

オンライン資料 1

barabasi.com/NetworkScienceBook

このウェブサイトは教科書そのものや、各章で触れられている動画、ソフトウェア、インタラクティブツールへのアクセスを提供している。他にもネットワーク科学の授業で用いたスライドや本の中のデータセットも提供している。

クティブツール、データセットやデータソース—に紐づけられたたくさんのオンライン教材が示されている。これらの教材は <http://barabasi.com/NetworkScienceBook> のウェブサイトからアクセス可能である。

このウェブサイトは私がネットワーク科学の授業で用いた PowerPoint のスライドも含んでおり、このスライドの内容が、この教科書に反映されている。ネットワークを教える人であれば誰でもこのスライドを使ってもらって構わない。また、各自の授業に合わせて改変いただいて結構である。教育に関する限り、このスライドを用いることについて私に問い合わせる必要はない。

ネットワーク科学の始まりが実証的であることから、この教科書も実際のネットワークの分析にかなり重点を置いている。それゆえ、いろいろなネットワーク上の特徴の検証のために、研究上よく用いられている 10 個のネットワークを取り上げている。これら 10 個は、ネットワーク科学が取り上げる、社会、生物、技術、情報システム上のネットワークであり、その幅広さを示す意味でも選ばれた。オンライン抄録は、この教科書を通じて取り上げられるこれら 10 個のデータを提示している。これらのデータに、ネットワーク科学のさまざまなツールが適用されていく。

最後に、英語と異なる言語で教えている人たちのために、ウェブサイトの方も翻訳プロジェクトが進んでいる。

*** 日本語サイトについて説明を追加する ***

ネットワーク科学の授業

私はこれまでに二つの異なるやり方でネットワーク科学を教えてきた。一つめは物理学、計算機科学と工学を学ぶ学部と大学院生を対象とした一学期分の授業である。もう一つは三週間で二単位分の授業で、経済学や社会科学を学ぶ学生向けの授業である。このテキストはこれらの教育経験に基づいて作られている。一学期の授業の方では、教科書のすべてを対象とし、発展的課題に入っている証明や逸話なども扱った。短い方の授業では、主なセクションだけを扱い、発展的課題と次数相関を扱った第 7 章を省いた。

両方の授業で重要なのは、次に述べる宿題や研究プロジェクトである。

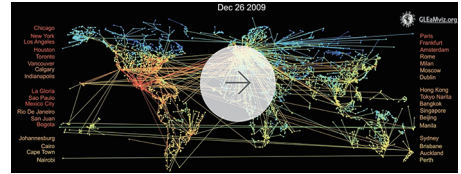
割を得ることになった。それは、感染経路や時間発展がピークに達する数か月前に正確に予測された最初の感染症であった（オンライン資料 1.1）[33]。これが可能となったのは、ウイルスの拡散における移動ネットワークの役割がわかってきたという本質的な進歩のおかげである。

2000 年以前では、感染についてのモデル化の主流は、社会物理的に同じ区分の中にいる人はその他の人すべてを感染させることができると仮定する、区分化に基づくモデルであった。ネットワークを基本とするモデル化の枠組みが出現することによって、根本的な変化が生じ、新しいレベルの予測が可能となった。今日では、感染症予測は、ネットワーク科学における最も活発な応用分野の一つで、インフルエンザの蔓延やエボラ熱の閉じ込めを予見している [33, 34]。感染症予測は、本書で取り扱うような生物学的、デジタル的、社会的なウイルス（ミーム）の蔓延を予測するモデルを可能とするいくつかの基本的な結果の源泉でもある。

これらの進歩のインパクトは、感染症学を超えた事例からも知ることができる。2010 年の 1 月に、携帯電話を通じたウイルス蔓延の発生の必要条件を、ネットワーク科学のツールは予測した [35]。最初の携帯電話感染の大流行は、中国において 2010 年の秋に始まり、ほぼ予測通りのシナリオに従って、毎日 30 万以上の携帯電話を感染させた。

1.5.5 神経科学：脳のマッピング

ヒトの脳は、数千億のニューロンが互いに結合しており、ネットワーク科学の観点から最も解明が進んでいないネットワークの一つである。その理由は単純である。ニューロンがつながる様子を示したマップがないためである。完全にニューロンのマップが手に入るのは、わずか 302 ニューロンからなる **C エレガンス線虫**についてだけである。哺乳類の脳についての詳細なマップは脳科学に革命を引き起こし、多くの神経的な脳の病気を理解し治療することが可能になるであろう。脳の研究はネットワーク科学の最も活発な研究分野の一つとなっている [36]。そのようなマップができれば革新的な影響が起こり得るという可能性に促されて、米国国立衛生研究所では、哺乳類の脳の正確なニューロンレベルのマップの作製を目指して 2010 年にコネクトームプロジェクトを開始した



オンライン資料 1.1

H1N1 大流行の予測

2009 年の H1N1 大流行の実時間予測の最初の成功例である [33]。世界中の交通機関ネットワークの構造と動態を説明するデータに依拠したプロジェクトは、通常期待されるインフルエンザの 1 月から 2 月の流行時期とは異なり、H1N1 が 2009 年 10 月にピークを迎えると予測した。これは、ワクチン接種の時期が 2009 年 11 月では遅すぎて、結局のところ大流行の結果にほとんど影響を与えないということを示した。このプロジェクトの成功は、人類にとって特に重要な領域における進歩を促進するためにネットワーク科学が有効であることを示した。

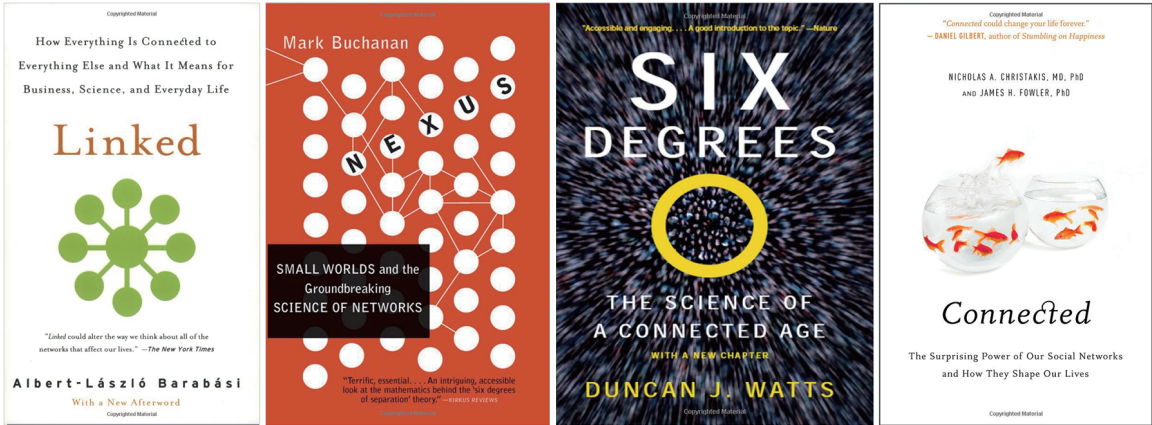
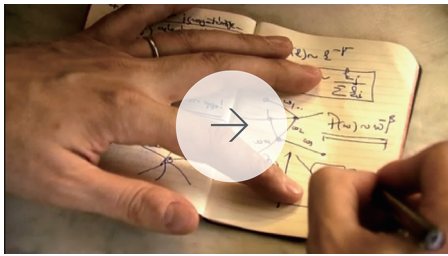


図 1.11 広範なインパクト

これら広く読まれている 4 冊の書籍は 20 か国語以上に翻訳され、ネットワーク科学を一般の読者向けに解説している [50, 51, 52, 53]。



オンライン資料 1.2 コネクティッド

Annamaria Talas 監督の、受賞ドキュメンタリー映画コネクティッドの予告編。これはネットワーク科学の導入となっており、俳優ケビン・ペーコンと数人のよく知られたネットワーク科学者を主役としている。

け入れられ大衆による認知がなされる 1920 年代までほぼ使われることはなかった。

本図では、これらの用語とネットワークという単語の使用の比較により、ネットワークの使用は 1980 年代に急激に増加し、**進化**や**量子**をはるかに超えていることを示している。ネットワークという用語にはさまざまな使用方法があるが、劇的な増加はネットワークの社会的認知の増加を反映している。

進化論、量子力学、ネットワークによってもたらされた進歩の間に共通することがある。それらは知的本質や知識体系を伴う重要な科学分野であるだけでなく、理論的枠組みを提供していることにある。実際のところ、遺伝学の現在の進歩は進化論の上に築かれており、量子力学は化学から電子工学までの現代科学の広い範囲におよぶ進歩の基盤を提供している。同様に、ネットワーク科学は、社会ネットワークから薬品設計までの広い範囲の科学的問題について、新規なツールや視点を提供することを**可能にする基盤**である。

