子の場合にみられるような状況とは異なり特異的な振る舞いを示す異常拡散現象を呈するモデルを提案した。今後の展望として、特に複雑ネットワーク上の情報拡散について、感染症拡大の程度をはかるために重要な指標である実効再生産数を今回の異常拡散モデルに対して理論的に計算することや、実データを用いた実証的な検討を行うことなどが挙げられる。

謝辞 本研究の内容に関して様々なご意見をくださった東京大学大学院数理科学研究科の山本宏子様から心から御礼申し上げます。

参考文献

- [1] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.
- [2] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proceedings of the national academy of sciences*, 101(11):3747–3752, 2004.
- [3] E. M. Stein. *Singular integrals and differentiability properties of functions*, volume 2. Princeton university press, 1970.