



HOKKAIDO UNIVERSITY

Title	複雑ネットワークにおける長距離次数相関
Author(s)	藤木, 結香
Degree Grantor	北海道大学
Degree Name	博士(工学)
Dissertation Number	甲第14427号
Issue Date	2021-03-25
DOI	https://doi.org/10.14943/doctoral.k14427
Doc URL	https://hdl.handle.net/2115/81184
Type	doctoral thesis
File Information	Yuka_Fujiki.pdf



令和2年度博士論文

複雑ネットワークにおける 長距離次数相関

Long-range degree correlations in complex networks

北海道大学工学院

応用物理学専攻 数理物理工学研究室

藤木 結香

目次

第 1 章	序論	3
第 2 章	複雑ネットワークの構造的特徴	7
2.1	ネットワークの数学的表現	7
2.2	次数分布関数とスケールフリー性	8
2.3	次数相関	10
2.4	スモールワールド性	15
2.5	フラクタル性	19
2.5.1	現実世界のフラクタル・スケールフリー・ネットワーク	20
2.5.2	臨界ネットワークのフラクタル性とスケールフリー性	21
2.5.3	決定論的フラクタル・スケールフリー・ネットワーク・モデル	21
第 3 章	ネットワーク上の諸現象と隣接次数相関	25
3.1	パーコレーション転移と隣接次数相関	25
3.2	同期現象と隣接次数相関	31
3.3	感染症拡散と隣接次数相関	33
第 4 章	隣接次数相関の限界	38
4.1	(u, v) -flower のリワイヤリング	38
4.2	ネットワーク上の諸現象	41
4.3	本論文の目的	42
第 5 章	長距離次数相関の定式化	44
5.1	確率関数による一般的記述	44
5.2	相関が無いネットワーク	47
5.2.1	平均場近似	49
5.2.2	木構造無限サイズ・ネットワークにおいて	55
5.3	現実ネットワークの評価と新たな評価指標	57

第 6 章	内因性長距離次数相関	61
6.1	外因性長距離次数相関	61
6.2	平均場近似	64
6.3	現実のネットワークの長距離次数相関と新たな評価指標	65
第 7 章	結論	71
	謝辞	74
付 録 A	スケールフリー・ネットワークの次数モーメント	75
付 録 B	決定論的フラクタル・スケールフリー・ネットワークの特徴量	77
B.1	特徴量の計算方法	77
B.2	具体例	83
付 録 C	平均場近似の精度	85
付 録 D	現実世界の複雑ネットワーク	88
	参考文献	91

第1章 序論

我々の身の回りの複雑系を構成する要素間の関係はネットワークという概念を用いることで極めて一般的かつ簡便に抽象化することができる。航空路線網，人間関係，神経網などの極めて広い領域に渡る複雑系は複雑ネットワークとして表される [1–5]。現実世界を記述する多くのネットワークでは，ノードから出るエッジの本数（次数）がノードごとに大きく異なり，次数の分布がべき的な裾の減衰を見せるスケールフリー性と呼ばれる性質が見られる [6]。こうした強い次数不均一性をもつネットワークでは，ノード対の次数どうしの相関（次数相関）がネットワークの構造やネットワーク上のダイナミクスに強い影響を与える [7, 8]。中でも特によく調べられている次数相関は，隣接するノード間の次数相関（隣接次数相関）である。隣接次数相関の研究はこれまで盛んに行われており [9–17]，クラスター性 [18–22]，コミュニティ構造 [23]，平均ノード間距離 [24]，フラクタル性 [25–28] など，複雑ネットワークの構造的な性質と密接な関係があることに加え，感染症の拡散現象 [29–32, 41]，同期現象 [33–40]，ゲーム理論 [42–45]，ネットワーク故障 [46–50] といったネットワーク上の数々のダイナミクスに影響を与えることが明らかにされている。

しかしながら，一般的な次数相関が隣接次数相関によって完全に記述されるわけではない。隣接次数相関のような局所相関によってネットワークの大域構造を全て表すことができないことは明らかである。実際，Orsini ら [51] は現実世界の複雑ネットワークから抽出された局所的な特徴量（次数列，クラスター係数，隣接次数相関）を再現する数理モデルを作り，これらの局所の特徴量によってのみネットワークのどのような性質が再現されるのかを調べたが，最短経路長分布，媒介中心性，コミュニティ構造などの大域的性質が再現されることはなかった。局所的な特徴量では記述することができない大域的性質には隣接以上に離れたノードの間の次数相関である長距離次数相関がある。高い次数を有するノード（ハブ）どうしの距離はネットワーク上のダイナミクスにも直接的に影響を与える重要な物理量であり [52–55]，長距離次数相関への関心が近年高まっている。

これまでも複雑ネットワークにおける長距離次数相関の定式化の試みはいくつか提案されてきた。Rybski ら [56] は時系列解析のアナロジーを用いて，最短経路上のノードの次数揺らぎに対する Hurst 指数を長距離次数相関を特徴づける指標として提案している。また，Mayo ら [57] や Arcagni ら [58] は，長距離次数相関を定量化するために，ノード間距離 l だけ離れたノード対の次数に対する Pearson 相関係数 r_l を提案している。Mayo らの研究によ

り、社会的ネットワークでは大きな l に対する r_l が負になる傾向があることが明らかになっている。さらに、Allen-Perkins ら [59] は隣接以上の次数相関を表す他の指標として、エッジ両端のノードの「隣接ノードの次数の和」を確率変数とした Pearson 相関係数 $\tilde{r}$ を提案しており、エッジ両端の次数に対する Pearson 相関係数が負となる工学的ネットワークや生物学的ネットワークを、 $\tilde{r}$ の正負に従ってさらに細かく分類できることを示している。これらの先行研究の例からも判るように、これまでに提案されている指標は長距離次数相関の特定の側面を定量化することを目的としたものとなっている。しかしながら、ネットワーク構造と長距離次数相関の関係を系統的に研究するためには、あらゆる種類の長距離次数相関を特徴づけるための一般的な定式化が不可欠である。ネットワークの構造的特徴だけでなく、隣接次数相関がそうであったように、長距離次数相関はネットワーク上の様々なダイナミクスにも強い影響を与えるものと考えられる。そうした影響を解明するためにも長距離次数相関を一般的に記述する方法を確立することは極めて重要である。しかし、第5章で述べる通り、長距離次数相関の一般論を展開するためには有限サイズ・ネットワークを扱うことが本質的に重要となるため、未だ一般論の構築に成功した研究は報告されていない。そのため、複雑ネットワークにおける長距離次数相関を一般的に解析するための枠組みを構築することが急務の課題となっている。

そこで本論文では、有限サイズおよび無限サイズのネットワークに対し、長距離次数相関を解析するための一般的なフレームワークを提案する [60]。そのため、まずはじめに、ノード間距離 l だけ離れた次数 k のノードと次数 k' のノードの間の相関を完全に記述するために、 k , k' , l を確率変数とした同時確率と、4つの条件付き確率を導入する。これら5種の確率関数は任意の距離にあるノード対の次数相関の情報を全て有しており、隣接次数相関に関する情報もこの確率関数の $l = 1$ における特別な場合として内包される。また、5種の確率関数は互いに独立ではなく、Bayes の定理およびいくつかの和則を用いて互いから互いを導出することができる。本論文では、これら5種の確率関数間の関係、および隣接次数相関を記述する確率関数との関係を明らかにする。加えて、対象とするネットワークの長距離次数相関の有無を判断するため、まず基準とすべき「いかなるスケールにおいても次数相関が無いネットワーク（長距離無相関ネットワーク）」を規定し、長距離無相関ネットワークに対する5種の確率関数を平均場近似と局所木構造仮定を用いて解析的に導出する。これにより、対象となるネットワークに対する5種の確率関数と、同じノード数および同じ次数分布を有する長距離無相関ネットワークの確率関数を比較することで、与えられたネットワークが長距離次数相関を有するか否かを判別することができる。また、これらの確率関数を使って長距離次数相関を定量評価する指標を定義することにより、ネットワークが有する長距離次数相関がどのような相関であるかを明らかにすることができる。さらに、現実世界に実在する複雑ネットワークの長距離次数相関を本手法で解析することで、本論文で提案する解析

手法の有用性を確認する。ただし、隣接次数相関のあるネットワークでは、たとえ長距離に及ぶ相互作用が存在しなかったとしても、隣接相関の効果が積み重なることによって次隣接以上離れたノード間に見かけ上の相関（外因性長距離次数相関）が生じうる。長距離次数相関とネットワーク特性の関係を理解するためには、このような外因性長距離次数相関と、長距離に及ぶ直接的相互作用の帰結としての本質的な長距離次数相関（内因性長距離次数相関）を区別しなければならない。そこで本論文では、外因性長距離次数相関のみを純粋に有するネットワークである、すなわち隣接次数相関以外には完全にランダムなネットワーク（隣接相関ランダム・ネットワーク）を用いて、内因性長距離次数相関と外因性長距離次数相関の識別を行う [61]。まず、長距離無相関ネットワークに対する 5 種の確率関数を解析計算した近似理論を拡張することで、隣接相関ランダム・ネットワークに対する 5 種の確率関数を解析的に計算する。与えられたネットワーク G に対する確率関数と、 G と同じノード数、同じ次数分布、および同じ隣接次数相関を有する隣接相関ランダム・ネットワークに対する確率関数を比較することで、 G の示す長距離次数相関が外因性の相関なのか内因性の相関なのかを区別することができる。本論文では、現実世界を記述する多様な複雑ネットワークの長距離次数相関を系統的に解析するため、外因性・内因性の程度や次数相関の種類を定量判別するための複数の指標を導入した。これらの指標を用いて、社会学、経済学分野から工学、生物学分野に至る 21 種類のネットワークを調べた。その結果、一部のネットワークを除くほぼ全てのネットワークが長距離次数相関を有しており、そのうち大多数が内因性の長距離次数相関を有していることが明らかになった。また、次数どうしの相関のしかたはネットワークの種類によって異なり、例えば低次数ノードどうしが長距離反発するネットワーク（タンパク質折り畳み構造ネットワーク）や、高次数ノード間に反発傾向が見られるネットワーク（WWW）が存在することが明らかとなった。

本論文の構成は以下の通りである。まず第 2 章では、複雑ネットワークの基本的な性質とそれらを特徴づける特徴量について簡単に説明する。第 3 章では、隣接次数相関と隣接次数相関がネットワーク上の諸現象に与える影響に関する従来研究を詳しく説明する。第 4 章では隣接次数相関のみを指定することで再現できるネットワーク構造の限界について説明する。本章ではさらに、スケールフリー性とフラクタル性を併せ持つネットワークに見られるハブどうしの長距離に渡る反発相関が、隣接次数相関では説明され得ないことを明らかにした上で、本論文の目的を述べる。本論文の主題となる長距離次数相関の一般的な定式化は第 5 章にて行う。本章ではまず、第 2 章で述べた隣接次数相関の一般的記述方法を拡張する形で長距離次数相関を特徴づける 5 種の確率関数を導入する。次に、長距離無相関ネットワークを規定し、長距離無相関ネットワークに対する確率関数の関数形を平均場近似と局所木構造仮定を用いて解析的に計算する。また、これら 2 つの近似の妥当性についてもここで確認する。さらに、提案した解析手法を現実世界のネットワークに適用することで、これらのネッ

トワークが示す長距離次数相関を解析し、本解析手法の有用性を確認する。第6章では、隣接次数相関によって引き起こされる外因性長距離次数相関と内因性長距離次数相関とを識別する手法を提案し、長距離次数相関のより詳細な性質に関する議論を展開する。そのため、まず隣接相関ランダム・ネットワークに対する確率関数を平均場近似と局所木構造仮定を用いて解析的に計算する。本章では、長距離次数相関の多様な側面を定量化する数多くの指標も提案される。さらに、これらの結果を用いて、多数の現実ネットワークの長距離次数相関を調べることで、現実世界の複雑系がいかに多様性に溢れているかを示す。本論文全体の結論と今後の課題については第7章で述べる。

第2章 複雑ネットワークの構造的特徴

系の構成要素をノード、要素間相互作用をエッジとした抽象表現であるネットワークは極めて汎用性が高く、人文社会学的な問題から理工学的な問題までありとあらゆる分野における系の構造が複雑ネットワークとして表される [1–5]。例えば物理的な実態としてのエッジを有するものであれば、都市を結ぶ交通路線網や有線で接続されるルーターどうしのネットワーク（インターネット）が挙げられる。また物理的なつながりが存在せずとも、情報の交換や貨物の往来をエッジと考えることができ、人間関係、航空路線網、銀行間取引、WWWなどはそのような複雑ネットワークの例として挙げることができる。さらに抽象的なものでは、生体内における2種類のタンパク質間の相互作用の有無をエッジで表したタンパク質相互作用ネットワークや、自然界における生物種間の被捕食関係をエッジで表現した食物網ネットワークなども複雑ネットワークに含まれる。ネットワーク科学という研究領域では、このような一見すると全く共通点の無い系を包括的に取り扱うことができる。それが可能である所以は、複雑ネットワークが表す元々の系が極めて多種多様であるにもかかわらず、系のトポロジーに注目することでスケールフリー性やスモールワールド性、フラクタル性といった普遍的性質が確認できるからである。本章では、多くの複雑ネットワークに共通するこれらの構造的特徴について説明する。

2.1 ネットワークの数学的表現

ネットワークは「グラフ」とも呼ばれ、離散的な要素（ノード）の集合 V と要素間の対関係（エッジ）の集合 E の組み合わせ $G(V, E)$ として定義される。グラフとネットワークという言葉の使い分けは厳密でなく、数学的な場面ではグラフ、数学以外（物理学や社会学、生物学など）の分野ではネットワークがよく用いられる傾向にあるが、同義の言葉であるため本論文では使い分けは行わない。以後、特に断りが無い場合、全ノード数 $|V|$ を N 、全エッジ数 $|E|$ を M と書くこととする。エッジに重み付けや方向付けも行わない。また、重複するエッジ（多重エッジ）や同じノードを接続するエッジ（自己ループ）は存在しないか、または無視できるほど少ないものとする。

ノード i と j の間にエッジ (i, j) が存在するとき「 i と j は隣接する」または「 i は j の隣接ノードである」という。ネットワーク G にエッジ (i, j) が存在するとき行列成分が $a_{ij} = 1$

となり、存在しないとき $a_{ij} = 0$ となるような正方行列 $A = \{a_{ij}\}$ を $\mathcal{G}$ の隣接行列という。隣接行列 A によってネットワーク $\mathcal{G}$ の構造は完全に記述される。2つの隣接行列 A と A' の間に $A' = PAP^T$ の関係が成り立つような置換行列 P が存在するとき、それぞれの隣接行列が表すネットワーク $\mathcal{G}$ と $\mathcal{G}'$ は「同じトポロジー」であるという。

2.2 次数分布関数とスケールフリー性

ノード i の隣接ノードの個数をノード i の次数といい、任意抽出したノードの次数が k である確率 $P(k)$ を次数分布という。図 2.1 の (A)-(C) はそれぞれ、俳優共演関係 ($N = 225,583$, $M = 8,302,734$) [63], WWW ($N = 685,230$, $M = 6,649,470$) [62], インターネットにおける AS 単位の通信ネットワーク ($N = 26,475$, $M = 53,381$) [62] の次数分布を表す。これらのネットワークをはじめとして、現実世界の数多くの複雑ネットワークの次数分布は高次数領域においてべき関数 $P(k) \propto k^{-\gamma}$ に従い、指数 γ の値は概ね $2 \lesssim \gamma \lesssim 3$ の範囲にある [4]。実際、図 2.1 の (A)-(C) のネットワークも $k \gg 1$ での $P(k)$ はべき的に振る舞い、俳優共演関係ネットワークでは $\gamma = 2.3$, WWW では $\gamma = 2.3$, インターネットでは $\gamma = 2.0$ となっている。この性質を「スケールフリー性」といい、べき的な次数分布を有するネットワークはスケールフリー・ネットワークと呼ばれる [6]。次数分布 $P(k)$ が 3 以下という低いスケールフリー指数 γ で裾を引くことは、ノードのほとんどが次数の小さなノードである一方で、非常に高い次数をもつノード（ハブ）もまた無視できない割合で存在していることを意味する。低い指数 γ のスケールフリー・ネットワークでは、熱力学極限 $N \rightarrow \infty$ において次数モーメント

$$\langle k^m \rangle = \sum_k k^m P(k), \quad (2.1)$$

がしばしば発散する（条件の詳細は付録 A を参照）。これによって後述する 2.3 節の Pearson 相関係数の下限値や、3 章のパーコレーション転移点、同期現象・感染症拡散現象の転移点がゼロに収束する。このことは、スケールフリー・ネットワークの大きな特徴の一つである。ただし、次数モーメントの発散は $k \rightarrow \infty$ においても次数分布が $k^{-\gamma}$ に比例する場合であり、取り得る次数に事実上の上限がある場合は、べき分布に上限 κ をおき、 $P(k) \propto k^{-\gamma} e^{-k/\kappa}$ とすることもある。上限 κ が有限であるとき次数モーメントは発散しない。(2.1) 式で $m = 1$ とした $\langle k \rangle$ は平均次数であり、 $P(k)$ の形に依らず

$$\langle k \rangle = \frac{2M}{N}, \quad (2.2)$$

で与えられる。

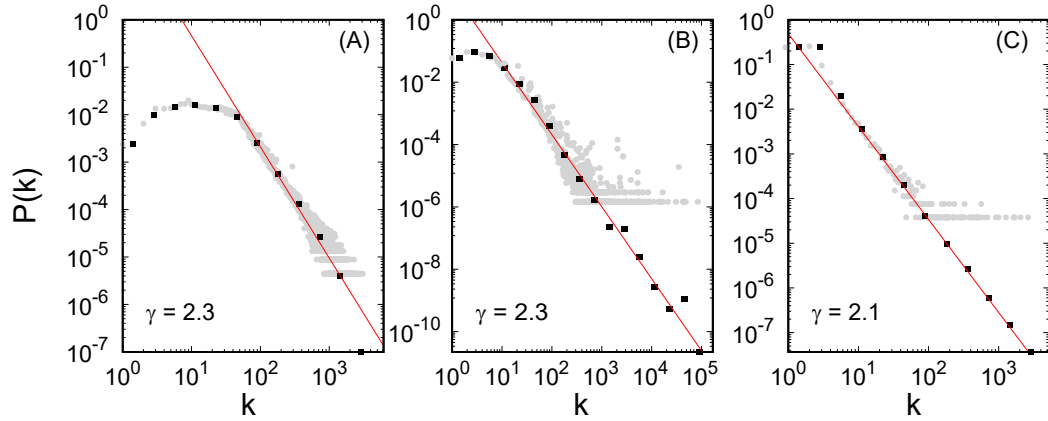


図 2.1: 現実世界のスケールフリー・ネットワークの次数分布. 灰色の丸は次数分布 $P(k)$ の値, 黒の四角は k の或る区間で平均化された $P(k)$ の値を表す. 赤の直線の傾きがスケールフリー指数 γ である. 左から (A) は俳優共演関係 ($N = 248,243$, $M = 8,302,734$), (B) は WWW ($N = 685,230$, $M = 6,649,470$), (C) は AS 単位のインターネット ($N = 26,475$, $M = 53,381$) の次数分布を表す.

スケールフリー性の起源は, 時間経過とともにノード数が増加する「成長」と, 次数の大きいノードがより多くの次数を獲得しやすくなる「優先選択」によって説明できる. このことは, これらのルールを支柱に据えた数理モデルである Barabási-Albert (BA) モデル [6] によって実証される. BA モデルの形成アルゴリズムは以下の通りである.

1. 時刻 $t = 0$ の初期状態として, ノード数 $N = m_0$ の単連結なネットワーク $\mathcal{G}(t = 0)$ を用意する.
2. 各時刻 t において, ネットワーク $\mathcal{G}(t)$ に m 本のエッジをもつノードを 1 つ加える. m 本のエッジのつながる先はネットワーク $\mathcal{G}(t)$ のノードから, その次数に比例した確率 $\Pi_i = k_i / \sum_j k_j$ にしたがってランダムに選ぶ.
3. 時刻を $t \rightarrow t + \delta t$ とする.
4. 所望のネットワーク・サイズに成長するまで 2 と 3 を繰り返す.

BA モデルの次数分布関数は連続体近似や, マスター方程式などを解くことによって解析的に導出できる [6, 64]. 結果のみを述べると, BA モデルの次数分布関数は高次数領域で $P(k) \propto k^{-3}$ となる. BA モデルのスケールフリー指数は $\gamma = 3$ 以外の値をとり得ないが, BA モデルの考え方を拡張したモデルは多数存在し, 特に Dorogovtsev ら [64] が提案するモデルは, 新たなパラメータ k_0 ($-m < k_0 < \infty$) を加えた優先選択確率 $\Pi_i = (k_i + k_0) / (\sum_j k_j + k_0)$ を採用することによって BA モデルと全く同じアルゴリズムで $\gamma = 3 + k_0/m$ ($2 < \gamma < \infty$) のスケールフリー・ネットワークを生成することができる.

2.3 次数相関

スケールフリー・ネットワークをはじめとして、大きな次数揺らぎをもつネットワークにおいては、次数どうしの相関関係（次数相関）に基づくネットワークの分類が大きな意味をもつ。中でも、隣接するノードの次数どうしの相関関係（隣接次数相関）についてはこれまで精力的な研究が行われてきた。その結果、社会的ネットワークにおいては同次数ノードどうしが隣接しやすくなる「正の隣接次数相関」が現れる傾向があり、工学的または生物学的ネットワークにおいては異なる次数のノードどうしが隣接しやすくなる「負の隣接次数相関」が現れる傾向があることが知られている [12]。また、隣接次数相関の正負がコミュニティ構造やフラクタル性といったネットワークの大域的な性質と密接な関係にあることなども明らかになっている [23, 25–28]。本論文の主題である長距離次数相関は、隣接次数相関を拡張とした概念であるため、以下で少し詳しく隣接次数相関に関する一般論を説明する。

ネットワークの隣接次数相関に関する情報は、隣接する 2 ノードの次数 k と k' を確率変数とした同時確率 $P_{\text{nn}}(k, k')$ 、および条件付き確率 $P_{\text{nn}}(k|k')$ に集約されている。エッジ数 M をもつネットワークにおいて、エッジの端点は $2M$ 個ある。同時確率 $P_{\text{nn}}(k, k')$ は、任意抽出したエッジの片端にあるノードの次数が k であり、かつ同じエッジの反対側の端にあるノードの次数が k' である同時確率として定義される。以下に同時確率 $P_{\text{nn}}(k, k')$ の一般的性質を述べる。定義から明らかなように同時確率 $P_{\text{nn}}(k, k')$ は k と k' に関して対称な確率であり、 $P_{\text{nn}}(k, k') = P_{\text{nn}}(k', k)$ の関係が成り立つ。また、任意抽出したエッジの片端の次数が k である確率 $Q_{\text{nn}}(k)$ は、 $P_{\text{nn}}(k, k')$ を用いて、

$$Q_{\text{nn}}(k) = \sum_{k'} P_{\text{nn}}(k, k'), \quad (2.3)$$

と表される。全ノード数が N であるネットワークにおいて、次数 k のノードの数は $N_k = NP(k)$ であり、次数 k のノードを端点とするエッジの片割れ（スポーク）の総数は $kN_k = kNP(k)$ 本と考えられる。全エッジ数が M のときスポークは全部で $2M$ 本あるため、確率 $Q_{\text{nn}}(k)$ は $P(k)$ を用いて、

$$Q_{\text{nn}}(k) = \frac{kNP(k)}{2M} = \frac{kP(k)}{\langle k \rangle}, \quad (2.4)$$

と求められる。ここで、(2.2) 式を用いた。ここで、エッジ両端の次数間に相関が全く無いネットワークに対する同時確率の関数形を考えることにしよう。このとき、エッジの片端の次数が k である事象と、もう一方の片端の次数が k' である事象は互いに独立である。したがって隣接次数相関が無いネットワークの同時確率 $P_{\text{nn}}(k, k')$ は、

$$P_{\text{nn}}(k, k') = Q_{\text{nn}}(k)Q_{\text{nn}}(k') = \frac{kk'P(k)P(k')}{\langle k \rangle^2}, \quad (2.5)$$

のように $Q_{nn}(k)$ と $Q_{nn}(k')$ の単純な積で表される．上記の式を満たすネットワークを隣接無相関ネットワークと呼ぶ．

隣接次数相関を完全に記述するもう一つの確率関数として条件付き確率 $P_{nn}(k'|k)$ がある． $P_{nn}(k'|k)$ は任意抽出したエッジの片端にあるノードの次数が k であるという条件の下での、エッジの反対側の端にあるノードの次数が k' である条件付き確率である．任意抽出したエッジの片端にあるノードの次数が k である確率が $Q_{nn}(k)$ であるから、 $P_{nn}(k, k')$ と $P_{nn}(k'|k)$ の間には、

$$P_{nn}(k, k') = Q_{nn}(k)P_{nn}(k'|k) = \frac{kP(k)}{\langle k \rangle} P_{nn}(k'|k), \quad (2.6)$$

の関係が成り立つ．隣接無相関ネットワークにおいては (2.5) 式が成り立つので、(2.6) 式より、

$$P_{nn}(k'|k) = Q_{nn}(k') = \frac{k'P(k')}{\langle k \rangle}, \quad (2.7)$$

となる [30]．

隣接次数相関は、同程度の次数のノードどうしが隣接しやすい正の隣接次数相関と、異なる次数のノードどうしが隣接しやすい負の隣接次数相関に大別することができる．確率 $P_{nn}(k, k')$ と $P_{nn}(k'|k)$ は、隣接次数相関に関する全ての情報を保持してはいるものの、2変数関数であるから直接的に次数相関の内容を把握することは難しい．したがって多くの場合は変数 k と k' のどちらか、あるいはその両方の縮約が行われる．よく用いられる値は、次数 k のノードの隣接ノードの平均次数（平均隣接次数）、

$$k_{nn}(k) = \sum_{k'} k' P_{nn}(k'|k), \quad (2.8)$$

である．隣接無相関ネットワークに対しては (2.7) 式が成立するので、 $k_{nn}(k) = \langle k^2 \rangle / \langle k \rangle$ となり、 k に依存しない一定値をとることになる．それゆえ、関数 $k_{nn}(k)$ が k の増加関数であるか減少関数であるかによって、隣接次数相関の正負が判断できる．図 2.2 は図 2.1 と同じネットワークの平均隣接次数を表す．幾つかの複雑ネットワークでは、平均隣接次数がベキ関数 $k_{nn}(k) \propto \text{const} + k^\alpha$ の形をとる．このとき指数 α は隣接次数相関の正負とその強さを表す指標としての役割を果たす．ただし、図 2.2(B) から明らかなように、現実世界の複雑ネットワークに対して平均隣接次数 $k_{nn}(k)$ は必ずしも単調増加とは限らず、ベキ関数で表されるという保証も無い．そこで更なる変数の縮約を行い、エッジ両端のノードの次数間の相関係数によって隣接次数相関の正負とその強さを表す方法が主流となっている．

Newman [11] はエッジの両端の次数間の Pearson 相関係数 r を隣接次数相関の指標として提案した．Pearson 相関係数は、2つの確率変数がどれだけ線形関係に近いかを表す統計

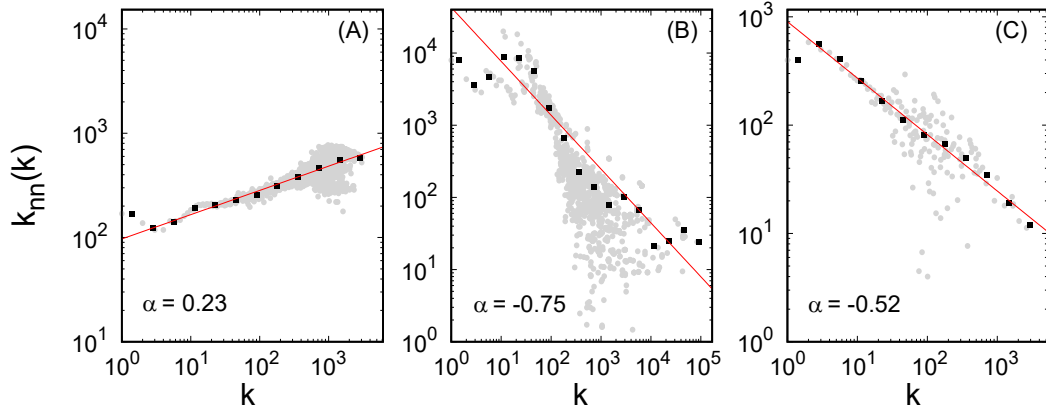


図 2.2: 現実世界の複雑ネットワークの平均隣接次数. 灰色の丸は平均隣接次数 $k_{nn}(k)$ の値, 黒の四角は k の或る区間で平均化された $k_{nn}(k)$ の値を表す. 赤の直線の傾きは $k_{nn}(k) \propto \text{const} + k^\alpha$ としたときのべき指数 α を表す. 左から (A) は俳優共演関係 ($N = 248,243$, $M = 8,302,734$), (B) は WWW ($N = 685,230$, $M = 6,649,470$), (C) は AS 単位のインターネット ($N = 26,475$, $M = 53,381$) の平均隣接次数を表す.

的指標であり, 共分散をそれぞれの確率変数の分散の積の平方根で割った値として定義されている. エッジ両端のノード i および j の次数 k_i および k_j を確率変数とすると, Pearson 相関係数は,

$$r = \frac{\frac{1}{M} \sum_{(i,j) \in E} k_i k_j - \left\{ \frac{1}{2M} \sum_{(i,j) \in E} (k_i + k_j) \right\}^2}{\frac{1}{2M} \sum_{(i,j) \in E} (k_i^2 + k_j^2) - \left\{ \frac{1}{2M} \sum_{(i,j) \in E} (k_i + k_j) \right\}^2}, \quad (2.9)$$

によって与えられる. なお, $\sum_{(i,j) \in E}$ はエッジ集合 E に対する和を表す. 上記の式は $P_{nn}(k, k')$ を用いて,

$$r = \frac{\sum_{k,k'} k k' P_{nn}(k, k') - \left\{ \sum_k k Q_{nn}(k) \right\}^2}{\sum_k k^2 Q_{nn}(k) - \left\{ \sum_k k Q_{nn}(k) \right\}^2}, \quad (2.10)$$

と表すこともできる. $Q_{nn}(k)$ は次数分布だけで決まるため, (2.10) 式の中で実際に隣接次数相関を反映しているのは, エッジの端点ノードの次数の積の期待値 $\sum_{k,k'} k k' P_{nn}(k, k')$ のみである. 以後この値を $[kk']_E$ と書くこととする. ここで (2.6) 式と (2.8) 式より,

$$[kk']_E = \sum_k \frac{k^2 P(k)}{\langle k \rangle} k_{nn}(k), \quad (2.11)$$

であるから, Pearson 相関係数 r は,

$$r = \frac{\sum_k k_{nn}(k) k^2 P(k) - \langle k^2 \rangle^2 / \langle k \rangle}{\langle k^3 \rangle - \langle k^2 \rangle^2 / \langle k \rangle}, \quad (2.12)$$

と表わされる. Pearson 相関係数は, 同程度の次数のノードどうしが隣接しやすいときには最大値を $r_{\max} = 1$ とする正の値をとり, 異なる次数のノードどうしが隣接しやすいときに

表 2.1: 現実のネットワークのノード数 N と, 隣接次数相関に関する Pearson 相関係数 r の値 [12].
 r の正負が隣接次数相関の正負に対応する.

分類	現実のネットワーク	ノード数 N	Pearson 相関係数 r
社会	共著関係 (物理学)	52 909	0.363
	共著関係 (生物学)	1 520 251	0.127
	共著関係 (数学)	253 339	0.120
	俳優共演関係	449 913	0.208
	会社役員	7 672	0.276
	学生の交際関係	573	-0.029
	E メールアドレス帳	16 881	0.092
工学	電力網	4 941	-0.003
	インターネット	10 697	-0.189
	WWW	269 504	-0.067
	ソフトウェア	3 162	-0.016
生物	たんぱく質相互作用	2 115	-0.156
	代謝ネットワーク	765	-0.240
	ニューラルネットワーク	307	-0.226
	海洋食物網	134	-0.263
	淡水食物網	92	-0.326

は最小値を $r_{\min} = -1$ とする負の値をとる. 現実のネットワークにおける Pearson 相関係数 r の値を表 2.1 にまとめておく. 表 2.1 からわかる通り, 隣接次数相関は社会的ネットワークにおいて正となりやすく, 工学的ネットワークおよび生物学的ネットワークにおいて負となりやすい.