このように、ネットワークは科学と社会の両方に例外的に大きいインパクトをもたらしたのであるから、我々はネットワークを理解して定量化するツールに習熟しなければならない。本書の以降の章は、この価値ある主題に充てられている。

第 2 章

グラフ理論

2.1 ケーニヒスベルクの橋	49	2.8 経路と距離	62
2.2 ネットワークとグラフ	50	2.9 連結性	67
2.3 次数, 平均次数, 次数分布	53	2.10 クラスタ係数	70
2.4 隣接行列	57	2.11 まとめ	71
2.5 現実のネットワークはスパース (疎) である	59	2.12 演習	74
2.6 重み付きネットワーク	60	2.13 発展的話題 2.A 大域的クラスタ係数	76
2.7 2部グラフ	60		

2.1 ケーニヒスベルクの橋

いつどこで誕生したかを歴史の中に正確に位置づけられる研究分野はほとんどない。しかし、ネットワーク科学の数学的な土台であるグラフ理論ではそれが可能である。その起源は 1736 年の東プロシアの首都で、当時商業で栄えたケーニヒスベルクに遡ることができる。忙しなく行きかう貨物船団の交易のおかげで、市の行政当局は、町を取り囲んでいるプレーゲル川に七つの橋を架けることができた。このうちの五つはプレーゲル川の 2 本の支流の間にある優美な島、クナイフォフと町の本土を結んでおり、2 本は川の 2 本の支流をまたいでいた (オンライン資料 2.1, 図 2.1)。この独特の構図から、同じ橋を二回渡ることなく七つの橋をすべて渡ることができるのか、という当時における難問が生み出された。多くの試みにも関わらず、誰もそのような経路を見つけることはできなかった。そのような経路は存在しないとの厳密な数学的証明を、スイス生まれの数学者、Leonard Euler が提示す



オンライン資料 2.1

ケーニヒスベルクの橋

ケーニヒスベルクの問題と Euler の解決策を紹介した短いビデオを見てみよう。

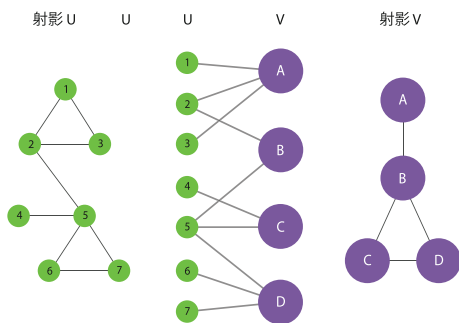


図 2.9 2部グラフ

U と V と二つのノードのセットをもつ 2 部グラフがある。U セットのノードは V セットのノードのみに直接結びつく。したがって U と U, V と V, などの結びつきは存在しない。上の図は 2 部グラフから作ることのできる二つの射影を示している。射影 U は、2 部グラフにおいて二つの U ノードが同じ V ノードにつながっているときに、その二つの U ノードをリンクすることによって得られる。射影 V は、2 部グラフにおいて二つの V ノードが同じ U ノードにつながっているときに、その二つの V ノードをリンクすることによって得られる。



オンライン資料 2.2

疾病ネットワーク

高解像度の疾病ネットワークをダウンロードして [55], *New York Times* によって作られたオンライン・インターフェースを用いてそのネットワークを調べてみよう。

ノードの間にリンクを張ったネットワークである。

あらゆる 2 部グラフについて、図 2.9 に示すように、二つの射影が可能である。最初の射影は、同一の V ノードにリンクがある二つの U ノードをつないだネットワークである。第二の射影は同一の U ノードにリンクがある二つの V ノードをつないだネットワークである。

ネットワーク理論では、さまざまな 2 部グラフを取り扱う。よく知られた例は、映画 (U) と役者 (V) の二つの集合から構成される、ハリウッドの役者の共演である。役者が映画に出演すると、それらの映画と役者の間にリンクが張られる。この 2 部グラフの一つの射影は、同じ映画に出演した役者がリンクされた**役者の共演ネットワーク**である。これが、表 2.1 に示されたネットワークである。もう一つの射影は、少なくとも一人の同じ役者が出演した二つの映画がリンクされた**映画ネットワーク**である。

医学では、2 部グラフの別の事例が知られている。ある病気の原因となる突然変異が生じ得る遺伝子をその病気とリンクした**疾病ネットワーク**を、図 2.10 に示す。

最後に、図 2.11 に示すように、**多部グラフ**、たとえば、料理、材料、およびそれに含まれる成分で構成される 3 部グラフを定義することもできる。

2.8 経路と距離

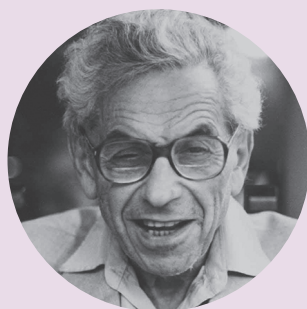
物理的な距離は、物理系の連結成分の間に働く相互作用を決めるのに主要な役割をはたす。たとえば、結晶中の二つの原子の間の距離、宇宙のなかの二つの銀河の間の距離は、これらの間に働く力を決める。

ネットワークにおいては、距離は少し難しい概念である。二つのウェブページ間の距離や、互いに知らない二人の個人間の距離とは、一体何を意味するのであろうか？ここでは物理的な距離は重要でない。二つのウェブページは、地球の反対側のコンピューターに置かれて、互いにリンクされているかもしれない。同時に、同じ建物に住む二人の個人が、互いに顔見知りでないこともありうる。

ネットワークにおいては、物理的な距離は**経路長**によって置き換えられる。**経路**とは、ネットワークのリンクに沿ってたどることができる道筋である。**経路長**とは、その経路に含ま

Box 3.2 ランダム・ネットワーク：簡単な歴史

(a)



(b)



ロシアからアメリカに移住した Anatol Rapoport (1911-2007) が、ランダム・ネットワークの研究をした最初の人物である。Rapoport はピアノの演奏者として成功するためには金持ちのパトロンが必要であることに気づき、その道を諦めた後に、数学に関心を持つようになった。数学者と生物学者がほとんどお互いに言葉を交わすことがなかった頃であったが、彼は数理生物学に集中することにした。Ray Solomonoff と 1951 年に書いた論文 [80] において、Rapoport はネットワークの平均次数を増やしていくと、ばらばらのノードが突然、巨大成分を含むグラフへと変化することを示した。

ランダム・ネットワーク研究は Pál Erdős と Alfréd Rényi の重要な仕事によって一つの高みに達した (図 3.2)。1959 年から 1968 年の間に出版された 8 本の論文 [3, 72, 73, 74, 75, 76, 77, 78] で、彼らは確率論と組み合わせ論をグラフ理論と融合し、数学の新たな領域としてランダム・グラフ理論を確立した。

ランダム・ネットワークの定義に関わるモデルは、Erdős と Rényi とは別に、Edgar Nelson Gilbert (1923-2013) によって導入されたが [79]、それは Erdős らが同じ課題について最初の論文を出版したのと同じ年のことであった。しかしながら、Erdős と Rényi の仕事のインパクトがあまりに圧倒的であったために、彼らがランダム・グラフ理論の正式な創始者だと考えられている。

「数学者はコーヒーを定理に変換する装置である」Alfréd Rényi (よく Erdős の言葉と考えられているがそれは間違いである)

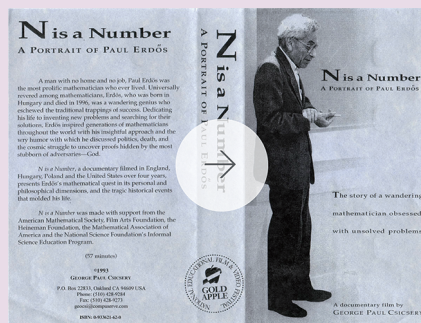
図 3.2 経ランダム・グラフ理論の創始者

(a) Pál Erdős (1913-1996)

ハンガリー出身で、類まれなる科学的業績と変わり者であることで知られる数学者。紛れもなく Erdős は数学史上で最も多くの論文を書いた数学者であり、500 人以上の数学者と共著論文を書き、エルデシュ数を生み出すものになった。彼の伝説的な人柄と偉大なる業績は伝記 [81, 82] やドキュメンタリー [83] に記されている (オンライン資料 3.1)。

(b) Alfréd Rényi (1921-1970)

ハンガリー出身で、組み合わせ論やグラフ理論、数論の発展に非常に大きい貢献をした数学者。彼の業績は数学にとどまらない：レニイ・エントロピーはカオス理論で広く使われている。また彼がその発展の一翼を担ったランダム・ネットワーク理論はネットワーク科学のまさに中心である。ハンガリーにおける数学研究の中心である Alfréd Rényi Institute of Mathematics にその名が残されている。



オンライン資料 3.1

「N は数」：Pál Erdős の人物像

1993 年に George Paul Csicsery の監督のもと制作された伝記ドキュメンタリー。Erdős の生涯や彼の科学的業績を垣間見ることができる [83]。

