

ネットワーク構造から探る人間行動の影響力

Influence power of man action for which it searches from network structure

大 崎 佑 一

OOSAKI Yuuich

1. はじめに

我々は、毎日数多くの意思決定をしている。その多くは、周りの人達がとる行動にも大きく左右され、そのために、全体として意図せぬ帰結を招くことが多い。そして、個人レベルでの行動には何の因果関係もない性質が全体として現れることを、創発という。創発現象の例としては、ある商品が爆発的にヒットしたり、風評などによって伝統ある企業が市場から急に閉め出されたりすることなどがある。

このような創発現象を引き起こすことになる大きな要因は、個人と個人の相互作用、特に、相互作用を規定する人々のつながり、すなわちネットワーク構造にその秘訣が隠されていると考えられる。

人と人のつながり方や、要素間のつながり方をネットワークとして扱い、そしてネットワークの性質や全体的に現れる現象を扱うのが“ネットワーク科学”である。創発現象を生むメカニズムを解明するために、ネットワーク科学に対する関心が高まっている。

人間行動に関わる研究は、規範的及び記述的なアプローチに分類できる。前者は、経済学、意思決定理論、ゲーム理論などに見られるアプローチで、合理的な行動とは何かや、個々に合理的に振る舞うとき、全体の挙動に現れる性質などを明らかにする。後者は、認知心理学や社会心理学、実験経済学などに見られるアプローチで、人間は、ある与えられた状況の下で、どのように行動するのかを明らかにする。

情報技術の深化の恩恵を受けて、人間行動に関する研究は、従来の実験室の中での行動研究から、現実の生活の中で観察される行動を分析する研究が進んでいる。特に、ネットワーク科学の手法を使い、大量なデータに基づく研究も盛んになりつつある。しかしながら、外部からの観測が困難なのが、個人は周りからどのような影響を受けるのか、と同時に、周りに対してどのような影響を与えているのかである。

本論文では、相互に影響をし合うことで現れる人間行動の全体的な性質を、人と人のつながり方（ネットワーク構造）に焦点をあてながら明らかにする。感染症などの流

行の感染爆発、新技術や文化などの普及、あるいは合意形成や相互学習のための情報交換などの問題を取り上げ、特に、人間行動を特徴づける、個人の内部属性と周りとのつながり方（ネットワーク構造）の特性から、全体の性質が定まることを示す。すなわち、人と人とのつながり方が背景にあり、個人の行動が周りに及ぼす影響とその個人が周りの人間から受ける影響の相対的な関係によって、全体に現れる普遍的な性質が特徴づけられることを明らかにする。

2. 相互作用とネットワーク構造

マスメディアや政治的キャンペーンは、大きな影響力をもっている。一方、マスメディアの力によらない、人々の口コミの持つ影響力も大きいものがある。また、うわさは人間社会にとって常に重要な役割を果たしてきており、本人の知らないうちに変な情報が流れたり、マスコミがうわさの流布に貢献したりする。

最近の研究では、うわさの伝播は、人と人のつながりが決め手になっていることが明らかになってきた。うわさの伝わり方を調査する上で、近年、インターネットが関心を集めている。インターネット上には、さまざまなうわさが流れており、インターネット上でのやり取りは、人間が直接伝え合う方法とは異なる。そして、つながり方（ネットワーク）の違いによって、うわさの伝わり方にどのような違いがあるのか、またどのような広がりを見せるのか、などが明らかになってきた。

社会生活をしていく上で、我々の行動は、意識することのない人を含めて、他の多くの人がとる行動によって影響を受けている。と同時に、自分の行動は、自分の見知らぬ人も含め、他の人に何らかの影響を与えている。

うわさの広まり方や、新しい製品の普及プロセスなどを解明していくには、相互に影響を及ぼす人間行動を探る必要があり、またそれだけでなく、人と人とのつながり方をモデル化する必要がある。

情報を交換し合う、あるいは情報を参集しながら相互に影響し合うとき、相互作用という。そして、個人の行動が相互作用する人数の大きさによって異なるとき、ネットワークからの影響を受けることになる。基本的な相互作用モデルは、以下のように分類される。

