

情報伝達に基づいた 有向重み付き複雑ネットワーク解析

鈴木 智 也^{†1}

自然界や社会に存在するネットワークには、スモールワールド現象を示すものが数多く存在する。スモールワールド現象とは、各ノードが局所的にクラスタ化している割には他の遠いノードと短いステップでアクセス可能という特徴を持ち、情報伝達のしやすさに関係している。この性質は Watts らによって、クラスタ係数と平均最短経路長により定義されている。平均最短経路長はダイクストラの方法を用いればネットワークの種類を問わず算出できるが、クラスタ係数においては有向グラフの場合、前処理なしでは計算できない。そこで本研究では、情報伝達の仕方を考慮することで前処理の方法を検討し、有向グラフでも適切にクラスタ係数を計算する手法を提案する。その妥当性を検証するために、Watts らが提案した Small-world ネットワークモデルを重み付き有向グラフに拡張し、本提案手法を適用した。さらに、今まで調査されてきた線虫の神経網などの実ネットワークについて本提案手法を適用したところ、従来よりも顕著にスモールワールド現象が確認された。

Analysis for Directed and Weighted Complex Networks on the Basis of Information Flow

TOMOYA SUZUKI^{†1}

To analyze a structure of natural or social networks, the small-world network structure is often discussed. In the small-world network, each node can be more highly clustered than the random network, and can access other nodes through fewer edges than the regular network. It is well known that the small-world network can effectively communicate and share information, and is defined by two measures: the cluster coefficient and the shortest path length. The shortest path length can be calculated by Dijkstra's algorithm, which can be applied to directed and/or weighted graphs. However, the cluster coefficient cannot be applied to directed networks. In this report, we propose a method to estimate the cluster coefficient for directed and weighted networks on the basis of how to propagate information sent by parent nodes. Then, we confirm the validity of the proposed method with the numerical network models which we proposed by

extending the Watts-Strogatz model to directed and weighted type. Moreover, we apply the proposed method to real directed networks discussed by previous studies, and can confirm the small-world phenomena of real networks more clearly from the viewpoint of directed and weighted networks.

1. はじめに

1967年のMilgramの実験により、実社会において他人と他人の間を仲介する人数はたかだか6人程度であることが示された。我々個人は、世界の人口に比べてきわめて少人数としか知り合いでないにもかかわらず、たかだか6人のつてをたどるだけで世界中の人々にアクセスできてしまうことは驚異的な事実である。しかし普遍的な事実は分かったものの、それを説明するために、人間関係のネットワーク構造をモデル化することは容易でなかった。なぜならば、従来のグラフ理論では、各ノード間の最短経路が短いことを表現するには(他者間の媒介者が少ないことに対応)、ランダムグラフが適当であった。しかし、それでは実社会で陽に存在する友達同士のコミュニティをまったく表現できない。一方、コミュニティを表現するにはレギュラーグラフが適切である。しかし、これでは各ノード間の最短経路が長くなってしまふ。このような矛盾の解消が課題であったが、1997年にWattsらは環状のレギュラーグラフの一部をランダムに張り替えることによって、最短経路性とコミュニティ性の両方を満足するシンプルなネットワークモデル(重みなし無向グラフ)を提案した¹⁾。また、ネットワークを特徴づける指標として平均最短経路長とクラスタ係数を用い、短い平均最短経路長と高いクラスタ係数を同時に有することをスモールワールド現象と呼んだ。このようにネットワークを特徴づける指標をわずか2つに厳選したことと、様々な実ネットワークからもこのスモールワールド現象が確認されたことにより、スモールワールドネットワーク研究のブームが到来した。その後、Barabásiらのスケールフリーネットワーク²⁾なども同様に重要なモデルと認知され、これら周辺概念を総合して「複雑ネットワーク」と呼ばれるようになった。

実社会に存在するネットワークの多さを考えれば、複雑ネットワーク研究の応用例が見えてくる。たとえば、うわさの広がり、伝染病の伝搬、経済活動、インターネット網、神経回路網、交通渋滞、地震の連鎖など広範囲の分野で研究が活発に行われている。しかし当然な

^{†1} 同志社大学理工学部情報システムデザイン学科

Department of Information System Design, Faculty of Science and Engineering, Doshisha University

がら、現実のネットワークは Watts らが提案したモデルのような重みなし無向グラフだけではない。人間関係のネットワークですら実際は有向グラフであり、脳内神経回路網では有向性のほかにもシナプス加重といった“重み”がある。しかし文献 1) のクラスタ係数では、有向または重みを持つネットワークを評価できないので、従来研究においては、重みや有向性を取り除くことでネットワークを簡略化し、重みなし無向グラフとして解析するが多い。しかし、その簡略化の仕方によって解析結果が変化することも少なくなく、さらに簡略化手法自体をも議論しないままでは、得られた解析結果の信頼性に疑問が残る。

このような問題を解決するために、近年、重み付きグラフ用のクラスタ係数を提案する研究が盛り上がりを見せている³⁾。提案者によってクラスタ係数の定義が異なるものの、定義に至った経緯や有用性が議論されているため、解析結果の信頼性を向上させることができる。さらには、Fagiolo によって有向グラフ用のクラスタ係数の計算方法が提案されている⁴⁾。この方法により、ネットワークのトポロジをグラフ理論的な一般性を持って特徴づけることができる。