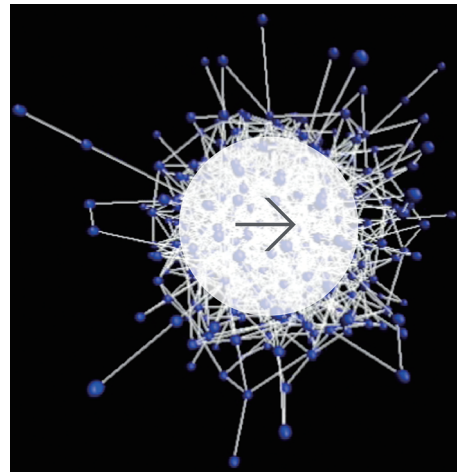
- 現実のネットワークにおける次数分布の幅はランダム・ネットワークモデルが予測するよりはるかに広い。この違いは標準偏差 σ_k (図 3.4) で捉えることができる。インターネットが仮にランダムであったとすれば、 $\sigma = 2.52$ と予想できる。測定値によると $\sigma_{\text{Internet}} = 14.14$ であり、ランダム・ネットワークモデルの予測よりもはるかに高い。この違いは図 3.6 に示されるネットワークに限ったことではない。表 2.1 に掲載されているすべてのネットワークにこの性質がみられる。

まとめると、ランダム・ネットワークモデルは現実のネットワークの次数分布を捉えることができないことが現実のデータとの比較からわかる。ランダム・ネットワークにおいてはほとんどのノードがよく似た次数をもち、ハブは現れない。一方で、現実のネットワークではかなりの数の次数の高いノードがあり、また、次数にも大きい幅が見られる。これらについては、第 4 章でさらに議論する。

3.6 ランダム・ネットワークの成長

この章の初めに取り上げたカクテルパーティ問題は、刻一刻と変化する過程を含んでいる。最初は孤立した N 人の招待客 (N 個のノード) である。その招待客がランダムに相手を選び知り合いとなることによって、そのノード間に徐々にリンクが張られていく。これは p が徐々に増加していくことに対応し、その p の増加はネットワークのトポロジーに驚くべき結果をもたらす (オンライン資料 3.2)。この過程を定量的に取り扱うため、まずネットワーク内の最大連結成分の大きさ N_G が平均次数 $\langle k \rangle$ によってどのように変化するかを調べよう。次に挙げる極端な二つの場合については容易に理解できる。

- $p = 0$ の場合：このとき平均次数 $\langle k \rangle = 0$ であるから、すべてのノードは孤立している。したがって、 $N_G = 1$ であり、 N が大きくなるにつれて $N_G/N \rightarrow 0$ である。
- $p = 1$ の場合：このとき平均次数 $\langle k \rangle = N - 1$ であるから、ネットワークは完全グラフであり、すべてのノードが一つの成分に属している。したがって、 $N_G = N$ であり $N_G/N = 1$ である。



オンライン資料 3.2

ランダム・ネットワークの成長

リンクを張る確率 p が大きくなるにつれてどのようにランダム・ネットワークの構造が変化するかを示すビデオ。小さい p のときはネットワーク内に巨大成分は存在しないが、 p が臨界値に達するや突然にそれが現れる様子が鮮明にわかる。

第 4 章

スケールフリーの性質

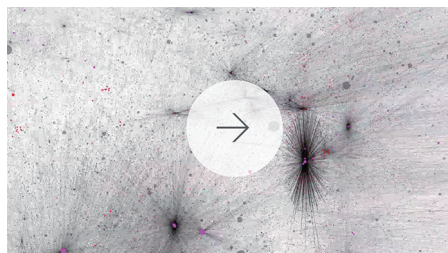
4.1 はじめに	123	には	148
4.2 べき則とスケールフリー・ネットワーク	125	4.9 まとめ	154
4.3 ハブ	129	4.10 演習	156
4.4 スケールフリーの意味するところ	132	4.11 発展的話題 4.A べき則	157
4.5 普遍性	137	4.12 発展的話題 4.B べき分布をプロットする	162
4.6 超スモールワールド性	142	4.13 発展的話題 4.C 次数分布のべき指数を推定する	165
4.7 べき指数の役割	145		
4.8 任意の次数分布をもつネットワークを作る			

4.1 はじめに

WWW は、文書をノードとし、ユニフォーム・リソース・ロケータ (URL) をリンクとするネットワークで、クリッカー一つで WWW 上の文書を行き交うことが可能である。WWW 上には推定 1 兆以上 ($N \approx 10^{12}$) の文書があり、その数は人間の脳のニューロン数 ($N \approx 10^{11}$) をも超え、人類が構築した最大のネットワークである。

我々の普段の生活における WWW の重要性は極めて大きいものである。それと同じように、ネットワーク理論の発展において WWW が果たした役割も非常に大きい。WWW によって多くのネットワークの基礎的な特性の発見が促進され、WWW はネットワークを研究する標準的なテスト基盤となった。

クローラーというソフトウェアを用いると WWW のネットワーク図をマッピングすることができる。クローラーはあるウェブ文書からスタートして、その文書のリンク (URL) を識



オンライン資料 4.1

WWW にズームイン

スケールフリー性の発見につながった WWW からの標本にズームインしていくオンラインビデオを見てみよう [9]。これは表 2.1 に含まれ、図 4.1 にも示されているネットワークで、その特徴量がこの本を通してテストされることになる。

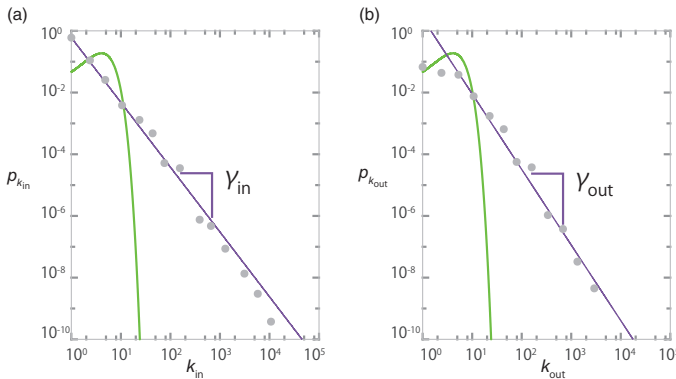


図 4.2 WWW の度数分布

Albert らの 1999 年の研究 [9] でマッピングされた WWW 標本における (a) 入次数分布と (b) 出次数分布。度数分布は、べき則があれば直線となる両対数 (log-log) 軸でプロットされている (オンライン資料 4.2)。プロット中の各点は実測値、直線は入次数については指数 $\gamma_{in} = 2.1$ 、出次数については指数 $\gamma_{out} = 2.45$ のべき分布に従う直線を表わす。図中の緑の線は WWW 標本の平均次数 $\langle k_{in} \rangle = \langle k_{out} \rangle = 4.60$ をもつポアソン分布を示す。

ネットワークとの不可解な違いがあることがわかる。ランダム・ネットワークにおいては、ハブと呼ばれる非常に大きい数のリンクをもつノードは事実上存在しないのである。一方、図 4.1 では、少ない数のリンクをもつ多数のノードが、比較的少数のハブと共存している。

本章では、このようなハブが WWW 特有のものではなく、ほとんどの現実のネットワークにおいて観察されることを示していこう。このことは、スケールフリー性と呼ばれるネットワークの組織化原理の特徴の一つである。そこで、我々は実際に観察できるネットワークの度数分布を検証することにより、スケールフリー・ネットワークの存在を明らかにする。本章で紹介している分析と実証結果は、この本で解説しているモデルの基礎となるものである。実際に、コミュニティから拡散過程までいかなるネットワークの興味深い性質であれ、ネットワークの度数分布の観点から考察しなければならない。

4.2 べき則とスケールフリー・ネットワーク

もし WWW がランダム・ネットワークであるとする、WWW の文書の次数はポアソン分布に従うはずである。しかし、図 4.2 が示すように、ポアソン分布は WWW の度数分布にうまく適合しないことがわかる。一方、両対数スケールで見ると直線を描いており、これは WWW の度数分布は

$$p_k \sim k^{-\gamma}. \quad (4.1)$$

で精度よく近似できることを示している。方程式 (4.1) はべき分布と呼ばれ、指数 γ はべき指数である (Box 4.1)。方程式 (4.1) の対数をとると

オンライン資料 4.2

Fitting a Power Law

※翻訳もれのため、原著の英文を入力します。

The algorithmic tools to perform the fitting procedure described in this section are available at <http://tuvalu.santafe.edu/~aaronc/powerlaws/>.