<均質なモデル>

- (1) 全数モデル：各人が他の全員と相互作用するモデルを、全数モデルという。これは全連結グラフ上での相互作用になる。
- (2) 局所モデル：一次元または二次元の格子上に配置された近傍の人と相互作用するとき、局所モデルという。

全数モデルや局所モデルでは、個人が相互作用する人数は全員同じであるので、均質なモデルという。

<非均質なモデル>

- (1) ランダムモデル：各人が、相互作用する相手を他の全員の中からある一定の割合

でランダムに選ぶとき、ランダムモデルという。このとき、個人が相互作用する人数は同じではなく、ポアソン（指数）分布などを用いる。

- (2) スモールワールドモデル：局所モデルのように、近傍の相手と相互作用することを基本とするが、その一部を他の全員の中から一定の割合でランダムに選ぶことを、スモールワールドモデルという。
- (3) スケールフリーネットワーク：相互作用する人数（次数）が個人によって大きく異なり、その次数分布がべき分布になるとき、スケールフリーネットワークという。

3. 感染爆発とネットワーク構造

人類の最大の脅威は、感染症であろう。例えば、天然痘で亡くなる人の数は、戦前には毎年100万人を超えていたという。また、第一次世界大戦時にヨーロッパを中心に起こった「スペイン風邪」と呼ばれたインフルエンザは、1年間に4,500万人の命を奪った。

現在は、SARSや鳥インフルエンザなどの感染症の爆発的な感染が懸念されている。特に、交通手段の発達に伴うグローバル化により、世界中にあつという間に広がる可能性が高まっている。

感染症の流行は、次のようにモデル化される。最初に起こることは、感染の可能性のある集団（免疫のない人たち）(S)の中の何人かが感染する。感染者は、近くの他の人に移し、その人がさらに周囲の人に移すということを繰り返すことで、感染者の数(I)が増えていく。一方で、感染者は次第に治癒していく。この場合、治癒した感染者は免疫を持っているので感染者と接触しても発病しなく、免疫を持った人の数(R)が増えていく。このようにモデル化したのを、SIRモデルという。このモデルは、病原体の持つ固有の特徴などにあまり依存しないで、人と人との相互作用を考慮した、感染症の拡散に関する基本モデルであり、拡散の速さや終息条件などが明らかになる。

特に、初期感染人口 $S(0)$ がある閾値よりも小さいとき、感染症は広まらないことが知られている。ここで、感染率を β 、治癒率を γ で表すとき、感染症の広がりを判断する上で重要なパラメータは、一人の感染者が完治するまでに感染させる人数を表す基本再生産数 R_0 である。そして、 $R_0 = S(0)\beta/\gamma$ が1より大きいとき、感染症は流行し、 R_0 が1未満のとき流行爆発は発生しなく、感染症は自然に減衰する。これは、「1人の感染者が未感染集団に入ってきたとき、伝染性をもつ期間において平均何人に感染症を引き起こすか」を表す基本再生産数は、最初1人の患者が、全員未感染の状態において、何人の二次の患者を造り出せるかという意味を持っている。

個人の移動空間の広まりや個人差を考慮した場合、人と人とのつながりを考慮した、社会ネットワーク上での感染症の広まりに関する研究が注目されている。各人が接触する人数(k)に個人差があり、それを次数分布として分布関数 $p(k)$ で表すと、基本再生産数 R_0 は次式で求まる。

$$R_0 = T(\langle k^2 \rangle / \langle k \rangle - 1) \quad (3.1)$$

ここで、 $\langle k \rangle$ と $\langle k^2 \rangle$ は次数分布関数の平均と自乗平均を表す。ここで、 $R_0 = 1$ の条件を満たすことになる以下の値が閾値となる。

$$T_c = \langle k \rangle / (\langle k^2 \rangle - \langle k \rangle) \quad (3.2)$$