表 2.1 において、負の隣接次数相関を示すネットワークのネットワーク・サイズ N に注目すると、 N が大きいネットワークほど r の絶対値が小さくなる傾向が見られる。この傾向は (2.10) 式より得られる、

$$r = \frac{[kk']_E \langle k \rangle^2 / \langle k^2 \rangle^2 - 1}{\langle k^3 \rangle \langle k \rangle / \langle k^2 \rangle^2 - 1} > \frac{-1}{\langle k^3 \rangle \langle k \rangle / \langle k^2 \rangle^2 - 1}, \quad (2.13)$$

の不等式を用いて解析的に説明することができる。次数分布関数 $P(k) \propto k^{-\gamma}$ をもつネットワーク・サイズ N のスケールフリー・ネットワークに対し、連続体近似の下で評価される上式右辺の分母第一項は、

$$\frac{\langle k^3 \rangle \langle k \rangle}{\langle k^2 \rangle^2} \propto \begin{cases} N^0, & (\gamma > 4) \\ \log N, & (\gamma = 4) \\ N^{(4-\gamma)/(\gamma-1)}, & (3 < \gamma < 4) \\ \frac{N^{\frac{1}{2}}}{(\log N)^2}, & (\gamma = 3) \\ N^{(\gamma-2)/(\gamma-1)}, & (2 < \gamma < 3) \\ \log N, & (\gamma = 2) \\ N^0, & (\gamma < 2) \end{cases} \quad (2.14)$$

となる。この結果より、 $2 \leq \gamma \leq 4$ のスケールフリー・ネットワークでは、 $N \rightarrow \infty$ の極限において (2.13) 式右辺の分母が発散し、 r は 0 以上の値しか取れなくなることがわかる。それゆえ、いかに明確に高次数ノードと低次数ノードが隣接しやすいネットワークに対しても、ネットワーク・サイズ N が大きくなると Pearson 相関係数は負になり得ないことになる。また、(2.14) 式から明らかなように、 r は N に依存するようになる。そのため、サイズ N の異なるネットワークの隣接次数相関の強さを r によって比較することができないという問題が生じる。現実ネットワークのスケールフリー指数が $2 \lesssim \gamma \lesssim 3$ の範囲にあることを考えると、Pearson 相関係数のこの問題点は隣接次数相関の指標として致命的である [65, 66]。

この問題を解決するために、Spearman 相関係数 ρ という指標が提案されている [17, 67]。Spearman 相関係数は、確率変数の順位どうしの Pearson 相関係数として定義される。したがって、隣接次数相関の指標としての Spearman 相関係数 ρ は、(2.9) 式における次数 k_i と k_j を次数の順位 r_{k_i} と r_{k_j} に置き換えることによって、

$$\rho = \frac{\frac{1}{M} \sum_{(i,j) \in E} r_{k_i} r_{k_j} - \left\{ \frac{1}{2M} \sum_{(i,j) \in E} (r_{k_i} + r_{k_j}) \right\}^2}{\frac{1}{2M} \sum_{(i,j) \in E} (r_{k_i}^2 + r_{k_j}^2) - \left\{ \frac{1}{2M} \sum_{(i,j) \in E} (r_{k_i} + r_{k_j}) \right\}^2}, \quad (2.15)$$

で与えられる量となる。Pearson 相関係数の下限値を表す (2.13) 式に対応して、Spearman

相関係数の下限値を表す不等式は,

$$\rho \geq \frac{-1}{\frac{\frac{1}{2M} \sum_{(i,j) \in E} (r_{k_i}^2 + r_{k_j}^2)}{\left[\frac{1}{2M} \sum_{(i,j) \in E} (r_{k_i} + r_{k_j}) \right]^2} - 1}, \quad (2.16)$$

で与えられる. エッジ数 M が十分大きいとき,

$$\frac{1}{2M} \sum_{(i,j) \in E} (r_{k_i}^2 + r_{k_j}^2) \propto M^2 \quad (M \gg 1), \quad (2.17)$$

$$\frac{1}{2M} \sum_{(i,j) \in E} (r_{k_i} + r_{k_j}) \propto M \quad (M \gg 1), \quad (2.18)$$

となるため, (2.15) 式右辺の分母はネットワーク・サイズに依存しない定数となる. 本論文では, Spearman 相関係数 ρ を隣接次数相関を表す指標として用いることとする. 次数の順位付けの方法はいくつか提案されているが, 本論文では次数 k のノード全体から出るスポークの総数 $kN_k = kNP(k)$ を用いて,

$$r_k = \sum_{k'=k_{\min}}^{k-1} k'N_{k'} + \frac{1 + kN_k}{2}, \quad (2.19)$$

を次数 k の順位 r_k と定義する. ここで N_k は次数 k のノードの個数である. この順位付けの方法を用いた Spearman 相関係数は,

$$\rho = \frac{\frac{1}{M} \sum_{(i,j) \in E} r_{k_i} r_{k_j} - \left(\frac{2M+1}{2} \right)^2}{\frac{1}{2M} \sum_{(i,j) \in E} (r_{k_i}^2 + r_{k_j}^2) - \left(\frac{2M+1}{2} \right)^2}, \quad (2.20)$$

となる.

2.4 スモールワールド性

隣接するノード対は1本のエッジで仲介されるが, 隣接ノードの隣接ノードもまた2本のエッジで仲介されると言える. ノード対間を仲介するエッジの組み合わせを経路といい, 経路に含まれるエッジの本数を経路長という. 多くの場合, ノード対間には複数の経路が考えられるが, 一般にノード i とノード j の間の距離 l_{ij} は最短経路の長さを指す. なお, 同じノードの間の距離 ($i = j$ の場合) は $l_{ij} = 0$ と定義する. ノード対間に経路が存在するとき, これらのノード対は「連結である」といい, 互いに連結するノードの集合を連結成分という. 一つのネットワーク内で最も多くのノードを含む連結成分をそのネットワークの最大連結成分といい, その大きさはしばしばネットワーク・サイズ N に対する相対サイズ S に

よって表される．全ノードが一つの連結成分となる ($S = 1$ となる) ネットワークを単連結ネットワークと呼ぶ．単連結ネットワークでは，全てのノード対に渡るノード間距離の平均である平均ノード間距離

$$\langle l \rangle = \frac{2}{N(N-1)} \sum_{i>j} l_{ij}, \quad (2.21)$$

が定義できる．単連結でない場合は，ノード間距離を定義できるノード対のみに渡る平均をとることが多い．ネットワークの差し渡しの長さは，平均ノード間距離 $\langle l \rangle$ または最長のノード間距離 $l_{\max}$ (ネットワーク直径ともいう) によって特徴づけられる．

ノード集合 V とエッジ集合 E からなるグラフ $\mathcal{G}(V, E)$ に対して， V と E の部分集合 $V' \subseteq V$ ， $E' \subseteq E$ から成るグラフ $\mathcal{G}'(V', E')$ を $\mathcal{G}(V, E)$ の部分グラフという．特に，1 次元的にノードが連なる $|V'| = n+1$ ， $|E'| = n$ の部分グラフを長さ n の経路 (n -path) といい， $(n-1)$ -path の端点ノードどうしにエッジが張られた $|V'| = |E'| = n$ の部分グラフを長さ n のサイクル (n -cycle) という．

いくつかの現実のネットワークには，共通の隣接ノードをもつノードどうしが隣接しやすくなる傾向が見られる．こうした 3 体の局所的相関を特徴づけるために，3-cycle の多寡を 2-path と比較して定量化する指標が 2 種類存在する．本論文では平均クラスター係数 C ，大域クラスター係数 C' と呼び分けることとするが，多くの場合はどちらも同じくクラスター係数と呼ばれる．平均クラスター係数 C はノードごとの局所的なクラスター係数 C_i の全ノードに渡る平均値である． C_i はノード i の隣接ノードどうしに張られるエッジの本数 M_i を用いて，

$$C_i = \frac{2M_i}{k_i(k_i - 1)}, \quad (2.22)$$

と定義される．大域クラスター係数 C は社会学では推移性とも呼ばれ

$$C' = \frac{3 \times (\text{ネットワークに含まれる 3-cycle の数})}{(\text{ネットワークに含まれる 2-path の数})}, \quad (2.23)$$

によって定義される [68]．ただし，ネットワークにおける多体の局所的相関は 3 体に限らない． d 次元立方格子などでは 3-cycle が存在しない代わりに 4-cycle が数多く存在する．4-cycle が数多く存在するネットワークの例としては，論文と研究者をどちらをノードとしても定義しうる共著関係ネットワークのように，ノードの種類が 2 種類あるようなネットワークが挙げられる．したがって，一般の n に対する n -cycle の多寡を定量化する値であるサイクル係数 [69] もしばしば用いられる．

平均ノード間距離 $\langle l \rangle$ とネットワーク・サイズ N の間に，

$$\langle l \rangle \propto \log N, \quad (2.24)$$

表 2.2: 現実世界のスケールフリー・ネットワークの例と、そのネットワークが有する特徴量（ネットワーク・サイズ N ，平均次数 $\langle k \rangle$ ，スケールフリー指数 γ ，平均ノード間距離 $\langle l \rangle$ ，平均クラスター係数 C ）の値．参考として、同じ次数列をもつ以外には完全ランダムにノード間にエッジが張られるネットワーク（一般化ランダムグラフともいう [68, 70, 71]）における平均ノード間距離 $\langle l \rangle_{\text{RG}}$ と平均クラスター係数 C_{RG} を併記する．値は文献 [4] より引用．

ネットワーク	N	$\langle k \rangle$	γ	$\langle l \rangle$	C	$\langle l \rangle_{\text{RG}}$	C_{RG}	文献
WWW	203,549,046	10.5	2.1	16.18	0.11	8.40	5.14×10^{-8}	[73]
共著関係	1,520,251	18.1	2.5	4.6	0.60	5.20	1.19×10^{-5}	[74]
俳優共演関係	449,913	113.4	2.3	3.48	0.78	3.01	2.52×10^{-4}	[75]
単語同時出現	478,773	74.2	2.7	2.63	0.687	3.41	1.54×10^{-4}	[76]
ソフトウェア	1,989	3.98	2.85	3.2	0.08	5.38	2.00×10^{-3}	[77]
航空路線網	3,880	9.7	1.8	4.37	0.3	3.85	2.50×10^{-3}	[78]

の関係が成り立ち、さらに平均クラスター係数 C と N の間にも、

$$C \propto N^0, \quad (2.25)$$

の関係が成り立つ性質をスモールワールド性という [79]．表 2.2 からわかる通り、スケールフリー性とならび、スモールワールド性もまた多くの現実ネットワークに見られる性質である．広義には C に関する (2.25) 式の性質を無視し、2.5 節にて後述するフラクタル性に相反する (2.24) 式の性質だけをスモールワールド性と呼ぶこともある．ここでは (2.24) 式と (2.25) 式を共に満たす、狭義のスモールワールド性について述べる．

全てのノード対に一定確率でエッジが存在するランダム・ネットワーク（Erdős-Rényi ランダムグラフ [72]）では、ノード数 N ，平均次数 $\langle k \rangle > 1$ に対して $\langle l \rangle$ と C は、

$$\langle l \rangle \approx \frac{\log N}{\log \langle k \rangle}, \quad (2.26)$$

$$C = \langle k \rangle / N, \quad (2.27)$$

となることが知られている．(2.24) 式は成立しているが (2.25) 式が成り立たないため、Erdős-Rényi ランダムグラフはスモールワールド性をもたないことが判る．一方で、単位胞に三角形を含む d 次元立方格子のような完全に規則的なネットワークでは、明らかに、

$$\langle l \rangle \propto N^{1/d}, \quad (2.28)$$

$$C \propto N^0, \quad (2.29)$$

である．それゆえ、今度は (2.25) 式は成立しているものの (2.24) 式が成立していないため、規則的ネットワークもスモールワールドではない．では、どのようなネットワークにおいて

(2.24) 式と (2.25) 式が両立するのだろうか。1998 年に提案された Watts-Strogats (WS) モデル [79] はスモールワールド性の起源が「規則性」と「わずかながらの不規則性」にあることを示唆する。WS モデルによるネットワークは、ノード数 N 、平均次数となる偶数 $\langle k \rangle$ 、エッジのつながぎ替え確率 p をパラメータとして、以下のアルゴリズムから作られる。

1. 円環上に N 個のノードを並べる。
2. 各ノードと左右 $\langle k \rangle/2$ -近傍までのノードの間にエッジを張る。
3. 全てのエッジに対して、つながぎ替え確率 p に従い、片側のノードを全ノードの中からランダムに選びなおす。

図 2.3 の左の図は p の値を変化させたときのネットワークの構造を表す。 $p = 0$ ではエッジのつながり方が規則的であり、 $p = 1$ におけるネットワークはほとんど元の構造を残さないランダムなネットワークとなる。右図は平均ノード間距離 $\langle l \rangle$ (図中では $L(p)$) とクラスター係数 $C(p)$ の p 依存性を表している。右図より $p \sim 0.01$ の領域におけるネットワークは明らかに大きな C と小さな $\langle l \rangle$ を両立しており、スモールワールド性の定義式 (2.24), (2.25) を満足することが期待できる。実際、このモデルにおいて $C(p)$ の p 依存性は、比較的簡単に計算できる。 $C(p = 0)$ に存在する 3-cycle がつながぎ替えによって破壊されない確率が $(1 - p)^3$ であるから、

$$C = \frac{2\langle k \rangle - 6}{4\langle k \rangle - 4}(1 - p)^3 \propto N^0, \quad (2.30)$$

となり、 $p \approx 1$ の場合を除き (2.25) 式を満足する。 $\langle l \rangle(p)$ の計算は $C(p)$ に比べて複雑であるためここでは説明しないが、結論だけ述べると、

$$\langle l \rangle \approx \begin{cases} \frac{N}{2\langle k \rangle}, & \left(\frac{N\langle k \rangle p}{2} \ll 1 \right), \\ \frac{\log(N\langle k \rangle p)}{\langle k \rangle^2 p}, & \left(\frac{N\langle k \rangle p}{2} \gg 1 \right), \end{cases} \quad (2.31)$$

となることが知られている。 $N\langle k \rangle p/2 = pM$ であることを踏まえると、上の式からは「つながぎ替えられるエッジの本数が 1 のオーダーを上回るとき」に (2.24) 式が満たされることになる。以上から、 $1 \gg p \gg 1/M$ のとき WS モデルのネットワークはスモールワールド性を示すことが確かめられる。図 2.3 の左図の中央は、パラメータ p がこの領域にあるときのネットワークである。図からもわかる通り、 $1 \gg p \gg 1/M$ 領域におけるネットワークが規則的な構造を保ちながらもわずかに不規則なエッジ (ショートカット・エッジ) を有するネットワークであることから、スモールワールド性の起源が「規則性」と「わずかながらの不規則性」であることが理解できる。

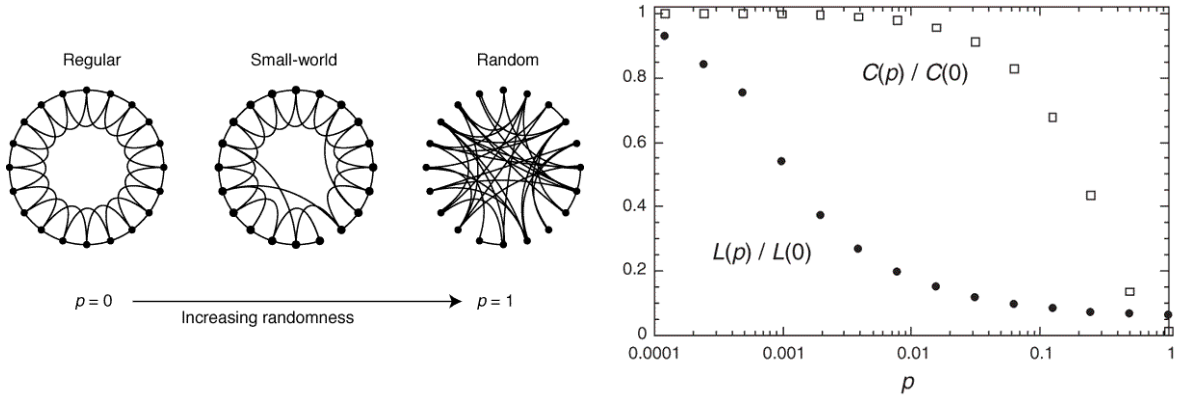


図 2.3: WS モデルにおいて、つなぎ替え確率 p に伴って変化するネットワークの構造（左）と、平均ノード間距離 $\langle l \rangle$ （図中では $L(p)$ ）およびクラスター係数 $C(p)$ の p 依存性（右）．右図の数値計算結果は、ノード数 $N = 1,000$ 、平均次数 $\langle k \rangle = 10$ のネットワーク 20 サンプルに対する平均値である．図は文献 [79] より転用．

2.5 フラクタル性

ネットワークに限らず、フラクタルな系とは一般に特徴的な長さスケールの存在しない系を指す．ユークリッド空間に埋め込まれたフラクタル構造には特徴的なユークリッド距離が存在しない．これに対してネットワークにおけるフラクタル性とは特徴的なノード間距離が存在しない性質を指す．一般にネットワーク・サイズ N はネットワーク直径 $l_{\max}$ または平均ノード間距離 $\langle l \rangle$ の増加関数となるが、フラクタル・ネットワークにおいてこの関数形に特徴的なノード間距離が含まれることは許されない．したがって、ネットワーク・サイズ N と平均ノード間距離 $\langle l \rangle$ の間の関係は、特徴的スケールをもたない唯一の関数形であるべき関数によって表されなければならない．すなわち、フラクタル・ネットワークとは、

$$\langle l \rangle \propto N^{1/d_f}, \quad (2.32)$$

の関係式が成り立つネットワークと定義される．(2.32) 式の代わりにネットワーク直径 $l_{\max}$ と N の間の同様の関係によってフラクタル性が定義されることもあるが本質的な違いは無い．指数 d_f をネットワークのフラクタル次元という．ここで 2.4 節で述べたように、スモールワールド性を有するネットワークでは、

$$N \propto e^{\langle l \rangle / l_e}, \quad (2.33)$$

となるため、明らかに特徴的スケール l_e を含む．このように、広義の意味でのスモールワールド性はフラクタル性と相反する性質である．現実世界の多くのネットワークは、距離とネットワーク・サイズの関係からスモールワールド的なネットワークとフラクタル的なネット

ワークの2種に大別される．現在までに知られているフラクタル・ネットワークには，生物系の遺伝子制御ネットワークや代謝ネットワーク，俳優共演関係，WWW などがある [80–82]．

ネットワークがフラクタル性を有するか否かを判定する代表的な方法に，クラスター成長法とボックス・カウンティング法がある．クラスター成長法の考え方では，あるノードからノード間距離 l_C 以内にある平均的なノード数 $N_C(l_C)$ が，

$$N_C(l_C) \propto l_C^{d_f}, \quad (2.34)$$

を満たす場合，ネットワークはフラクタル性を有することになり， l_C の指数関数で表される場合はスモールワールド性を有することになる．一方，ボックス・カウンティング法では，ネットワークの被覆に基づいてフラクタル性とスモールワールド性を区別する．直径 l_B の部分グラフ（ボックス）でグラフ全体を被覆する際に最小限必要となるボックス数 $N_B(l_B)$ が

$$N_B(l_B) \propto l_B^{-d_f}, \quad (2.35)$$

を満たすならば，ネットワークはフラクタルである．これに対し， N_B が l_B とともに指数関数的に減少するならばスモールワールド性を有することになる．クラスター成長法とボックス・カウンティング法ではフラクタル次元 d_f の定義が異なるため，それぞれの方法で求めた d_f が同じ値をとるとは限らない．しかしながら，特殊な場合を除きほとんどの複雑ネットワークにおいてクラスター成長法とボックス・カウンティング法から求められるフラクタル次元 d_f は同じ値となる [83]．ボックス・カウンティング法でフラクタル解析をする際に問題となるのは，グラフを被覆するボックスの最小数をどのように見つけ出すかである．これまでグラフの最小被覆のために様々なアルゴリズムが提案されてきている [83, 84]．

2.5.1 現実世界のフラクタル・スケールフリー・ネットワーク

ごく少量のショートカットエッジの存在によりネットワークがスモールワールド化するという事実は，フラクタル的なネットワークが現実世界において安定的に存在することの困難さを示唆する．しかしこの認識は，複雑ネットワークのフラクタル性が発見された 2005 年当時において，「複雑ネットワークにおけるフラクタル性の有無は大域的スケールである平均ノード間距離 $\langle l \rangle$ がノード数のべき乗に比例するか否かによって判別される」と考えられていたことによる．実際は，一つのネットワーク内でも注目するスケールによってノード間距離とノード数の関係は変化する．相関長以下のスケールにおける平均ノード間距離 $\langle l \rangle$ はノード数のべき乗 N^{1/d_f} に比例する一方，相関長より大きいスケールにおいて対数 $\log N$ に比例するという構造的クロスオーバー現象が生じる [83]．したがって，現実ネットワーク

においても十分小さいスケールにおいてはフラクタル性が認められることは少なくない。表 2.3 にまとめられたネットワークは、スケールによってフラクタル性とスモールワールド性のどちらの性質も有する現実世界の複雑ネットワークの例である。

2.5.2 臨界ネットワークのフラクタル性とスケールフリー性

複雑ネットワークにおけるフラクタル性の起源は未だ完全には明らかになっていない。ネットワーク上のダイナミクスとネットワーク構造の間の相互作用による自己組織化臨界性からのアプローチも試みられてはいる [50]。また、Erdős-Rényi ランダムグラフや一般化ランダムグラフをはじめとするランダム・ネットワークにおいて、臨界確率 $f_c = 1 - \langle k \rangle / (\langle k^2 \rangle - \langle k \rangle)$ でノード（またはエッジ）を削除したとき、大域連結成分が木構造のフラクタル・ネットワークとなることが、平均場近似を用いた解析計算によって明らかにされている [87–89]。ノード（またはエッジ）を確率 p で占有させるサイト（またはボンド）パーコレーションにおいて、最大連結成分の次数分布は元のネットワークの次数分布 $P(k)$ を用いて

$$P_{GCC}(k) = \frac{P(k)(1 - x^k)}{1 - \sum_k P(k)x^k}, \quad (2.36)$$

と与えられる [90]。ここで x はあるエッジをたどった先が最大連結成分に到達しない確率であり、臨界点より上で $x \rightarrow 1$ であるため $P_{GCC}(k)$ はおよそ $kP(k)$ に比例する。したがって元のネットワークの次数分布が $P(k) \propto k^{-\gamma}$ であるならば、最大連結成分はスケールフリー指数 $\gamma - 1$ のフラクタル・スケールフリー・ネットワークとなる。 $\gamma < 3$ の場合は、ランダム・ネットワークのパーコレーション臨界点は熱力学極限で 0 に収束し、臨界点近傍における最大連結成分の相対サイズも 0 となってしまうため、最大連結成分の性質について議論することができない。 $\gamma > 3$ におけるフラクタル次元は、連結成分のノード数 n と平均ノード間距離 $\langle l \rangle$ の関係から $\gamma > 4$ のとき $d_f = 2$ 、 $4 > \gamma > 3$ のとき $d_f = (\gamma - 2)/(\gamma - 3)$ である。

2.5.3 決定論的フラクタル・スケールフリー・ネットワーク・モデル

現実のスケールフリー・ネットワークにおいてどのようにフラクタル性が発現するのかは未だ明らかになっていないが、入れ子構造となるようにネットワークを成長させることで、フラクタル性を有するスケールフリー・ネットワークを作成できる。最も有名なフラクタル・スケールフリー・ネットワーク (FSN) モデルである (u, v) -flower [91] と Song-Havlin-Makse (SHM) モデル [80] によって生成されるネットワークは、そのようなネットワークの典型例である。例えば、整数 u と v ($1 \leq u \leq v$) をパラメータとする (u, v) -flower では、 $u = 1$ のときスモールワールド性を、 $u \geq 2$ のときフラクタル性を示すスケールフリー・ネットワークが生成される。 (u, v) -flower のアルゴリズムは以下の通りである。

表 2.3: 現実世界の FSN のスケールフリー指数 γ と短いスケール領域におけるフラクタル次元 d_f .
ただし, フラクタル性は単連結なネットワークに対して定義される性質であるため, フラクタル次元は全て最大連結成分に対して計算されたものである. フラクタル次元の計算には Compact-Box-Burning 法 [84] を用いた.

分野	ネットワーク	N_{LCC}	M_{LCC}	γ	d_f	ref.
社会	俳優共演関係 (1)	81,860	3,664,547	1.68	5.11	[85]
	俳優共演関係 (2)	225,226	8,300,988	2.34	4.91	[63]
	E メール	33,696	180,811	1.88	4.67	[85]
	SNS (Petstar)	1,788	12,476	1.95	3.30	[85]
	SNS (Twitter)	22,322	31,823	1.87	3.90	[85]
	ブログ引用関係	1,222	16,714	0.93	4.39	[85]
	論文引用関係	365,154	1,721,981	2.92	3.84	[85]
工学	インターネット (1)	10,515	21,455	2.13	4.66	[86]
	インターネット (2)	26,475	53,381	2.08	4.69	[85]
	ソフトウェア	30,817	213,208	2.23	5.16	[85]
	WWW (1)	654,782	6,581,871	2.31	2.36	[85]
	WWW (2)	255,265	1,941,926	2.28	2.34	[85]
	WWW (3)	325,729	1,090,108	2.47	3.77	[85]
	電力網	4,941	6,594	3.26	2.33	[85]
	航空路線網 (1)	1,572	17,214	1.41	3.93	[85]
	航空路線網 (2)	500	2,980	1.80	3.46	[85]
生物	タンパク質折り畳み構造	132,167	228,967	2.86	2.10	[86]
	代謝系	453	2,025	2.08	3.26	[85]
	PIN (Reactome)	5,973	145,778	1.04	2.71	[85]

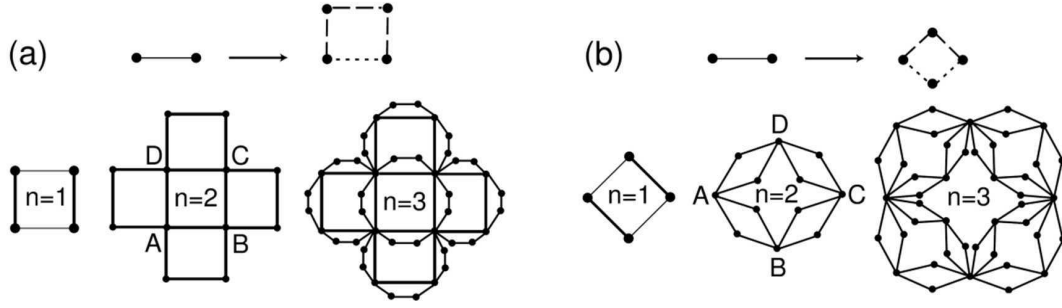


図 2.4: (u, v) -flower の成長の様子. (a) は $(1, 3)$ -flower, (b) は $(2, 2)$ -flower の第 1 世代 ($n = 1$) から第 3 世代 ($n = 3$) までの成長の様子を表している. 図は文献 [91] より転用.

1. ノード数 $w = u + v$ の w -cycle を用意する. これを (u, v) -flower の第 1 世代とする.
2. 第 n 世代の (u, v) -flower の全エッジを, 元のエッジと同じ端点をもつ u -path と v -path に置き換える. これを第 $n + 1$ 世代の (u, v) -flower とする.
3. 操作 2 を必要な世代数に達するまで繰り返す.

図 2.4 は, 上記のアルゴリズムにより作られた $(1, 3)$ -flower と $(2, 2)$ -flower それぞれの第 1 世代から第 3 世代までを表している. (u, v) -flower は, 手順 1 において第 1 世代となる w -cycle や, 手順 2 においてエッジと置き換えられる u -path および v -path をより自由を選ぶことで一般化することができる. 一般化によって作られる決定論的 FSN モデルのアルゴリズムは以下の通りである.

1. 小規模ネットワーク G_0 を第 0 世代のネットワークとする. この G_0 をイニシエータと呼ぶ.
2. 第 r 世代ネットワーク G_r の全エッジを, ジェネレータと呼ばれる小規模ネットワーク G_1 へと置き換えることで第 $r + 1$ 世代ネットワーク G_{r+1} を得る. ただし, G_1 は連結グラフである. G_1 には 2 つのルートノードが指定されており, 2 つのルートノードに関して G_1 は対称である. また, ルートノードの次数 κ は $\kappa \geq 2$ であり, ルートノード間の距離 λ は $\lambda \geq 2$ である. G_r のエッジを G_1 で置き換える際, エッジ両端のノードが G_1 のルートノードに重なるように置き換えるものとする.
3. 手順 2 を十分な回数繰り返す.

このようにして形成されるネットワークは, スケールフリー性とフラクタル性を併せ持つ. そのスケールフリー指数 γ やフラクタル次元 d_f , さらにはクラスター係数や次数相関などの統計的性質もイニシエータ G_0 とジェネレータ G_1 の性質から解析的に計算される (付録 B 参照). また, 既存の FSN モデルである (u, v) -flower [91] や SHM モデル [80] は, 図 2.5

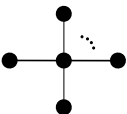
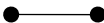
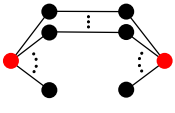
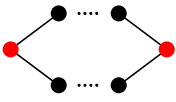
	SHM model	(u,v)-flower
Initiator		
Generator		

図 2.5: (u, v) -flower と SHM モデルは，ジェネレータとイニシエータをそれぞれ表の通りに選んだ場合に相当する．赤のノードはルートノードである．

のようにイニシエータとジェネレータを選ぶことにより一般化された決定論的 FSN モデルの特別な場合として含まれる．

第3章 ネットワーク上の諸現象と隣接次数 相関

次数揺らぎの強いスケールフリー・ネットワークでは、次数どうしの相関関係（次数相関）がネットワークの性質を決める重要な要素となる．特に「隣り合う次数どうしの相関関係」である隣接次数相関は、現実世界のほとんど全ての複雑ネットワークが有している性質であることや、物理的意味が明快でありかつ解析的な扱いも比較的容易であることにより、これまで精力的に研究されてきた．その結果、隣接次数相関がネットワーク上の様々な物理現象に強い影響を与えることが明らかにされている．本章では、それらの中で最も基本的なネットワーク上の物理現象であるパーコレーション転移、同期現象、感染症拡散と隣接次数相関との間の関係について述べる．

3.1 パーコレーション転移と隣接次数相関

単一エッジまたは複数のエッジから成る経路によってつながり合ったノードの集合を連結成分といい、ネットワーク全体に匹敵する大きさの連結成分を巨大連結成分という．また、ネットワーク内に巨大連結成分が存在するという性質を、ネットワークの大域的連結性という．タンパク質相互作用ネットワーク、航空路線網、インターネットなどのネットワークが十分にその機能を果たすことができるのは、ネットワーク構造が大域的連結性を有しているからである．しかしながら、現実世界ではしばしば災害や事故によってノードやエッジの一部に欠損が生じる．このような場合においても大域的連結性が失われにくいネットワークは頑強なネットワークと呼ばれる．頑強性が高ければ、エッジやノードが多少失われたとしてもネットワーク全体が機能停止に陥るような事態にはならない．一方で頑強なネットワークが好ましくない状況も存在する．機密情報の漏洩・拡散や感染症の蔓延は頑強なネットワークでは防ぎにくい．このように、ネットワークの頑強性はネットワーク科学の実学的な役割の根幹を担っている．頑強性の問題を数理的に扱うのがネットワーク上のパーコレーション問題である．ノードやエッジを確率 f で取り除いたときにネットワークが大域的連結性を失うパーコレーション臨界点 f_c の値は、まさに頑強性の定量的指標となる． f_c が大きければネットワークは頑強と考えられる．ネットワーク上のパーコレーション問題を考

えるもう一つの重要な意義は、臨界指数による普遍性クラスへの分類である。 d 次元格子や Bethe 格子上でのパーコレーションと同様に、ネットワーク上でのパーコレーションでは巨大連結成分サイズに関するスケーリング則が表れ、臨界指数が定義できる。系の詳細によらない普遍的な量である臨界指数は、自由度が極めて高い複雑ネットワークのトポロジーとその上で生じる物理現象を大域的な対称性に基づき分類する。ネットワーク・トポロジーがネットワーク上のパーコレーションへ与える影響についてはこれまでもよく研究されており [7, 46, 47, 92, 93], 2 章で紹介したようなネットワークの様々な構造的特徴がパーコレーションの振る舞いに大きな影響を与えることが明らかになっている。その中でも特に重要な構造的特徴が隣接次数相関である。

パーコレーション問題にはランダムにエッジが削除されるボンド・パーコレーション問題とランダムにノードが削除されるサイト・パーコレーション問題の 2 種類がある。ノードまたはエッジの削除割合を高くしていくと、いずれは巨大連結成分が消滅することになる。無限に大きなネットワークにおけるこの変化はパーコレーション転移と呼ばれ、転移点となる削除率の近傍において系は臨界的に振る舞う。どちらのパーコレーション問題に対しても巨大連結成分サイズを計算するための効率的な数値計算手法が確立されている。特定の有限サイズ・ネットワークが与えられたとき、そのネットワークに対する非後退行列の固有値問題を解くことによって巨大連結成分サイズを計算することができる [95]。また、高速かつ高精度な数値計算手法である Newman-Ziff 法 [94] も知られている。しかしながら数値計算では有限系しか扱えないため、転移点 (臨界点) やその周りで定義される臨界指数を厳密に計算することはできない。本章では与えられた次数分布と隣接次数相関を有することを除いて一切の構造的相関をもたない無限に大きなネットワークを考え、そこでのパーコレーション問題を解析的に解くことにする。このようなネットワークの平均構造は短いループをもたない局所木構造であると考えられるため、局所木構造仮定と平均場近似を用いて、臨界点 f_c と臨界指数を解析的に導くことができる。

はじめに簡単のため隣接次数相関が無い場合を考え、次に隣接次数相関の効果を理論に取り入れていくことにする。エッジ (またはノード) の削除確率を f とすると、確率 $p = 1 - f$ はエッジ (またはノード) の占有確率となる。任意にノードを一つ選び、さらにそのノードから出るエッジを一本任意に選んだとき、エッジの先に巨大連結成分が存在しない確率 u を導入する。サイト・パーコレーションにおいては、エッジの (元のノードと反対側の) 先のノードは確率 $1 - p$ で削除されている。ボンド・パーコレーションにおいては、たどろうとしているエッジが確率 $1 - p$ で削除されている。どちらの場合もエッジをたどった先に巨大連結成分は存在しない。また、エッジをたどった先のノード (次数 k のノード i_k とする) が巨大連結成分に属していないことも考えられる。この確率は、 i_k から出る $k - 1$ 本のエッジのいずれの先にも巨大連結成分が存在しない確率に等しく、局所木構造仮定のもとでは u^{k-1}

で与えられる．エッジをたどった先のノードの次数が k である確率は $kP(k)/\langle k \rangle$ であるから，確率 u は自己無撞着方程式

$$u = 1 - p + p \sum_k \frac{kP(k)}{\langle k \rangle} u^{k-1}, \quad (3.1)$$

を解くことによって求められる．図 3.1 は (3.1) 式の解の値を示す概念図である．図より明らかな通り，いかなるエッジの先にも巨大連結成分が存在しないことを意味する自明な解 $u = 1$ が常に存在する．また p の変化によって解 u は連続的に変化するため，ネットワーク上のパーコレーション転移は連続転移である．(3.1) 式の解 u を用いて，あるノードが巨大連結成分に所属する確率 S は

$$1 - S = 1 - p + p \sum_k P(k) u^k, \quad (\text{サイト・パーコレーションの場合}) \quad (3.2)$$

$$1 - S = p \sum_k P(k) u^k, \quad (\text{ボンド・パーコレーションの場合}) \quad (3.3)$$

から求められる．

自己無撞着方程式 (3.1) 式を $u = 1$ の周りで展開すると

$$1 - u \approx (1 - u)p \sum_k \frac{k(k-1)P(k)}{\langle k \rangle} = (1 - u)pb, \quad (3.4)$$

となる．ここで b は分枝因数 (branching factor) $b = \langle k^2 \rangle / \langle k \rangle - 1$ である．非自明な解 $u < 1$ が現れる条件は $pb > 1$ であるから，パーコレーション臨界点は $p_c = 1/b$ と求められる．この結果は以下のように解釈できる．熱力学極限 $N \rightarrow \infty$ におけるランダム・ネットワークは木構造ネットワークである．ランダムな木構造ネットワークにおいてエッジを 1 歩，2 歩，3 歩，4 歩とたどっていくとき，到達しうるノードの数の期待値は 1 個， b 個， b^2 個， b^3 個と指数関数的に増えていく．エッジ（またはノード）が確率 $1 - p$ で削除されている場合， n 歩目に到達しうるノード数の期待値は $(pb)^{n-1}$ となるため， pb が 1 より大きいときネットワークはパーコレートする．

ネットワークが隣接次数相関を有するとき，エッジをたどった先のノードの次数はエッジのもう片方の端点の次数に依存する．隣接次数相関が有るとき，自己無撞着方程式 (3.1) は隣接次数相関を記述する条件付確率 $P_{nn}(k|k')$ を用いて

$$u_{k'} = 1 - p + p \sum_k P(k|k') u_k^{k-1}, \quad (3.5)$$

となる．ここで確率 u_k は，次数 k のノードから出るエッジを任意に選んだとき，そのエッジの先に巨大連結成分が存在しない確率である．(3.1) 式と同様に (3.5) 式にも自明な解

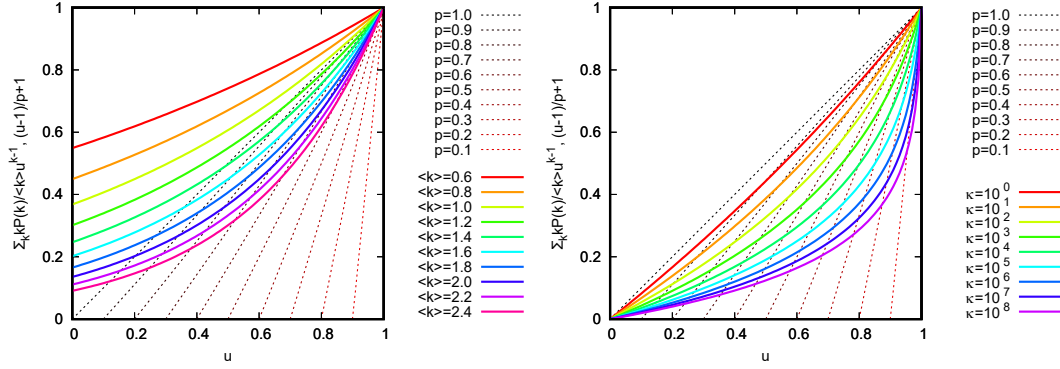


図 3.1: 次数分布が $P(k) = e^{\langle k \rangle} \langle k \rangle^k / k!$ である場合 (左) と $P(k) \propto k^{-2} e^{-k/\kappa}$ である場合 (右) における自己無撞着方程式 (3.1) の解の値を示す概念図. 横軸は u , 縦軸は関数 $f(u) = \sum_k u^{k-1} k P(k) / \langle k \rangle$ および $(u-1)/p+1$ である. 曲線と点線の交点が (3.1) 式の解を与える.