第 5 章

バラバシ・アルバート・モデル

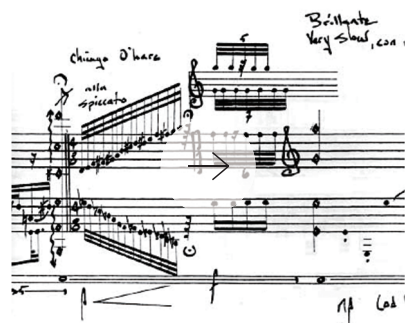
5.1 導入	177	5.9 優先的選択の起源	195
5.2 成長と優先的選択	178	5.10 ネットワーク直径とクラスター係数	202
5.3 バラバシ・アルバート・モデル	180	5.11 まとめ	203
5.4 次数ダイナミクス	184	5.12 演習	205
5.5 次数分布	186	5.13 発展的话题 5.A 次数分布の導出	207
5.6 成長または優先的選択がない場合	188	5.14 発展的话题 5.B 非線形優先的選択	210
5.7 優先的選択を測る	190	5.15 発展的话题 5.C クラスター係数	213
5.8 非線形優先的選択	192		

5.1 導入

ハブは、ランダム・ネットワークとスケールフリー・ネットワークの最も顕著な違いを表している。WWW 上では、それらは google.com または facebook.com のような例外的に数多くのリンクをもつウェブサイトであり、代謝ネットワークにおいては、それらは ATP または ADP のような例外的に数多くの化学反応が関与している分子である。これらのハブの存在、さらに関連するスケールフリー性をもつネットワーク形状の存在は、二つの基本的な疑問を提起する：

- なぜ、WWW または細胞という異なる系は、類似したスケールフリー構造に収束するのか？
- なぜ、エルデシュ・レニイのランダム・ネットワークモデルは、実際のネットワークで観察されるハブとべき法則を再現できないか？

スケールフリー性を示す系の性質、起源と範囲のような根



オンライン資料 5.1

スケールフリー・ソナタ

Michael Edward Edgerton が演奏する「ピアノのためのソナタ」を聞いてみよう。スケールフリー・ネットワークを表現したものである。



オンライン資料 5.2

スケールフリー・ネットワークの発現

パラバシ・アルバート・モデルによるスケールフリー・ネットワークの成長とハブの発現を示すビデオ。Dashun Wang の好意による。

要がある。これが、次の節の主題である。

5.4 次数ダイナミクス

スケールフリー性の発現を理解するために、我々はパラバシ・アルバート・モデルの時間発展に焦点を当てる必要がある。一つのノードの次数が時間に伴ってどう変化するか調べるところから始めよう [11]。

このモデルにおいて、新しいノードがネットワークに追加されるたびに、既存のノードはその次数を増やすことができる。この新しいノードは、すでに系に存在する $N(t)$ ノードの m とつながる。これらリンクのうちの一つがノード i につながる確率は式 (5.1) によって与えられる。

次数 k を、成長プロセスにおいて実現されうる多くのネットワークの期待値として、連続実変数を用いて近似してみよう。既存のノード i が新しいノードからつながる単位時間当たりのリンク数は、

$$\frac{dk_i}{dt} = m\Pi(k_i) = m \frac{k_i}{\sum_{j=1}^{N-1} k_j}. \quad (5.3)$$

である。係数 m は、新たなノードが一つ増えるたびに m 個のリンクが加えられることを意味している。それゆえに、ノード i は m 回リンク先として選ばれる可能性がある。このように、式 (5.3) の分母では新たに加えられるノードを除くネットワーク上のすべてのノードについて和をとるため、

$$\sum_{j=1}^{N-1} k_j = 2mt - m. \quad (5.4)$$

となる。ゆえに式 (5.3) は

$$\frac{dk_i}{dt} = \frac{k_i}{2t-1}. \quad (5.5)$$

のようになる。 t の値が大きい場合、分母における (-1) は無視できるので、

$$\frac{dk_i}{dt} = \frac{1}{2} \frac{k_i}{t}. \quad (5.6)$$

となる。式 (5.6) を積分して、ノード i が時間 t_i にネットワークに追加され m 個のリンク張ることを意味する式 $k_i(t_i) = m$ を用いることによって、

はその次数のみによって決まると仮定していた。適応度の効果を取り入れるため、優先選択はノードの適応度 η と次数 k の積によって決まると仮定しよう。そのようなモデルは、ビアンコーニ・バラバシ・モデルまたは適応度モデルと呼ばれて、次の二つのステップで記述される [166, 167] :

成長

各時間ステップで、 m 本のリンクと適応度 η_j をもつノード j が一つ付加される。ここで、適応度 η_j は与えられた適応度分布 $\rho(\eta)$ に従って生成されたランダムな数である。いったん割り当てられた適応度は変化しない。

優先選択

新しく付加されたノードが他のノード i にリンクを張る確率は

$$\Pi_i = \frac{\eta_i k_i}{\sum_j \eta_j k_j}. \quad (6.1)$$

のようにノード i の次数 k_i と適応度 η_i の積に比例する。

式 (6.1) では、確率 Π_i の k_i 依存性は、より次数の大きいノードが見えやすく、すなわちリンクを張られやすいことを意味し、 η_i 依存性は、同じ次数をもつノードであっても、より大きい適応度をもつノードの方がリンクを張られやすいことを意味している。したがって、式 (6.1) で表される確率によれば、まだリンク数の小さい若いノードであっても他のノードに比べて大きい適応度をもっていれば、急速にリンク数を大きくしていくことが可能である。

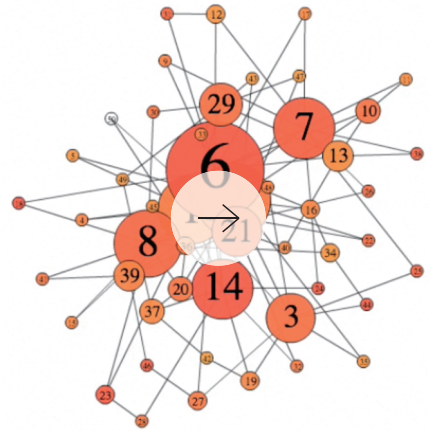
6.2.1 次数成長

それぞれのノードの次数がどのように成長していくかについては、連続体理論を用いて調べることができる。式 (6.1) によれば、ノード i の次数は、単位時間当たり

$$\frac{\partial k_i}{\partial t} = m \frac{\eta_i k_i}{\sum_j \eta_j k_j}. \quad (6.2)$$

に従って変化する。次数 k_i の時間発展は適応度に依存する指数 $\beta(\eta_i)$ を用いたべき則

$$k(t, t_i, \eta_i) = m \left(\frac{t}{t_i} \right)^{\beta(\eta_i)}. \quad (6.3)$$



オンライン資料 6.1

ビアンコーニ・バラバシ・モデル

この映像では、無作為に与えられた適応度をもつてノードが追加されることによって成長していくネットワークの成長の様子を見ることができる。適応度はノードの色によって示される。新しいノードは式 (6.1) による一般化された優先選択によって既存のノードにリンクを張っていく。ノードの成長率はその適応度に比例することになる。ノードの大きさは次数に比例しており、時間の経過に従って、最も大きい適応度をもつノードが最も大きいハブになっていく。このビデオは Dashun Wang の好意による。

ズ気体がある臨界温度よりも低温にすることにより生ずる (Box6.4)。ネットワークにおいては、式 (6.16) の温度 β_T は単なるダミー変数であり、次数分布 p_k のようなトポロジ的に重要な諸量から最終的には消えてしまう。したがって、ボーズ・アインシュタイン凝縮が起こるかどうかは適応度分布 $\rho(\eta)$ の形のみ依存することになる。あるネットワークがボーズ・アインシュタイン凝縮を起こすためには、適応度分布は条件

$$\int_{\eta_{\min}}^{\eta_{\max}} \frac{n\rho(\eta)}{1-\eta} d\eta < 1$$

を満たす必要がある。ボーズ・アインシュタイン凝縮を起こす適応度分布の一つとして、

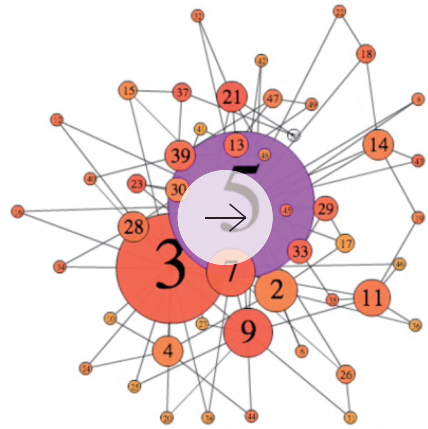
$$\rho(\eta) = (1-\zeta)(1-\eta)^\zeta, \quad (6.22)$$

がある。ここでは、 ζ を変えることによって、ボーズ・アインシュタイン凝縮を起こすことができる (図 6.9 参照)。実際、式 (6.20) が解をもつか否かは $\rho(\eta)$ の形によって決まるエネルギー分布 $g(\varepsilon)$ に依存する。特に、もし式 (6.22) が与えられた $g(\varepsilon)$ に対して非負の解をもたなければ、ボーズ・アインシュタイン凝縮が起き、全粒子のうちの有限部分が最低のエネルギー準位に落ち込む (オンライン資料 6.2)。