人と人とのつながりを表すのが、ネットワークの次数分布である。その分布関数を $p(k)$ で表す時、(3.2) 式によって閾値 T_c が求まり、その値が1以上のとき、感染爆発が起こり、1未満ならば、感染は広まることなく局所的なものになる。

ここで、平均次数 $\langle k \rangle$ が等しい次数分布を持つ、べき分布のスケールフリーネットワーク及びランダムネットワークの閾値を比較する。べき分布の分散は大きな値になることから、(3.2) 式によって求まる閾値は小さな値になる。一方で、ランダムネットワークの分散は小さいことから、スケールフリーネットワークの閾値はランダムネットワークの閾値よりも小さく、感染爆発が起こり易いことがわかる。

感染症の爆発的な流行を未然に防ぐには、基本再生産数をできるだけ小さく抑えるだけでなく、拡散速度との戦いでもある。それによって、新しい感染症が発見されたときの有効なワクチンの開発体制や、あらかじめ用意しておくべきワクチンの量、そして重点的に備蓄しておくべき地域の決定などの政策決定に重要になる。

ここで、 N 人存在する社会における人と人とのつながりを隣接行列として表す。個人 i と個人 j とつながりがあるときは1、そうでないときは0の値をとる a_{ij} を要素として持つ $N \times N$ の行列 $A = [a_{ij}]$ を隣接行列という。感染症の広まりの速さを予測するうえで、隣接行列 A が持つ2番目に大きい固有値 λ_2 (第2固有値) が重要である。そして、人間(ノード)の感染率を β 、そのノードの治癒率を γ とするとき、

$$\begin{aligned} (1) \quad & \beta/\gamma < 1/\lambda_2 \text{ ならば、感染症は広まらない} \\ (2) \quad & \beta/\gamma > 1/\lambda_2 \text{ ならば、感染症は広まる} \end{aligned} \quad (3.3)$$

ことが知られている。

例えば、平均次数が $\langle k \rangle$ のランダムネットワークの第2固有値は、 $\lambda_2 = \langle k \rangle$ で、スター構造のネットワークは、 $\lambda_2 = \sqrt{d}$ (d は最大次数)、そして、スケールフリーネットワークはさらに大きい第2固有値を持つことが知られている。このことから、人とのつながりを表す次数が個人によって大きく異なるネットワークの下では、感染爆発が起こり易いことがわかる。

(3.3) 式のように、感染率と治癒率の基本感染パラメータと人々のつながり方を表す隣接行列の固有値の関係によって、感染爆発の条件が求まることから、感染(コンピュータウイルスや噂などを含めて)の広まりを予測した、効果的な予防策を練る上で重要な知見である。そして、感染爆発を抑止するには、隣接行列の固有値を調べながら人々のつながり方を適切に制御することである、という知見が得られる。

基本都市間の感染モデルと、都市あるいは地域内感染モデルがある。前者は、交通網、特に航空路の発達により、感染の爆発が世界レベルで、しかも短期間での拡散を予測するためのモデルである。例えば、14世紀に欧州で猛威をふるったペストは、欧州内に広まるのに10年以上の年数を要したが、昨今のSARSは数カ月で、世界中に拡散した。この背景にあるのが、世界的に航空網が整備されたことで、世界中の人々がつながりを持つ可能性が増えたことである。

一方で、都市や地域内部での感染モデルは、人々の行動パターンに着目した研究である。例えば、子どもたちは学校などで、成人は通勤経路、職場、あるいは商業地域で接触する可能性が他の年代の人たちよりも高い。そのことで、どのような広まりを見せるのかが明らかになってきた。特に、基本再生産数 R_0 が同じでも、地域内での人々の行ききするパターンが異なる二つの地域では、感染の広まりが大きく異なる。従って、人々のつながり方が感染症の広まりを左右することがわかってきた。

4. 技術や文化の普及とネットワーク構造の関係

毎日の生活で、うわさに接する機会は誰にも多い。自分が聞いたうわさは誰かに伝え、そして、1人ひとりが伝道者になることで、うわさは広まっていくことになる。私達にとって、うわさを伝え合うことは日常的な活動の一部になっているが、なぜ、私達はうわさを伝え合うのかの疑問は、まだ十分には明らかになっていない。