$u_1 = u_2 = \dots = 1$ が存在する. 自明な解の周りで展開された (3.5) 式は

$$1 - u_{k'} \approx p \sum_k P(k|k') (k-1) (1 - u_k), \quad (3.6)$$

と近似できる. ここで, Λ_1^B を分枝行列 (branching matrix)

$$B_{k,k'} = \sum_k P(k|k') (k-1), \quad (3.7)$$

の最大固有値とすると, パーコレーション臨界点は非自明な解が存在する条件 $p\Lambda_1^B > 1$ より $p_c = 1/\Lambda_1^B$ と求められる.

最大固有値 Λ_1^B には上限値と下限値が存在し,

$$\sum_k \frac{[k_{nn}(k) - 1] k (k-1) P(k)}{\langle k^2 \rangle - \langle k \rangle} \leq \Lambda_1^B \leq \max_k k_{nn}(k) - 1, \quad (3.8)$$

の範囲に収まる [47]. 隣接次数相関が無いとき, $k_{nn}(k) = b+1$ であるから下限値と上限値はどちらも分枝因数 b となり, 上下から挟まれた Λ_1^B の値は $\Lambda_1^B = b$ と定まる. この結果にある種の隣接次数相関をわずかに加えると, その隣接次数相関の強さの 1 次近似の範囲で Λ_1^B は (3.8) 式の下限と一致し,

$$\Lambda_1^B \approx \sum_k \frac{[k_{nn}(k) - 1] k (k-1) P(k)}{\langle k^2 \rangle - \langle k \rangle} = b + \frac{r}{b} \left(\frac{\langle k^3 \rangle}{\langle k \rangle} - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2} \right), \quad (3.9)$$

となる. ここで r は (2.9) 式で導入した Pearson 相関係数である. これより, 弱い隣接次数相関による Λ_1^B への寄与は Pearson 相関係数 r によって表されることがわかる. Pearson 相関係数は隣接次数相関が無いとき $r = 0$, 正の隣接次数相関が有るとき $0 < r \leq 1$, 負の隣

接次数相関が有るとき $-1 \leq r < 0$ となる．このため，それぞれの場合における Λ_1^B の大小関係は

$$\Lambda_1^B(\text{負相関}) < \Lambda_1^B(\text{無相関}) < \Lambda_1^B(\text{正相関}), \quad (3.10)$$

となる．したがって，その逆数であるパーコレーション臨界点 p_c は

$$p_c(\text{負相関}) > p_c(\text{無相関}) > p_c(\text{正相関}), \quad (3.11)$$

となり，正の隣接次数相関によってネットワークは頑強になり，負の隣接次数相関によってネットワークは脆弱になることがわかる．

隣接次数相関はパーコレーション臨界点だけでなく臨界指数へも影響を与える．次数分布が Poisson 分布 $P(k) = e^{-\langle k \rangle} \langle k \rangle^k / k!$ に従う Erdős-Rényi ランダムグラフにおいて，パーコレーション臨界点は $p_c = 1/\langle k \rangle$ である．このときパーコレーション臨界点近傍で以下のスケーリング則が成立する．

$$S \sim (p - p_c)^\beta, \quad (\beta = 1, p > p_c) \quad (3.12)$$

$$n(s) \sim s^{-\tau}, \quad (\tau = 5/2, p = p_c) \quad (3.13)$$

$$\langle s \rangle \sim |p - p_c|^{-\gamma}, \quad (\gamma = 1) \quad (3.14)$$

ここで， $n(s)$ は任意抽出したノードが有限サイズ s の連結成分に属する確率， S は巨大連結成分に属する確率（すなわち $1 - \sum_s n(s)$ ）， $\langle s \rangle$ は有限サイズの連結成分に所属するときの連結成分サイズの期待値 $\langle s \rangle = \sum_s s n(s)$ である．Noh [46] は，グラフハミルトニアン¹を

$$H(\mathcal{G}) = -\frac{J}{2} \sum_{i,j} a_{ij} k_i k_j, \quad (3.15)$$

をとした指数ランダムグラフ集合 [96] を用いて，無相関ネットワーク ($J = 0$)，正相関ネットワーク ($J = 1$)，負相関ネットワーク ($J = -1$) の，それぞれの場合における臨界点と臨界指数を数値的に求めている．図 3.2 はエッジを確率 p で占有するボンド・パーコレーションにおける S と $\langle s \rangle$ の 1,000 サンプル平均である．図 3.2(a) からは，3 種のネットワークの臨界点 p_c が大小関係 (3.11) を満たすことが確認できる．一方，図 3.2(b) からは，正相関ネットワーク ($J = 1$) においてはいかなる p においても $\langle s \rangle$ の発散が生じないという特異な現象が見られる．これらの数値計算結果を有限サイズスケーリングの関係式

$$S(p, N) = N^{-\beta/\bar{\nu}} \mathcal{P} \left[(p - p_c) N^{1/\bar{\nu}} \right], \quad (3.16)$$

$$\langle s \rangle(p, N) = N^{-\gamma/\bar{\nu}} \mathcal{S} \left[(p - p_c) N^{1/\bar{\nu}} \right], \quad (3.17)$$

¹無限個のグラフから成る集合において，特定のグラフ $\mathcal{G}$ の出現確率分布が $P(\mathcal{G}) \propto e^{-H(\mathcal{G})}$ で与えられるとき， $H(\mathcal{G})$ をその集合のグラフハミルトニアンという

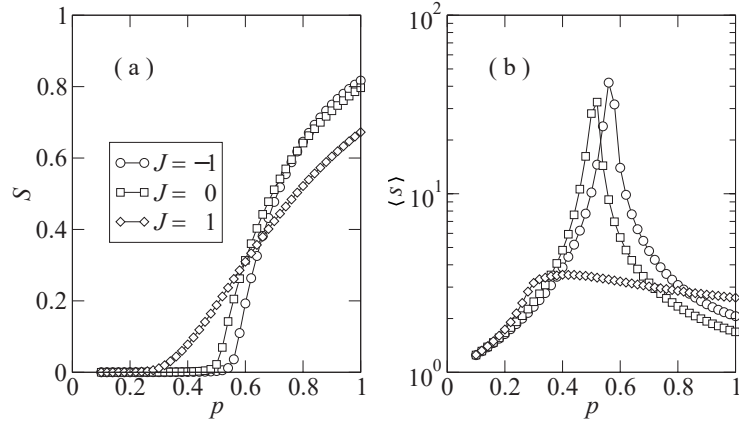


図 3.2: 無相関ネットワーク ($J = 0$), 正相関ネットワーク ($J = 1$), 負相関ネットワーク ($J = -1$) のそれぞれにおける, 確率 p でエッジが占有されたときの (a) 最大連結成分の相対サイズ S と (b) 有限サイズの平均連結成分サイズ $\langle s \rangle$ の数値計算結果. ネットワーク・サイズは $N = 80,000$, 1,000 サンプルに対して平均をとっている. 文献 [46] より引用. ただし本論文の用法に合わせるために記号の書き換えを行った.

表 3.1: 次数相関を有する指数ランダムグラフの臨界指数 β , $\bar{\nu}$, τ の値. 値は文献 [46] より引用

	β	$\bar{\nu}$	τ	γ
正相関 ($J = 1$)	—	—	3.9	—
無相関 ($J = 0$)	1	3	5/2	1
負相関 ($J = -1$)	1	2.9(1)	—	1

に適用することによって求められる臨界指数を表 3.1 にまとめた. ただし正相関ネットワーク ($J = 1$) に関してはいかなる p においても S や $\langle s \rangle$ に臨界的振る舞いが見られないため, 臨界指数 β , $\bar{\nu}$, γ を求めることはできない. 表 3.1 の結果は, 負相関ネットワーク上と無相関ネットワーク上のパーコレーションが同じ平均場のクラスに属することを示唆するとともに, 正相関ネットワーク上のパーコレーションは平均場のクラスとは全く異なる普遍性クラスに属する (あるいは臨界現象が消失する) ことを意味する.

Goltsev ら [47] らは, 以下の 3 つの条件を満たす場合には隣接次数相関が臨界指数 β に影響を及ぼさないことを示した.

1. 分枝行列 B の最大固有値 (第 1 固有値) Λ_1^B と $\langle k^2 \rangle$ がどちらも発散する. またはどちらも発散しない.
2. B の第 2 固有値 Λ_2^B の値が有限である.
3. B の第 1 固有ベクトルの成分が $k \rightarrow \infty$ 極限でゼロではない有限の値となる.

条件 2 は強い正の隣接次数相関が有るとき成立せず, 条件 3 は強い負の隣接次数相関が有る

表 3.2: 次数分布が $P(k) \propto k^{-\gamma}$, 隣接次数相関が $k_{\text{nn}}(k) \propto \text{const} + k^\alpha$ に従うスケールフリー・ネットワークにおける, 分枝行列の最大固有値 Λ_1^B , 第2固有値 Λ_2^B , およびパーコレーション臨界指数 β [47]. 強い正相関ネットワークは (3.18) 式で与えられるモデル, 強い負相関ネットワークは (3.19) 式で与えられるモデルを用いている.

隣接次数相関	γ	Λ_1^B	Λ_2^B	β
無相関または 弱相関	$\gamma > 4$	有限	有限	1
	$3 < \gamma < 4$	有限	有限	$1/(\gamma - 3)$
	$2 < \gamma < 3$	∞	有限	$1 + 1/(3 - \gamma)$
強い正相関	$\gamma > 3 + \alpha$	∞	∞	$1 + (2 - \gamma + \alpha)/\alpha$
	$2 < \gamma < 3 + \alpha$	∞	∞	$1 + 1/(3 - \gamma)$
強い負相関	$\gamma > 4 - 3\alpha$	有限	0	1
	$3 - 2\alpha < \gamma < 4 - 3\alpha$	有限	0	$(1 - \alpha)/(\gamma - 3 + 2\alpha)$
	$2 < \gamma < 3 - 2\alpha$	∞	0	$1 + (1 - \alpha)/(3 - \gamma - 2\alpha)$

とき成立しない. Goltsev らはさらに, 解析的に分枝行列 B の固有ベクトルを求めることができる2種類のモデルを用いて, 上の条件を満たさないような強い隣接次数相関を有するネットワークの臨界指数を具体的に求めている. 1つ目のモデルは, 隣接次数相関が

$$P^{(+)}(k, k') = \delta_{k, k'} [f(k) + 1] g^2(k) P^2(k) + (1 - \delta_{k, k'}) g(k) P(k) g(k') P(k'), \quad (3.18)$$

によって表されるネットワークである. ここで $f(k) > 0$ は任意の関数であり, $g(k)$ は規格化条件によって決まる関数である. $f(k) \rightarrow 0$ で無相関ネットワークの極限, $f(k) \rightarrow \infty$ で正の隣接次数相関の極限 (すなわち $r \rightarrow 1$) に至る. 2つ目のモデルは Vázquez ら [97] によって提案された

$$P^{(-)}(k, k') = \frac{1}{\langle k \rangle} [(1 - h(k')) k' P(k') (1 - \delta_{k', 1}) \delta_{k, 1} + (1 - h(k)) k P(k) (1 - \delta_{k, 1}) \delta_{k', 1}] \\ + \frac{1}{a \langle k \rangle^2} h(k') k' P(k') h(k) k P(k) (1 - \delta_{k', 1}) (1 - \delta_{k, 1}), \quad (3.19)$$

であり, 任意の減少関数 $h(k)$ に対して負の隣接次数相関を示す. ここで, $a = \sum_{k \geq 1} h(k) k P(k) / \langle k \rangle$ である. 両モデルにおいて, 次数分布が $P(k) \propto k^{-\gamma}$ に従うとき平均隣接次数は $k_{\text{nn}}(k) \propto \text{const} + k^\alpha$ の形をとる. これら2のモデルに対し, スケールフリー指数 γ と隣接次数相関を表す α の値によって臨界指数 β は表 3.2 のように決まる.

3.2 同期現象と隣接次数相関

エッジでつながれたノード上の振動子が相互作用しあい, 互いに位相を近づけることによって, やがてネットワーク全体が1つの振動子のようにふるまうようになる現象をネット

ワーク上の同期現象という。自然界においては蛍の明滅やコンサート会場での拍手などがこれにあたる。また、機能的ネットワークにおいては、電力網上の交流電流や神経細胞の発火リズムなどに同期現象が見られる。電力網上では、位相がそろわなければ交流電流は流れない。また、てんかん発作は脳神経の発火リズムが脳全体で同期してしまうことによって生じられている [98]。これらの同期現象において、相互作用の様子を表すモデルは複数提案されているが、最も有名なモデルが蔵本モデルである [99,100]。ネットワーク上の蔵本モデルでは、ノード i を自然振動数 ω_i の振動子とみなす。ネットワーク $\mathcal{G}$ の構造が隣接行列 $\{a_{ij}\}$ で与えられるとき、隣接するノードどうしは互いに位相を近づけようとする。この時間発展の様子を蔵本モデルでは、

$$\frac{d\theta_i}{dt} = \omega_i + K \sum_j a_{ij} \sin(\theta_j - \theta_i) \quad (3.20)$$

と表わす。係数 K を結合強度という。(3.20) 式は、 $R_i e^{i\Psi_i} = \sum_j a_{ij} e^{i\theta_j}$ とおくことで

$$\frac{d\theta_i}{dt} = \omega_i + K R_i \sin(\Psi_i - \theta_i), \quad (3.21)$$

とかける。すなわちノード i は、隣接ノードの位相の平均 Ψ_i に対して常に自身の位相 θ_i を強さ $K R_i$ で近づけようとする。同期の程度は秩序変数に相当する量

$$R(t) e^{i\Psi(t)} = \frac{1}{N} \sum_j e^{i\theta_j}, \quad (3.22)$$

によって定量化される。結合強度が臨界点 K_c よりも大きいとき、定常状態 ($t \rightarrow \infty$) において有限割合のノードの位相が同期した状態となり $R(t) > 0$ となる。

Restrepo ら [39] は隣接次数相関が有るときの臨界点 K_c を、同じ次数をもつノード集合をひとまとまりとした平均場近似を用いて求めている。結果のみを記述すると、次数 k の自然振動数分布 $g(\omega|k)$ がローレンツ分布

$$g(\omega|k) = \frac{\Delta/\pi}{[\omega - \omega_0(k)]^2 + \Delta^2}, \quad (3.23)$$

に従うとき、臨界結合強度は

$$K_c = 2\Delta/\Lambda_1^C, \quad (3.24)$$

となる。ここで $\omega_0(k)$ は次数 k のノードの中心自然振動数、 Δ は半値半幅、 Λ_1^C は連結行列 (connectivity matrix)

$$C_{kk'} = k P_{nn}(k'|k), \quad (3.25)$$

の最大固有値である．なお，連結行列 C と (3.7) 式の分枝行列 B は，次数 $\{1, 2, \dots\}$ を成分とした対角行列 D によって

$$D^{-1}C = B(D - I), \quad (3.26)$$

と関係づけられている．隣接次数相関が弱い場合，連結行列 C の最大固有値は，(2.10) 式で与えられるエッジ両端の次数に対する Pearson 相関係数 r を用いて

$$\Lambda_1^C \approx [kk']_E \frac{\langle k \rangle}{\langle k^2 \rangle} = b + 1 + \frac{r}{b + 1} \left(\frac{\langle k^3 \rangle}{\langle k \rangle} - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2} \right), \quad (3.27)$$

と近似される．したがって，正相関ネットワーク，負相関ネットワーク，無相関ネットワークのそれぞれの連結行列の最大固有値は

$$\Lambda_1^C(\text{負相関}) < \Lambda_1^C(\text{無相関}) < \Lambda_1^C(\text{正相関}), \quad (3.28)$$

となる．臨界結合強度 K_c は Λ_1^C の逆数に比例するから，結局，

$$K_c(\text{負相関}) > K_c(\text{無相関}) > K_c(\text{正相関}), \quad (3.29)$$

となる．この結果から，正の隣接次数相関が有るとネットワークは同期しやすくなり，負の隣接次数相関が有るとき同期しにくくなることがわかる．

3.3 感染症拡散と隣接次数相関

人から人へと次々と広がっていく伝染病のように，ネットワーク上で情報や状態が隣接ノードから隣接ノードへと拡散していく様子を数理的に記述するために，いくつかの感染症モデルが提案されている [8]．これまでに提案された感染症のモデルの多くは，SIR モデルまたは SIS モデルと呼ばれる 2 つのモデルを出発点としている．前者の「SIR」は“Susceptible-Infected-Recovered”の略称であり，各ノードは“S” (susceptible, 感染可能)，“I” (infected, 感染済)，“R” (recovered, 回復済) の 3 状態のいずれかの状態をとる．状態 I のノードと状態 S のノードをつなぐエッジは，単位時間 δt が経過するごとに確率 $\beta \delta t$ で S 状態のノードを I 状態に変える．これは感染者との接触によって伝染病が拡大する様子を表しており，パラメータ β を感染率という．状態 I のノードは単位時間 δt が経過するごとに確率 $\mu \delta t$ で R 状態へと移行する．これは時間経過によって感染者が自然回復または死亡することで感染拡大能力を失ったことを意味している．このときのパラメータ μ を回復率という．SIR モデルにおいては一度 R 状態になった（回復または死亡した）ノードは二度と S 状態や I 状態にはならない．現実世界において，このことは天然痘，麻疹・風疹，結核，コレラなどに罹患した感染者が，回復後に獲得する免疫によって一定期間の間は再感染することがない事実

に対応する．後者の「SIS」は“Susceptible-Infected-Susceptible”を意味し，各ノードの状態は“S”（susceptible, 感染可能）または“I”（infected, 感染済）の2状態のいずれかとなる．SIRモデルと同様に，状態Iのノードと状態Sのノードをつなぐエッジは，単位時間 δt が経過するごとに確率 $\beta\delta t$ で状態SのノードをI状態に変える．しかし感染後の経過はSIRモデルと異なる．状態Iのノードは単位時間 δt が経過するごとに確率 $\mu\delta t$ で元のS状態に戻る．そのため，一度感染を経験したノードでも再感染する可能性が有る．このモデルは感染によって獲得する免疫が弱く再感染を繰り返す恐れのある梅毒や性器クラミジア，遺伝子の変化が速く免疫の有効期間が短いインフルエンザウイルスなどの特徴をとらえている．さらに，感染症拡散モデルの多くは，感染症だけでなく文化や流行の伝播のモデルと解釈することもできる．例えばSISモデルに基づいてある映画の流行現象を考えると，状態Sはその映画に興味をもっていない状態，状態Iはその映画を高く評価したことで周囲へ視聴を促している状態ととらえられる．状態Iのノードは時間経過とともに映画への関心が薄れて状態Sへ移行するが，その後も周囲のI状態ノードから刺激を受けることで再び状態Iへと変化することが考えられる．感染症の拡散においても，映画の流行においても，現実世界ではしばしば国や地球規模に及ぶ大流行（パンデミック）が起きている．流行がパンデミックに至るのか，それとも一部の値域だけにとどまるのかは，感染率や回復率の大きさだけでなく，ネットワークの構造がどのようなものかによっても大きく異なる．本節では，隣接次数相関が感染症拡散に与える影響についてSISモデルとSIRモデルを用いて議論する．

どちらのモデルでも流行の規模は $t \rightarrow \infty$ におけるS状態のノードの数によって定量化される．時刻 t におけるS状態，I状態，R状態のノード数のネットワーク・サイズ N に対する比を，それぞれ $\rho^S(t)$ ， $\rho^I(t)$ ， $\rho^R(t)$ とする．SIRモデルにおいてもSISモデルにおいても時間発展には終状態があり，一度 $\rho^I(t) = 0$ となった後は感染イベントも治癒イベントも生じない．ただし，熱力学極限 $N \rightarrow \infty$ におけるSISモデルでは $\rho^I(t)$ が有限となる平衡状態が存在する．初期状態を $\rho^I(0) \ll 1$ （SIRモデルにおいてはさらに $\rho^R(0) = 0$ ）としたとき，与えられたネットワークにおいて感染症がどの程度まで広がるかは，感染率 β と回復率 μ の比 $\lambda = \beta/\mu$ によって決まる．無限に大きなネットワークにおける感染の様子は，最終的な流行の規模 $J = 1 - \lim_{t \rightarrow \infty} \rho^S(t)$ の期待値が無小となる収束相と，有限値をとる流行相とに分類することができ，感染症拡散がどちらの相に属するかは λ と臨界点 λ_c との大小関係によって決まる．以下では，SIRモデルとSISモデルのそれぞれの場合における λ_c について説明する．

SIRモデルの終状態における $J = 1 - \lim_{t \rightarrow \infty} \rho^S(t)$ は，終状態に至るまでの時間に1本のエッジが感染を媒介する確率 T をエッジの存在確率 p とみなすことで，ボンド・パーコ

レーションにおける連結成分サイズ S に対応付けることができる． T と λ の間には

$$T = \frac{\lambda}{1 + \lambda}, \quad (3.30)$$

の関係式が成り立つことが知られている．したがって SIR モデルとボンド・パーコレーションとの対応関係は $p = \lambda/(1 + \lambda)$ である．パーコレーション臨界点 p_c は分枝行列 (3.7) の最大固有値 Λ_1^B の逆数，すなわち $p_c = 1/\Lambda_1^B$ で与えられるから，SIR モデルの臨界点は $\lambda_c = 1/(\Lambda_1^B - 1)$ と求められる．したがって，弱相関近似における λ_c と隣接次数相関との関係は，(3.9) 式より，

$$\lambda_c(\text{負相関}) > \lambda_c(\text{無相関}) > \lambda_c(\text{正相関}), \quad (3.31)$$

となる．この結果より，免疫の効果が大きい SIR 型の感染症は正の隣接次数相関によって流行しやすくなり，負の隣接次数相関によって収束しやすくなることがわかる．また，SIR モデルとボンド・パーコレーションの対応関係が $p = \lambda/(1 + \lambda)$ によって表されるから，

$$\left. \frac{dp}{d\lambda} \right|_{p=p_c} = \frac{1}{(1 + \lambda_c)^2}, \quad (3.32)$$

となる．これより， $p = p_c + dp$ ， $\lambda = \lambda_c + d\lambda$ とすると，

$$p - p_c \approx \frac{\lambda - \lambda_c}{(1 + \lambda_c)^2}, \quad (3.33)$$

を得る．すなわち， $(p - p_c) \propto (\lambda - \lambda_c)$ であるから，SIR モデルの臨界指数はパーコレーションに対するそれと等しい．

SIS モデルは SIR モデルとは異なり，パーコレーション過程へ対応付けることはできない．次数 k のノードが時刻 t において状態 I にある確率を $\rho_k^I(t)$ とすると，SIS モデルの時間発展は同じ次数をもつノードを区別しないという平均場近似の下で

$$\frac{d\rho_k^I(t)}{dt} = -\rho_k^I(t) + \lambda k[1 - \rho_k^I(t)] \sum_{k'} P_{nn}(k'|k) \rho_{k'}^I(t), \quad (3.34)$$

と表される．これを線形化すると

$$\frac{d\rho_k^I(t)}{dt} \approx -\rho_k^I(t) + \lambda \sum_{k'} C_{kk'} \rho_{k'}^I(t) \quad (3.35)$$

を得る．ここで， C は (3.25) 式で定義される連結行列である．感染が収束した状態 $\rho_1^I = \rho_2^I = \dots = 0$ が安定であるための条件は，行列 $I - \lambda C$ の最大固有値が正であることである．これより， C の最大固有値を Λ_1^C としたとき臨界点は $\lambda_c = 1/\Lambda_1^C$ となり， $\lambda = K/2\Delta$ とするこ

表 3.3: 隣接次数相関のあるネットワーク上における物理現象の臨界点. Λ_1^B は (3.7) 式で定義される分枝行列の最大固有値, Λ_1^C は (3.25) 式で定義される結合行列の最大固有値である.

パーコレーション	同期現象	SIR モデル	SIS モデル
$1/\Lambda_1^B$	$2\Delta/\Lambda_1^C$	$1/(\Lambda_1^B - 1)$	$1/\Lambda_1^C$

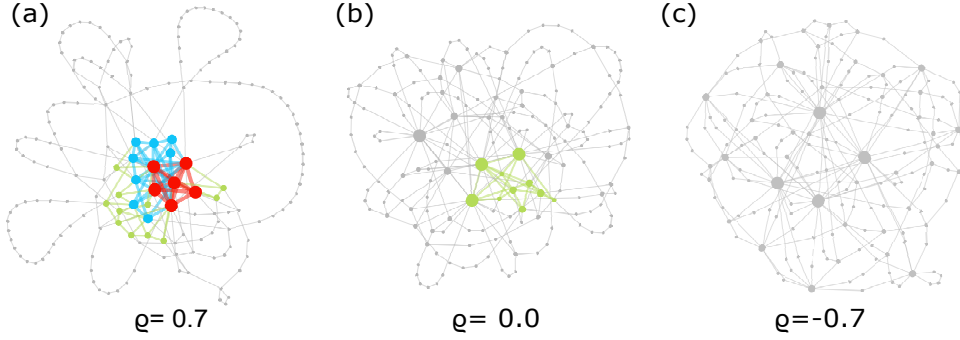


図 3.3: (a) 正相関ネットワーク ($q = 0.7$), (b) 無相関ネットワーク ($q = 0.0$), (c) 負相関ネットワーク ($q = -0.7$) のそれぞれにおける, ネットワーク構造. 緑色のノードは 3 コア, 青色ノードは 4 コア, 赤色のノードは 5 コアに属する.

とにより 3.2 節で述べた同期現象と対応付けられる. 連結行列 C の最大固有値と隣接次数相関との関係は (3.27) 式の通りであるから, 臨界点 λ_c は

$$\lambda_c(\text{負相関}) > \lambda_c(\text{無相関}) > \lambda_c(\text{正相関}), \quad (3.36)$$

となる. これより, SIR モデルと同様に, 正の隣接次数相関によって感染が起こりやすく, 負の隣接次数相関によって感染が起こりにくくなることがわかる.

表 3.3 に隣接次数相関を有するネットワーク上のパーコレーション, 同期現象, 感染症拡散の臨界点をまとめた. いずれの現象においても, 正の隣接次数相関を有するネットワークにおいて無秩序相から秩序相への相転移は起こりやすく, 負の隣接次数相関を有するネットワークでは起こり難い. 一方で秩序極限においては, 図 3.2(a) に見られるように, 負相関ネットワークにおける秩序変数の方が正相関ネットワークのそれよりも大きい傾向にある. より正確に言うと, 正の隣接次数相関を有するネットワークは秩序相に入りやすいが, 秩序相に入った後の秩序変数の成長はゆるやかである. この傾向の理由として k コアの存在が挙げられる. ネットワーク $\mathcal{G}$ に対し, 次数 k 以下のノードの削除を繰り返したとき, 最後まで残った成分を $\mathcal{G}$ の k コアという [101, 102]. 図 3.3 の 3 つのネットワークは, 同じ次数列を有する (a) 正相関ネットワーク, (b) 無相関ネットワーク, (c) 負相関ネットワークである. 緑色のノードは 3 コア, 青色ノードは 4 コア, 赤色のノードは 5 コアに属する. 正の隣接次数相関が強いとき, 図 3.3(a) のように高次数ノードによる k コアが生じる. このような高次数ノードどうしによって強固に連結した部分グラフを「リッチクラブ」という. 大

い次数 k のコアは完全グラフのように壊れにくく、 k コア内部では振動子の同期や感染症の拡散が起こりやすくなっている。正の隣接次数相関を有するネットワークに秩序が表れやすい理由は、強固なリッチクラブが存在しているためである。高次数ノードがリッチクラブをつくる一方で、図 3.3(a) を見ると、低次数ノードどうしの連結によってつくられる疎な部分グラフがリッチクラブから遠く離れたネットワークの外殻に出現している。この外殻は平均次数が低く、ノード破壊やエッジ破壊によって容易く分裂し、同期や拡散も起こりにくくなっている。このため、高次の k コアが秩序化することで秩序変数が有限になったとしても、残りの低次数外殻ノードがなかなか秩序化しないため、秩序変数の増加の速さは緩慢となる。また、外殻に押しやられた低次数ノードが 1 次元構造や臨界木構造をつくるために、正相関ネットワークの直径は大きくなる傾向がある。一方、図 3.3(c) のような強い負の隣接次数相関を有するネットワークでは、高次数ノードは低次数ノードによって包囲されやすく、界面活性剤によってミセルコロイドとなり分離する油のように高次数ノードどうしは離れる傾向にある。そのためリッチクラブは出現せず、秩序化し難くなる。しかし、一旦秩序化が起これば、秩序化を遅くする低次数外殻ノードが存在しないため、秩序化は急速に進行するようになる。また、低次数ノードが高次数ノードの周囲に引き込まれるため、低次数ノードどうしの結合による 1 次元構造や臨界木構造は生じず、ネットワーク直径は小さくなる傾向がある。

第4章 隣接次数相関の限界

前章で述べた通り，隣接次数相関の重要性について疑いの余地は無い．ただし複雑ネットワークの全ての次数相関が隣接次数相関によって説明されるわけではない．隣接だけでなく「次隣接以上に離れたノードどうしの相関」もまたネットワークの構造やダイナミクスと密接な関係にある．本章では，隣接次数相関だけではネットワークの特徴を説明しきれないこと，および長距離に渡って離れたノード間の次数相関（長距離次数相関）を考えることが重要であることを具体例を挙げて示す．

4.1 (u, v) -flower のリワイアリング

隣接次数相関の研究はこれまでも盛んに行われており [9–17]，クラスター性 [18–20, 22]，コミュニティ構造 [23]，平均ノード間距離 [24]，フラクタル性 [25, 26, 28] などの，ネットワーク構造に関する基本的な性質と密接な関係にあることがわかっている．これに加えて，ネットワーク上での情報および感染症の拡散現象 [29–32, 41]，各ノードに置かれた振動子の同期 [33–40]，各ノードをエージェントとした戦略ゲーム [42–45]，ノード削除の下でのネットワークの頑強性 [46–50] などのネットワーク上の数々のダイナミクスに対しても隣接次数相関は強い影響を与える．

しかしながら隣接次数相関が一般的な次数相関を完全に記述するわけではない．本論文では，フラクタル・スケールフリー・ネットワーク・モデルである (u, v) -flower を例として取り上げることで，隣接次数相関によってネットワークを特徴づけることの限界を示す [103]．そのため，まずエッジのつなぎ替え操作（リワイアリング [104]）により作られる， (u, v) -flower と同じ隣接次数相関をもつネットワーク（隣接相関ランダム・ネットワーク）と，元の (u, v) -flower の性質を比較することで，隣接次数相関がフラクタル性を生じさせるか否かを調べた．つなぎ替え操作のルールは以下の通りである．

1. ネットワーク内からランダムに 2 本のエッジを選び，図 4.1 のようにそれぞれの片端を交換するという方法をとる．
2. どちらの端を交換するかはランダムに選ぶ．このとき，選ばれた 2 つの端の次数が同じである場合のみつなぎ替えを実行することとする．なお，つなぎ替えの結果として多

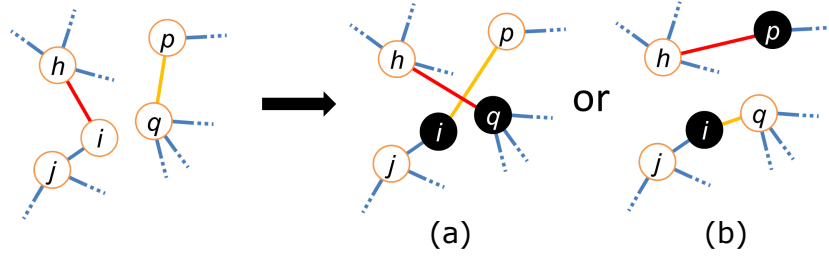


図 4.1: エッジのつながり替え操作の概念図. ランダムに選ばれた 2 本のエッジ (h, i) と (p, q) のそれぞれの片端を交換する. どちらの端を交換するかもランダムに選ぶため, つなぎ替えによってできるエッジが中央図のように (h, q) と (i, p) なる確率と, 右図のように (h, p) と (i, q) となる確率は, それぞれ $1/2$ である. ただし, ノード n の次数を k_n としたとき, (a) のパターンの場合は $k_i = k_j$ または $k_h = k_p$, (b) のパターンの場合は $k_i = k_p$ または $k_h = k_j$ でなければならない.