まとめると、適応度分布の正確な形が成長するネットワークのトポロジーを決定する。一様な適応度分布ではスケール・フリーのトポロジーとなるのに対して、ある $\rho(\eta)$ ではボーズ・アインシュタイン凝縮が起こる。もし、ネットワークがボーズ・アインシュタイン凝縮を起こすと、一つ、あるいは、数個のノードがほとんどのリンクを独占することになる。その結果、スケール・フリー状態を作り出す適者繁栄の過程が、勝者独占の現象へと変わってしまう。ボーズ・アインシュタイン凝縮はこのようにネットワークの構造に明白な影響を与えるため、それが存在するなら見逃すことはほとんどない。ボーズ・アインシュタイン凝縮はスケールフリー・ネットワークを特徴づけるハブの階層構造を破壊し、ネットワークを星状のハブとスポークから成るトポロジーへと変えてしまう (図 6.9)。

6.5 成長ネットワークモデル

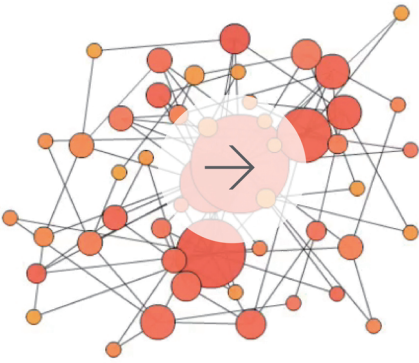
バラバシ・アルバート・モデルはスケール・フリー性が出現



オンライン資料 6.2

ネットワークにおけるボーズ・アインシュタイン凝縮

このビデオは、紫色のノード一つが他のノードよりもずっと大きい適応度をもつネットワークの時間発展を示すものである。この一つのノードがほとんどのリンクを独占してしまい、系はボーズ・アインシュタイン凝縮を起こす。ビデオは Dashun Wang の好意による。



オンライン資料 8.1

ノード故障時のスケールフリー・ネットワーク

スケールフリー・ネットワークの頑健性を図示するために、[オンライン資料 4.1](#) でパラバシ・アルパート・モデルによって作成したスケールフリー・ネットワークを用いる。無作為にノードを選択し、一つ一つ除去していく。動画が示すように、かなり多くのノードを除去しているにもかかわらず、ネットワークは断片化しない。ビデオ作製は Dashun Wang による。

連結成分から構成される。式 (8.4) のモロイ・リード基準は、巨大連結成分のありなしで表されるネットワークの全体性を、 $\langle k \rangle$ と $\langle k^2 \rangle$ に関係づける。これは、あらゆる次数分布 p_k について成り立つ。

式 (8.4) の妥当性を確かめるために、ランダム・ネットワークにこれを適用してみよう。 $\langle k^2 \rangle = \langle k \rangle (1 + \langle k \rangle)$ の場合、

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{\langle k \rangle (1 + \langle k \rangle)}{\langle k \rangle} = 1 + \langle k \rangle > 2 \quad (8.5)$$

または

$$\langle k \rangle > 1. \quad (8.6)$$

が満たされれば、ランダム・ネットワークは巨大連結成分をもつことになる。この主張は、巨大連結成分が存在するための必要条件 (3.10) と一致する。

8.3.2 臨界しきい値

[図 9.7](#) で確認される頑健性の数学的背景を理解するために、どのしきい値でスケールフリー・ネットワークの巨大連結成分が消えるかを確認する。モロイ・リード基準を任意の次数分布のネットワークに適用することにより、臨界しきい値は

$$f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}. \quad (8.7)$$

のようになることがわかった ([発展的話題 8.C](#))。式 (8.7) の最も重要な主張は、臨界しきい値 f_c が、次数分布 p_k によって決まる $\langle k \rangle$, $\langle k^2 \rangle$ のみに依存する点にある。

ランダム・ネットワークの分断しきい値を計算することにより、式 (8.7) の効用を明確にしよう。 $\langle k^2 \rangle = \langle k \rangle (\langle k \rangle + 1)$ を使用することにより、

$$f_c^{\text{ER}} = 1 - \frac{1}{\langle k \rangle}. \quad (8.8)$$

を得る ([発展的話題 8.D](#))。したがって、ランダム・ネットワークの密度が大きいほど、 f_c は高くなる。言い換えると、そのネットワークを分断するためにはより多くのノードを除去する必要がある。さらに、式 (8.8) から f_c は常に有限であるので、ランダム・ネットワークはある一定のノードを取り除くと分断することがわかる。

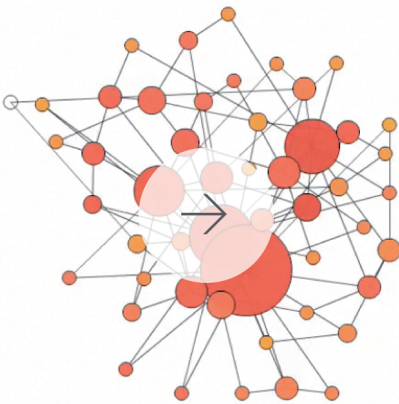
式 (8.7) より、[図 8.7](#) で確認された高い頑健性の起源を理解

への依存性は**発展的**話題 8.A で説明する。

- 頑健性の高まりは、ノードを除いたときだけでなく、リンクを除いたときにも生じる (図 8.10)。

まとめると、本節では、無作為なノードの故障に対する頑健性の検証を通じて、現実のネットワークの基本的な特性を確認した。式 (8.7) は、ネットワークの分断しきい値は、ネットワークの次数分布で決定される $\langle k \rangle$ と $\langle k^2 \rangle$ によることを示している。そのため、スケールフリー・ネットワークを分断するためには、すべてのノードを取り除かなければならず、これらのネットワークが無作為な故障に対して高い頑健性を有していることが確認できた。

この高い頑健性は、高い $\langle k^2 \rangle$ の値に起因する。現実のネットワークのほとんどの $\langle k^2 \rangle$ が無作為に期待されるものよりも大きいことを考慮すると、高い頑健性は多くのネットワークの一般的な特性である。この高い頑健性の原因は、無作為なノードの故障がネットワークの全体性を維持するうえで限定的な役割しか担わない数多くの小さいノードにおもに影響を及ぼすことに拠っている。



オンライン資料 8.2

攻撃下のスケールフリー・ネットワーク

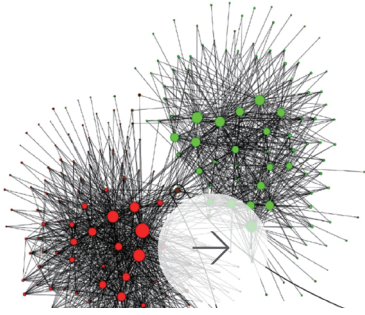
攻撃時は、ネットワークに最も大きいダメージを与えようとする。そうするためには、次数が高いノードから除去していけばよい。動画が示しているように、少数のハブを除去するだけでスケールフリー・ネットワークは分断した成分へと変わる。無作為なノードの故障時にネットワークが分断しないことを示しているオンライン資料 8.1 と比較をしてみると、その違いは明らかである。ビデオ作製は Dashun Wang による。

8.4 攻撃耐性

ハブが有するスケールフリー・ネットワークを維持する役割を考慮すると、ノードを無作為に取り除くのではなく、ハブを取り除いたらどうなるのかという疑問が生ずる。最初に最も次数が高いノードを取り除き、次々に次数が高い順に取り除いていくことを考える。通常このような特定の順序でノードが故障する可能性はゼロに近い。しかし、これはネットワークのトポロジーを理解し、ハブを攻撃することができ、ネットワークを故意に活動不能にしようとするようなネットワークへの**攻撃**を模倣している [12]。

一つのハブを除去しても他のハブがネットワークを維持することができるため、ネットワークが断片化される可能性は低い。しかし、いくつかのハブを除去すると、莫大な数のノードが分断されていく (オンライン資料 8.2)。攻撃が継続すると、ネットワークはたちまち小さいクラスターに分かれるだろう。

スケールフリー・ネットワークにおけるハブの除去の影響



オンライン資料 9.1

モジュラリティに基づくアルゴリズム

広く使われている、モジュラリティ最大化によってコミュニティを検出するアルゴリズムがいくつかある。

最適化された貪欲アルゴリズム

疎な行列では、データ構造を工夫することにより、貪欲アルゴリズムの計算の複雑さは $O(N \log^2 N)$ まで減る [310]。http://cs.unm.edu/~aaron/research/fastmodularity.htm を参照のこと。