これまでの研究では、人々を取り巻く不安な状況がうわさを伝えさせる要因とされてきた。我々は、自分がどのような状況にいるのか大まかには把握しているが、新たな情報を知ると、今まで把握していたことが変化して説明できないような状況に置かれるために、不安になってしまい、そこで、他人達と情報を交換して新しい状況を把握しようとする。このとき得た情報は、同じように不安な状態に置かれた人から次の人へまた伝わっていく。すなわち、人々が共通して不安な状態にあり、それらを軽減できる確実な情報が得られないときは、できるだけ適切な情報を知ろうとして、うわさが伝えられていくが、不安が解消すると再び状況を把握することができるために、うわさは自然に消滅していく。

情報の伝搬や爆発的な流行など、社会現象を分析する上で根底になるのが、個人の選択行為である。

人と人のつながりが重要になるのは、あることを選択した結果もたらされる効用や利得が、同じ行動を選択する他の人の人数などに影響されるときである。例えば、新しい技術が開発されたとしても、初期段階では、それを採用している人は少ないことから、たとえその技術が優れているからといって社会に普及するとは限らない。新技術の普及や文化などの伝搬の問題を扱うとき、根底となる個人間の相互作用は、次のような二人によるゲームとして定式化することができ、また相互依存的な状態での各人の選択肢を戦略とよび、各人は、次の二つの戦略を選択肢として持っている。

戦略1：新しい手法（新方式）を採用する。

戦略2：今までの手法（旧方式）を継続する。

ここで、新方式を採用した場合のコストを c 、利得を a とする。新方式を採用するにはコスト c が伴い、お互いに新方式を採用するときは、コストを上回る利得 a を得る。ただし、新方式を採用するのが自分一人のとき、新方式を採用することの利便性は得られることなく、コストだけを負担することになる。旧方式を継続するときは、現状維持として、利得ゼロとする。

利得とコストの間には、 $a > c > a/2$ の関係があり、お互いに新方式を採用すればコストを上回るメリットをもたらす。しかし、新方式の採用にはコストを伴うため、リスクを伴う、という構造になっている。このとき、新方式をパレート優越戦略、旧方式をリスク優越戦略という。新しい方式の方が優れているからといっても、他の人がそれを採用しないのであれば、新方式を採用することのメリットは無い。そして、多くの人が旧方式を継続するのであれば、旧方式を継続することが良いため、新方式は普及しないことになる。このことから、新方式を普及させるには、初期の段階で採用する人の割合を、ある一定以上にすることが必要であり、その臨界値をチッピングポイントという。

たとえ初期段階で採用する人が極めて少なくても、新方式を採用するメリットが十分に高ければ普及し、逆にコストが高ければ、普及しないことになる。このように、新方式採用に伴うメリットとコストの比によって、普及するかどうかが決まる。そして、利得に対するコストの相対比 $\theta = c/a$ がある閾値よりも小さいとき、新方式は普及し、それ以上の時は普及しないことになる閾値が存在する。

Lopez-Pintado は、格子モデルの下での閾値 (θ_H)、べき分布の下での閾値 (θ_{SF})、そして指数分布のネットワークの下での閾値を (θ_E) 比較し、

$$\theta_H < \theta_{SF} < \theta_E \quad (4.1)$$

の関係にあることを示した。このことから、ネットワーク外部性が存在し、他の人も同じ選択をすることでメリットが高まるような新技術や商品が最も普及し易いのは、人と人のつながり方がスケールフリーネットワークではなく、指数分布をもつネットワークであることがわかる。

感染症などの拡散は、前節で示したように、スケールフリーネットワークのように、ハブノードが存在するネットワークが最も早く、その拡散を防止する上でも、ハブノードが重要になる。一方で、技術や文化の普及など、その普及において個人レベルで費用対効果に基づく意思決定が伴う場合には、ハブノードが存在するようなネットワークや格子ネットワークのような均質なモデルより、多くの人から影響を受けると同時に、一人ひとりが同じような伝道者となって多くの人に影響を及ぼすノードが多数存在するような、指数分布のようなネットワーク構造のときである。