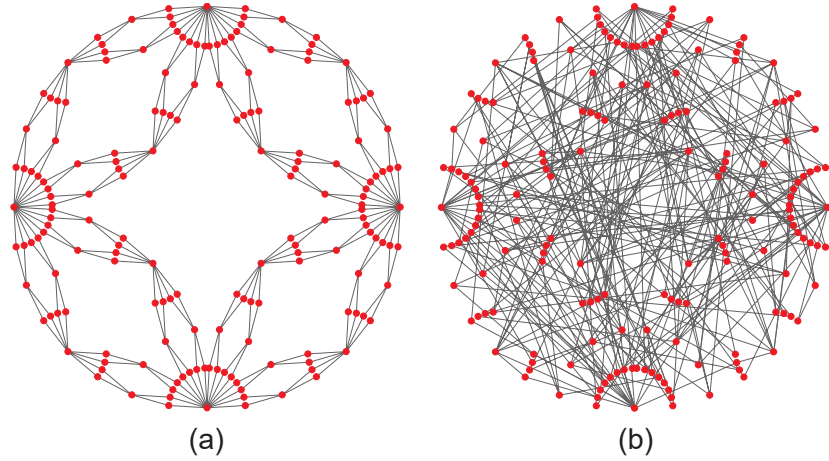


図 4.2: 第 4 世代 $(2, 2)$ -flower (a) と, その隣接相関ランダム・ネットワーク (b).

重エッジや自己ループが形成される場合は, 2 本のエッジを選ぶところからやり直すこととする.

3. 手順 2 を元のネットワークの大域的構造が消失するまで繰り返す.

図 4.2 の (a) と (b) は $u = 2$, $v = 2$ における (u, v) -flower, すなわち $(2, 2)$ -flower と, 上記のエッジのつながり替え操作によって得られる $(2, 2)$ -flower と同じ隣接次数相関をもつ隣接相関ランダム・ネットワークである. 2 つのネットワークは全く同じ隣接次数相関を有しているが, 図 4.3 に示すように元の (u, v) -flower が $d_f = 2$ のフラクタル性を示すのに対し, エッジのつながり替え後のネットワークでは最小被覆数 $N_B(l_B)$ がボックス直径 l_B に対して指数的に減衰し, フラクタル性が消失している.

$(2, 2)$ -flower とその隣接相関ランダム・ネットワークの大域的構造の違いをさらに詳しく調べるため, 高次数ノード (ハブ) どうしの距離分布を計算した. この分布関数 $\tilde{P}(l)$ は, 上

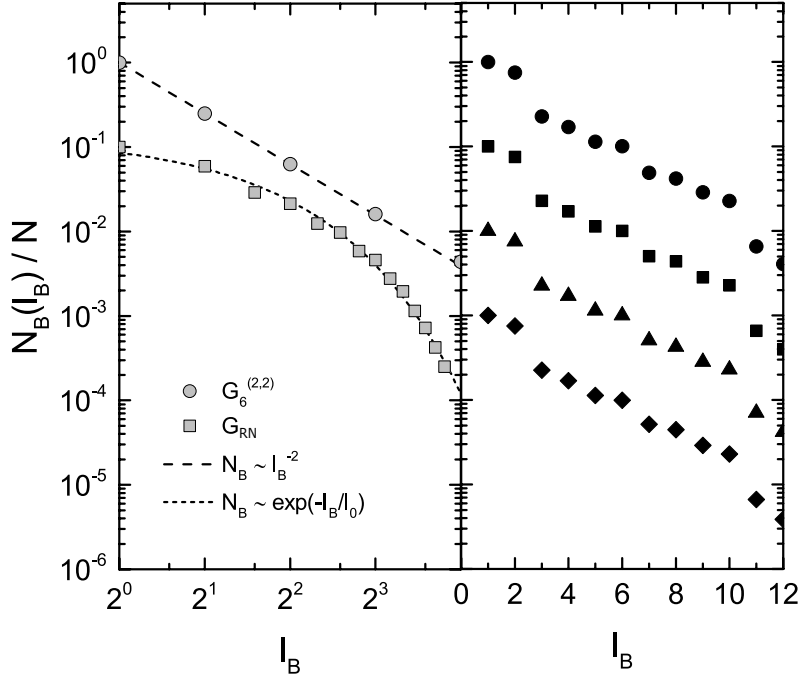


図 4.3: Compact-Box-Burning 法 [84] によるフラクタル解析の結果. 縦軸はネットワークを被覆するサブグラフ (ボックス) の最小数 N_B , 横軸はボックス直系 l_B である. 灰色の丸は第 6 世代 (2,2)-flower ($G_6^{(2,2)}$) に対する結果を示す. 灰色の四角は第 6 世代 (2,2)-flower と同じ次数列を有するランダム・ネットワーク (G_{RN}) に対する結果を示す. 黒のシンボルは第 6 世代 (2,2)-flower の隣接相関ランダム・ネットワークのうち, 4 つの異なるトポロジーを有するサンプルに対する結果を示す. ただし, 結果は見やすさのために縦軸方向にずらして描画している.

位 2% の高次数を有するノード対のノード間距離が l となる確率を与えるもので

$$\tilde{P}(l) = C \sum_{k,k'}' P(l|k,k') k P(k) P(k') \quad (4.1)$$

で定義される. ここで $P(l|k,k')$ は次数 k と k' のノード間の距離分布であり, k と k' が上位 2% の大きな値であるときのみ和はとられる. また $P(k)$ は次数分布関数であり, C は規格化定数である. 図 4.4 は, 第 6 世代の (2,2)-flower ($G_6^{(2,2)}$), その隣接相関ランダム・ネットワーク (G_{NNCRN}), および (2,2)-flower の次数列のみを保持したランダム・ネットワーク (G_{RN}) に対して $\tilde{P}(l)$ を比較したものである. この図から, 元の (2,2)-flower におけるハブは互いに遠く離れあい, ハブ間に長距離に渡る反発傾向が認められるのに対し, G_{NNCRN} や G_{RN} にはそのような傾向が無いことがわかる. 図 4.4 のインセットに示す通り, 同様の傾向は現実世界のフラクタル・スケールフリー・ネットワークである WWW においても見られる. これらの事実より, 隣接次数相関はフラクタル性を誘起しないこと, また, フラクタル・ネットワークに見られるハブ間の長距離反発のような「隣接以上に離れたノード間の次

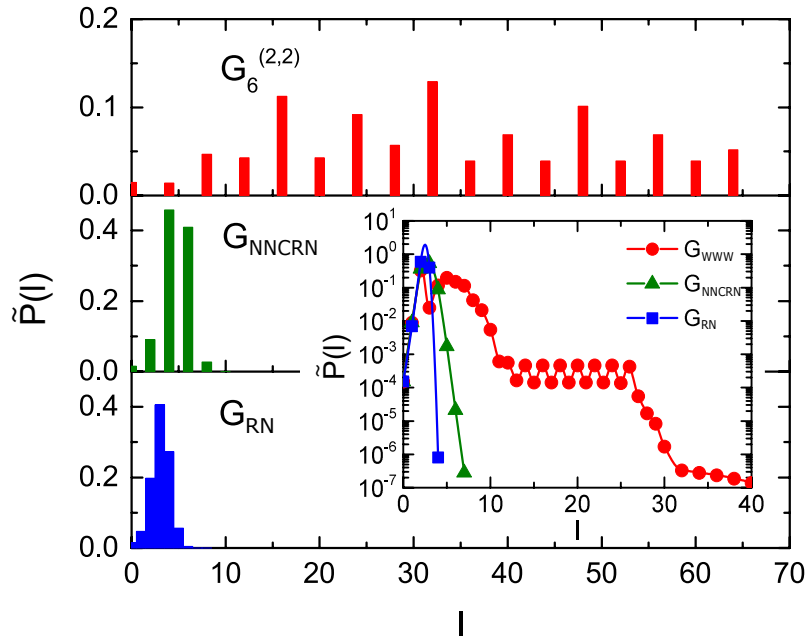


図 4.4: 上段のパネルは第 6 世代 $(2, 2)$ -flower ($G_6^{(2,2)}$) における上位 2% の高次数ノードどうしの中の距離分布 $\tilde{P}(l)$ である。中段のパネルは $G_6^{(2,2)}$ の隣接相関ランダム・ネットワーク (G_{NNCRN}) に対する結果を表し、下段のパネルは $G_6^{(2,2)}$ の次数列のみを保持したランダム・ネットワーク (G_{RN}) に対する結果を表す。インセットは WWW (ノード数 $N = 325,729$, エッジ数 $M = 1,090,108$, [85]) に対する同様の結果である。赤の丸は WWW (G_{WWW})、緑の三角形はその隣接相関ランダム・ネットワーク (G_{NNCRN})、青の四角形は G_{WWW} の次数列のみを保持したランダム・ネットワーク (G_{RN}) のにおける $\tilde{P}(l)$ の値を示す。曲線は図の見やすさのための近似曲線である。なお、隣接相関ランダム・ネットワーク G_{NNCRN} とランダム・ネットワーク G_{RN} に対する結果は 100 サンプルに渡る平均値である。

数相関」が隣接次数相関では説明できないことがわかる。

類似の研究は Orsini ら [51] によっても行われている。彼らは現実世界の複雑ネットワークの局所的な特徴量 (クラスター係数, 隣接次数相関) を保持してランダム化した場合, 元のネットワークのどのような性質が再現されるのかを調べている。その結果, ノード間距離分布, 媒介中心性, コミュニティ構造などの大域的性質が再現されることはないことが明らかとなった。

4.2 ネットワーク上の諸現象

3.1 節ではランダム・ネットワークと隣接相関ランダム・ネットワーク上でのパーコレーション臨界点を導出した。しかしながら, 次隣接以上のスケールにおいて相関を有するネットワークの臨界点はこれらの結果とは異なった値となる。特に, (u, v) -flower 上のサイト・パーコレーションでは, 有限区間の占有確率 p において臨界相が出現することが知られてい

る [105]. (u, v) -flower は強い負の隣接次数相関を有するが、負の隣接次数相関を有するネットワークが常に臨界相をもつわけではないので、臨界相の起源は負の隣接次数相関ではない。このような特異な現象の背景には、隣接次数相関よりも長距離に渡る構造的相関の影響があると考えられる。すなわち、隣接次数相関だけでネットワークの頑強性を特徴づけることはできないことになる。

同様の結論はネットワーク上の感染症拡散ダイナミクスからも得られる。Boguñá ら [53] は、SIS モデルに特定の次数を有するノード間の距離の効果を取り入れた有効モデルを提案し、次数 k と k' を有するノード間の平均距離が

$$l(k, k') \approx 1 + \frac{\log[N\langle k \rangle / kk']}{\log \kappa}, \quad (4.2)$$

によって近似されるスモールワールド的ネットワークと、

$$l(k, k') \approx 1 + \alpha N^{1/d_f}, \quad (4.3)$$

によって近似されるフラクタル・ネットワークの臨界点を調べた。その結果、スモールワールド・ネットワークの熱力学極限における臨界点は 0 へ収束するのに対し、長距離に渡る次数相関をもつフラクタル・ネットワークでは熱力学極限においても有限の臨界点を有することが明らかになった。したがって感染症拡散ダイナミクスを論じる際にも、隣接次数相関だけを考えるだけでは不十分であることになる。その他にも、高次数ノード（ハブ）間の距離がネットワーク上の様々なダイナミクスに対して直接的に影響を与える重要な要素であることがわかっている [52–55]。

4.3 本論文の目的

以上の議論から明らかなように、複雑ネットワークの大域的構造や、そこでの多様なダイナミクスを隣接次数相関だけで特徴づけることは不可能であり、隣接を超えた長距離に渡る次数相関、すなわち長距離次数相関を考慮することが極めて重要となる。そのため、これまでも複雑ネットワークにおける長距離次数相関の定式化の方法はいくつか提案されてきた。Rybski ら [56] は時系列解析のアナロジーを用いて、最短経路上のノードの次数揺らぎに対する Hurst 指数を長距離次数相関の指標として提案している。また、Mayo ら [57] は、(2.9) 式のエッジ両端のノードの次数に対する Pearson 相関係数を拡張し、ノード間距離 l だけ離れたノード対の次数に対する Pearson 相関係数 r_l を、

$$r_l = \frac{4\langle kk' \rangle_l - \langle k + k' \rangle_l^2}{2\langle k^2 + k'^2 \rangle_l - \langle k + k' \rangle_l^2}, \quad (4.4)$$

により定義した．ここで， $\langle f(k, k') \rangle_l$ はノード間距離 l 離れたノード対全体に渡る $f(k, k')$ の平均を意味する．同様の値は Mayo たちとは独立に Arcagni ら [58] によっても提案されている．また，(2.8) 式の平均隣接次数 $k_{nn}(k)$ を拡張し，次数 k のノードからノード間距離 l だけ離れたノードに対する平均次数（平均 l 次隣接次数） $k_l(k)$ も長距離次数相関の指標として提案されている．これらの指標を用いて Mayo らは複数の現実ネットワークの長距離次数相関を調べており，社会的ネットワークにおいては大きい l に対する r_l が負になる傾向があることを明らかにしている．隣接以上の次数相関を表す別の方向性から定義される Pearson 相関係数として，Allen-Perkins ら [59] は，エッジ両端のノードの「隣接ノードの次数の和」を確率変数とした Pearson 相関係数 $\tilde{r}$ を提案している．彼らはこの指標を用いて，エッジ両端の次数に対する通常の Pearson 相関係数が負となるネットワークを， $\tilde{r}$ の正負に従いさらに分類できることを示している．ただし，これらの指標は長距離次数相関の特定の側面を定量化することを目的としたものであり，長距離次数相関を完全に特徴づけるものではない．ネットワークの大域的構造やネットワーク上の諸現象と長距離次数相関との関係を統一的に理解するには，あらゆるタイプの長距離次数相関を系統的に記述することのできる数学的枠組みと汎用性の高い解析手法が不可欠である．しかしながら，長距離次数相関を一般的に論じる際に避けて通ることのできない有限サイズ・ネットワークを解析的に扱うことが難しいため複雑ネットワークの長距離次数相関の一般論は未だ確立していない．そのため，複雑ネットワークの構造的特徴やネットワーク上のダイナミクスが長距離次数相関と密接に関係していることは明白であるにもかかわらず，それらの間の関係を解明する系統的な研究は現在のところ報告されていない．本論文の目的は，ノード対間の距離 l ，ノード対の次数 k, k' を変数とした同時確率および条件付き確率を用いて，極めて一般的な長距離次数相関の定式化を行うことである．

第5章 長距離次数相関の定式化

長距離次数相関に関する系統的な議論を行うためには、まずはじめにネットワークの長距離次数相関の記述方法を確立しなければならない。ネットワークの隣接ノード間の次数相関（隣接次数相関）は、2.3節で述べたようにエッジ両端のノードの次数 k と k' を変数とする確率関数 $P_{nn}(k, k')$ および $P_{nn}(k'|k)$ によって完全に記述されるものであった。隣接するノードの次数が互いに独立なネットワークは隣接次数相関の無いネットワーク（隣接無相関ネットワーク）と呼ばれる。与えられたネットワーク G における隣接次数相関の有無は、 G と、 G に対応する隣接無相関ネットワークとの間で、確率関数 $P_{nn}(k, k')$ （または $P_{nn}(k'|k)$ ）を比較することによって判断することができる。また、隣接次数相関を評価するための関数である平均隣接次数 $k_{nn}(k)$ や Pearson 相関係数 r は、これらの確率関数を用いて定義されている。すなわち、隣接次数相関の系統的な研究の原点は確率関数 $P_{nn}(k, k')$ および $P_{nn}(k'|k)$ であると言える。本章ではこのアナロジーを長距離次数相関へと拡張し、長距離次数相関を完全に記述する確率関数を起点とした、複雑ネットワークの長距離次数相関を解析するための一般的フレームワークを構築する。

5.1 確率関数による一般的記述

ノード間距離 l だけ離れた次数 k と k' のノード対の間の次数相関は、同時確率 $P(k, k', l)$ によって完全に記述される。確率 $P(k, k', l)$ は任意抽出したノード対の次数が k と k' であり、かつそのノード間距離が l である同時確率である。この $P(k, k', l)$ を用いてさらに下記の4つの条件付き確率を考えることができる。

$$P(l|k, k') = \frac{P(k, k', l)}{\sum_l P(k, k', l)}, \quad (5.1a)$$

$$P(k'|k, l) = \frac{P(k, k', l)}{\sum_{k'} P(k, k', l)}, \quad (5.1b)$$

$$P(k, k'|l) = \frac{P(k, k', l)}{\sum_{k, k'} P(k, k', l)}, \quad (5.1c)$$

$$P(k', l|k) = \frac{P(k, k', l)}{\sum_{k', l} P(k, k', l)}. \quad (5.1d)$$

これら 4 つの条件付き確率の意味を同時確率 $P(k, k', l)$ の意味とともに表 5.1 にまとめた。隣接次数相関が同時確率 $P_{nn}(k, k')$ 、または条件付き確率 $P_{nn}(k'|k)$ によって表されたことと同様に、ネットワークの長距離次数相関は $P(k, k', l)$ だけでなく、(5.1) 式の 4 つの条件付き確率によっても記述することができる。全 5 種の確率関数はそれぞれ規格化条件、

$$1 = \sum_{k, k', l} P(k, k', l), \quad (5.2a)$$

$$1 = \sum_l P(l|k, k'), \quad (5.2b)$$

$$1 = \sum_{k'} P(k'|k, l), \quad (5.2c)$$

$$1 = \sum_{k, k'} P(k, k'|l), \quad (5.2d)$$

$$1 = \sum_{k', l} P(k', l|k), \quad (5.2e)$$

を満たす。なお、距離 l に関する和については、同一ノードのノード対に対する $l = 0$ と、ノード対間に経路が存在しないことを意味する便宜的な距離 l_∞ も含めるものとする。注意すべきことはネットワーク・サイズ N が無限大である場合、すなわち熱力学極限 $N \rightarrow \infty$ において、有限の l に対する同時確率 $P(k, k', l)$ と条件付き確率 $P(l|k, k')$ 、 $P(k', l|k)$ (l の分布関数を含む確率関数) が常にゼロになることである。そのため、無限系におけるこれらの確率関数は意味をもたない。これに対し、 l を条件の側にもつ $P(k'|k, l)$ 、 $P(k, k'|l)$ は無限サイズ・ネットワークにおいても有限の値をとることができるため、無限系においても意味のある確率関数として定義される。

同時確率 $P(k, k', l)$ を用いると、次数分布 $P(k)$ とノード間距離分布 $R(l)$ はそれぞれ、

$$P(k) = \sum_{k', l} P(k, k', l), \quad (5.3)$$

$$R(l) = \sum_{k, k'} P(k, k', l), \quad (5.4)$$

と表される。ここで、条件付き確率 $P(k, k'|l)$ を用いて、距離 l だけ離れたノード対を任意に選んだときの片方の次数が k である確率、

$$Q(k|l) = \sum_{k'} P(k, k'|l), \quad (5.5)$$

を導入する。これら $P(k)$ 、 $R(l)$ 、 $Q(k|l)$ は同時確率 $P(k, k', l)$ に対して、

$$\sum_{k'} P(k, k', l) = Q(k|l)R(l), \quad (5.6)$$

$$\sum_l P(k, k', l) = P(k)P(k'), \quad (5.7)$$

表 5.1: 長距離次数相関を記述する同時確率および条件付き確率とその意味

確率関数	確率関数の意味
$P(k, k', l)$	任意抽出したノード対の一方の次数が k , もう一方の次数が k' であり, ノード間距離が l である同時確率.
$P(l k, k')$	次数が k と k' のノード対を任意抽出したとき, ノード対間の距離が l である条件付き確率. 次数 k と k' をもつノード対の間のノード間距離分布とも解釈できる.
$P(k' k, l)$	任意抽出した次数 k のノードからノード間距離 l だけ離れたノードの次数が k' である条件付き確率. 次数 k をもつノードの l 次隣接ノードに対する次数分布とも解釈できる.
$P(k, k' l)$	ノード間距離 l のノード対をランダムに選んだとき, 次数が k と k' である条件付き確率.
$P(k', l k)$	任意抽出したノード対の一方の次数が k であるとき, もう一方の次数が k' であり, かつノード対間の距離が l である条件付き確率.

の和則を与える. 上記の和則と (5.1) 式から, 表 5.1 の確率関数どうしは $P(k)$, $R(l)$, $Q(k|l)$ を介して,

$$P(k', l|k) = P(k')P(l|k, k'), \quad (5.8)$$

$$P(k, k'|l) = Q(k|l)P(k'|k, l), \quad (5.9)$$

$$P(k)P(k', l|k) = Q(k'|l)R(l)P(k|k', l), \quad (5.10)$$

$$P(k)P(k')P(l|k, k') = R(l)P(k, k'|l), \quad (5.11)$$

と関係づけられる. 特に関係式 (5.10) と (5.11) は, 事象 A と事象 B に対するベイズの定理 $P(A|B) = P(B|A)P(A)/P(B)$ の直接的な帰結となっている.

隣接次数相関を特徴づける同時確率 $P_{\text{nn}}(k, k')$ と条件付き確率 $P_{\text{nn}}(k'|k)$ は, 長距離次数相関を特徴づけるために導入された上記の確率関数の $l = 1$ に対する特別な場合として含まれており,

$$P_{\text{nn}}(k, k') = P(k, k'|l = 1), \quad (5.12)$$

$$P_{\text{nn}}(k'|k) = P(k'|k, l = 1), \quad (5.13)$$

と表される. 同様に, 任意抽出したエッジの一方の端点ノードに対する次数分布 $Q_{\text{nn}}(k)$ は,

$$Q_{\text{nn}}(k) = Q(k|l = 1), \quad (5.14)$$

として $Q(k|l)$ の特別な場合として記述される．したがって， $l = 1$ における (5.9) 式は隣接次数相関の議論における (2.6) 式に帰着する．このように，長距離次数相関を記述する確率関数は，隣接次数相関を特徴づける $P_{nn}(k, k')$ ， $P_{nn}(k|k')$ の自然な拡張となっている．

上記のような隣接次数相関と長距離次数相関との間の対応関係を踏まえると，これまで隣接次数相関のために導入されてきたいくつかの指標を，長距離次数相関を特徴づける指標へと簡単に拡張することができる．例えば (2.9) 式のエッジ両端のノードの次数に対する Pearson 相関係数 r は距離 l だけ離れたノード対の両端の次数に対する Pearson 相関係数 r_l へ，

$$r_l = \frac{\sum_{k,k'} kk' P(k, k'|l) - \{\sum_k k Q(k|l)\}^2}{\sum_k k^2 Q(k|l) - \{\sum_k k Q(k|l)\}^2}, \quad (5.15)$$

のように拡張できる．また，(2.8) 式で定義される平均隣接次数 $k_{nn}(k) = \sum_{k'} k' P_{nn}(k'|k)$ の拡張として，次数 k のノードからノード間距離 l にあるノードの平均次数（平均 l 次隣接次数）を，

$$k_l(k) = \sum_{k'} k' P(k'|k, l), \quad (5.16)$$

によって定義することもできる．指標 r_l と $k_l(k)$ は Mayo ら [57] によって提案された指標と同じものであるが，表 5.1 の確率関数を用いることによって，単なる隣接次数相関の指標の拡張ではなく，長距離次数相関のための全く新しい指標を考えることもできる．そのような例については，5.3 節および 6.3 節で後述する．

5.2 相関が無いネットワーク

5.1 節では複雑ネットワークの長距離次数相関を記述する最も基本的な確率関数を導入した．本節では，与えられたネットワークにおける長距離次数相関の有無を判断するための基準となる確率関数の関数形を求める．すなわち，長距離次数相関が無いネットワークにおいて表 5.1 の確率関数がどのような関数形になるのかについて考える．隣接次数相関の議論でも，隣接次数相関が無いネットワーク（隣接無相関ネットワーク）における $P_{nn}(k, k')$ や $P_{nn}(k|k')$ の形が相関の有無の判断基準となっているので，そこでの議論を参照してみよう．「隣接次数相関が無い」とは，エッジ両端にあるノードの片方の次数が k である事象と，もう片方の次数が k' である事象が，互いに独立であることを意味する．したがって 2.3 節では，隣接無相関ネットワークにおける同時確率 $P_{nn}(k, k')$ を $Q_{nn}(k)$ と $Q_{nn}(k')$ の関数の積 $Q_{nn}(k)Q_{nn}(k')$ として与えた．この考えを拡張すると，長距離無相関ネットワークとは「任意の距離 l に対してノード間距離 l だけ離れたノード対の次数どうしが互いに独立であるネッ

トワーク」であると考えられる．この条件を式で表すと，長距離無相関ネットワークが満たすべき関係式は，

$$P_0(k, k'|l) = Q_0(k|l)Q_0(k'|l), \quad (5.17)$$

である．ここで， $P_0(k, k'|l)$ および $Q_0(k|l)$ はそれぞれ長距離無相関ネットワークにおける $P(k, k'|l)$ および $Q(k|l)$ である．本論文では以後，長距離無相関ネットワークに対する関数や指標であることを，添え字 “0” を付加することで表現する．(5.9) 式が長距離無相関ネットワークに対しても成立する一般的な関係式であることを踏まえると，この式と (5.17) 式より

$$P_0(k'|k, l) = Q_0(k'|l), \quad (5.18)$$

を得る．上式は (5.17) 式と等価であり，これを長距離無相関ネットワークの定義とみなすこともできる．

隣接次数相関を記述する同時確率 $P_{nn}(k, k')$ が隣接無相関ネットワークに対して $P_{nn}(k, k') = kk'P(k)P(k')/\langle k \rangle^2$ という簡単な関数形で表される一方で，長距離無相関ネットワークにおける条件付き確率 $P_0(k, k'|l)$ を求めることは容易ではない．なぜならば， $Q_{nn}(k)$ が隣接次数相関の有無に依存しないのとは異なり， $Q(k|l)$ はネットワークの長距離次数相関によって変化するためである．加えて先に述べたように，長距離次数相関を記述する確率関数のうちのいくつかは，無限サイズのネットワークにおいては意味を失う．無限に大きな連結成分をもつようなネットワークにおいては，平均ノード間距離 $\langle l \rangle$ が発散するため，有限なノード間距離 l を変数とするような分布関数は常にゼロとなってしまう．このことは，単にノード間距離分布 $R(l)$ が定義できないだけでなく，同時確率 $P(k, k', l)$ や条件付き確率 $P(l|k, k')$ および $P(k', l|k)$ も定義できないことを意味している．したがってネットワークの長距離次数相関の議論は，本質的にネットワークの有限性を前提としていることとなる．

長距離次数相関の議論においてネットワークの有限性が前提とすると，我々は極めて重要な事実直面する．すなわち，長距離無相関ネットワークの定義 (5.17) を厳密に満たす有限サイズのネットワークは存在しないと言う事実である．なぜならば，エッジのつながり方が完全にランダムなネットワーク（Erdős-Rényi ランダムグラフや一般化ランダムグラフ）であっても，ネットワーク・サイズが有限である場合は有限サイズ効果による不可避的で非本質的な次数相関が生じるからである．実際，図 5.1 に示すように，有限サイズの Erdős-Rényi ランダムグラフに対して数値的に計算された (5.17) 式の左辺と，同じく数値的に計算された (5.17) 式の右辺とは大きな l に対して一致しなくなる．このような非本質的な相関は，自己ループや多重エッジが許されない場合， $l = 1$ のスケールでも生じている．ただし，ネットワーク・サイズが十分に大きい場合はこのような相関の影響は無視できるほど小さくなるため，十分に大きな Erdős-Rényi ランダムグラフは便宜的に隣接無相関とみなされる．同様

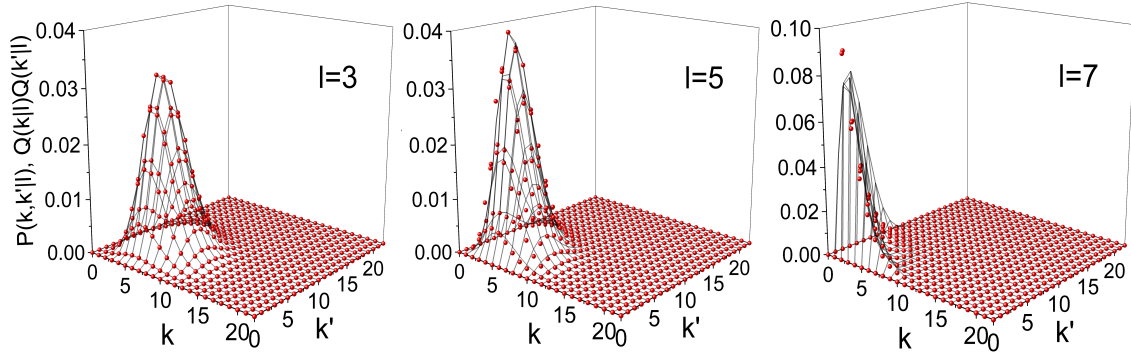


図 5.1: Erdős-Rényi ランダムグラフ ($N = 1,000$, $\langle k \rangle = 5.0$, $\langle l \rangle = 4.47$) に対する数値計算結果. 黒のワイヤフレームは (5.17) 式の左辺 $P(k', k|l)$, 赤のシンボルは (5.17) 式の右辺 $Q(k|l)Q(k'|l)$ を表す. 値は 100 サンプルにわたり平均化されている.

に, 十分に大きなランダム・ネットワークでは, 平均ノード間距離よりも十分小さいスケール $l \ll \langle l \rangle$ で (5.17) 式を近似的に満たすことになる. このことは図 5.1 から明らかである. したがって, 熱力学極限 $N \rightarrow \infty$ では厳密な意味での長距離無相関ネットワークに至るものと考えられる. このことから, ランダム・ネットワークは, 有限サイズ効果に由来する非本質的な相関以外に長距離次数相関をもたないネットワークであると結論づけられる. 現実ネットワークのサイズが有限であることを考えると, 与えられたネットワーク $\mathcal{G}$ の確率関数と, $\mathcal{G}$ と同じ次数列 (すなわち同じノード数および同じ次数分布) を有するランダム・ネットワークの確率関数を比較することで, $\mathcal{G}$ の長距離次数相関を評価することが実用上の観点からは有用である. 本論文では, このようなランダム・ネットワークを $\mathcal{G}$ に対する長距離無相関ネットワークとみなすことにする. 次節では, 有限サイズの長距離無相関ネットワークの確率関数を平均場近似と局所木構造仮定を用いて解析的に計算する方法を説明する. この近似理論は非常に高い精度で確率関数を計算することができるものの, 漸化式を逐次的に計算する方法であるため, 確率関数を閉じた形の関数形で表すことができない. これに対し, 5.2.2 節では, 無限に大きな単連結ネットワークが木構造を取ると仮定した場合, 無限ネットワークにおいても意味を失わない確率関数 $P(k, k'|l)$ および $P(k'|k, l)$ は閉じた形に簡単化できることを示す.