Louvain のアルゴリズム

モジュラリティ最適化アルゴリズムの計算の複雑さは $O(L)$ である [282]。ゆえに数百万個のノードをもつネットワークにおいても、図 9.1 のようにコミュニティを検出することができる。このアルゴリズムは発展的課題 9.C で説明する。https://sites.google.com/site/findcommunities/ を参照。

計算の複雑さ

それぞれの ΔM の計算は一定の時間で行われるため、貪欲アルゴリズムのステップ 2 には $O(L)$ の計算時間が必要である。どのコミュニティを統合させるかを決めたのちは、行列の更新は最悪のケースでも $O(N)$ で行われる。このアルゴリズムでは $N-1$ 回コミュニティを統合させる必要があるため、複雑さは $O((L+N)N)$ 、あるいは疎なグラフでは $O(N^2)$ になる。最適化することによってアルゴリズムの複雑さは $O(N \log^2 N)$ へと減少する (オンライン資料 9.1)。

9.4.3 モジュラリティの限界

コミュニティを特定する上でモジュラリティは重要な役割を果たすが、その限界についても理解しておくことが必要である。

分解能の限界

モジュラリティの最大化によって、小さいコミュニティは強制的に大きいものになってしまう。実際に、コミュニティ A と B を一つのコミュニティに統合すると、ネットワークのモジュラリティは

$$\Delta M_{AB} = \frac{l_{AB}}{L} - \frac{k_A k_B}{2L^2}, \quad (9.13)$$

のように変化する (発展的課題 9.B)。ここで、 l_{AB} は、次数の合計が k_A のコミュニティ A のノードと、次数の合計が k_B のコミュニティ B のノードとを結び合わせるリンクの数である。A と B とが異なるコミュニティであれば、 M が最大化される時これらは分かれたままであるはずである。しかし、次に見るとおり、常にそういうわけではない。

$(k_A k_B)/2L < 1$ の場合を考える。二つのコミュニティの間にリンクが最低でも一つあるならば ($l_{AB} \geq 1$)、式 (9.13) により $\Delta M_{AB} > 0$ となり、モジュラリティを最大化するために A と B を統合させなければならない。単純化するため $k_A \sim k_B = k$ と仮定すると、コミュニティの次数の合計が

$$k \leq \sqrt{2L} \quad (9.14)$$

を満たすとき、A と B を一つのコミュニティに統合することによって、モジュラリティは高くなる。式 (9.14) を満たさない場合は、A と B は異なるコミュニティとなる。これはモ

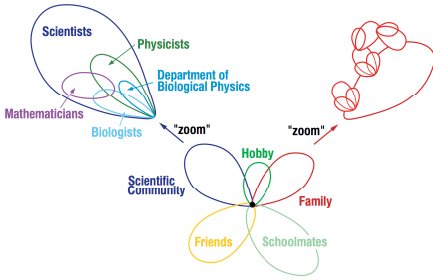
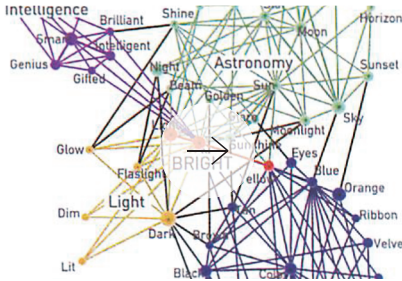


図 9.19 重なり合うコミュニティ

重なり合うコミュニティという概念を導入した、Tamas Vicsek 自身を取り巻くコミュニティの模式図。研究者コミュニティを拡大すると彼の科学的な関心の特徴づけるコミュニティが入れ子状になり重なり合っている様子が現れる。文献 [311] による。



オンライン資料 9.2

Cfinder

CFinder というソフトウェアを使えば、重なり合うコミュニティを見つけ出すことができる。CFinder は www.cfinder.org からダウンロードできる。

にも属しているだろう (図 9.19)。それらのコミュニティは、それぞれがまた他のいくつかのコミュニティに属する個人からなっており、その結果、入れ子になり重なり合った複雑なコミュニティのウェブができあがる [311]。重なり合うコミュニティは社会的なシステムに限定されたものではない。同じ遺伝子が多くの疾病に関与していることはしばしばあり、異なる機能不全からなる疾病のモジュールが互いに重なり合っていることを示唆している [29]。

入れ子になったコミュニティ構造が存在することは、社会学者 [321] やグラフ分割に興味をもつ工学者の間では以前からよく知られていたのであるが、ここまで議論されたアルゴリズムでは、それぞれのノードは単一のコミュニティに属するというものになっていた。転回点となったのは、Tamas Vicsekらの研究である [311, 322]。彼らは重なり合うコミュニティを見つけ出すアルゴリズムを提案し、ネットワーク科学コミュニティの関心をこの問題へと向けさせた。この節では、重なり合うコミュニティを検知する、クリーク・パーコレーションとリンク・クラスタリングという、二つのアルゴリズムを議論する。

9.5.1 クリーク・パーコレーション

クリーク・パーコレーション・アルゴリズムはしばしば CFinder と呼ばれる (オンライン資料 9.2)。このアルゴリズムはコミュニティを重なり合うクリークの統合として捉える。

- 二つの k -クリークは、それらが $k-1$ 個のノードを共有しているならば隣接していると考え (図 9.20b)。
 k -クリークコミュニティとは隣接するすべての k -クリークの統合によって得られる最大連結部分グラフのことである。
- ある特定の k -クリークから到達できない k -クリークは、別の k -クリークコミュニティに属している (図 9.20c,d)。

CFinder アルゴリズムはすべてのクリークを見つけ出してから、大きさ $N_{clique} \times N_{clique}$ のクリーク・クリーク重なり合い行列 O を作る。ここで N_{clique} はクリークの数であり、要素 O_{ij} はクリーク i とクリーク j で共有されているノード数である (図 9.39)。CFinder アルゴリズムの典型的な出力を図 9.21 に示す。この図は *bright* という英単語のコミュニティ構造で

(9.45) の第 1 項はコミュニティ間の移動を記述するために必要なビット数の期待値である。ここで、 q は、ランダム・ウォーカーが与えられたステップ中にコミュニティを移る確率である。

第 2 項は、コミュニティ内の移動を記述するために必要なビット数の期待値である。ここで、 $H(P_c)$ はコミュニティ内の移動についてのエントロピーであり、コミュニティ i から出ていくときの退出コードも含まれている。

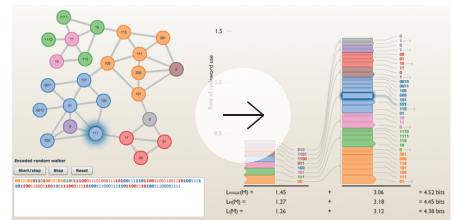
マップ方程式中の個々の項やネットワーク内のランダム・ウォーカーの動きを記述する確率の計算はいささか込み入っているが、文献 [318, 319, 320] には詳しい記述がある。オンライン資料 9.3 では式 (9.45) の背後にあるメカニズムやその使い方がよくわかるようなインタラクティブなツールを紹介している。

最後に、この評価関数 \mathcal{L} は、ネットワークのある特定のコミュニティ分割について与えられる値であることを注意しておく。最良の分割を見つけるためには、可能な分割すべてを調べて \mathcal{L} を最小化しなければならない。この最小化のために最もよく使われるのは、Louvain のアルゴリズムのステップ I と II である。具体的には、個々のノードを別のコミュニティとして出発する。そして、隣接するノードを一つのモジュールに統合することにより \mathcal{L} が減少するなら、その統合を採用することとする。それぞれの統合の後に式 (9.45) を用いて \mathcal{L} の値をアップデートする。これをすべてのノードについて系統的に繰り返す。得られたコミュニティ構造はさらに新しいコミュニティへと統合され、一つのパスが終わる。その後、ノード数が集約されたこの新しいネットワークについて、またこのアルゴリズムを繰り返す。

計算の複雑さ

Infomap の計算量は、マップ方程式 \mathcal{L} を最小化するために必要な手順数によって決まる。Louvain の手順を用いる場合には、計算量は Louvain のアルゴリズムと同じであり、たかだか $O(L \log L)$ 、あるいは、疎なグラフについては $O(N \log N)$ の程度である。

まとめると、Louvain のアルゴリズムと Infomap は、高速なコミュニティ検出アルゴリズムである。また、いくつかのベンチマークを用いて測定された、二つのアルゴリズムの正



オンライン資料 9.3

Infomap のマップ方程式

マップ方程式の背後のメカニズムをいきいきと視覚化したものとして、<http://www.tp.umu.se/~rosvall/livemod/mapequation/> がある。