個人の行動が周りに及ぼす影響と、個人が周りから受ける影響の相対的な関係がより大事になるのは、次節で取り上げるコンセンサス問題である。

5. コンセンサス問題

複数のエージェントやシステムが分散して保有している情報を、どのようにして全体で共有できるかの問題は、多くの分野にも共通する本質的な課題である。並列分散系では、この問題を放送 (broadcast) 方式や、ゴシップ (gossip) 方式に基づく“情報散布”として扱う。そして、情報通信技術の進歩に伴って、この問題の研究は幅広い展開を見せている。また、並列分散システムを構成する各サブシステムの状態変数が同じ値に収束するように制御する問題は、“コンセンサス問題”として知られている。

情報散布問題は、コンピュータ科学、コンセンサス問題は制御理論の分野でそれぞれ別々に研究されているが、これらは自律分散系の制御問題として共通点が多い。特に、サブシステム (エージェント) 同士のつながり方 (系の根底にあるネットワーク構造) によって、系の安定性や収束性の善し悪しが決まるといふ共通の性質を持っている。

コンセンサス問題は、次のように定式化される。エージェント i に関連するシステムダイナミクスを

$$\dot{x}_i = u_i \quad (5.1)$$

と記述する。(5.1) 式の \dot{x}_i をエージェント i の状態量 (例えば位置情報)、 u_i をエージェント i に加えられる入力とする。N体のエージェントで構成される系 (マルチエージェントシステム) において、それぞれの状態変数が

$$x_1 = x_2 = \dots = x_N \quad (5.2)$$

と同じ値に収束するとき、エージェント間でコンセンサスが達成されるという。具体的には、すべてのエージェントの状態変数の値が、それぞれの初期値の平均値

$$a = \frac{1}{n} \sum_i x_i(0) \quad (5.3)$$

に収束するかどうかを調べる問題である。ここで、次の入力をもつダイナミクスを考える。

$$\dot{x}_i = \sum_{j \in N_i} a_{ij} (x_j(t) - x_i(t)) \quad (5.4)$$

そして、エージェント i の近傍を N_i で表す。また、 a_{ij} は、エージェント i がエージェント j と連結されているときは1、連結されていないときは0の値をとる。これらの要素をもつ隣接行列を、 $N \times N$ の行列 $A = [a_{ij}]$ で表し、この隣接行列が表すネットワークをグラフ G という。

コンセンサス問題は、個々のエージェントが (5.4) 式の下、近傍のエージェントとの間で情報交換を繰り返すとき、(5.2) 式を満たす収束性を調べる問題である。

Olfati-Saberらは、エージェント間に形成される動的なネットワークに着目して、コンセンサス問題をスペクトル分析の観点から明らかにした。式 (5.4) のダイナミクスをまとめて、

$$\dot{x} = -Lx \quad (5.5)$$

と表す。ここで、 $x = \{x_1, \dots, x_N\}$ である。式 (5.4) において、エージェント間のつながり方をグラフ G で表したが、そのラプラシアン行列 L を、次数行列 D と、隣接行列 A を用いて次式で定義する。

$$L \equiv D - A \quad (5.6)$$

グラフ G が連結グラフのとき、すなわち系を構成するエージェント同士の連結性が確保されるとき、(5.5) 式で定義されるダイナミクスは安定性の条件を満たし、すべてのエージェントの状態変数は同じ値に収束する。また、式 (5.5) のダイナミクスの収束性は、ラプラシアン行列 L の固有値からわかる。すなわち、2 番目に大きい固有値と最も大きい固有値の比、 λ_N / λ_2 (代数的連結係数という) によって収束速度を左右し、代数的連結係数が小さいほど、式 (5.5) のダイナミクスの収束性は高いことが知られている。