5.2.1 平均場近似

有限サイズの長距離無相関ネットワークに対して表 5.1 の確率関数を導出することは隣接無相関ネットワークの場合のように簡単ではない. しかし 5 つある確率関数のうちの一つでも関数形が得られれば関係式 (5.1) より他の確率関数は直ちに導出できる. そこで次数 k と

k' をもつノード対間の距離分布 $P(l|k, k')$ に注目する．長距離次数相関の議論を目的としたものではないが，ランダム・ネットワークの距離分布 $R(l)$ を近似的に計算する手法は幾つか提案されている [68, 90, 106, 107]．それらの研究の中には，次数 k と k' をもつノード対間の距離分布を計算するものもある．特に，平均場近似と局所木構造仮定を用いた Melnik と Gleeson の理論 [107] は最も近似の精度が高く，数値計算とも非常によく一致する．本論文では Melnik らの考え方を利用し，条件付き確率 $P_0(l|k, k')$ を導出することにする．ただし，Melnik らの理論では $P(l|k, k')$ における k と k' に対する対称性の扱いや，ネットワークの非連結成分の扱いなど，いくつか誤りがあるため，それらを修正した理論を以下で詳細に述べる．

まず，任意抽出した次数 k と k' のノード対間の距離が l 以下である確率 $\rho(l|k, k')$ を導入する．この $\rho(l|k, k')$ を用いると $P_0(l|k, k')$ は，

$$P_0(l|k, k') = \begin{cases} \rho(0|k, k') & (l = 0), \\ \rho(l|k, k') - \rho(l-1|k, k') & (l \geq 1), \\ 1 - \lim_{l \rightarrow \infty} \rho(l|k, k') & (l = l_\infty), \end{cases} \quad (5.19)$$

と表される．なお，ノード対間の距離が $l = 0$ であるとは「ノード対が同じノードであること」を意味する．また， $l = l_\infty$ は非単連結のネットワークにおける便宜的な距離であり，ノード対間の距離が $l = l_\infty$ であることは「ノード対間に経路が存在しないこと」を意味する．したがって $P(l|k, k')$ の規格化条件は，

$$\sum_{l=0}^{\infty} P_0(l|k, k') + P_0(l_\infty|k, k') = 1 \quad (5.20)$$

である．Melnik と Gleeson の論文 [107] では $\rho(l_{\max}|k, k')$ をノード対がどちらも最大連結成分（直径 $l_{\max}$ ）に所属する確率とし，

$$P_0(l|k, k') = \frac{\rho(l|k, k') - \rho(l-1|k, k')}{\rho(l_{\max}|k, k')} \quad (5.21)$$

を最大連結成分の $P_0(l|k, k')$ としているが，これは誤りである．以下で説明する平均場理論は，最大連結成分に限った距離分布の計算を行っておらず，非連結ノード対を含めた全ての成分に関する平均的なノード対に対する $P_0(l|k, k')$ の計算となっている．実際，最大連結成分に対しては $P_0(l=1|k=1, k'=1) = 0$ でなければならないが，(5.20) 式から計算される $P_0(l=1|k=1, k'=1)$ はゼロとならない．

ノード間距離 $l \geq 1$ において確率 $\rho(l|k, k')$ に対して関係式

$$\rho(l|k, k') = 1 - [1 - \rho(0|k, k')] [1 - \bar{q}(l-1|k, k')]^k, \quad (5.22)$$

を立てる．右辺の $\rho(0|k, k')$ は、「任意抽出した次数 k と k' のノードを i_k と $j_{k'}$ としたとき、 i_k が $j_{k'}$ 自身である確率」を表す． $\bar{q}(l-1|k, k')$ は、「 i_k の隣接ノードから任意に 1 つを選んだとき、そのノードから $j_{k'}$ まで i_k を通らない長さ $l-1$ 以下の経路が存在する確率」を意味する． i_k のある隣接ノードから $j_{k'}$ まで i_k を通らない長さ $l-1$ 以下の経路を有する確率は i_k の全ての隣接ノードで等しくかつ互いに独立であるから「 i_k のいずれの隣接ノードから $j_{k'}$ までの長さ l 以下の経路が存在しない確率」はネットワークが局所的にはサイクルの無い木構造ネットワークであるという木構造仮定のもとで $1 - \bar{q}(l-1|k, k')$ の k 乗の形で書ける．したがって (5.22) 式が示すおおよその意味は、「 i_k が $j_{k'}$ から距離 l 以内にある確率」(左辺) が、「 i_k が $j_{k'}$ 自身であるか、あるいは k 個の隣接ノードのいずれかから $j_{k'}$ まで i_k を通らない長さ $l-1$ 以下の経路が存在する確率」(右辺) と等しいことである．ここで $\bar{q}(l|k, k')$ は、新たな確率 $q(l|k, k')$ を用いて、

$$\bar{q}(l|k, k') = \sum_{k''} P_{nn}(k''|k) q(l|k'', k'), \quad (5.23)$$

から計算される．確率 $q(l|k'', k')$ は、「任意に選んだノード $j_{k'}$ と任意に選んだエッジ e の端点 $h_{k''}$ (次数 k'') に対して、 $h_{k''}$ から $j_{k'}$ まで e のもう片方の端点 i を通らない長さ l 以下の経路が存在する確率」を意味している．ここノード i の次数を k とすると、 $P_{nn}(k''|k) q(l|k'', k')$ は「任意に選んだ次数 k のノード i_k の隣接ノードの次数が k'' であり、かつその隣接ノードから $j_{k'}$ まで i_k を通らない長さ l 以下の経路が存在する同時確率」を意味することになる．それゆえ、 k'' に対する和 $\sum_{k''} P_{nn}(k''|k) q(l|k'', k')$ 、すなわち (5.23) 式の右辺は、「任意に選んだノード $j_{k'}$ と、任意に選んだエッジ e の端点ノード i_k に対して、 e のもう片方の端点から $j_{k'}$ まで i_k を通らない長さ l 以下の経路が存在する確率」となり、 $\bar{q}(l|k, k')$ の定義と一致する．一方、(5.22) 式の成立根拠と同様の理由から、 $l \geq 1$ における $q(l|k, k')$ と $\bar{q}(l-1|k, k')$ の間には、(5.22) 式とよく似た、

$$q(l|k, k') = 1 - [1 - \rho(0|k, k')] [1 - \bar{q}(l-1|k, k')]^{k-1}, \quad (5.24)$$

の関係が成り立つ． $l = 0$ に対する $q(0|k, k')$ は「任意に選んだノード $j_{k'}$ と、任意に選んだエッジの端点 i_k が同じノードである確率」となるが、この確率は次数 k と k' をもつ 2 つのノードを任意抽出したときに同じノードが選ばれる確率 $\rho(0|k, k')$ に等しい． $\rho(0|k, k')$ は $k \neq k'$ においてゼロであり、 $k = k'$ においては次数 k をもつノードの数 $NP(k)$ の逆数となるから、

$$q(0|k, k') = \rho(0|k, k') = \frac{\delta_{kk'}}{NP(k)}, \quad (5.25)$$

である．図 5.2 は関係式 (5.22)、(5.24)、(5.26) の意味を表す概念図である．

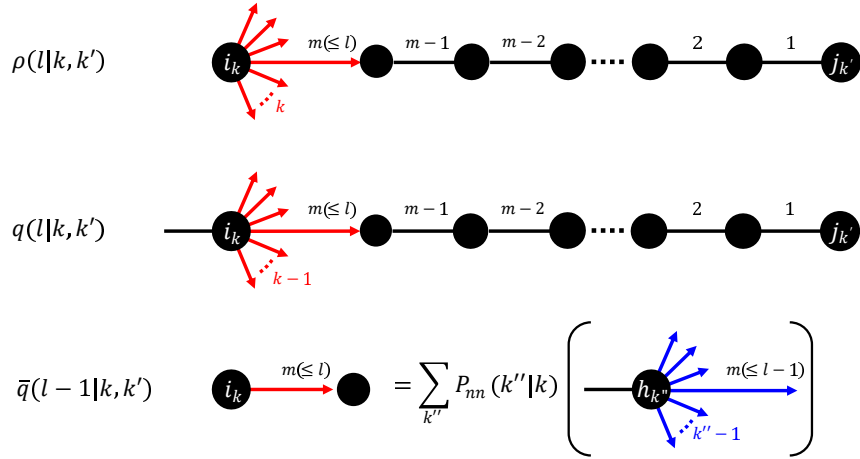


図 5.2: 上から順に関係式 (5.22), (5.24), (5.26) の意味を表す概念図. 赤の矢印の先端にあるノードから, $j_{k'}$ まで矢印の根本のノード i_k を通らない長さ $l-1$ 以下の経路が存在する確率は $\bar{q}(l-1|k, k')$ である. (上) 任意に選んだ2つのノードを i_k と $j_{k'}$ としたとき, i_k から $j_{k'}$ まで長さ l 以下 (図中では長さ m) の経路が存在する確率 $\rho(l|k, k')$ は, i_k が $j_{k'}$ 自身であるか, あるいは k 個の隣接ノード (赤矢印の先) のうち少なくとも1つから $j_{k'}$ まで長さ $l-1$ 以下の経路が存在する確率に等しい. (中) 任意に選んだエッジの端点 i_k から任意に選んだノード $j_{k'}$ まで, エッジのもう片方の端点を通らない長さ l 以下の経路が存在する確率 $q(l|k, k')$ は, i_k が $j_{k'}$ 自身であるか, あるいはエッジのもう片方の端点を除いた $k-1$ 個の i_k の隣接ノードのうち, 少なくとも1つから $j_{k'}$ まで i_k を通らない長さ $l-1$ 以下の経路が存在する確率に等しい. (下) あるノード i_k の任意に選んだ隣接ノードから, 任意に選んだノード $j_{k'}$ まで i_k を通らない長さ $l-1$ 以下の経路が存在する確率 $\bar{q}(l|k, k')$ は, 選ばれた隣接ノード ($h_{k''}$ とする) の次数が k'' でありかつ $h_{k''}$ から $j_{k'}$ までの距離が $l-1$ 以下である同時確率に対して k'' の和をとった周辺確率として与えられる.

次数相関の無いネットワークに対しては (2.5) 式が成り立つので, (5.23) 式は

$$\bar{q}(l|k, k') = \sum_{k''} \frac{k'' P(k'')}{\langle k \rangle} q(l|k'', k'), \quad (5.26)$$

となる. このとき $\bar{q}(l|k, k')$ は k には依存しないため, 以後はこの量を $\bar{q}(l|k')$ とかく. (5.24) 式の両辺に $kP(k)/\langle k \rangle$ をかけ k に対する和をとると, (5.26) 式から $\bar{q}(l|k')$ の l に対する漸化式,

$$\bar{q}(l|k') = 1 - G_1 [1 - \bar{q}(l-1|k')] + \frac{k'}{N\langle k \rangle} [1 - \bar{q}(l-1|k')]^{k'-1}, \quad (5.27)$$

が得られる. ここで, 関数 $G_1(x)$ は隣接ノードの残り次数分布の母関数であり, $G_1(x) = \sum_k x^{k-1} kP(k)/\langle k \rangle$ により定義される. $l=0$ における (5.25) 式から得られる初期条件,

$$\bar{q}(0|k') = \frac{k'}{N\langle k \rangle}, \quad (5.28)$$

の下、漸化式 (5.27) を逐次的に解くことによって任意の $l (\geq 1)$ に対する $\bar{q}(l|k')$ が得られる。この $\bar{q}(l|k')$ と (5.25) 式を使うと、(5.22) 式から $\rho(l|k, k')$ が求められ、さらに (5.19) 式より $P_0(l|k, k')$ が得られる。他の確率関数 $P_0(k, k', l)$, $P_0(k'|k, l)$, $P_0(k, k'|l)$, $P_0(k, l|k')$ は 5.1 節の議論から計算される。具体的には、まず、(5.7) 式と (5.1a) 式を使って、同時確率 $P_0(k, k', l)$ を

$$P_0(k, k', l) = P_0(l|k, k')P(k)P(k') \quad (5.29)$$

と計算する。同時確率が求められれば他の条件付き確率は (5.1) 式から計算することができる。

上記の理論は平均場近似を用いている。そのため、スケールフリー性のような強い次数不均一性をもつネットワークに適用する場合には注意が必要である。その定義から確率 $\rho(l|k, k')$ は k と k' に対し対称な関数でなければならないが、(5.22) 式より計算される $\rho(l|k, k')$ は $l \geq 1$ において k と k' に対して非対称となる。この事実は (5.22) 式からすぐに確かめることができ、実際 $l = 1$ かつ $k \neq k'$ において、

$$\begin{aligned} \rho(1|k, k') &= 1 - \left(1 - \frac{k'}{N\langle k \rangle}\right)^k \\ &= \frac{kk'}{N\langle k \rangle} - \frac{1}{2} \frac{k(k-1)k'^2}{(N\langle k \rangle)^2} + \dots, \end{aligned} \quad (5.30)$$

のように N^{-2} のオーダーの非対称性が表れる。この非対称性は、次数 k と k' をもつノードの隣接ノードに対する平均場近似の精度の差に起因する。平均場近似は隣接ノードの性質に大きな揺らぎが無いことを仮定しているが、スケールフリー・ネットワークのような次数揺らぎの大きなネットワークでは、たとえ同じ次数 k を有するノードであっても隣接するノードの次数揺らぎが大きいために隣接ノードからの寄与には大きな不均一性が生じる。揺らぎによる平均場近似の誤差は $l = 1$ においてはわずかであるが、大きな l に対しては、複数回に渡る (5.27) 式の計算によって逐次誤差が増幅されていくため大きくなる。そのような揺らぎは隣接ノードに渡る平均化によって抑制されるため、隣接ノードに関する平均場近似は、隣接ノード数の小さな（すなわち次数の小さな）ノードに対してよりも、次数の高いノードに対しての方が高い精度をもつ。本計算においては、 $\bar{q}(l|k')$ を次数 k' の始点ノードから (5.27) 式にしたがって逐次的に計算するので、 $k < k'$ のときには、高次数ノードを始点ノードとして計算される $\rho(l|k, k')$ の方が $\rho(l|k', k)$ よりも高い精度をもつことになる。したがって、 $k < k'$ のとき計算される $P_0(l|k, k')$ の精度は、 $\rho(l|k, k')$ から計算される値の方が $\rho(l|k', k)$ から計算される値よりも真の値に近くなる。図 5.3 は、ランダムグラフに対する $P_0(l|k, k')$ の理論計算（実線と破線）と数値計算結果（シンボル）を比較したものである。次数 $k = 2$ のノードと次数 $k' = 10$ のノードの間の距離を $\rho(l|k = 2, k' = 10)$ から計算した場合（実線）と、

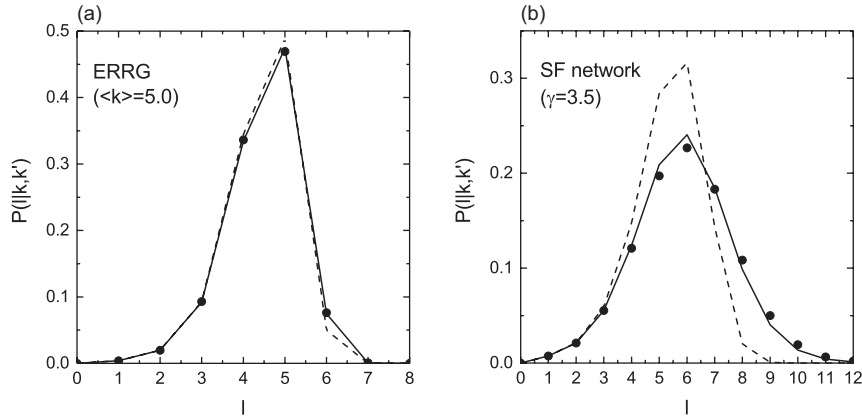


図 5.3: 次数 2 と次数 10 をもつノード対の間のノード間距離分布 $P(l|k, k')$. シンボルはランダムグラフに対して数値計算で得られた $P(l|k=2, k'=10) [= P(l|k=10, k'=2)]$ である. 実線および破線は解析計算の結果を表し, 実線は $\rho(l|k=2, k'=10)$ から, 破線は次数 $\rho(l|k=10, k'=2)$ から計算された $P_0(l|k, k')$ を表している. (a) は $\langle k \rangle = 5.0$ で $N = 1,000$ の Erdős-Rényi ランダムグラフに対する結果であり, 数値計算は 100 サンプルに渡る平均である. (b) は次数分布関数が $P(k) \propto k^{-3.5}$ となるような次数列をもつ $N = 1,000$ の一般化ランダムグラフに対する結果であり, 10,000 サンプルに渡る平均をとっている.

$\rho(l|k=10, k'=2)$ から計算した場合 (破線) を比較すると, 明らかに $\rho(l|k=2, k'=10)$ から計算した実線の方が数値計算であるシンボルに近い値を示す. これらのことから, 本論文では $k < k'$ に対する $\rho(l|k', k)$ を $\rho(l|k, k')$ を計算することにより与える.

平均場近似の精度は, 臨界点 $\langle k^2 \rangle / \langle k \rangle = 2$ 近傍においても低下する. $\rho(l|k, k')$ の計算では次数 k' の始点ノードが属する連結成分のサイズについても十分に揺らぎが小さいことを想定しているが, 臨界点近傍においては連結成分サイズの揺らぎ (標準偏差) が発散するため, 近似の精度が著しく低下する. 実際どの程度の誤差が生じるかについては, 付録 C にて示す. なお, (5.22) 式の $[1 - \bar{q}(l|k, k')]^k$ や (5.24) 式の $[1 - \bar{q}(l|k, k')]^{k-1}$ のように複数の隣接ノード全てが特定の条件を満たす確率を, 各隣接ノードがその条件を満たす確率の積の形で表すことができるのは, ネットワーク中のサイクルが無視できるほど少なく, 隣接ノード間で事象が独立とみなせる局所木構造仮定が成り立つことを要請する. ただし, 実際にはかなりサイクルの多いネットワークに対しても近似の精度は十分に保たれることが, 付録 C で示した数値計算結果から確かめられる.

近似計算の精度を確認すべく, ランダムグラフに対して数値的に計算された $P_0(l|k, k')$, $P_0(k'|k, l)$, $P_0(k, k'|l)$ と理論計算結果とを比較した. 図 5.4 に $l = 4$ における各種確率関数の k および k' 依存性を示す. ワイヤフレームとシンボルはそれぞれ解析計算と数値計算の結果を表す. 上段は平均次数 $\langle k \rangle = 5.0$, ノード数 $N = 1,000$ の Erdős-Rényi ランダムグラフに対する結果である. 数値計算結果は 100 サンプルにわたり平均化されている. 有限サ

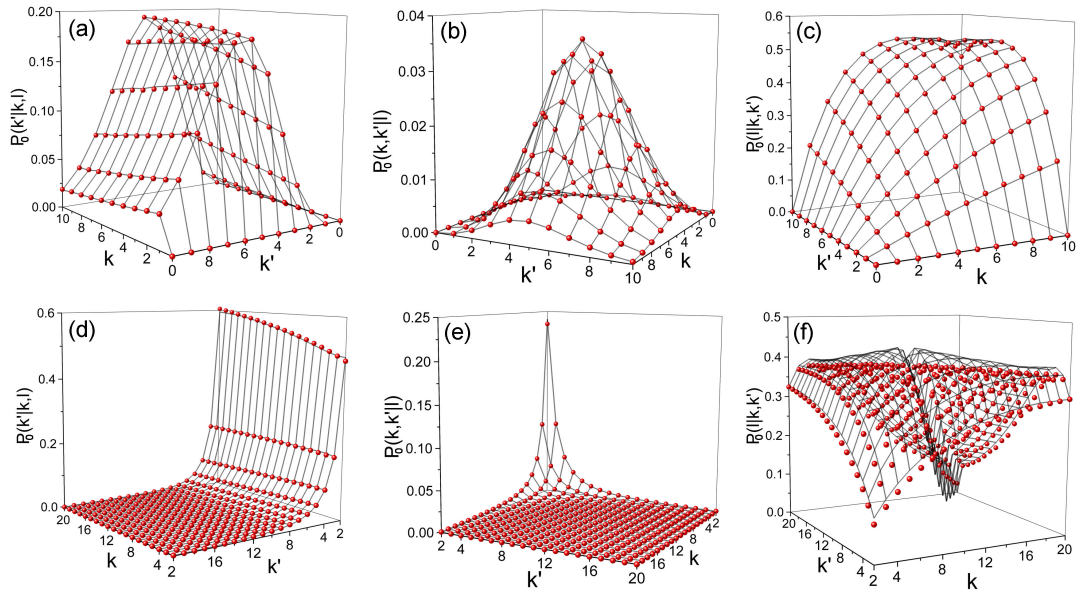


図 5.4: 条件付き確率 $P_0(k'|k, l)$, $P_0(k, k'|l)$, $P_0(l|k, k')$ ($l = 4$, $N = 1,000$) の k および k' 依存性. ワイヤフレームは解析計算の結果, シンボルは数値計算の結果を表す. 上段 [(a)-(c)] は Erdős-Rényi ランダムグラフ ($\langle k \rangle = 5.0$), 下段 [(d)-(f)] は次数分布 $P(k) \propto k^{-3}$ に従うスケールフリー性をもつ一般化ランダムグラフに対する結果である.

イズであっても近似の精度が保たれることが分かるようにあえて比較的小さいサイズのネットワークを用いている. 平均ノード間距離は $\langle l \rangle = 4.5$ である. 下段はノード数 $N = 1,000$ の一般化ランダムグラフに対する結果である. ネットワークはスケールフリー性を有しており, 次数分布関数は $2 \leq k \leq 50$ において $P(k) \propto k^{-3}$, それ以外の次数 k では $P(k) = 0$ とした. 数値計算結果は 10,000 サンプルにわたり平均化されている. 平均次数, 平均ノード間距離はそれぞれ $\langle k \rangle = 3.1$, $\langle l \rangle = 5.4$ である. 上下段ともに解析計算の結果は数値計算の結果を非常によく再現し, 近似理論が有限サイズのネットワークにおいても高い精度を保持することが認められる.

5.2.2 木構造無限サイズ・ネットワークにおいて

平均場理論からは精度の良い確率関数を解析的に求めることができるが, その計算方法は漸化式を逐次的に解くというものであり, 閉じた形での関数形を得ることはできない. 一方で, 木構造をもつ無限サイズの単連結ネットワークの場合は, $\langle l \rangle$ の発散によっていくつかの確率関数が定義できなくなってしまうものの, 次数 k からノード間距離 l にあるノードの次数分布 $P_0(k'|k, l)$ を厳密に, かつ閉じた形で求めることができる.

木構造をもつネットワークではノード対間の経路が一意に定まるため, 次数 k のノード

i からノード間距離 l にあるノード j に隣接し、ノード i から距離 $l-1$ にあるノードは一意に定まる．このノードを h としその次数を k'' とする．長距離無相関ネットワークは $l=1$ における隣接次数相関も持たないため、ノード h の隣接ノードが次数 k' をもつ確率は $P_{\text{nn}}(k'|k'') = k'P(k')/\langle k \rangle$ となる．したがって、ノード i が次数 k をもつ確率 $P_0(k'|k, l)$ は、 $l \geq 1$ に対して

$$P_0(k'|k, l) = \sum_{k''} P_{\text{nn}}(k'|k'') P_0(k''|k, l-1) = \frac{k'P(k')}{\langle k \rangle}, \quad (5.31)$$

と計算される． $l=0$ においては $P_0(k'|k, l=0) = \delta_{kk'}$ であるから、

$$P_0(k'|k, l) = \begin{cases} \delta_{kk'} & (l=0), \\ \frac{k'P(k')}{\langle k \rangle} & (l \geq 1), \end{cases} \quad (5.32)$$

となり．ここで、 $l \geq 1$ において $P_0(k'|k, l)$ は l に依存しないことに注意する．また、(5.17) 式と (5.18) 式より、

$$Q_0(k|l) = \frac{kP(k)}{\langle k \rangle} \quad (l \geq 1), \quad (5.33)$$

$$P_0(k, k'|l) = \frac{kk'P(k)P(k')}{\langle k \rangle^2} \quad (l \geq 1), \quad (5.34)$$

を得る．すなわち、木構造を有する無限サイズの長距離無相関ネットワークにおいては、 $P_0(k, k'|l)$ と $P_0(k'|k, l)$ は l に依存せず、それぞれ隣接次数相関が無いときの $P_{\text{nn}}(k, k')$ と $P_{\text{nn}}(k'|k)$ と同じ関数形で表されることになる．

上記のように求められた無限サイズの長距離無相関ネットワークにおける $P_0(k'|k, l)$ および $P_0(k, k'|l)$ の表式は、無限サイズのネットワークだけでなく、有限サイズの長距離無相関ネットワークにおいても $l \ll \langle l \rangle$ の領域において近似的に成立する．これは、有限サイズであっても N が十分大きいネットワークにおいてはサイクルの存在が無視できるほど疎になり、 $l \ll \langle l \rangle$ 領域において局所的な木構造が現れるためである．図 5.5 は、 $l \ll \langle l \rangle$ において (5.32) 式および (5.34) 式が成立していることを数値計算との比較から確認したものである．ノード数 $N = 1,000$ 、平均次数が $\langle k \rangle = 5.0$ 、平均ノード間距離が $\langle l \rangle = 4.46$ の Erdős-Rényi ランダムグラフに対して、 $l \leq 3$ の $P(k'|k, l)$ と $P(k, k'|l)$ が、それぞれ (5.32) 式および (5.34) 式でよく近似されていることが分かる．5.2.1 節で議論した平均場理論は、(5.17) 式を満たす厳密な意味での長距離無相関ネットワークに対する理論ではなく、有限サイズのランダム・ネットワーク [68] に対し、条件付き確率 $P(l|k, k')$ を計算するものであった．しかし有限ランダムグラフが $l \ll \langle l \rangle$ の領域においては、(5.17) 式が近似的に成立することを反映し、ネットワークが有限サイズであっても $l \ll \langle l \rangle$ であれば (5.32)-(5.34) 式が成り立つことになる．

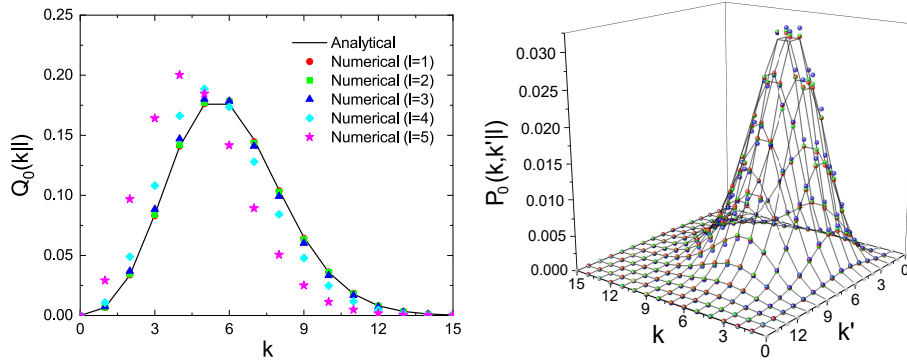


図 5.5: $N = 1,000$, $\langle k \rangle = 5.0$ の Erdős-Rényi ランダムグラフに対する (a) $Q_0(k|l)$ と (b) $P_0(k, k'|l)$. (a) のシンボルは, $l = 1, 2, \dots, 5$ に対する $Q_0(k|l)$ の数値計算結果を表しており, 実線は無限系に対する理論式 (5.33) を示す. (b) の赤, 緑, 青のシンボルは, それぞれ $l = 1, 2, 3$ に対する $P_0(k, k'|l)$ の数値計算結果であり, 黒のワイヤフレームは無限系に対する理論式 (5.34) を示す. 数値計算は 100 サンプルに渡る平均である.

5.3 現実ネットワークの評価と新たな評価指標

前節までの議論より, 長距離次数相関の有無は与えられたネットワークに対する確率関数 P と, そのネットワークに対応した長距離無相関ネットワークの確率関数 P_0 とを比較することにより知ることができる. しかしながら, ネットワークが長距離次数相関をもつ場合, 本章で導入した 5 種の確率関数が 3 変数関数であるため, それらの関数形から, ネットワークにどのような種類の長距離次数相関が潜んでいるのかを直感的に理解することは容易ではない. そこで変数の縮約を行うことで, その意味が直感的に理解しやすい関数や相関係数のような指標へと落とし込み, 与えられたネットワークの長距離次数相関の評価を行う.

どの変数をどのように縮約するか of のやり方は無数に考えられるが, 一つの方性としては隣接次数相関の議論においてすでにその有用性が認められた関数や指標を長距離次数相関に拡張するというものがある. 隣接次数相関は, エッジ両端のノードの次数 k と k' を変数とする確率関数 $P_{nn}(k, k')$ および $P_{nn}(k'|k)$ によって完全に記述されるものであった. これらの確率関数を用いて平均隣接次数 $k_{nn}(k)$ や, エッジ両端の次数に対する Pearson 相関係数のような量が定義される. こうした指標を用いることで, 複雑ネットワークの隣接次数相関の系統的な議論が可能となる. このような指標を隣接以上にノード間距離が離れたノード対間の次数相関, すなわち長距離次数相関へと拡張することが可能なことは (5.15) 式や (5.16) 式ですでに示した. ここではいくつかの現実世界の複雑ネットワークを例にとりあげ, 実際に平均 l 次隣接次数 $k_l(k)$ による長距離次数相関の評価を行うことで, その有用性を実証する. 対象とするネットワークは Gnutella ネットワーク [62], Google+ ネットワーク [85], Twitter ネットワーク [85] である. Gnutella ネットワーク ($\langle k \rangle = 7.4$) は, Peer-to-Peer

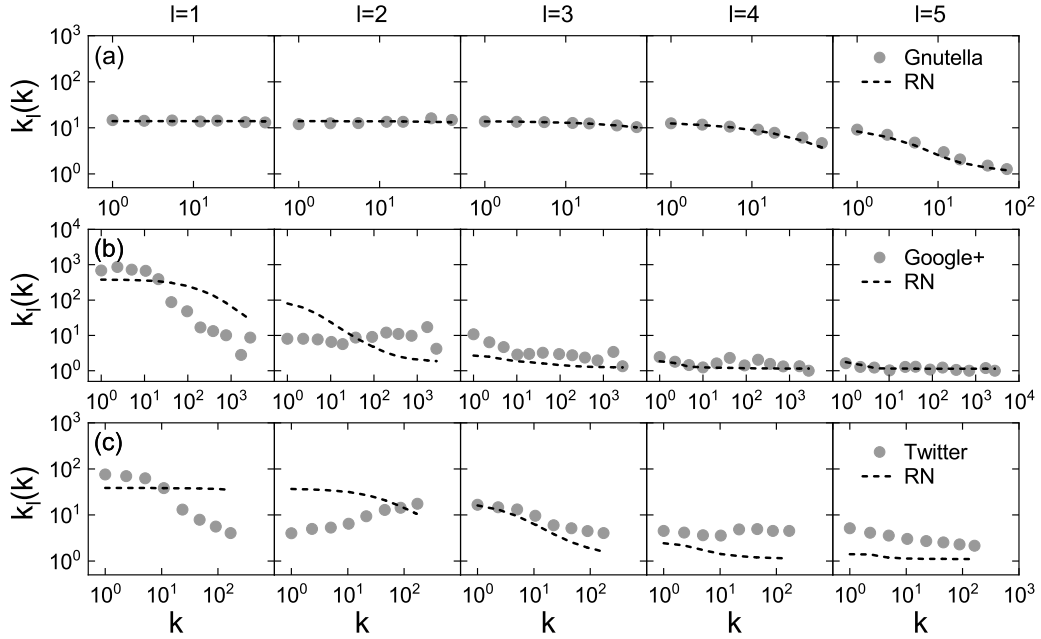


図 5.6: 左から $l = 1, 2, \dots, 5$ のそれぞれの l における現実世界の複雑ネットワークの平均 l 次隣接次数 $k_l(k)$. 上から (a)Gnutella ネットワーク, (b)Google+ ネットワーク, (c)Twitter ネットワークに対する結果である. 灰色のシンボルは実ネットワークとから数値的に計算された $k_l(k)$ を表し, 黒の破線はそれぞれの実ネットワークと同じ次数列を有する長距離無相関ネットワークの $k_l(k)$ である.

ファイル共有サービスである Gnutella を利用する 10,876 のホストをノードとした通信ネットワークである. Google+ ネットワーク ($\langle k \rangle = 3.3$) は, SNS である Google+ を利用する 23,628 のユーザ間での交友関係ネットワークである. Twitter ネットワーク ($\langle k \rangle = 2.8$) は, 同じく SNS である Twitter を利用する 23,370 のユーザ間でのフォロー/フォロワー関係を表している. 平均ノード間距離は, Gnutella ネットワーク, Google+ ネットワーク, ならびに Twitter ネットワークのそれぞれに対して $\langle l \rangle = 4.64, 4.03, 6.30$ である. なお, これらのネットワークを解析するにあたり, エッジの重さや方向は無いものとした. 図 5.6 のシンボルは実ネットワークにおける平均 l 次隣接次数 $k_l(k)$ ($l = 1, 2, \dots, 5$) の実測値を示す. 黒の破線は, 各々の実ネットワークと同じ次数列を有する長距離無相関ネットワークの $k_l(k)$ であり, 5.2.1 節の平均場近似によって解析的に得られた $P_0(k|k', l)$ から計算されたものである. 図 5.6(a) から明らかなように, Gnutella ネットワークの $k_l(k)$ (シンボル) と対応する長距離無相関ネットワークの $k_l(k)$ は, 全ての l において一致している. このことは, Gnutella ネットワークが長距離にわたり無相関なネットワークであることを示している. 一方で, 図

5.6 の Google+ ネットワーク (b) と Twitter ネットワーク (c) の $k_l(k)$ は、少なくとも $l \leq 3$ の領域において、長距離無相関ネットワークに対する $k_l(k)$ と大きく異なっている。これは Google+ ネットワーク (b) と Twitter ネットワーク (c) が長距離次数相関を有することを意味する。6 章で詳しく述べるが、図 5.6 の 3 種のネットワークに加えて 18 種の実ネットワークを解析した結果、Gnutella ネットワークのような無相関なネットワークはごく少数であり、ほぼ全ての実ネットワークは何かしらの長距離次数相関を有していることがわかった。

ノード対の片側ノードの次数に対して平均操作を行っている平均 l 次隣接次数 $k_l(k)$ では、どのような次数をもつノードどうしにどのような長距離次数相関が表れているのかわからない。そこで、次数 k と k' を有するノード対間の平均ノード間距離

$$\langle l(k, k') \rangle = \sum_l l \tilde{P}(l|k, k') \quad (5.35)$$

をまず定義する。ここで $\tilde{P}(l|k, k')$ は、任意に選んだノード対が同じ連結成分に所属し、かつ次数が k と k' であるときに、ノード間の距離が l である条件付き確率

$$\tilde{P}(l|k, k') = \frac{P(l|k, k')}{\sum_{l=1}' P(l|k, k')} \quad (5.36)$$

である。 $\sum_l'$ は、 $l = l_\infty$ を含まない l に関する和を意味する。さらに、解析対象のネットワークと同じ次数列を有する長距離無相関ネットワークに対しても同様の平均ノード間の距離 $\langle l(k, k') \rangle_0$ を求め、両者の比

$$\langle l(k, k') \rangle_{\text{res}} = \frac{\sum_l l \tilde{P}(l|k, k')}{\sum_l l \tilde{P}_0(l|k, k')}, \quad (5.37)$$

をとることにより、新たな指標 $\langle l(k, k') \rangle_{\text{res}}$ を導入する。この相対平均ノード間距離が 1 より大きいか小さいかを調べることにより、長距離無相関ネットワークに比べてどのような次数をもつノードどうしが近距離に位置しやすいのか、あるいは長距離にわたり離れやすいのかを明らかにすることができる。図 5.7 は、特徴的な長距離次数相関を有する現実世界の複雑ネットワークである (a) Twitter ネットワーク [85], (b) タンパク質折り畳み構造ネットワーク [86], (c) WWW (ノートルダム大学のドメイン内) [85] の相対平均ノード間距離 $\langle l(k, k') \rangle_{\text{res}}$ を示している。これらのネットワークでは、ほとんど全ての次数 k, k' において $\langle l(k, k') \rangle_{\text{res}}$ が 1 よりも大きい。この事実は、現実ネットワークの長距離次数相関がどのようなノード対間に対してもその距離を増加させる傾向があることを示唆する。特に (b) タンパク質折り畳み構造ネットワークにおいては、低次数ノード間の長距離に渡る反発傾向が強く、(a) Twitter ネットワークと (c) WWW においては高次数ノード間の長距離反発傾向が確認できる。さらに、その大きさに関しても、(c) WWW の方が (a) Twitter ネットワークにくらべ強いハブ間反発傾向を有することがわかる。図 5.7 の例からもわかる通り、現実世界の

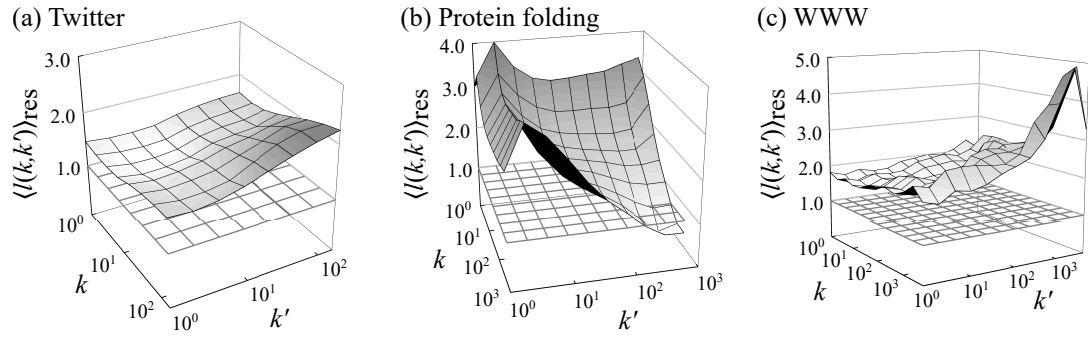


図 5.7: (a)Twitter ネットワーク, (b) タンパク質折り畳み構造ネットワーク, (c)WWW (ノートルダム大学のドメイン内) に対する (5.37) 式で定義された相対平均ノード間距離 $\langle l(k, k') \rangle_{\text{res}}$. 2つの水平軸は対数スケールにおけるノード対の次数 k, k' である. $\langle l(k, k') \rangle_{\text{res}} = 1$ における平面メッシュは, ネットワークが無相関であったときの相対平均ノード間距離を表す.