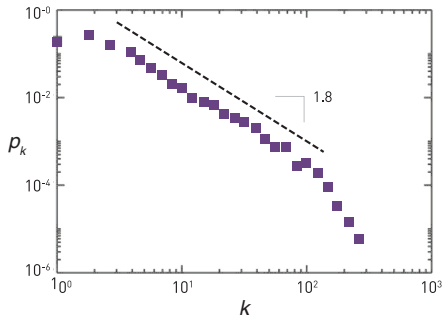
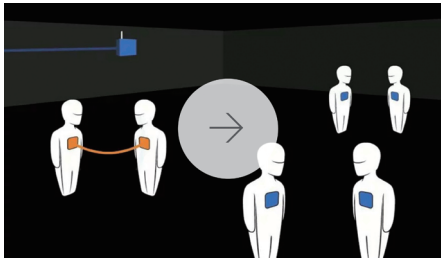


図 10.15 航空輸送ネットワーク

航空輸送ネットワークの次数分布は指数 $\gamma = 1.8 \pm 0.2$ のべき分布で近似される。この図は、世界の空港とその間の直行便リストを含む 2002 年発行の国際航空輸送協会のデータベースを使って作成された。その結果として得られたネットワークは、全世界の航空輸送の 99% を占める上位空港 $N = 3,100$ をノードとし、それらを結ぶ直行便 $L = 17,182$ をリンクとする重み付きのグラフである。文献 [370] 参照。



オンライン資料 10.1

RFID を介したネットワークの探索

このビデオは、RFID 技術と、それをを用いた社会的相互作用のマッピングを紹介している。

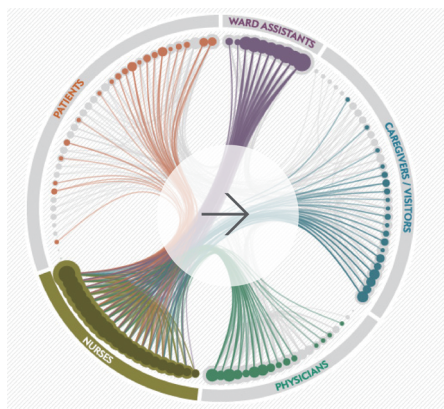
トワークは次数指数 $\gamma = 1.8$ のスケールフリー・ネットワークである。このように指数が低い値となるのは、空港間に直行便が複数あり、ネットワークが単純ではないためである。同様のべき分布は、リンクの重みの分布にも見られ、空港間の旅客量は通常それほど多くないが、いくつかの空港間の旅客量は驚異的であることを示している。10.5 節で議論するように、この不均一さは特定の病原体の感染では主要な役割を担う。

局所的接触パターン

多くの空気感染症は、人と人との対面の相互作用によって感染していく [372, 373, 374, 375]。これらの相互作用のパターンは、電波による個体認識デバイス (RFID) [373, 375]、携帯電話を用いた計量社会学用バッジ [376]、または、その他の無線テクノロジー [377] を用いてモニターできる。

RFID は、近くにいるバッジをつけた他人を感知するデジタルバッジであり、人どうしの相互作用を捉えることを目的としてさまざまな環境下で配布されてきた (オンライン資料 10.1)。例を挙げると、3 か月間のある科学展示会の 14,000 人以上の訪問客や 3 日間の会議の 100 人以上の参加者などである [373]。図 10.16 は、RFID によって明らかにされた、2 日間の高校生およびその教員間の相互作用を捉えたネットワークである。以下のような、いくつかの発見があった：

- RFID タグは、同じバッジをつけた互いに向き合う個人間の相互作用のみを感知するので、感知された接触の数は制限されている。したがって、一般的にこの研究において図示される接触ネットワークは指数次数分布をもつ。
- 個々の対面の相互作用の持続時間は、数桁以上にわたってべき分布に従う。したがって、大部分の接触は短時間である、しかし少数ではあるが長時間にわたる接触、それは病原体の感染のような重大な結果をもたらすバースト性の時間パターンが観察される [378] (10.5 節参照)。
- 二人が一緒に過ごした累積時間を表すリンクの重みは、べき分布に従う。したがって、大部分の人はほんの少数の人としか時間を過ごさない。これはまた感染のパターンに対して重要な意味をもっている (10.5 節参照)。
- 空気感染するほとんどの病原体では、空間的に近いところ



オンライン資料 10.2

院内感染

現在の抗生物質に対して耐性をもつ細菌は、地球規模の衛生にとって重要な脅威である。そのような細菌は、特に病院と保健施設内で感染する。この *Scientific American* の Interactive Feature では、いくつかの病院における細菌蔓延を追跡した様子を描写している。

し、コンピューター内で発見される他の電子メールアドレスに、自分自身のコピーを送信する。したがって、この場合の接触ネットワークは電子メールネットワークであり、表 4.1 で議論したように、スケールフリー性をもっている [113]。その他のコンピューター・ウイルスはいろいろな通信プロトコルを利用して、インターネットのもつ相互接続のパターンを反映したネットワーク上で感染するが、それもまたスケールフリー性をもつ (表 4.1)。最後に、いくつかのマルウェアは IP アドレスをスキャンし、完全連結ネットワーク上を感染する。

携帯電話ウイルス

携帯電話ウイルスは、MMS とブルートゥースを介して感染する (図 10.2)。MMS ウイルスは、電話のコンタクトリストで見つかるすべての電話番号に自分自身のコピーを送信する。すなわち、MMS ウイルスは、携帯電話通信の背後にある社会ネットワークを利用するのである。表 4.1 に示されるように、携帯電話ネットワークは、大きいべき指数をもつスケールフリー・ネットワークである。携帯電話ウイルスはブルートゥースによって感染することもあり、BT 接続を用いて物理的に近い場所にあるすべての感染可能な電話に自分自身をコピーする。上述のように、この位置ネットワークについても、非常に次数の不均一性が大きい [35]。

要約すると、この十年ほどの技術的進歩によって、性的関係から近接性をもとにした接触ネットワークにいたるまで、生物学的あるいはデジタルウイルスの感染を可能にするさまざまなネットワークの構造を図示できるようになった (オンライン資料 10.2)。電子メールネットワーク、インターネット、あるいは性的ネットワークなどの、これら多くのネットワークは、スケールフリー性をもつ。他のネットワークにおいては、共同研究ネットワークのように、次数分布は単純なべき則には従わないかもしれないが、それでも $\langle k^2 \rangle$ は大きく、次数の不均一性は顕著である (表 4.1, 表 4.4)。これは、前節で得られた分析結果は、大部分のネットワーク上の病原体の感染について、直接的な重要性をもっていることを意味する。すなわち、背後にある接触ネットワークの次数の不均一性のおかげで、感染力がそれほど強くないウイルスでさえ集団内で感染することが容易になるのである。



オンライン資料 10.3

社会ネットワーク上での「感染」

「あなたの友人が肥満であるならば、あなたの肥満の危険性は45パーセントより大きい。…あなたの友人の友人が肥満であるならば、あなたの肥満の危険性は25パーセントより大きい。…あなたの友人の友人の友人、それはあなたが多分知りもしないだろう誰かだろうが、その人が肥満であるならば、あなたの肥満の危険性は10パーセントより大きい。あなたの友人の友人の友人の友人まで考えて、ようやく、その人の体の大きさとあなた自身の体の大きさが何の関係もなくなる。Nicholas Christakisによる、社会ネットワーク上での健康状態の「感染」パターンの説明を見てみよう。

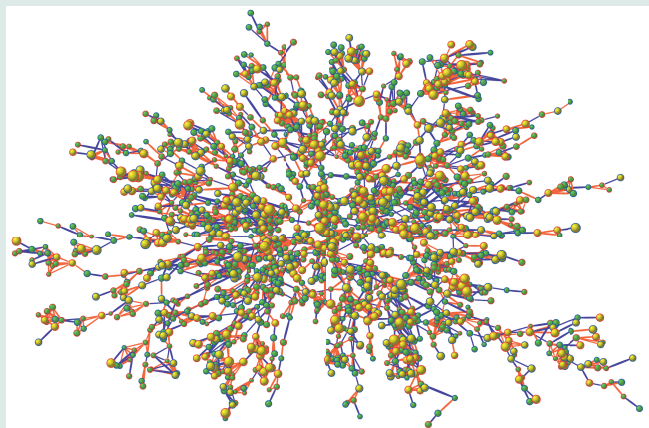
図 10.21 肥満のウェブ

フラミンガム心臓研究所に所属している2,200人の友人関係を調べたネットワーク内の最大連結成分。各々のノードは個人を表す。青い境界線のノードが男性で、赤い境界線のノードは女性である。各々のノードの大きさはその人のBMIに比例し、黄色で塗られたノードはBMI > 30の肥満の人である。紫色のリンクは友人または婚姻関係で、オレンジ色のリンクはたとえば兄弟などの家族関係である。肥満の人の作るクラスターと肥満でない人の作るクラスターがネットワーク内に見とれる。この分析結果は、これらのクラスターは「類は友を呼ぶ」傾向、すなわち、体の大きさが同じ人どうしが友人になりやすいことから説明できないことを示唆している。これは、複雑な感染過程によって、社会ネットワークのリンクに沿って肥満の「感染」が起きていることの証拠である。文献 [393] による。