(5.4) 式において、要素 a_{ij} を実数値としてエージェント i と j の連結の強さを表す問題に拡張することもできる。そのとき、各ノード i (エージェント) に入るフローと、そこから出ていくフローの大きさが次式を満たす時、

$$\sum_{i \neq j}^N a_{ij} = \sum_{j \neq i}^N a_{ji} \quad (5.7)$$

グラフ G は“バランスがとれたネットワーク”という。そして、 G が“バランスがとれたネットワーク”のとき、(5.4) 式の下で近傍のエージェントとの間で情報交換を繰り返す場合には、(5.2) 式を満たす収束性が保証される。

技術や文化の普及などでは、周りから強く影響を受けると同時に、一人ひとりが同じような伝道者となって多くの人に影響を及ぼすノードが多数存在し、そして指数分布のようなネットワーク構造のとき、効率的に広まる。一方で、合意形成や相互学習の問題では、一人一人が等しく周りから影響を受けると同時に、同じような周りに影響を及ぼすようなネットワーク構造が効率的である。

このような性質をもつネットワーク構造では、格子ネットワークが、各ノードは同じ次数をもつバランスのとれたグラフであるため、コンセンサス形成上、優れている。また、格子モデルを根底にして形成されるスモールワールドネットワークは、いくつかのノード間にショートカットを持つ隣り合った近傍のノード同士の規則的なリンク以外に、遠く隔てたノードとランダムな方法でリンクし、いくつかのショートカットを持つ。このとき、局所ネットワークの代数的連結係数と比較すると、スモールワールドネットワークの代数的連結係数はさらに小さくなり、コンセンサス問題のアルゴリズムの収束が速いことが知られている。

6. まとめ

感染症などの流行拡散、新技術や文化などの普及、そして相互学習のための情報交換などの問題を取り上げ、各ノードの内部属性とネットワーク構造との関係から、現れる全体的な性質を特徴づけられることを明らかにした。いずれのケースも、全体の性質が大きく異なることになる相転移点が存在し、そのような相転移は、人間の内部属性と他の人とのつながり方（ネットワーク構造）の特性から求まる。そして、感染症などは、感染したノード（人間）から未感染ノードへ伝搬する一方的な流れである。このため、多くのノードとつながっているハブノードが存在するネットワークの下で最も感染が広まる。

一方で、新技術や文化などの普及は、各ノードでは個人の損得勘定に基づくと意思決定が伴う。このため、他から影響を受けやすく、と同時に、他のノードに影響を与えやすいノードが多く存在する指数分布のようなネットワークが伝搬効率が高い。

一方で、合意形成の問題では、ハブノードや中心となるノードが存在しなく、相互に等しく情報を交換することになる格子ネットワークが効率的であることを示した。

謝辞

本論文の作成にあたり、ご指導を頂きました防衛大学校情報工学科の生天目章教授に、心より感謝申し上げます。

【参考文献】

- (1) Ball. P, "The Physical Modeling of Human Social Systems", Complexes, Vol.1, pp.190-206 (2003) .
- (2) Ball. P, "Critical Mass", Farrar, Straus and Giroux (2004) .
- (3) Binmore.k, "Natural Justice", Oxford Univ. Press (2005) .
- (4) Buchanam. M, "The Social Atoms", Reed Elsevier Inc (2007) .
- (5) Lopetz-Pintado. D, "Contagion and coordination in random networks", Int. Journal of Game Theory, Vol. 34, pp.371-382 (2006) .
- (6) Namatame.A, "Adaptation and Evolution in Collective Systems", World Scientific (2006) .
- (7) 生天目章, 「ゲーム理論と進化ダイナミクス」, 森北出版 (2003) .
- (8) Meyers. Pourbohloul, Newman.M.E.J, "Network theory and SARS: predicting outbreak diversity", Journal of Theoretical Biology, 232, pp.71-81 (2005) .
- (9) Colizza. Barrat, A, Barthelemy.M, and Vespignanic, "The role of the airline transportation network in the prediction and predictability of global epidemics ", PNAS, vol.103, pp.2015-2020 (2006) .

(平成20年10月31日受付、平成20年10月31日受理)