複雑ネットワークはそれぞれが多様な長距離次数相関を有している. その評価は, 長距離次数相関を記述する確率関数の変数を縮約した平均 l 次隣接次数 $k_l(k)$ や, 長距離 Pearson 相関係数 r_l , 相対ノード間距離 $\langle l(k, k') \rangle_{\text{res}}$ などの指標を調べることにより可能となる.

第6章 内因性長距離次数相関

隣接次数相関を有するネットワークでは、たとえ長距離に渡るノードどうしの反発傾向などが無くとも、隣接の効果が積み重なることによって次隣接以上に離れたノード間にみかけの長距離次数相関が現れることがある。複雑ネットワークの構造やダイナミクスと長距離次数相関との関係について議論を行うためには、隣接次数相関によって引き起こされるこのような外因性の長距離次数相関と、それ以外の内因性の長距離次数相関とを区別する必要がある。本章では、隣接次数相関以外の相関をもたないネットワーク（nearest neighbor correlated random network, NNCRN）が純粋な外因性長距離次数相関のみを有することに着目し、任意の隣接次数相関に起因する外因性長距離次数相関を表す確率関数を求め、それを用いて与えられたネットワークが有する内因性長距離次数相関と外因性長距離次数相関を識別する手法を開発する。また、内因性長距離次数相関の強さを表す指標を導入し、これらを利用することで、現実世界の複雑ネットワークが示す長距離次数相関を評価する。

6.1 外因性長距離次数相関

隣接次数相関は次隣接以上の長距離次数相関に影響を及ぼす。図 6.1 は負の隣接次数相関を有するネットワークの内部構造を表す概念図である。この図から明らかなように、異なる次数を有するノードどうしが隣接しやすい傾向が有るとき偶数距離だけ離れあったノード間の次数相関は正となり、奇数距離だけ離れあったノード間の次数相関は負となる長距離次数相関が生じる。このような隣接次数相関由来の長距離次数相関（外因性長距離次数相関）は、図 6.1 の例のような負の隣接次数相関由来のものだけではない。正負を問わず極端に強い隣接次数相関は、ネットワークを長距離構造（大域的構造）であるコミュニティ構造へと

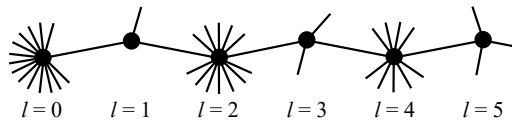


図 6.1: 強い負の隣接次数相関のあるネットワークにおいては、同程度の次数のノードが偶数距離、異なる次数のノードが奇数距離だけ離れあう傾向がある。

に至らしめる [23]. 特に、強い正の隣接次数相関を有するネットワークに現れるコミュニティ構造はオニオンライク構造と呼ばれ [23, 48], このとき高次数ノードどうしが結合しあったコミュニティと低次数ノードどうしが結合しあったコミュニティは互いに長距離に渡って離れあう. これもまた外因性長距離次数相関の一例である. このような長距離次数相関は、フラクタル・ネットワークなどに特有なハブ間反発のように隣接次数相関からは誘起され得ない長距離次数相関（内因性長距離次数相関）とは本質的に異なるものである. 長距離次数相関がネットワーク上のパーコレーションや同期現象、拡散現象等へ及ぼす影響を明らかにするためには、与えられたネットワークの長距離次数相関が外因性のものか内因性のものかを判別できなければならない. しかしながら 5 章で述べた長距離無相関ネットワークとの比較からは、長距離次数相関の有無は判別できても、外因性や内因性を判別することはできない. これを判別するため、純粋な外因性長距離次数相関のみを有するネットワークである隣接相関ランダム・ネットワーク (nearest neighbor correlated random network, NNCRN) を導入する. NNCRN は与えられた隣接次数相関を有し、かつそれ以外においては完全にランダムにノードどうしが結合するネットワークである.

NNCRN はランダム・ネットワークの拡張であり、ランダム・ネットワークと同様にランダムなエッジのつなぎ替え [104] によって数値的に生成される. ランダム・ネットワークの生成では各ノードの次数が保たれるようにランダムなエッジのつなぎ替えが行われるが、NNCRN の生成では与えられた隣接次数相関が保存されるようにランダムなエッジのつなぎ替えを行う. 例えば、以下のようなやり方により Erdős-Rényi ランダムグラフから NNCRN を作成することができる. まず Erdős-Rényi ランダムグラフのエッジを 2 本ランダムに選び、隣接ノードの次数間の Spearman 相関係数 ρ が上昇（または下降）するときのみ両エッジの端点を入れ替えてつなぎ替えを行う. この操作を繰り返し、Spearman 相関係数 ρ が目的の値に到達したときにエッジのつなぎ替えを止める. こうしてできたネットワーク $\mathcal{G}$ は所定の大きさの隣接次数相関を有することになるが、長距離次数相関を制御するような操作を行っていないので、 $\mathcal{G}$ には外因的な相関だけでなく内因性長距離次数相関が含まれている可能性が有る. これを排するため、ネットワーク $\mathcal{G}$ から同一次数のノードをランダムに 2 つ選び、それらから出るエッジをそれぞれ 1 本ずつランダムに選ぶ. これら 2 本のエッジの端点を入れ替える操作を十分な回数繰り返し替えることで、 $\mathcal{G}$ の隣接次数相関 $P_{nn}(k, k')$ と次数分布を変えることなくランダム化することができる. それゆえ、こうしてできたネットワークは Erdős-Rényi ランダムグラフと同じ次数分布関数を有し、 $\mathcal{G}$ と同じ隣接次数相関をもつ NNCRN となる. 図 6.2 は上記の手順によって作られた NNCRN の長距離次数相関を記述する確率関数 $P(k, k'|l)$ を表している. ネットワーク・サイズは $N = 1,000$, 平均次数は $\langle k \rangle = 5.0$ とした. 確率関数 $P(k, k'|l)$ は 100 サンプルにわたり平均化されている. 上段のパネルは正の隣接次数相関 ($\rho = 0.8$) を有する NNCRN に対する結果である. 左から順

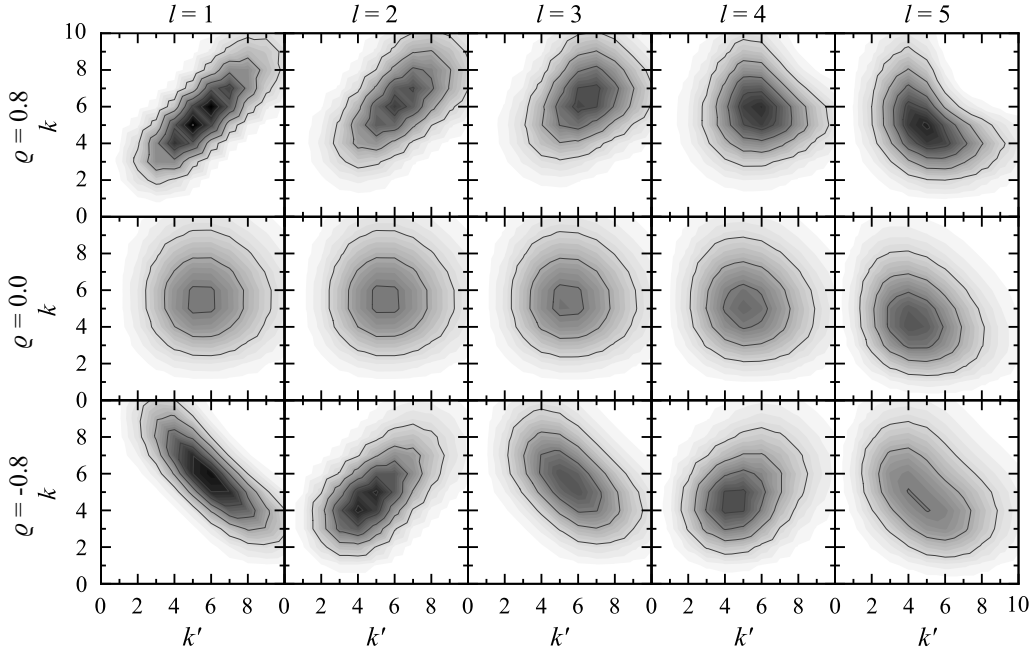


図 6.2: 距離 l だけ離れたノード対の次数が k と k' である確率 $P(k, k'|l)$ の数値計算結果. 等高線図の横軸が k' , 縦軸が k である. 左から順に $l = 1, 2, \dots, 5$ に対する $P(k, k'|l)$ を示している. 上段は正の隣接次数相関を有する NNCRN ($\rho = 0.8$, $r = 0.76$), 中段は長距離無相関ネットワーク ($\rho = 0.0$, $r = 0.0$), 下段は負の隣接次数相関を有する NNCRN ($\rho = -0.8$, $r = -0.75$) に対する結果である. ここで, ρ と r は隣接するノードの次数を変数とした Spearman 相関係数と Pearson 相関係数である. いずれのネットワークもサイズ $N = 1,000$, 平均次数 $\langle k \rangle = 5.0$ の Erdős-Rényi ランダムグラフから生成されたものであり, 結果は 100 サンプルにわたり平均化されている.

に距離 $l = 1, 2, \dots, 5$ だけ離れたノード対の次数分布 $P(k, k'|l)$ を等高線図によって表している. 隣接次数相関と長距離次数相関の関係式 (5.12) から自明ではあるが, ノード対間距離が 1 の場合における $P(k, k'|l = 1)$ からは明らかに次数 k と k' の間に強い正の相関が表れていることがわかる. $l = 1$ で見られる正の相関は l が大きくなるにつれ徐々に鈍化するものの, $l = 3$ 程度までは $l = 1$ と同じく対角方向に偏った確率関数 $P(k, k'|l = 1)$ が確認できる. 一方, $l = 5$ においては, $l \leq 3$ における正の相関の反動として次数間に弱い負の相関が見られる. 下段のパネルは負の隣接次数相関 ($\rho = -0.8$) を有する NNCRN に対する結果である. 図 6.1 の例からも予想された通り, 次数相関の正負がノード間距離 l が 1 つ増加するごとに反転している. このような相関符号の交互変化は, 反強磁性体におけるスピン秩序や木構造グラフの 2 色カラーリング問題における彩色秩序にも見られるものである. 図 6.2 の距離 $l \geq 2$ で見られる正または負の次数相関は隣接次数相関によって誘起されたものであり, 外因性長距離次数相関に他ならない.

6.2 平均場近似

与えられたネットワーク $\mathcal{G}$ の確率関数と、 $\mathcal{G}$ に対応したランダム・ネットワーク（長距離無相関ネットワーク）の確率関数を比較することにより、 $\mathcal{G}$ における長距離次数相関の有無を判別することができることは5章で述べた。そのため、同章ではランダム・ネットワークの確率関数を平均場近似により求めた。同様に、 $\mathcal{G}$ の確率関数と $\mathcal{G}$ に対応した NNCRN の確率関数を比較することで、 $\mathcal{G}$ の長距離次数相関が外因的なものか内因的なものかを判別することができる。本節では、5.2.2 節で展開した平均場理論とほとんど同じ議論により、与えられた隣接次数相関を有する NNCRN の確率関数を計算することができることを示す。以降では、NNCRN の長距離次数相関を表す確率関数には、添え字 “1” をつけることにする。

5 種の確率関数の中でまず $P_1(l|k, k')$ に着目するのは、ランダム・ネットワークに対する計算と同じである。また、距離分布 $P_1(l|k, k')$ が累積分布関数 $\rho(l|k, k')$ によって (5.19) 式のように表されることも、隣接次数相関の有無とは無関係であるから、ランダム・ネットワークと同様に $P_1(l|k, k')$ は

$$P_1(l|k, k') = \begin{cases} \rho(0|k, k') & (l = 0), \\ \rho(l|k, k') - \rho(l-1|k, k') & (l \geq 1), \\ 1 - \lim_{l \rightarrow \infty} \rho(l|k, k') & (l = l_\infty), \end{cases} \quad (6.1)$$

と表される。さらに、木構造仮定の下で (5.22)-(5.25) 式は隣接次数相関の有無とは無関係に成り立つ。ただし、隣接次数相関があるとき $P_{nn}(k''|k)$ を $k''P(k'')/\langle k \rangle$ と書くことはできないため、(5.26) 式は成り立たない。それゆえ、ランダム・ネットワークにおいて $\bar{q}(l|k, k')$ が k には依存しない事実を用いて (5.24) 式から導出した漸化式 (5.27) 式は NNCRN に対して成り立たない。また、同じ理由で (5.25) 式から導出した (5.27) 式も成り立たない。 $P_1(l|k, k')$ を計算するためには、隣接次数相関の有無とは無関係に成り立つ初期条件 (5.25) である

$$q(0|k, k') = \rho(0|k, k') = \frac{\delta_{kk'}}{NP(k)} \quad (6.2)$$

から、(5.23) 式と (5.24) 式の連立漸化式

$$\bar{q}(l|k, k') = \sum_{k''} P_{nn}(k''|k) q(l|k'', k'), \quad (6.3)$$

$$q(l|k, k') = 1 - [1 - \rho(0|k, k')] [1 - \bar{q}(l-1|k, k')]^{k-1}, \quad (6.4)$$

を逐次的に解く必要がある。

上記の近似理論が高い近似精度を有することを確認するため、図 6.2 で扱ったものと同じ NNCRN に対する確率関数を理論計算と数値計算で比較した。図 6.3 は、 $l = 5$ にお

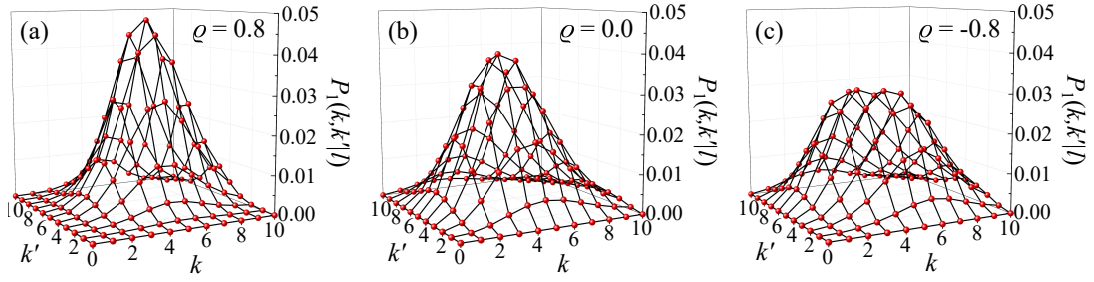


図 6.3: 次数 k と k' を変数としたときの確率関数 $P_1(k, k'|l=5)$. 赤のシンボルは数値計算結果, 黒のワイヤフレームは解析計算結果を表す. ネットワークは図 6.4 と同じ NNCRN であり, それぞれの Spearman 相関係数は (a) $\rho = 0.8$, (b) $\rho = 0.0$, (c) $\rho = -0.8$ である.

ける $P_1(k, k'|l)$ に対する結果であり, それぞれのパネルは, Spearman 相関係数 ρ の値が (a) $\rho = 0.8$, (b) $\rho = 0.0$, (c) $\rho = -0.8$ であるときの結果を示している. 図 6.3 の赤のシンボルは $P(k, k'|l=5)$ の数値計算結果, 黒のワイヤフレームは解析計算結果を表しており, 両者はよく一致している. 長距離次数相関を記述する確率関数は 5 種類あるが, それぞれの関数は (5.1) 式の数学的關係によって互いに厳密に結び付けられているため, $P(k, k'|l)$ の計算精度は他の 4 つの確率関数の計算精度と等しく, 平均場近似により非常に高精度で計算することができる. ただし, $P_0(k, k', l)$ に対する平均場近似と同様, NNCRN がパーコレーション臨界点の近くにある場合は近似の精度が低下する. しかしながら, 現実世界の複雑ネットワークを解析する際, 同じ隣接次数相関をもつ NNCRN がパーコレーション臨界点近傍にあることは極めて稀であり, 実用上この制約が問題になることはない.

6.3 現実のネットワークの長距離次数相関と新たな評価指標

5 章ではネットワークが有する長距離次数相関の記述するための確率関数を導入し, 与えられたネットワーク $\mathcal{G}$ と同じ次数列を有する長距離無相関ネットワークとの比較によって $\mathcal{G}$ における長距離次数相関の有無を解析するためのフレームワークを構築した. 6.2 節ではさらに, $\mathcal{G}$ の長距離次数相関が隣接次数相関に由来する外因性長距離次数相関であるか, そうではなく $\mathcal{G}$ の大域的構造に固有の内因性長距離次数相関であるかを判断するための手法を提案した. 本節では, これらの方法を総合的に用いて, 現実世界の複雑ネットワークの長距離次数相関について, 外因性および内因性の議論も含めた詳細な解析を行う.

図 6.4 は, 図 5.6 で扱ったものと同じ現実ネットワークの平均 l 次隣接次数 $k_l(k)$ である. ただし図 5.6 の結果に加えて, それぞれの実ネットワークと同じ隣接次数相関を有する NNCRN の $k_l(k)$ が赤の実線で示されている. NNCRN に対する $k_l(k)$ は, 6.2 節で述べた平均場近似から計算された $P_1(k'|k, l)$ を使って (5.16) 式から求められたものである. 図 6.4(a) の Gnutella

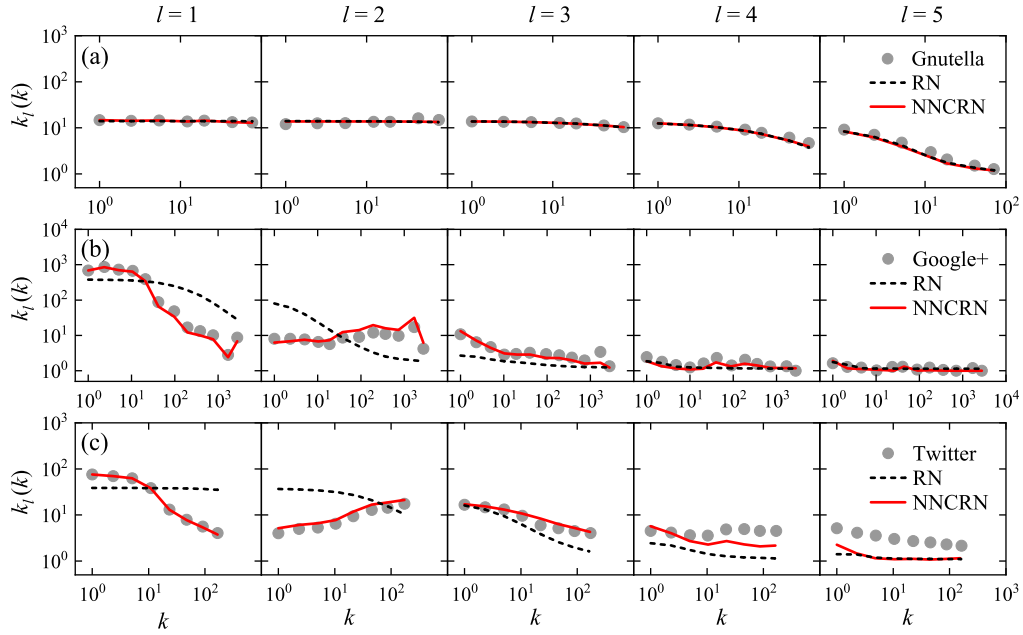


図 6.4: 左から $l = 1, 2, \dots, 5$ における現実世界の複雑ネットワークの平均 l 次隣接次数 $k_l(k)$. 上から (a)Gnutella ネットワーク, (b)Google+ ネットワーク, (c)Twitter ネットワークに対する結果である. 灰色のシンボルは実ネットワークから数値的に計算された $k_l(k)$, 黒の破線はそれぞれの実ネットワークと同じ度数分布を有するランダム・ネットワークの $k_l(k)$ を平均場近似を用いて計算した結果, 赤の実線はそれぞれの実ネットワークと同じ隣接次数相関を有する NNCRN の $k_l(k)$ を平均場近似を用いて計算した結果を表す.

ネットワークは 5.3 節で示されたように長距離無相関なネットワークであるから, 内因性長距離次数相関も外因性長距離次数相関も持たない. これは, 図 6.4(a) において実ネットワーク, ランダム・ネットワーク, NNCRN の全ての結果が重なっている事実とも符合する. 図 6.4(b) の Google+ ネットワークに対する結果は, ランダム・ネットワークの $k_l(k)$ である黒の破線とは一致しないものの, $l = 1, 2, \dots, 5$ の全てにおいて NNCRN に対する $k_l(k)$ とはよく一致する. したがって, Google+ ネットワークには長距離次数相関が見られるものの, それは隣接次数相関によって誘起された外因性長距離次数相関であると結論づけられる. 一方, 図 6.4(c) に示した Twitter ネットワークの $k_l(k)$ は, 対応するランダム・ネットワークのものとも NNCRN のものとも領域 $l \geq 4$ において重ならない. このことは, Twitter ネットワークに見られる長距離次数相関が隣接次数相関には由来しない内因性長距離次数相関であることを示唆する. これらの例より明らかなように, 現実世界の複雑ネットワークは長距離次数相関の観点から 3 つのタイプに分類される. すなわち, 長距離無相関ネットワーク, 外因的な長距離次数相関のみを有するネットワーク, および内因性長距離次数相関を有する

ネットワークの3つである。

さらに数多くの実ネットワークの長距離次数相関を調べるために、新たな指標を導入する。そのため、まず2つの一般的なネットワーク $\mathcal{G}$ と $\mathcal{G}'$ の長距離次数相関の相違度を表す指標として、それぞれの長距離次数相関を記述する同時確率分布 $P(k, k', l)$ と $P'(k, k', l)$ に対して換算距離

$$d(P, P') = \sqrt{\frac{\sum_{k, k'} \sum'_{l=1} [P(k, k', l) - P'(k, k', l)]^2}{\sum_{k, k'} \sum'_{l=1} [P(k, k', l)^2 + P'(k, k', l)^2]}} \quad (6.5)$$

を定義する。ここで、ノード間距離 l に対する和 $\sum'_l$ は $l = l_\infty$ を含まないものとする。この指標の値域は $0 \leq d(P, P') \leq 1$ であり、同時確率分布どうしが完全に一致するときに最小値 $d(P, P') = 0$ をとり、同時確率分布どうしに全く重なりが無いときに最大値 $d(P, P') = 1$ をとる。この確率関数間の距離 $d(P, P')$ を用いることによって、与えられたネットワーク $\mathcal{G}$ が有する長距離次数相関の強さは、 $\mathcal{G}$ に対する同時確率分布 $P(k, k', l)$ と、 $\mathcal{G}$ と同じ次数列を有するランダム・ネットワークの $P_0(k, k', l)$ との間の換算距離 $d(P, P_0)$ によって定量化される。また、 $\mathcal{G}$ の内因性長距離次数相関の強さは、 $\mathcal{G}$ に対する $P(k, k', l)$ と、 $\mathcal{G}$ と同じ隣接次数相関を有する NNCRN の $P_1(k, k', l)$ の間の換算距離 $d(P, P_1)$ によって定量化される。これら2つの換算距離 $d(P, P_0)$, $d(P, P_1)$ を用いることで、ネットワーク $\mathcal{G}$ の長距離次数相関を詳細に分析することができる。 $d(P, P_0) \approx d(P, P_1) \approx 0$ ならば、ネットワーク $\mathcal{G}$ にあらゆる意味での次数相関が無いことになる（長距離無相関）。 $d(P, P_0) \neq 0$ ではあるが $d(P, P_1) \approx 0$ ならば、次数相関は有るものの相関は隣接次数相関由来のものに限られていることになる（外因性長距離次数相関）。また、 $d(P, P_0) \neq 0$ かつ $d(P, P_1) \neq 0$ なら、 $\mathcal{G}$ には隣接次数相関以外の理由で現れる長距離次数相関（内因性長距離次数相関）が存在することになる。特に、 $d(P, P_1) \approx 1$ であれば、 $\mathcal{G}$ には強い内因性長距離次数相関が存在することになる。ランダム・ネットワークと同じ構造で内因性長距離次数相関が存在することは無いので、 $d(P, P_0) \approx 0$ で $d(P, P_1) \neq 0$ の組み合わせになることはない。ただし、 $\mathcal{G}$ に対応するランダム・ネットワークと NNCRN では最大ノード間距離が異なるので $0 < d(P, P_0) < d(P, P_1)$ となることはあり得る。

現実世界の複雑ネットワーク 21 種類に対して $d(P, P_0)$, $d(P, P_1)$ を調べた結果を表 6.1 にまとめる。 $d(P, P_1)$ の値が小さい上から4つのネットワーク（Gnutella ネットワークと3種類の AS レベルのインターネット）を除き、多くの実ネットワークの $d(P, P_1)$ の値は大きく、内因性長距離次数相関が強いことが見て取れる。表 6.1 からわかる通り、 $d(P, P_0)$ の値が大きいネットワークは $d(P, P_1)$ の値も大きい。このことは、実ネットワークに見られる強い長距離次数相関の起源として内因性の要因が支配的であることを意味する。この表で注目すべきもう一つの点は、3種類の AS レベルのインターネットに対する $d(P, P_0)$ の値がほとんど

0 であるにもかかわらず，隣接次数相関指標である Spearman 相関係数が絶対値の大きな負の値をとっていることである． $d(P, P_0)$ が 0 に近いランダムなネットワークであっても，次数の分布に強い不均一性が有るとき「構造的負相関」と呼ばれる負の隣接次数相関が表れることが知られている [104]．例えば，たとえエッジのつながり方が完全にランダムであっても，次数分布が $P(k) = [\delta_{k,N-1} + (N-1)\delta_{k,1}]/N$ のような星形グラフであれば，どのようなネットワーク・サイズ N であっても Spearman 相関係数は $\varrho = -1$ となる．実際，表 6.1 の “Internet (AS level) (1)” ($N = 10,515$) の最大次数は $k_{\max} = 2,277$ であり，そのノー

表 6.1: 現実世界の複雑ネットワークの基本的な特徴量（ノード数 N ，平均次数 $\langle k \rangle$ ，Spearman 相関係数 ϱ ）と長距離次数相関の強さおよび内因性長距離次数相関の強さの指標である $d_0 = d(P, P_0)$ および $d_1 = d(P, P_1)$ の値．さらに多くのサンプルに対する結果は付録 D を参照．

Network	N	$\langle k \rangle$	ϱ	d_0	d_1	Ref.
Gnutella	10,876	7.4	0.00	0.11	0.13	[62]
Internet (AS level) (1)	10,515	4.1	-0.55	0.17	0.25	[86]
Internet (AS level) (2)	26,475	4.0	-0.53	0.21	0.25	[85]
Internet (AS level) (3)	22,963	4.2	-0.51	0.19	0.26	[108]
Google+	23,628	3.3	-0.74	0.63	0.31	[85]
Email	36,692	10.0	-0.01	0.37	0.40	[62]
WWW (1)	415,624	11.4	-0.32	0.87	0.41	[85]
Brightkite	58,228	7.4	0.22	0.54	0.41	[85]
Facebook	63,731	25.6	0.38	0.73	0.43	[85]
Coauthor (1)	18,771	21.1	0.35	0.74	0.48	[85]
YouTube	1,134,890	5.3	-0.08	0.63	0.50	[85]
Coauthor (2)	12,006	19.7	0.72	0.89	0.52	[62]
Actor	82,583	88.8	0.34	0.73	0.57	[86]
Coauthor (3)	23,133	8.1	0.26	0.71	0.61	[62]
Twitter	23,370	2.8	-0.74	0.82	0.67	[85]
WWW (2)	325,729	6.7	-0.11	0.90	0.81	[85]
Internet (router level)	192,244	6.3	0.27	0.89	0.82	[109]
WWW (3)	255,265	15.2	-0.20	0.95	0.88	[85]
WWW (4)	685,230	19.4	-0.17	0.96	0.92	[85]
Protein folding	132,167	3.5	0.51	0.97	0.95	[86]
Amazon	334,863	5.5	-0.19	0.96	0.96	[85]

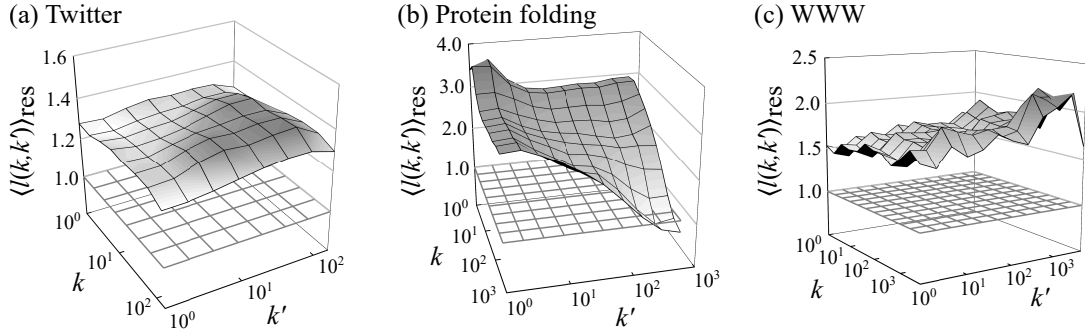


図 6.5: 縦軸は、現実世界の複雑ネットワークに対する、同じ次数分布を有する長距離無相関ネットワークと比較したときの、次数 k と k' を有するノード間距離の相対比 $\langle l(k, k') \rangle_{\text{res}}$ である。2つの水平軸は対数スケールにおけるノード対の次数 k と k' である。対象となるネットワークは左から (a) Twitter ネットワーク, (b) タンパク質折り畳み構造ネットワーク, (c) ノートルダム大学をドメインとする WWW である。 $\langle l(k, k') \rangle_{\text{res}} = 1$ 平面のメッシュは、ネットワークが内因性長距離次数相関をもたないときの相対ノード間距離を表す。

ドは全体の 20% ものノードと接続している。そのようなネットワークの構造は星形に近い。このように、Pearson 相関係数や Spearman 相関係数による次数相関の評価では、次数の不均一性に由来する構造的な見かけの相関と、真の意味での次数相関を区別することができないが、これらの相関係数と換算距離 $d(P, P_0)$ および $d(P, P_1)$ を組み合わせて評価することによって両者を区別することが可能となる。

5.2 節では、与えられたネットワーク $\mathcal{G}$ の次数 k と k' のノード対の平均ノード間距離 $\langle l(k, k') \rangle$ と、 $\mathcal{G}$ に対応するランダム・ネットワークの $\langle l(k, k') \rangle$ を比較することで、 $\mathcal{G}$ がどのような長距離次数相関を有しているのかを評価できることを述べた。また、このような評価方法を現実世界の複雑ネットワークに適用することで、図 5.7 に示したように、実ネットワークの長距離次数相関が平均的にノード間距離を増加させること、Twitter ネットワークと WWW では高次数ノードどうし、タンパク質折り畳み構造ネットワークでは低次数ノードどうしのノード間距離がより大きくなる傾向が強いことが明らかになった。相対平均ノード間距離 $\langle l(k, k') \rangle_{\text{res}}$ の定義を少し変更したうえで、この評価方法の考え方を内因性長距離次数相関をもつネットワークに適用すると、そのネットワーク内にどのような種類の内因性長距離次数がはたらいっているかを知ることができる。(5.37) 式で定義される相対平均ノード間距離 $\langle l(k, k') \rangle_{\text{res}}$ は、ランダム・ネットワークの平均ノード間距離 $\sum_l \tilde{P}(l|k, k')$ でリスケールされているが、ここでは対応する NNCRN の平均ノード間距離でリスケールすることとする。すなわち、

$$\langle l(k, k') \rangle_{\text{res}} = \frac{\sum_l l \tilde{P}(l|k, k')}{\sum_l l \tilde{P}_1(l|k, k')}, \quad (6.6)$$

と定義する．図 6.5 は図 5.7 と同じ実ネットワークに対して (6.6) 式の相対平均ノード間距離をプロットしたものである．図 5.7 の結果と同様に，どのネットワークに対する $\langle l(k, k') \rangle_{\text{res}}$ も平均的に 1 より大きくなる傾向が有る．このことから，内因性長距離次数相関が次数に依らずノード間距離を増加させることがわかる．図 5.7(a) は Twitter ネットワークの高次数ノード間に反発傾向が有ることを示していたが，図 6.5(a) における $\langle l(k, k') \rangle_{\text{res}}$ の増加の程度はノード対の次数 k, k' にそれほど依存していない．このことより，Twitter ネットワークにおける内因性長距離次数相関は全てのノード対間の距離を一定量引き伸ばしているだけで，高次数ノードどうしの反発傾向は隣接次数相関に由来していたことがわかる．また，図 6.5(b) に示したタンパク質折り畳み構造ネットワークに対する結果は，図 5.7(b) にも見られた低次数ノード間の強い反発傾向が内因的な相関に由来するものであることを物語っている．同様に，図 5.7(c) に示されている WWW における高次数ノード間の反発傾向も，図 6.5(c) から理解できるように，内因性長距離次数相関に起因したものである．このように，与えられたネットワーク G の長距離次数相関を， G と同じ次数列を有するランダム・ネットワーク，および G と同じ次数列および隣接次数相関を有する NNCRN の長距離次数相関と比較することによって， G の長距離次数相関に関する詳細な知見を得ることができる．