Box 10.2 我々の友人が我々を太らせる？

インフルエンザ、SARS、あるいはエイズのような病気は、病原体が伝達されることによって感染する。しかし、社会ネットワークも同様に非感染的な病気の蔓延を助けるといふことがあり得るだろうか？最近の調査は、それがあり得ることを示唆している。この調査では、社会ネットワークが、肥満、幸福、あるいは喫煙を止めるなどの行動パターンに影響を与えうるとの証拠が示された [403, 404]。

肥満は、肥満指数 (BMI) を用いて診断されるが、遺伝、食生活、運動など多数の要因が関わっている。実際の調査によると、友人関係もまた重要な役割を担っていることが示されている。5,209人の男性と女性から成る社会ネットワークの分析から、友人の一人が肥満であるならば、その人自身も次の2年から4年で体重が増加するリスクが57%増加することがわかった [403]。もし親友が肥満であるならば、その危険は3倍になる。実際、体重増加の可能性は171%にまで跳ね上がる (図 10.21)。肥満を伝達する「肥満病原体」などはないという事実にもかかわらず、実際上すべての点において、肥満はインフルエンザやエイズと同様の感染症であるかのように見えるのである。



感染症について報告する際の国際的なガイドラインである。その結果として、感染症の流行に付随するデータを系統的に入手する可能性が向上し、モデルへのパラメータ入力の実時間で可能になった。これらの開発結果の恩恵を最初に受けたのは2009年のH1N1ウイルスの発生で、このときは、感染症の流行を実時間で予測できた。

新しい病原体の発生は、次のようないくつかの重要な疑問をもたらす。

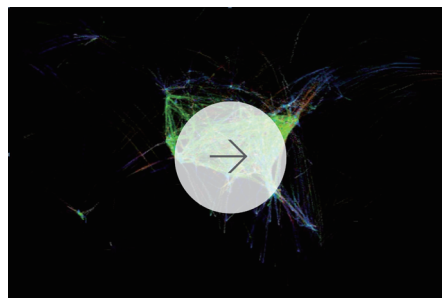
- 新しい病原体はどこで発生したのか？
- どこで新しい事象が発生するのか？
- 人口密集地域にいつ感染症は流行するのか？
- どれくらい多くの人が感染するのか？
- 流行を低減するために我々は何ができるのか？
- どのようにしてその病原体を根絶できるのか？

今日では、強力な感染症シミュレータ（オンライン資料 10.4）に、人口、移動、疫学に関するデータを入力しながら、これらの疑問に答えることができる [411, 412]。これらのツールでは、確率的メタポピュレーションモデル [413, 414, 415]、数百万人の移動と相互作用を捉えるエージェントベースシミュレーション [416]、などのアルゴリズムが使われる。この節では、これらのツールの機能を概観し、その開発におけるネットワーク科学の役割について説明する。

10.7.1 実時間予測

感染症予測の目的は、各々の大都市において週ごとの感染者の数を推定して病原体の実時間での伝播を予測することにある [416, 417]。ネットワーク科学に基づく実時間感染症予測の最初の成功事例は、世界中の人口と移動の高分解能データを入力として用いる確率論的なフレームワークである Global Epidemic And Mobility (GLEAM) モデルである [417] (図 10.26, オンライン資料 10.5)。GLEAM は、以下のようなネットワーク指向の計算モデルである：

- 世界中の各々の地理的位置をネットワークのノードに対応付ける。
- ノードの間の移動を表すリンクを旅客航空スケジュールの



オンライン資料 10.4

北米の航空機運航パターン

北米の実時間航空機運用は連邦航空局が発表するデータによる。この世界的な輸送ネットワークは、大陸間の病原体の感染に深くかかわっている。それ故に、運航スケジュールは感染予測の入力データとなっている。Aaron Koblin によって作られたこのビデオは、純粋に科学的な理解のためのものであるが、芸術分野ではデジタルアートとしても鑑賞されている。実際に、このビデオはニューヨーク近代美術館のメディア芸術のコレクションに収録されている。



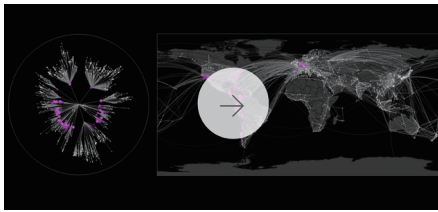
オンライン資料 10.5

GLEAM

感染予測を行う GLEAM ソフトウェアパッケージを説明するビデオ

図 10.31 有効距離

香港が感染源であるパンデミックの広がり。多数の感染者がいる地域を赤色のノードで示している。各々のパネルでは、有効距離表示（上）と通常の地理的表示（下）を用いて系の状態を比較している。地理的表示にみられる複雑な空間パターンは、有効距離表示では一定速度で外に向かって移動する円形の波となる（オンライン資料 10.6）。文献 [425] による。



オンライン資料 10.6

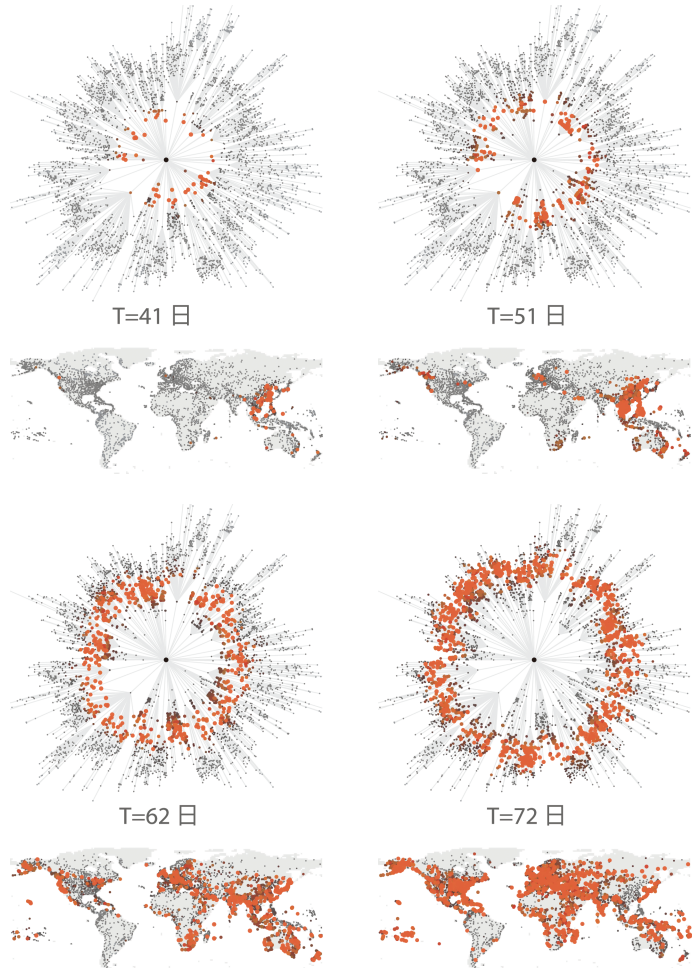
感染症の大流行

3 か所での発生から GLEAM で予測される病原体の感染。地理的な広がりのパターンの解釈は難しいが、有効距離を用いた表示では感染症の流行は通常の放射パターンを示す（図 10.31）。

観測された感染パターンから次のような疑問が出てくる。典型的な病原体が地球上に広がる速度はどれくらいなのだろうか？ その速度は、三つのパラメータに依存する：

1. 基本再生産数 R_0 、これはインフルエンザ型のウイルスでは 2 の近傍にある（表 10.2）。
2. 治癒率、これはインフルエンザでは約 3 日である。
3. 移動度、対象となる集団のなかで日中に移動する人の割合である。このパラメータは 0.01 から 0.001 の間にある。

これらのパラメータを用いて GLEAM を用いて、到達時間と感染源からの距離の相関を計算することによって、病原体が地球上に広がる速度をおよそ 250 から 300 km/日と見積もることができる。したがって、インフルエンザウイルスは大陸をスポーツカーまたは小型飛行機で移動することがわかる [425]。



に式 (10.31) では対数が使われている。

図 10.31 に図示されるように、感染症の発生源から各々の都市までの距離を式 (10.31) で表すと、病原体は円形の波面を形成する。これは、地理学的な空間でみた場合の感染症の蔓延の複雑なパターンとは好対照をなす。さらに、H1N1 の到達時間を物理的距離の関数として描いたときにはランダムな様子を示すが、到達時間は有効距離とは強い相関を示す（図 10.32）。したがって、我々は病原体の広がる速度を決めるために有効距離を使うことができる（オンライン教材 10.6）。

流行予測について驚きとともに歓迎すべきことは、それぞれの感染症モデルでは、移動データについて異なるもの（航空運行スケジュール [370] とドル紙幣の動き [369]）を用いても、また、感染症のパラメータ（治癒率、伝達率など）に