第7章 結論

本論文の目的は、複雑ネットワークにおける長距離次数相関の一般的定式化を行うことであった。特に、あらゆるタイプの長距離次数相関を記述することのできる数学的枠組みと汎用性の高い解析手法を提案することを目指した。そこで本論文では、ノード対間の距離 l 、ノード対の次数 k, k' を変数とした同時確率 $P(k, k', l)$ 、および4つの条件付き確率 $P(l|k, k')$ 、 $P(k, l|k')$ 、 $P(k|k', l)$ 、 $P(k, k'|l)$ によってネットワークの長距離次数相関を完全に記述される事実に着目し、これらの確率関数を用いて極めて一般的な長距離次数相関の定式化を行った。また、隣接次数相関の効果によって引き起こされる外因性の長距離次数相関と、隣接次数相関によっては生じ得ない内因性の長距離次数相関の判別方法を提案した。これにより、ネットワークの大域的構造やネットワーク上の諸現象と長距離次数相関との関係を明らかにするための系統的な研究が可能になった。

まず、第5章では、ネットワークの長距離次数相関を完全に記述する同時確率および4つの条件付き確率を導入するとともに、これらの5種の確率関数間に成り立つ一般的関係を Bayes の定理を用いて明らかにした。次に、長距離次数相関の有無を判断するための基準となる、いかなるノード間距離 l においても次数相関を持たないネットワーク（長距離無相関ネットワーク）に対する5種の確率関数の関数形を、ネットワーク・サイズが無限と有限のそれぞれの場合に対して明らかにした。有限サイズの長距離無相関ネットワークに対する確率関数の計算では、Melnik ら [107] の提案するノード間距離分布の計算手法を利用して、条件付き確率 $P_0(l|k, k')$ を解析的に計算する方法を提示した。残りの4つの確率関数は確率関数間の一般的関係から計算することができる。解析計算の結果は、Erdős-Rényi ランダムグラフとスケールフリー性をもつ一般化ランダムグラフのそれぞれに対して、数値的に求められる結果に非常によく一致することが確かめられた。これにより、与えられたネットワーク G に対する確率関数と、 G と同じノード数および同じ次数分布を有する長距離無相関ネットワークに対する確率関数を比較することで、 G が長距離次数相関を有するか否かを判断することができるようになった。さらに、本論文で提案した解析手法の有用性を確認するため、現実世界の複雑ネットワークの長距離次数相関の評価を行った。しかし、5種の確率関数は長距離次数相関に関する完全な情報を有している反面、3変数関数であるため扱いが容易ではない。そこで、次数 k のノードから長さ l の最短経路をもつノードの平均次数 $k_l(k)$ や、次数 k と k' を持つノード対間の平均ノード間距離 $\langle l(k, k') \rangle$ を導入し、確率関数の差異を明瞭

化した．導入した指標を基に現実世界の複雑ネットワークの長距離次数相関の評価を行った結果，タンパク質折り畳み構造ネットワークには低次数ノードどうしが，WWW と Twitter ネットワークには高次数ノードどうしが長距離に渡り互いに離れあうような長距離次数相関が見られることが明らかになった．

第 6 章では，第 5 章で定式化した長距離次数相関の数学的枠組みを用いて，次隣接以上に離れたノード間に隣接次数相関の効果が積み重なることによって生じる見かけ上の長距離次数相関（外因性長距離次数相関）と，それ以外の長距離次数相関（内因性長距離次数相関）を判断するための手法を提案した．はじめに，Erdős-Rényi ランダムグラフに対して隣接次数相関を強めるようなエッジのつながり替えを行うことで，実際にネットワークに外因性の長距離次数相関が生じることを数値的に実証した．次に，外因性長距離次数相関のみを純粋に有するネットワークとして，隣接次数相関以外には完全にランダムなネットワーク（NNCRN，隣接相関ランダム・ネットワーク）を導入し，長距離無相関ネットワークと同様の平均場近似と局所木構造仮定を用いて，有限サイズの NNCRN に対する 5 種の確率関数を導出した．これにより，与えられたネットワーク G に対する確率関数と， G と同じノード数，同じ次数分布，同じ隣接次数相関を有する NNCRN に対する確率関数とを比較することで， G の示す長距離次数相関が外因性の相関であるか内因性の相関であるかを区別することができるようになった．この方法を用いて，第 5 章において長距離次数相関の有無を調べた現実ネットワークと，それらのネットワークに対応する NNCRN の間で平均 l 次隣接次数 $k_l(k)$ を比較し，実際に長距離次数相関の内因性と外因性が判別できることを示した．さらに，数多くのネットワークの長距離次数相関を定量評価するために，長距離次数相関の強さを表す指標 d_0 と内因性長距離次数相関の強さを表す指標 d_1 を確率関数間の換算距離 $d(P, P')$ を用いて定義した．これらの指標を用い，社会学，経済学分野から工学，生物学分野に至る 21 種類のネットワークを調べた結果，一部のネットワークを除くほぼ全てのネットワークが長距離次数相関を有しており，そのうち大多数が内因性の長距離次数相関を有していることが明らかになった．最後に，現実世界のネットワークに現れる長距離次数相関に対して，内因性と外因性の判別も含めたより詳細な解析を行い，Twitter ネットワークに見られた高次数ノード間の長距離反発が外因性長距離次数相関であること，その一方で，タンパク質折り畳み構造ネットワークに見られる低次数ノードどうしの長距離反発，および WWW に見られる高次数ノード間の反発傾向が内因性の長距離次数相関であることを明らかにした．

本論文では長距離次数相関の解析方法の提供までに留まったが，今後の展望としては，頑強性，フラクタル性，同期現象，感染症拡散ダイナミクスをはじめ，複雑ネットワークの普遍的性質やネットワーク上のダイナミクスと長距離次数相関との関係を明らかにしていきたいと考えている．次数相関がネットワークの性質を決定する上で重要な要素となることは既によく知られており，例えば，頑強性 [47] や同期現象 [33] が隣接次数相関から大きい影響を

受ける。また、フラクタル・スケールフリー・ネットワーク（FSN）は強い負の隣接次数相関を有することが経験則として知られている。FSN の負の隣接次数相関は、ただの隣接の相関ではなく、高次数ノードどうしの長距離に渡る反発傾向と関連が深い。また、フラクタル性があるとき、ネットワーク上のランダム・ウォークでは異常拡散が見られ、パーコレーションでは臨界相が出現する。したがって、ネットワークのフラクタル性と長距離次数相関の関係解明は、現実世界のスケールフリー・ネットワークに表れるフラクタル性の起源を明らかにするためにも、長距離次数相関がネットワーク上のパーコレーションやネットワーク上のランダム・ウォークへ与える影響を明らかにするためにも重要である。また、今回導入した5種の確率関数は3変数関数であるため、直接的な比較から有益な情報を得ることは難しい。実用性の観点からは、ネットワークの長距離次数相関を定量化する新たな指標の開発が要請される。本論文ではノード間距離分布 $\langle l(k, k') \rangle$ 、平均 l 次隣接次数 $k_l(k)$ 、確率関数間の換算距離 d_0 および d_1 を用いたが、確率関数からは他にも有用な指標が定義できるだろう。適切な指標を選ぶことにより、様々なネットワーク特性と長距離次数相関の関係を明らかにすることができると思われる。

謝辞

本研究を遂行するにあたり，不肖の弟子であった筆者に対し細部に至る懇切丁寧なご指導をしてくださった矢久保考介教授に心より感謝いたします。また，共同研究者である水高将吾氏，高口太朗氏に深く御礼申し上げます。そして，矢久保考介教授，浅野泰寛准教授，小布施秀明助教，事務補佐員の富樫杏子様をはじめ，数理物理工学研究室の皆様には公私ともに大変お世話になりました。充実した大学生活を送ることができたのは皆様のおかげです。特に，羽部哲朗氏，水高将吾氏，渡邊章友氏，鈴木修氏，池谷聡氏をはじめ，博士課程の先輩方には研究生生活を送るうえで重要なご助言を数多く頂きました。ありがとうございました。今後も一層の精進を重ねて参ります。最後に，9年間にも及ぶ大学生活を支えてくれた家族に心より感謝いたします。

付 録 A スケールフリー・ネットワークの次数モーメント

ネットワークの次数分布が $P(k) = Ak^{-\gamma}$ ($\gamma > 1$) に従うとき、連続体近似を用いて次数の p 次モーメントを、

$$\langle k^p \rangle \simeq \int_{k_{\min}}^{k_{\max}} k^p AP(k) dk, \quad (\text{A.1})$$

として近似することができる。規格化定数 A は連続体近似の下での規格化条件 $\int_{k_{\min}}^{k_{\max}} P(k) dk = 1$ から、

$$A = \frac{-\gamma + 1}{k_{\max}^{-\gamma+1} - k_{\min}^{-\gamma+1}}, \quad (\text{A.2})$$

と求められるので、上の積分を実行すると

$$\langle k^p \rangle \simeq \begin{cases} \frac{-\gamma + 1}{k_{\max}^{-\gamma+1} - k_{\min}^{-\gamma+1}} \frac{k_{\max}^{p-\gamma+1} - k_{\min}^{p-\gamma+1}}{p - \gamma + 1}, & (p + 1 \neq \gamma) \\ \frac{-\gamma + 1}{k_{\max}^{-\gamma+1} - k_{\min}^{-\gamma+1}} (\log k_{\max} - \log k_{\min}), & (p + 1 = \gamma) \end{cases} \quad (\text{A.3})$$

となる。ここで、最大次数 $k_{\max}$ が N に依存することに注意が必要である。素朴に (A.3) 式から $k_{\max}$ の無限大極限をとると

$$\lim_{k_{\max} \rightarrow \infty} \langle k^p \rangle \simeq \begin{cases} \infty, & (p + 1 \geq \gamma) \\ \frac{-\gamma + 1}{p - \gamma + 1} k_{\min}^{p-1}, & (p + 1 < \gamma) \end{cases} \quad (\text{A.4})$$

となるが、ネットワーク・サイズ N が有限である限りは $k_{\max}$ も有限の値となる。有限サイズのネットワークにおいて最大次数 $k_{\max}$ とは、それより大きな次数をもつノードは存在しないか無視できるほど少ない次数であり、

$$\int_{k_{\max}}^{\infty} P(k) dk \simeq \frac{1}{N}, \quad (\text{A.5})$$

の関係から、

$$k_{\max} \simeq k_{\min} (N + 1)^{-1/(-\gamma+1)}, \quad (\text{A.6})$$

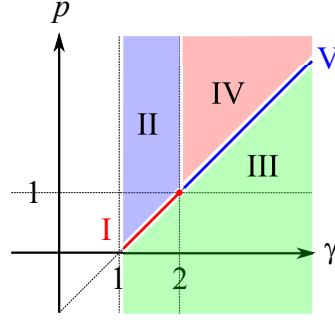


図 A.1: (A.8) 式から得られる相図.

と求められる．また次数が $N - 1$ を超えることはありえないため， N が十分大きいとき

$$k_{\max} \simeq \begin{cases} N & (\gamma \leq 2), \\ k_{\min} N^{-1/(-\gamma+1)} & (\gamma > 2), \end{cases} \quad (\text{A.7})$$

である．よって N が十分に大きいとき p 次モーメントは

$$\langle k^p \rangle \simeq \begin{cases} k_{\min}^{\gamma-1} (\gamma - 1) [\log N - \log k_{\min}], & (\text{I} : 1 < p + 1 = \gamma \leq 2) \\ k_{\min}^{\gamma-1} \frac{-\gamma+1}{p-\gamma+1} N^{p-\gamma+1}, & (\text{II} : \gamma < p + 1, 1 < \gamma \leq 2) \\ k_{\min}^p \frac{\gamma-1}{p-\gamma+1}, & (\text{III} : 1 < \gamma, p + 1 < \gamma) \\ k_{\min}^p \frac{\gamma-1}{p-\gamma+1} N^{(p-\gamma+1)/(\gamma-1)}, & (\text{IV} : 2 < \gamma < p + 1) \\ k_{\min}^{\gamma-1} \log N, & (\text{V} : 2 < \gamma = p + 1) \end{cases} \quad (\text{A.8})$$

となる．すなわち， $\langle k^p \rangle$ のネットワーク・サイズ依存性は γ の値によって異なり，特に， γ が $p + 1$ 以下であるとき無限サイズ・ネットワークの $\langle k^p \rangle$ は発散する．場合分けの結果を図 A.1 にまとめた．

付 録 B 決定論的フラクタル・スケールフリー・ネットワークの特徴量

本章では，決定論的フラクタル・スケールフリー・ネットワーク（FSN）モデルによって生成されるネットワークに対して，ネットワークの基本的特徴量であるノード数やエッジ数，次数列，スケールフリー指数，フラクタル次元，隣接次数相関，クラスター係数を計算する方法について説明する．さらに，既存モデルでは必ず値がゼロとなっていたクラスター係数が，適切なジェネレータ G_1 を選ぶことで有限の値となり得ることを具体例を用いて示す．

B.1 特徴量の計算方法

決定論的 FSN のトポロジーは一意的であるため，その構造に関する特徴量を解析的に求めることができる．導出は以下の通りである．

ノード数とエッジ数 ルートノードを除いたジェネレータのノード数を ν し，ジェネレータのエッジ数を μ とする．第 $r > 1$ 世代の FSN のエッジ数 M_r とノード数 N_r は，

$$M_r = \mu M_{r-1}, \quad (\text{B.1})$$

$$N_r = N_{r-1} + \nu M_{r-1}. \quad (\text{B.2})$$

の漸化式に従う．これを解くと，第 r 世代目におけるエッジ数 M_r とノード数 N_r はそれぞれ

$$M_r = \mu^r M_0, \quad (\text{B.3})$$

$$N_r = \nu M_0 \frac{1 - \mu^r}{1 - \mu} + N_0, \quad (\text{B.4})$$

となる．

フラクタル次元 エッジ数 M_r とノード数 N_r は世代数 r の増加にともない μ^r のオーダーで増加する．一方でイニシエータ（すなわち第 0 世代）に含まれるノード間の距離は第 0 世代に比べて λ^r 倍されており，ネットワーク直径や平均ノード間距離も同じく λ^r に比例して

増加する。したがってネットワーク・サイズ N とネットワークの特徴的長さ L のスケーリング則 $N \propto L^{d_f}$ から定義されるフラクタル次元は

$$d_f = \frac{\log \mu}{\log \lambda}, \quad (\text{B.5})$$

と求められる。

次数列 逆繰り込み操作によってノードの次数は κ 倍される。そのため、第 r 世代目のネットワーク $\mathcal{G}_r$ における次数 k のノードの数 $N_r(k)$ には、前の世代 $\mathcal{G}_{r-1}$ に存在した次数 k/κ のノードから $N_{r-1}(k/\kappa)$ だけの寄与がある。また、ジェネレータを構成するノードのうちルートノードではないもの（残りノードと呼ぶ）に次数 k のノード含まれる場合、エッジがジェネレータに置き換えられることによっても次数 k のノードは $M_{r-1}\nu p_k$ だけ生成される。ここで p_k は残りノードの次数分布である。両者の和をとり、世代 $r > 0$ における次数 k のノード数は

$$N_r(k) = N_{r-1}(k/\kappa) + M_{r-1}\nu p_k = N_0 q_{k/\kappa^r} + \nu M_0 \mu^r \sum_{n=0}^{r-1} \frac{p_{k/\kappa^n}}{\mu^{n+1}}, \quad (\text{B.6})$$

である。ここで q_k はイニシエータの次数分布である。なお $N_r(x)$, p_x , q_x は x が非整数であるとき 0 とする。

スケールフリー指数 以下ではイニシエータに属するノードはネットワーク全体に比べて無視できるほど少数であり、そのため、全ノードに対する次数分布 $P(k)$ はイニシエータに属さないノードの次数分布によって近似されるものとする。はじめに、イニシエータに属さず、第 1 世代以降に次数 k' で生成されたノードを母集団とした次数分布 $P_{k'}(k)$ を考える。この集団においてノードがとり得る次数は $k'\kappa^n$ (n は整数) である。 r 世代目において次数が $k'\kappa^{r-r'}$ であることは、そのノードが r' 世代目に生成されたノードであることを意味する。 r' 世代目に生成される次数 k のノード数は $N_{r'} - N_{r'-1}$ に比例し、したがって、

$$|P_{k'}(k'\kappa^{r-r'})dk| = \left| \frac{N_{r'} - N_{r'-1}}{N_r - N_0} dr' \right|, \quad (\text{B.7})$$

の関係から、

$$P_{k'}(k) = \frac{N_{r'} - N_{r'-1}}{N_r - N_0} \frac{dr'}{dk} = \frac{1 - \mu^{-1}}{1 - \mu^{-r}} \frac{k'}{\log \kappa} \left[\frac{k}{k'} \right]^{-\gamma}, \quad (\text{B.8})$$

を得る。ここで $\gamma = 1 + \log \mu / \log \kappa$ である。なお $N_{r'} - N_{r'-1} = \nu M_{r-1}$ と $r' = r - (\log k - \log k') / \log \kappa$ の関係を用いた。この結果は任意の k' に対して成り立つ。また、各ノードの生

成時の次数は残りノードの次数分布 p_k に従う．したがって，ネットワーク全体に対する次数分布 $P(k) = \sum_{k'} p_{k'} P_{k'}(k)$ は，

$$P(k) = k^{-\gamma} \frac{1 - \mu^{-1}}{1 - \mu^{-r}} \sum_{k'} p_{k'} \frac{k'^{\gamma+1}}{\log \kappa}, \quad (\text{B.9})$$

となる．これより， $\kappa \geq 2$ であるジェネレータ G_1 によって生成される決定論的 FSN は，スケールフリー指数 γ のスケールフリー・ネットワークであることがわかる．なおスケールフリー指数には $\gamma_{\min} = 2 + \log 2 / \log \kappa$ という下限が与えられる．これは，ルートノード間を直接つなぐエッジが存在しないという条件 $\lambda \geq 2$ によって，パラメータ間に $\mu \geq 2\kappa$ という大小関係が生じるためである．

図 B.1 における丸いシンボルは，決定論的 FSN によって生成することのできるネットワークのスケールフリー指数 γ とフラクタル次元 d_f の組み合わせを示している．それぞれのシンボルの色はジェネレータのエッジ数 μ の大きさを表す．黒の実線はフラクタルかつスケールフリーな臨界木構造ネットワークのフラクタル次元とスケールフリー指数の関係を表す．四角と星形のシンボルは表 2.3 にある現実世界の FSN における d_f と γ をプロットしたものであり，四角のシンボルは次数分布が $P(k) \propto k^{-\gamma}$ で近似できるネットワーク，星形のシンボルは次数分布が $P(k) \approx k^{-\gamma} e^{-k/\kappa}$ で近似できるネットワークに対応する．図から明らかとなおり，決定論的 FSN は極めて広い領域にわたる γ および d_f を有するネットワークを生成することができる．

次数の 1 次モーメントと 2 次モーメント 次数 k の 1 次モーメントは握手の定理 $\langle k \rangle = 2M/N$ から，

$$\langle k \rangle = 2\mu^r \left[\nu \frac{1 - \mu^r}{1 - \mu} + \frac{N_0}{M_0} \right]^{-1}, \quad (\text{B.10})$$

と直ちに求められる．2 次のモーメントの計算にはルートノードを除くジェネレータの次数分布 p_k およびイニシエータの次数分布 q_k が必要であり，

$$\langle k^2 \rangle = \sum_{r'=1}^r \frac{\nu M_0 \mu^{r'-1}}{N_r} \sum_k p_k (k \kappa^{r-r'})^2 + \frac{N_0}{N_r} \sum_k q_k (k \kappa^r)^2, \quad (\text{B.11})$$

より求められる．ここで $s = \kappa^2 / \mu$ と置き，ジェネレータ G_1 によって決まる $\langle k_p^2 \rangle = \sum_k p_k k^2$ とイニシエータ G_0 によって決まる $\langle k_q^2 \rangle = \sum_k q_k k^2$ を用いて式を整理すると世代数 r が十

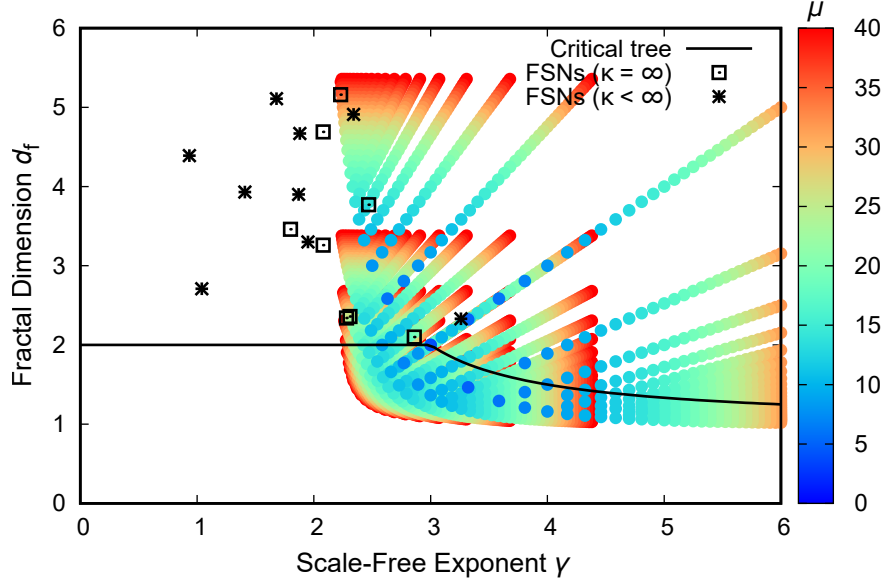


図 B.1: 様々な FSN のフラクタル次元 d_f とスケールフリー指数 γ . 丸いシンボルは決定論的 FSN の取り得る値を表し、それぞれのシンボルの色はジェネレータのエッジ数 μ の大きさを表す. 黒の実線はフラクタルかつスケールフリーな臨界木構造ネットワークにおける d_f と γ の関係を表す. 四角と星形のシンボルは表 2.3 にある現実世界の FSN における d_f と γ をプロットしたものであり、四角のシンボルは次数分布が $P(k) \propto k^{-\gamma}$ で近似できるネットワーク、星形のシンボルは次数分布が $P(k) \approx k^{-\gamma} e^{-k/\kappa}$ で近似できるネットワークに対応する.

分大きいとき

$$\begin{aligned} \langle k^2 \rangle &= \frac{s^r \langle k \rangle}{2} \left[\langle k_p^2 \rangle \frac{\nu}{\kappa^2} \sum_{r'=0}^{r-1} s^{-r'} + \langle k_q^2 \rangle \frac{N_0}{M_0} \right] \\ &\simeq \begin{cases} \frac{\langle k \rangle \langle k_p^2 \rangle s \nu}{2 \kappa^2 (1-s)}, & (s < 1) \\ \frac{\langle k \rangle \langle k_p^2 \rangle \nu}{2 \kappa^2} (r-1), & (s = 1) \\ \frac{s^r \langle k \rangle}{2} \left[\frac{\langle k_p^2 \rangle s \nu}{\kappa^2 (s-1)} + \langle k_q^2 \rangle \frac{N_0}{M_0} \right], & (s > 1) \end{cases} \quad (\text{B.12}) \end{aligned}$$

となる. ここで $\gamma = 1 + \log \mu / \log \kappa$ より, s を $\mu^{(3-\gamma)/(\gamma-1)}$ と書き直すと, 次数の 2 次のモーメント $\langle k \rangle$ が熱力学極限 $n \rightarrow \infty$ で発散する条件 $s > 1$ はスケールフリー・ネットワークの 2 次モーメントが熱力学極限で発散する条件 $\gamma < 3$ と一致する.

隣接次数相関 逆繰り込み操作によって全てのエッジはジェネレータに置き換えられる. したがって, $\mathcal{G}_r$ において次数 k と k' のノードをつなぐエッジの本数 $M_r(k, k')$ には, 新規ノードどうしに張られるエッジから $M_{r-1} \mu(k, k')$ の寄与がある. ここで, $\mu(k, k')$ はジェネレー

タに含まれる次数 k と k' の残りノードどうしをつなぐエッジの本数である。また、前の世代 $\mathcal{G}_{r-1}$ から存在するノードの隣接ノードは逆繰り込み操作によって全て新規ノードになる。 $\mathcal{G}_{r-1}$ で次数が k/κ であったノードは $\mathcal{G}_r$ では k 個のノードと隣接することになるが、それらの隣接ノードのうち次数 k' のノードの数は、ジェネレータに含まれる次数 k' の残りノードとルートノードとをつなぐエッジの本数を μ'_k として、 $\mu'_k k/\kappa$ と書ける。さらに $k \neq k'$ であるとき、前の世代から存在するノードの次数が k であり新規ノード側の次数が k である場合と、その逆の場合が考えられる。これらの事実から、 $\mathcal{G}_r$ において次数 k と k' のノード対をつなぐエッジの数は一つ前の世代のエッジ数と次数列を用いて

$$M_r(k, k') = M_{r-1}\mu(k, k') + \frac{1}{1 + \delta_{k,k'}} \left[\mu_{k'} \frac{k}{\kappa} N_{r-1}(k/\kappa) + \mu_k \frac{k'}{\kappa} N_{r-1}(k'/\kappa) \right], \quad (\text{B.13})$$

と表される。次数列の表式 (B.6) より $\mathcal{G}_{r-1}$ において次数 k/κ を有するノードの数は

$$N_{r-1}(k/\kappa) = N_0 q_{k/\kappa^r} + \nu M_{r-1} \sum_{n=1}^{r-1} \frac{p_{k/\kappa^n}}{\mu^n}, \quad (\text{B.14})$$

であるから (B.13) より

$$\begin{aligned} M_r(k, k') = M_{r-1}\mu(k, k') + \frac{N_0}{1 + \delta_{k,k'}} & \left[\mu_{k'} \frac{k}{\kappa} q_{k/\kappa^r} + \mu_k \frac{k'}{\kappa} q_{k'/\kappa^r} \right] \\ & + \frac{\nu M_{r-1}}{1 + \delta_{k,k'}} \sum_{n=1}^{r-1} \left[\frac{\mu_{k'} k}{\mu^n \kappa} p_{k/\kappa^n} + \frac{\mu_k k'}{\mu^n \kappa} p_{k'/\kappa^n} \right], \end{aligned} \quad (\text{B.15})$$

を得る。なお $\mu(x, x')$, μ_x は x, x' が非整数のとき 0 とする。 $\mathcal{G}_r$ の隣接次数相関は「任意に選んだエッジの片方の次数が k であり、かつもう片方の次数が k' である同時確率 $P_r(k, k')$ 」によって記述されるが、 $P_r(k, k')$ と $M_r(k, k')$ は一般に

$$P_r(k, k') = \frac{1 + \delta_{k,k'}}{2M_r} M_r(k, k'), \quad (\text{B.16})$$

の関係で結び付けられる。上の式と (B.13), (B.15) より決定論的 FSN の隣接次数相関を表す確率関数は

$$\begin{aligned} P_r(k, k') = \frac{1 + \delta_{k,k'}}{2\mu} \mu(k, k') + \frac{N_0}{2M_r} & \left[\mu_{k'} \frac{k}{\kappa} q_{k/\kappa^r} + \mu_k \frac{k'}{\kappa} q_{k'/\kappa^r} \right] \\ & + \frac{\nu}{2\mu} \sum_{n=1}^{r-1} \left[\frac{\mu_{k'} k}{\mu^n \kappa} p_{k/\kappa^n} + \frac{\mu_k k'}{\mu^n \kappa} p_{k'/\kappa^n} \right], \end{aligned} \quad (\text{B.17})$$

と求められる。これを用いて、FSN の隣接次数相関の強さを表す Pearson 相関係数 (2.9) の値は

$$r = \frac{\sum_{k,k'} k k' P_r(k, k') - \langle k^2 \rangle^2 / \langle k \rangle^2}{\langle k^3 \rangle / \langle k \rangle - \langle k^2 \rangle^2 / \langle k \rangle^2}, \quad (\text{B.18})$$

から求められる。

クラスター係数 ジェネレータの残りノードのクラスター係数の平均値を C_g とする．世代 $r-1$ から r への逆線り込み操作によって新たに生成されたノードは，第 r 世代のネットワークにおいては一度もエッジをジェネレータによって置き換えられた経験がないため，クラスター係数の平均値は C_g と等しい．したがって $N_r - N_{r-1}$ 個存在するこれらの新生ノードからの平均クラスター係数 C に対する寄与は $C_g(N_r - N_{r-1})/N_r$ である．また， r' 世代目 ($r' < r$) における次数 k' の新生ノードは r 世代目に次数 $k = k'\kappa^{r-r'}$ へと成長する．このようなノードが属する三角形の個数は，ジェネレータに含まれる三角形のうちルートノードが属するものの数を t_b としたとき $t_b k$ である．したがって全体での平均クラスター係数は

$$C = \frac{N_r - N_{r-1}}{N_r} C_g + \sum_{r'=1}^{r-1} \frac{N_{r'} - N_{r'-1}}{N_r} \sum_k p_k \frac{2t_b \kappa^{-1}}{(k\kappa^{r-r'} - 1)} + \frac{N_0}{N_r} \sum_k q_k \frac{2t_b \kappa^{-1}}{(k\kappa^r - 1)}, \quad (\text{B.19})$$

となる．なお p_k はルートノードを除くジェネレータの次数分布であり， q_k はイニシエータの次数分布である． $\langle k_p^h \rangle = \sum_k p_k k^h$ ， $\langle k_q^h \rangle = \sum_k q_k k^h$ と置き式を整理していくと世代数 r が十分大きいとき

$$C = \frac{\nu \langle k \rangle}{2\mu} C_g + \frac{\langle k \rangle \nu t_b}{\mu \kappa} \sum_{h=0}^{\infty} \langle k_p^{-h} \rangle \frac{1 - (\mu \kappa^h)^{1-r}}{\mu \kappa^h - 1} + \frac{\langle k \rangle N_0 t_b}{M_0 \kappa} \sum_{h=1}^{\infty} \langle k_q^{-h} \rangle (\mu \kappa^h)^{-r}, \quad (\text{B.20})$$

を得る．これより， $\mathcal{G}_r$ の平均クラスター係数 C はジェネレータに三角形がある限り必ず有限となることがわかる．一方で，大域クラスター係数 C' はジェネレータに含まれる三角形の数 τ を用いて

$$C' = \frac{3M_{r-1}\tau}{N_r \sum_k P(k)k(k-1)/2} = \frac{3\tau}{\mu} \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}, \quad (\text{B.21})$$

と表わされる．次数の2次モーメントの計算 (B.12) では $s = \kappa^2/\mu \geq 1$ であるとき熱力学極限で $\langle k^2 \rangle$ は発散することが示されていた．大域クラスター係数は2次モーメントを分母にもつことから，熱力学極限で C' が有限の値をとる条件の一つに $s < 1$ があることがわかる．

ハブ間距離 世代 $r' - 1$ から世代 r' へのエッジの置き換え操作によって生成された次数 k' のノードに注目する．第 r 世代目のネットワーク $\mathcal{G}_r$ ($r > r'$) において，このノードの次数は $k'\kappa^{r-r'}$ である．同世代 r' に生成されたノードの次数は $\kappa^{r-r'}$ 程度であり，より古いノードの次数はその κ 倍程度かそれ以上，より若いノードの次数はその κ^{-1} 倍程度かそれ以下となる．一方で第 r' 世代に存在していた2つのノードの間の距離は第 r 世代においては $\lambda^{r-r'}$ 倍されるため，第 r' 世代より前に生成されたノードどうしの間の距離は $\lambda^{r-r'}$ 以下になり得ない．すなわち次数が $\kappa^{r-r'}$ 程度より大きいノードどうしは必ず距離 $\lambda^{r-r'}$ 以上離れあう．これは決定論的 FSN モデルによって作られるネットワークに特有のハブ間の反発傾向である．ノードが生成されてから経過した世代数 $r - r'$ と次数の関係から，次数 k のノードは生

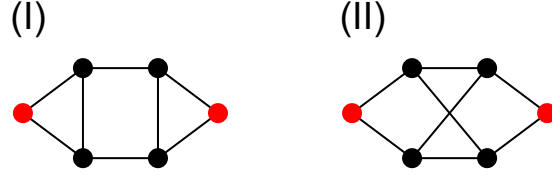


図 B.2: モデル I のジェネレータ (左) とモデル II のジェネレータ (右). 赤のノードはルートノードである.

成後 $r - r' = \log k / \log \kappa$ 世代経過していると見積もられる. これを $\lambda^{r-r'}$ に代入することで「次数 k のノードが同程度以上の次数をもつノードと反発しあう距離」が $\Lambda(k) \propto k^{\log \lambda / \log \kappa}$ と求められる. ここで決定論的 FSN のフラクタル次元は $d_f = \log \mu / \log \lambda$, スケールフリー指数は $\gamma = 1 + \log \mu / \log \kappa$ であるから $\Lambda(k) \propto k^{(\gamma-1)/d_f}$ と表わすことができる.

B.2 具体例

一般化された決定論的 FSN モデルでは, ジェネレータ G_1 を適切に選ぶことにより, (u, v) -flower および SHM モデルと異なり, 有限のクラスター係数を有する FSN を作ることができる. 図 B.2 のモデル I とモデル II をジェネレータとして生成される決定論的 FSN は, 同じ次数列, フラクタル次元, 隣接次数相関を有する一方で, クラスター性に関しては値は全く異なる性質を示す. モデル I はジェネレータに三角形が含まれるため熱力学極限においても有限のクラスター係数を取り得るが, モデル II のジェネレータには三角形が含まれないため, 初期状態 G_0 をのぞいたあらゆる世代 r においてクラスター係数の値は常にゼロとなる. 図 B.3 はモデル I の平均クラスター係数 (B.19) と大域クラスター係数 (B.21) の世代数依存性である.

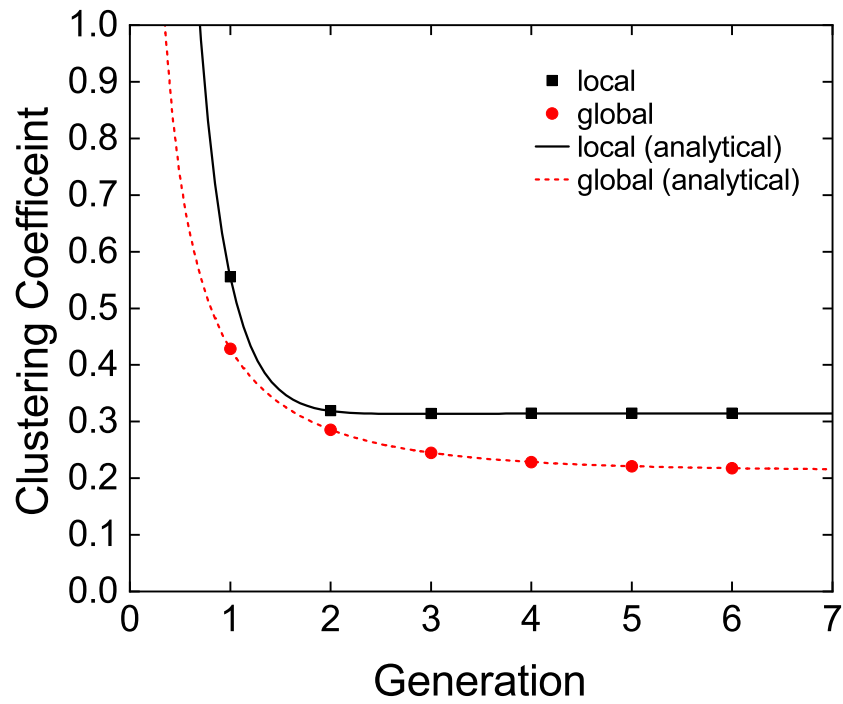


図 B.3: モデル I の平均クラスター係数と大域的クラスター係数。シンボルは数値計算結果，線は解析計算結果をそれぞれ表す。ジェネレータの特徴量は $\mu = 8$, $\nu = 4$, $\kappa = 2$, $p_k = \delta_{k3}$, $C_g = 1/3$, $t_b = 1$, $\tau = 2$ である。イニシエータの特徴量は $N_0 = 2$, $M_0 = 1$, $q_k = \delta_{k1}$ である。

付 録 C 平均場近似の精度

ここでは、第5章および第6章で展開した平均場理論の精度について詳しく述べる。(5.22)式と(5.24)式において用いられている局所木構造仮定はサイクルの多いネットワークでは成立しない。しかしながら、実際にはかなりサイクルの多いネットワークに対しても精度の良い近似となっている。例えば、図C.1(a)に示す通り、 $N = 1,000$ のErdős-Rényi ランダムグラフに対しては、 $\langle k \rangle \lesssim 200$ であれば、5.2.1節の理論と $R(l) = \sum_{k,k'} P(k)P(k')P_0(l|k,k')$ から求められるノード間距離分布 $R(l)$ は、数値計算結果とよく一致している。 $N = 3,000$ のErdős-Rényi ランダムグラフの場合だと、 $\langle k \rangle \lesssim 600$ に対して理論と数値計算の一致が望めることから、非常に大雑把に言って $\langle k \rangle \lesssim N/5$ 程度であれば局所木構造仮定は成立していると考えることができる。局所的サイクルの多寡を評価するには、平均クラスター係数 C が1つの指標となる。Erdős-Rényi ランダムグラフにおける平均クラスター係数は $C \approx \langle k \rangle / N$ となるから、局所木構造仮定が使えるランダムグラフの平均クラスター係数は $C \lesssim 1/5$ と見積もられる。複雑ネットワークの多くはスケールフリー性を示すが、次数分布関数が $P(k) \propto k^{-\gamma}$ で表されるような一般化ランダムグラフにおいて平均クラスター係数が $C \lesssim 1/5$ の範囲に留まるか否かはスケールフリー指数 γ に依存する。現実の複雑ネットワークのスケールフリー指数がおおよそ $2 \lesssim \gamma \lesssim 3$ の範囲に収まるのに対し、平均クラスター係数は $\gamma > 3$ において $C \propto N^{-1}$ 、 $2 < \gamma < 3$ において $C \propto N^{-(3\gamma-7)/(\gamma-1)}$ という N 依存性を見せる。したがって、 C が N に従い緩やかに増加する $\gamma < 7/3$ の範囲のスケールフリー指数をもつネットワークにおいては、近似の精度に注意する必要がある。

平均場近似はノードの所属する連結成分の大きさの揺らぎが十分に小さいことを仮定している。そのため、連結成分サイズの揺らぎが発散する臨界点近傍 $\langle k^2 \rangle / \langle k \rangle \simeq 2$ では近似に大きな誤差が生じる。図C.2は、このことをErdős-Rényi ランダムグラフとスケールフリー性を有する一般化ランダムグラフに対して具体的に示したものである。図C.2(a)のErdős-Rényi ランダムグラフに対する結果を見ると、臨界点である $\langle k \rangle = 1.0$ の近傍で、理論計算と数値計算は大きくずれている。図C.2(b)は、 $1 \leq k \leq 100$ において次数分布関数が $P(k) = ck^{-\gamma}$ (c は規格化定数)と振舞う一般化ランダムグラフに対する結果である。このネットワークは、

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{\sum_{k=1}^{100} k^{2-\gamma}}{\sum_{k=1}^{100} k^{1-\gamma}} = 2, \quad (\text{C.1})$$

を満たす γ 、すなわち $\gamma_c = 3.375$ が転移点であり、 $\gamma < \gamma_c$ において巨大連結成分を持つこ

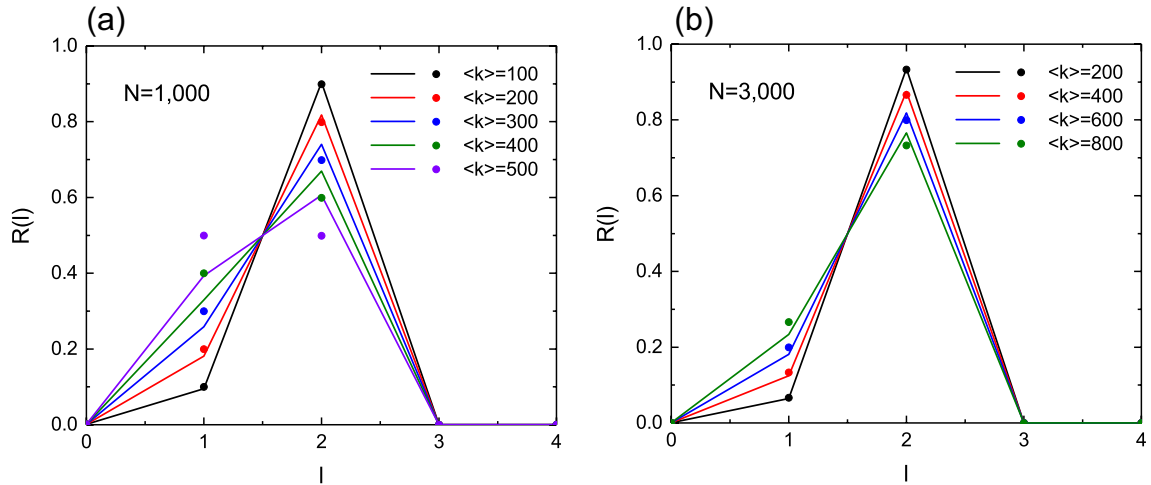


図 C.1: Erdős-Rényi ランダムグラフに対するノード間距離分布 $R(l)$. 色の違いは平均次数の異なるランダムグラフに対する結果を示している. 実線は5.2.1節の議論に基づく理論計算の結果, シンボルは数値計算結果 (10 サンプル平均) を表している. (a) はノード数が $N = 1,000$, (b) は $N = 3,000$ のネットワーク に対する結果である.

となる. 図 C.2(b) を見ると, スケールフリー・ネットワークにおいても $\gamma = \gamma_c$ 近傍で理論計算と数値計算の不一致が大きくなることが確認できる.

臨界点近傍における近似精度の低下を定量的に議論するために, 図 6.3 と同じ NNCRN に対して解析的に計算した $P_1^{\text{ana}}(k, k', l)$ と数値的に計算した $P_1^{\text{num}}(k, k', l)$ の差異を (6.5) 式で定義される確率関数間の換算距離を用いて定量化した. 結果は $\varrho = 0.8, 0.0, -0.8$ のそれぞれの場合に対して, $d(P_1^{\text{ana}}, P_1^{\text{num}}) = 0.027, 0.016, 0.013$ となった. 換算距離がゼロに近い値であることから, 平均場理論による近似計算の精度が極めて高いことが確認できる. しかし, 平均場近似は NNCRN がパーコレーション臨界点に近いとき精度が著しく低下することに注意しなければならない. 換算距離 $d(P_1^{\text{ana}}, P_1^{\text{num}})$ を平均次数 $\langle k \rangle$ の関数としたとき, $d(P_1^{\text{ana}}, P_1^{\text{num}})$ の最大値はパーコレーション臨界点 $\langle k \rangle_c$ 近傍で (またはわずかに $\langle k \rangle_c$ より大きな $\langle k \rangle$ で), 鋭いピークを作り現れる (表 C.1 参照). この原因は, パーコレーション臨界点付近においては連結成分サイズが極めて広く分布するにもかかわらず, 平均場理論が始点ノードの属する連結成分サイズの揺らぎを無視しているためである.

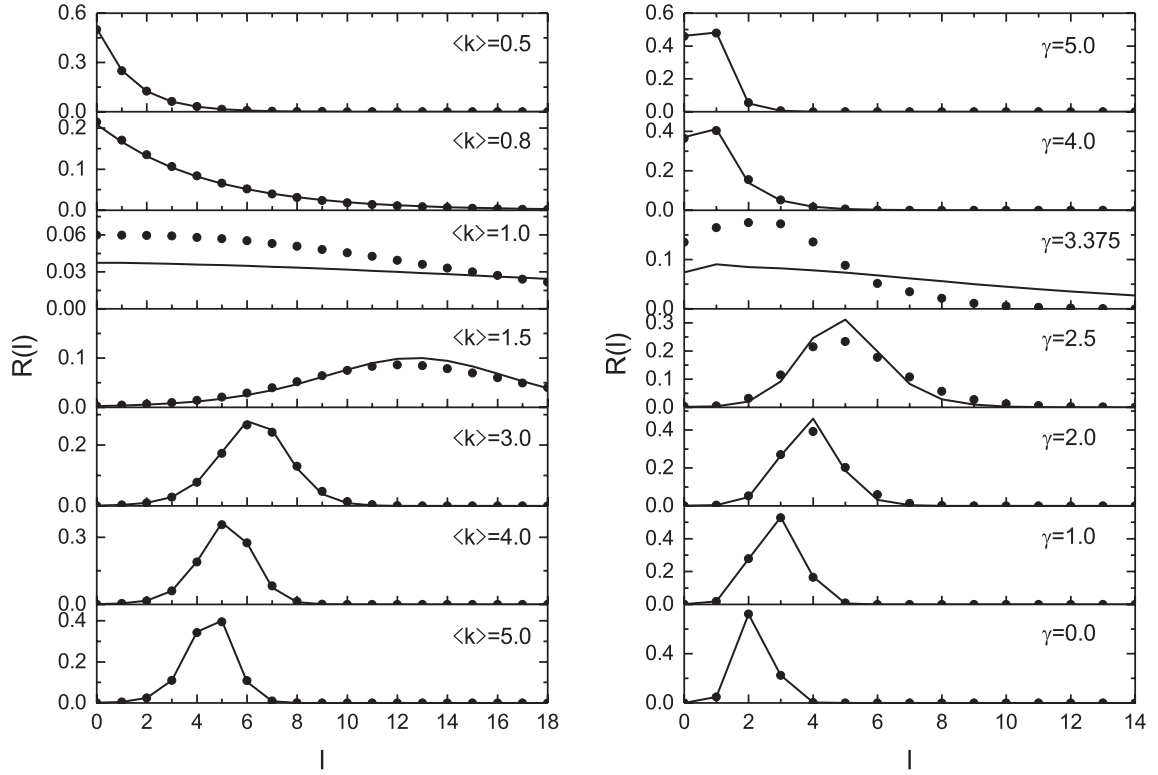


図 C.2: ノード間距離分布 $R(l)$ に対する理論計算（実線）と数値計算結果（シンボル）の比較. (a) は様々な平均次数 $\langle k \rangle$ を有する $N = 1,000$ の Erdős-Rényi ランダムグラフに対する結果であり, 数値計算は 100 サンプルに渡る平均を表している. (b) は最小次数 $k_{\min} = 1$ から最大次数 $k_{\max} = 100$ の間で次数分布関数が $P(k) \propto k^{-\gamma}$ となる $N = 1,000$ のスケールフリー・ランダムグラフに対する結果である. (b) の各パネルは γ の値の異なるネットワークに対する結果を示しており, 数値計算結果は 100 サンプルに渡る平均を表している.

表 C.1: 図 6.3 と同じ NNCRN に対してランダムにエッジを削除し, 換算距離 $d(P_1^{\text{ana}}, P_1^{\text{num}})$ を平均次数 $\langle k \rangle$ の関数としたときの最大値 $d_{\max}$ と, とパーコレーション臨界点における平均次数 $\langle k \rangle_c$. $\varrho = 0.8, -0.8$ の場合における臨界点 $\langle k \rangle_c$ は最大連結成分サイズから数値的に計算した値である. $\varrho = 0.0$ の場合に対しては, $\langle k \rangle_c = 1$ であることが知られている [1].

	$\varrho = 0.8$	$\varrho = 0.0$	$\varrho = -0.8$
$d_{\max}$	0.66	0.66	0.57
$\langle k \rangle_{\max}$	1.05	1.10	1.25
$\langle k \rangle_c$	0.8	1.0	1.1

付 録 D 現実世界の複雑ネットワーク

6 章では (6.5) 式で定義される確率関数間の換算距離を用いて、長距離次数相関の強さおよび内因性長距離次数相関の強さを表す指標 $d(P, P_0)$ および $d(P, P_1)$ を導入した．表 D.1 は、表 6.1 の現実ネットワークを含む 45 種類の複雑ネットワークに対して、ネットワークの種類とノード数、エッジ数、Spearman 順位相関係数、 $d(P, P_0)$ 、 $d(P, P_1)$ を調べた結果をまとめたものである． $d(P, P_0)$ を横軸、 $d(P, P_1)$ を縦軸として表 6.1 の値を図 D.1 にプロットした．図 D.1 から明らかなとおり、 $d(P, P_0)$ の値が大きいネットワークは $d(P, P_1)$ の値も大きい傾向がある．また、多くの実ネットワークにおいて $d(P, P_1)$ は $d(P, P_0)$ よりも大きく、内因性の長距離次数相関が支配的であることが見て取れる．

表 D.1: 図 D.1 で扱っている現実世界の複雑ネットワークにおける長距離次数相関の強さおよび内因性長距離次数相関の強さの指標である $d_0 = d(P, P_0)$ および $d_1 = d(P, P_1)$ の値. N はノード数, M はエッジ数, ρ は Spearman 相関係数の値である.

Label	分野	ノード	N	エッジ	M	ρ	d_0	d_1	ref.
actor	経済	俳優	82,583	共演関係	4,475,520	0.40	0.73	0.57	[85]
condmat	経済	研究者	23,133	共著関係	93,439	0.26	0.71	0.61	[62]
email	社会	人間	36,692	メールの送受信	183,831	-0.01	0.37	0.40	[85]
spires	経済	研究者	53,810	共著関係	4,573,868	0.66	0.97	0.68	[109]
actor98	経済	俳優	225,583	共演関係	8,302,734	0.30	0.67	0.54	[63]
as060722	工学	自律システム (AS)	22,963	回線	48,436	-0.51	0.19	0.26	[108]
as010316	工学	自律システム (AS)	10,515	回線	21,455	-0.55	0.17	0.25	[86]
gnutella	その他	ホスト	10,876	コネクション	39,994	0.00	0.11	0.13	[85]
router	工学	ルーター	192,244	リンク	609,066	0.27	0.89	0.82	[109]
linux	工学	ソースコード	30,834	内包関係	213,217	-0.35	0.39	0.26	[85]
neural	生物	神経	297	シナプス/ギャップ結合	2,148	-0.11	0.20	0.15	[108]
yeast03	生物	タンパク質	2,284	相互作用	6,646	-0.09	0.34	0.34	[85]
folding	生物	タンパク質構造	132,167	遷移	228,967	0.51	0.97	0.95	[86]
youtube	社会	SNS ユーザー	1,134,890	交友関係	2,987,624	-0.08	0.63	0.50	[85]
airport500	工学	空港	500	フライト	2,980	-0.20	0.35	0.22	[85]
berkstan	工学	ウェブサイト	685,230	ハイパーリンク	6,649,470	-0.17	0.96	0.92	[85]
stanford	工学	ウェブサイト	255,265	ハイパーリンク	1,941,926	-0.20	0.95	0.88	[85]
notredame	工学	ウェブサイト	325,729	ハイパーリンク	1,090,108	-0.11	0.90	0.81	[85]
baidu	工学	ウェブサイト	415,624	ハイパーリンク	2,374,044	-0.32	0.87	0.41	[85]
amazon	その他	商品	334,863	同時購入	925,872	-0.19	0.96	0.96	[85]
as000102	工学	自律システム (AS)	6,474	回線	12,572	-0.54	0.17	0.25	[85]
caida	工学	自律システム (AS)	26,475	回線	53,381	-0.53	0.21	0.25	[85]
celegans	生物	代謝物質	453	相互作用	2,025	-0.20	0.27	0.14	[85]
airtraffic	工学	空港/設備	1,226	推奨経路	2,408	0.02	0.56	0.54	[85]
citeseer	その他	論文	384,054	引用	1,736,145	-0.02	0.85	0.77	[85]
human	生物	タンパク質	3,023	相互作用	6,149	-0.09	0.40	0.37	[63]
wgoogle	工学	ウェブサイト	875,713	ハイパーリンク	4,322,051	-0.14	0.89	0.84	[85]
powergrid	工学	発電所/変圧器/変電所	4,941	高圧電線	6,594	-0.05	0.91	0.91	[85]
airportus	工学	空港	1,574	フライト	17,215	-0.07	0.28	0.26	[85]
reactome	生物	タンパク質	6,229	相互作用	146,160	0.50	0.82	0.63	[85]
software	工学	クラスファイル	6,434	依存関係	53,658	-0.54	0.78	0.06	[85]
brightkite	社会	SNS ユーザー	58,228	交友関係	214,078	0.22	0.54	0.41	[85]
hepph	経済	研究者	12,006	共著関係	118,489	0.72	0.89	0.52	[85]
hepth	経済	研究者	9,875	共著関係	25,973	0.29	0.70	0.61	[85]
hamster	社会	SNS ユーザー	1,858	交友関係	12,534	-0.05	0.41	0.33	[85]
twitter	社会	SNS ユーザー	23,370	交友関係	32,831	-0.74	0.82	0.67	[85]
blogs	社会	ブログサイト	1,224	ハイパーリンク	16,715	-0.21	0.23	0.20	[85]
astroph	経済	研究者	18,771	共著関係	198,050	0.35	0.74	0.48	[85]
gplus	社会	SNS ユーザー	23,628	交友関係	39,194	-0.74	0.63	0.31	[85]
facebook	社会	SNS ユーザー	63,731	交友関係	817,035	0.38	0.73	0.43	[85]
citepph	その他	論文	34,546	引用	420,877	0.13	0.72	0.65	[85]
citepth	その他	論文	27,769	引用	352,285	0.27	0.82	0.65	[85]
wikicat	その他	記事	1,853,493	同一カテゴリに所属	3,795,785	-0.62	0.91	0.82	[85]
wikihu	その他	記事	621,500	ハイパーリンク	29,680,468	0.13	0.65	0.24	[85]
wikise	その他	記事	185,262	ハイパーリンク	3,835,207	0.18	0.55	0.25	[85]

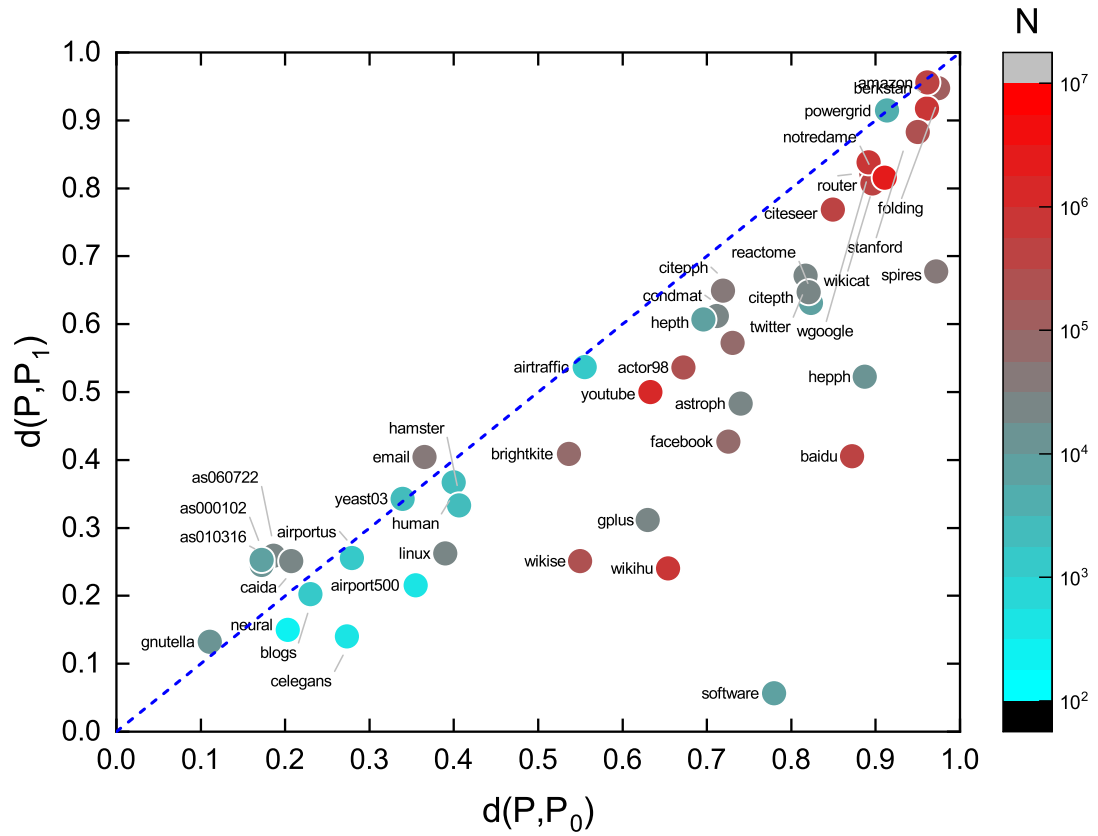


図 D.1: 表 D.1 にある現実ネットワークにおける長距離次数相関の強さおよび内因性長距離次数相関の強さ. 横軸は長距離次数相関の強さの指標である $d_0 = d(P, P_0)$, 縦軸は内因性長距離次数相関の強さの指標である $d_1 = d(P, P_1)$ としている. シンボルの色はネットワーク・サイズ N を表す.

参考文献

- [1] R. Albert, A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002)
- [2] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
- [3] 増田直紀、今野紀雄、「複雑ネットワークの科学」 産業図書株式会社 (2005)
- [4] 矢久保考介、「複雑ネットワークとその構造」 共立出版 (2013).
- [5] A.-L. Barabási, *Network Science* (Cambridge University Press, Cambridge, 2016).
- [6] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
- [7] S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, Rev. Mod. Phys. **80**, 1275 (2008)
- [8] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Rev. Mod. Phys. **87**, 925 (2015)
- [9] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001).
- [10] S. Maslov and K. Sneppen, Science **296**, 910 (2002).
- [11] M. E. J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).
- [12] M. E. J. Newman, Phys. Rev. E **67**, 026126 (2003).
- [13] J. Park and M. E. J. Newman, Phys. Rev. E **68**, 026112 (2003).
- [14] M. Catanzaro, G. Caldarelli, and L. Pietronero, Phys. Rev. E **70**, 037101 (2004).
- [15] M. A. Serrano, M. Boguñá, and R. Pastor-Satorras, Phys. Rev. E **74**, 055101 (2006).
- [16] B. Fotouhi and M. G. Rabbat, Eur. Phys. J. B **86**, 510 (2013).
- [17] N. Litvak and R. van der Hofstad, Phys. Rev. E **87**, 022801 (2013).
- [18] S. N. Soffer and A. Vázquez, Phys. Rev. E **71**, 057101 (2005).

- [19] M. A. Serrano and M. Boguñá, Phys. Rev. E **72**, 036133 (2005).
- [20] J. C. Miller, Phys. Rev. E **80**, 020901 (2009).
- [21] E. Estrada, Phys. Rev. E **84**, 047101 (2011).
- [22] R. van der Hofstad, A. J. E. M. Janssen, J. S. H. van Leeuwen, and C. Stegehuis, Phys. Rev. E **95**, 022307 (2017).
- [23] J. Menche, A. Valleriani, and R. Lipowsky, Phys. Rev. E **81**, 046103 (2010).
- [24] R. Xulvi-Brunet and I. M. Sokolov, Phys. Rev. E **70**, 066102 (2004).
- [25] S.-H. Yook, F. Radicchi, and H. Meyer-Ortmanns, Phys. Rev. E **72**, 045105(R) (2005).
- [26] C. Song, S. Havlin, and H. A. Makse, Nat. Phys. **2**, 275 (2006).
- [27] J. S. Kim, B. Kahng, D. Kim, Phys. Rev. E **79**, 067103 (2009).
- [28] Z.-W. Wei and B.-H. Wang, Phys. Rev. E **94**, 032309 (2016).
- [29] V. M. Eguíluz and K. Klemm, Phys. Rev. Lett. **89**, 108701 (2002).
- [30] M. Boguñá and R. Pastor-Satorras, Phys. Rev. E **66**, 047104 (2002).
- [31] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **90**, 028701 (2003).
- [32] T. Gross, Carlos J. Dommar D’Lima, and B. Blasius, Phys. Rev. Lett. **96**, 208701 (2006).
- [33] A. E. Motter, C. S. Zhou, and J. Kurths, Europhys. Lett. **69**, 334 (2005).
- [34] F. Sorrentino, M. di Bernardo, G. H. Cuéllar, and S. Boccaletti, Physica D **224**, 123 (2006).
- [35] M. Di Bernardo, F. Garofalo, and F. Sorrentino, Int. J. Bifurcation Chaos **17**, 3499 (2007).
- [36] M. D. LaMar and G. D. Smith, Phys. Rev. E **81**, 046206 (2010).
- [37] C. E. La Rocca, L. A. Braunstein, and P. A. Macri, Physica A **390**, 2840 (2011).

- [38] V. Avalos-Gaytan, J. A. Almendral, D. Papo, S. E. Schaeffer, and S. Boccaletti, *Phys. Rev. E* **86**, 015101 (2012).
- [39] J. Restrepo and E. Ott, *Europhys. Lett.* **107**, 60006 (2014).
- [40] S. Jalan, A. Kumar, A. Zaikin, and J. Kurths, *Phys. Rev. E* **94**, 062202 (2016).
- [41] J. Hindes and I. B. Schwartz, *Phys. Rev. E* **95**, 052317 (2017).
- [42] Z. Rong, X. Li, and X. Wang, *Phys. Rev. E* **76**, 027101 (2007).
- [43] Z. Rong and Z.-X. Wu, *Europhys. Lett.* **87**, 30001 (2009).
- [44] S. Devlin and T. Treloar, *Phys. Rev. E* **80**, 026105 (2009).
- [45] R. J. La, *IEEE/ACM Trans. Netw.* **25**, 2484 (2017).
- [46] J. D. Noh, *Phys. Rev. E* **76**, 026116 (2007).
- [47] A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, *Phys. Rev. E* **78**, 051105 (2008).
- [48] C. M. Schneider, A. A. Moreira, J. S. Andrade Jr., S. Havlin, and H. J. Herrmann, *Proc. Natl. Acad. Sci.* **108**, 3838 (2011).
- [49] T. Tanizawa, S. Havlin, and H. E. Stanley, *Phys. Rev. E* **85**, 046109 (2012).
- [50] S. Watanabe and Y. Kabashima, *Phys. Rev. E* **94**, 032308 (2016).
- [51] C. Orsini, M. M. Dankulov, P. Colomer-de-Simón, A. Jamakovic, P. Mahadevan, A. Vahdat, K. E. Bassler, Z. Toroczkai, M. Boguñá, G. Caldarelli, S. Fortunato, and D. Krioukov, *Nat. Commun.* **6**, 8627 (2015).
- [52] B. Tadić, S. Thurner, and G. J. Rodgers, *Phys. Rev. E* **69**, 036102 (2004).
- [53] M. Boguñá, C. Castellano, and R. Pastor-Satorras, *Phys. Rev. Lett.* **111**, 068701 (2013).
- [54] R. E. Boulos, A. Arneodo, P. Jensen, and B. Audit, *Phys. Rev. Lett.* **111**, 118102 (2013).
- [55] L. W. Swanson, O. Sporns, and J. D. Hahn, *Proc. Natl. Acad. Sci.* **113**, E5972 (2016).
- [56] D. Rybski, H. D. Rozenfeld, and J. P. Kropp, *Euro. Phys. Lett.* **90**, 28002 (2010).

- [57] M. Mayo, A. Abdelzaher, and P. Ghosh, *Comput. Soc. Netw.* **2**, 4 (2015).
- [58] A. Arcagni, R. Grassi, S. Stefani, and A. Torriero, *Eur. J. Oper. Res.* **262**, 708 (2017).
- [59] A. Allen-Perkins, J. M. Pastor, and E. Estrada, *Appl. Math. Comput.* **311**, 262 (2017).
- [60] Y. Fujiki, T. Takaguchi, and K. Yakubo, *Phys. Rev. E* **97**, 062308 (2018).
- [61] Y. Fujiki and K. Yakubo, *Phys. Rev. E* **101**, 032308 (2020).
- [62] <http://snap.stanford.edu/data/index.html>
- [63] <http://www.complex-networks.net/datasets.html>
- [64] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhim, **85**, 4633 (2000).
- [65] S. N. Dorogovtsev, L. Ferreira, V. Goltsev, J. F. F. Mendes, *Phys. Rev. E* **81**, 031135 (2010).
- [66] M. Raschke, M. Schläpfer, R. Nibali, *Phys. Rev. E* **82**, 037102 (2010).
- [67] W.-Y. Zhang, Z.-W. Wei, B.-H. Wang, X.-P. Han, *Physica A* **451**, 440 (2016).
- [68] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Phys. Rev. E* **64**, 026118 (2001).
- [69] H. J. Kim and J. M. Kim, *Phys. Rev. E* **72**, 036109 (2005).
- [70] A. Bekessy, P. Bekessy, J. Komlos, *Stud. Sci. Math. Hungar.* **7**, 343 (1972).
- [71] M. Molloy, B. Reed, *Random Struct. Algorithms* **6**, 161 (1995).
- [72] P. Erdős and A. Rényi, *Publ. Math.* **6**, 290 (1959).
- [73] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, *Comput. Netw.* **33**, 309 (2000).
- [74] M. E. J. Newman, *Proc. Natl. Acad. Sci. (USA)* **98**, 404 (2001).
- [75] L. A. N. Amaral, A. Scala, M. Barthélemy, H. E. Stanley, *Proc. Natl. Acad. Sci. (USA)* **97**, 11149 (2000).
- [76] R. F. i Cancho, and R. V. Solé, *Proc. Roy. Soc. London B* **268**, 2261 (2001).
- [77] S. Valverde, R. F. Cancho, and R. V. Solé, *Europhys. Lett.* **60**, 512 (2002).

- [78] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, Proc. Natl. Acad. Sci (USA) **101**, 3747 (2004).
- [79] D. J. Watts and S. H. Strogatz, Nature **393**, 440 (1998).
- [80] C. Song, S. Havlin, and H. A. Makse, Nature **433**, 392(2005).
- [81] L. K. Gallos, C. Song, H. A. Makse, Phys. Rev. Lett. **100**, 248701 (2008).
- [82] L. K. Gallos, H. A. Makse, M. Sigman, Proc. Natl. Acad. Sci. USA **109**, 2825 (2012).
- [83] F. Kawasaki, K. Yakubo, Phys. Rev. E **82**, 036113 (2010)
- [84] C. Song, L. K. Gallos, S. Havlin, H. A. Makse, J. Stat. Mech.: Theory Exp. (2007), P03006.
- [85] <http://konect.uni-koblenz.de>
- [86] <http://cosinproject.eu/data.html>
- [87] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, 2001, Phys. Rev. Lett. **86**, 3682.
- [88] R. Cohen, D. ben-Avraham, and S. Havlin, Phys. Rev. E **66**, 036113 (2002).
- [89] R. Cohen, S. Havlin, and D. ben-Avraham, in *Hand-book of Graphs and Networks*, edited by S. Bornholdt and H. G. Schuster (Wiley-VCH, Weinheim), p.85 (2003).
- [90] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Nucl. Phys. B **653**, 307 (2003).
- [91] H. D. Rozenfeld, S. Havlin, and D. ben-Avraham, New J. Phys. **9**, 175 (2007).
- [92] J. P. Gleeson, S. Melnik, and A. Hackett, Phys. Rev. E **81**, 066114 (2010)
- [93] M. A. Serrano, D. Krioukov, and M. Boguñá, Phys. Rev. Lett. **106**, 048701 (2011).
- [94] M. E. J. Newman and R. M. Ziff, Phys. Rev. E **64**, 016706 (2001).
- [95] B. Karrer, M. E. J. Newman, and L. Zdeborová, Phys. Rev. Lett. **113**, 208702 (2014)
- [96] J. Park and M. E. J. Newman, Phys. Rev. E **70**, 066117 (2004)
- [97] A. Vázquez and Y. Moreno, Phys. Rev. E **67**, 015101(R) (2003).

- [98] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, *Physics Reports* **424**, 4–5, 175–308 (2006).
- [99] Y. Kuramoto, in *Proceedings of the International Symposium on Mathematical Problems in Theoretical Physics, Lecture Notes in Physics*, Vol. 39, edited by H. Araki (Springer, Berlin, 1975).
- [100] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*, Springer Series in Synergetics, Vol. 19 (Springer, Berlin, 1984).
- [101] J. Chalupa, P. L. Leath, and G. R. Reich, *J. Phys. C* **12**, L31 (1979).
- [102] B. Bollobás, in *Graph Theory, and Combinatorics: Proceedings of the Cambridge Combinatorial Conference in Honour of Paul Erdős*, edited by B. Bollobás (Academic, New York), p.35 (1984).
- [103] Y. Fujiki, S. Mizutaka, and K. Yakubo, *Eur. Phys. J. B* **90**, 126 (2017).
- [104] S. Maslov, K. Sneppen, and A. Zaliznyak, *Physica A* **333**, 529, (2004).
- [105] T. Hasegawa, M. Sato, and K. Nemoto *Phys. Rev. E* **85**, 01710 (2012).
- [106] A. Fronczak, P. Fronczak, and J. A. Holyst, *AIP Conf. Proc.* **776**, 52 (2005).
- [107] S. Melnik and J. P. Gleeson, Preprint arXiv:1604.05521.
- [108] <http://www-personal.umich.edu/mejn/netdata>
- [109] <http://math.bu.edu/people/kolaczyk/datasets.html>