しかし、ネットワークのトポロジを情報伝達の仕方と関連づけて評価したい場合、グラフ理論的な一般性を重視するより、情報伝達の仕方をふまえた計算方法を議論する必要がある。たとえば、文献 4) の方法では、情報発信者と情報受信者、情報が流れる通路（発信者-受信者間および受信者-受信者間のエッジ）がすべて平等に扱われ、情報伝達における役割の違いが考慮されていない。そこで、本研究では、親ノードを情報発信者、子ノードを情報受信者と区別し、情報の伝達に特化した重み付き有向グラフ用のネットワーク評価手法を提案する。つまり、単にネットワークトポロジを評価するのではなく、親ノード（情報発信者）が発する情報が伝播するプロセスを考慮しつつネットワーク構造を定量化することを目的としている（2.2 節）。

本提案手法の特徴を確認するために、3 章では、従来の Watts らによって提案された無向かつ重みがないスモールワールドモデルを重み付き有向グラフに拡張し、スモールワールド現象の分析を行う。さらに、4 章では、従来より解析されてきた実ネットワーク（線虫の神経網⁵⁾ やブログの相互リンク網⁶⁾）のスモールワールド現象について再検討する。一連の実験を通じて、従来手法と本提案手法での解析結果を比較することにより、本提案手法の有用性を確認する。

2. 種々のネットワークのクラスタ係数

2.1 従来手法

いかなるネットワークも 2 次元平面上のグラフとして表現できる。各ノード番号を i, j, k などで表現すると、ノード間の結合強度を W_{ij}, W_{ik}, W_{jk} のように書く。もし有向グラフであれば $W_{i \rightarrow j}$ のように矢印を入れて書くことにする。一般に、 $W_{i \rightarrow j} \neq W_{j \rightarrow i}$ である。一方、重み付き無向グラフであれば、 $W_{ij} = W_{ji} \in [0, 1]$ であり、重みなし無向グラフであれば、 $W_{ij} = W_{ji} \in \{0, 1\}$ である。前章ではグラフの種類の違いによるネットワーク特徴量算出の問題点について説明したが、平均最短経路長は Dijkstra のアルゴリズムを用いれば、グラフの種類を問わず、いかなるグラフでも計算できる（重み付きグラフの場合、2 点間距離を $1/W_{i \rightarrow j}$ のように結合強度の逆数で与えるのが一般的）。しかし、クラスタ係数の算出に関しては、新しい考え方が必要となる。

Watts らが提案したクラスタ係数は以下のとおりである。

$$C_{WS}(k) = \frac{t}{nC_2} \quad (1)$$

ここで、ある親ノードのインデックスを k とし、親ノード k と結合している隣接ノード数を n としている。また、隣接ノードどうしが結合している数を t としている。ネットワーク全体のクラスタ係数を算出する場合は、親ノードに関して平均値をとる。つまり、

$$C_{WS} = \langle C_{WS}(k) \rangle_k \quad (2)$$

と書ける。もし解析対象のネットワークが重み付きグラフ ($W_{ij} \in [0, 1]$) であったら、ある閾値を設けて、重みなしグラフ ($W_{ij} \in \{0, 1\}$) に変換する必要がある。このようなパラメトリックな方法では、一般にパラメータによって解析結果が変わるので、採用したパラメータの妥当性について議論する必要がある。

さらに、従来提案された重み付き無向グラフ用のクラスタ係数を代表して Fernandez らのクラスタ係数⁷⁾ を紹介する。

$$C_{LF}(k) = \sum_{i \neq j \in N(k)} \frac{W_{ij}}{nP_2} \quad (3)$$

$$C_{LF} = \langle C_{LF}(k) \rangle_k \quad (4)$$

ここで、 $N(k)$ は親ノード k と結合している隣接ノード集合を意味する。この手法を用いても解析対象のネットワークが有向グラフ ($W_{i \rightarrow j} \neq W_{j \rightarrow i}$) の場合は、何らかの無向グラフへの変換を行う必要があり、その妥当性を示す必要がある。一例として、平均化 ($W_{ij} = \frac{W_{i \rightarrow j} + W_{j \rightarrow i}}{2}$)

や大きい方を採用 ($W_{ij} = W_{i \rightarrow j}$, if $W_{i \rightarrow j} > W_{j \rightarrow i}$) などの変換方法が考えられる。

さらに, Fagiolo らによって有向グラフ用のクラスタ係数⁴⁾が提案されている(式(5), (6)). 文献4)では様々なパターンのクラスタ係数を提案しているが, 本論文では, 最も代表的なものとして提案手法(次節)の比較を行う。

$$C_{GF}(k) = \frac{\sum_{j,h} (W_{i \rightarrow j}^{\frac{1}{3}} + W_{j \rightarrow i}^{\frac{1}{3}})(W_{i \rightarrow h}^{\frac{1}{3}} + W_{h \rightarrow i}^{\frac{1}{3}})(W_{j \rightarrow h}^{\frac{1}{3}} + W_{h \rightarrow j}^{\frac{1}{3}})}{2[d_i^{\text{tot}}(d_i^{\text{tot}} - 1) - 2d_i^{\leftrightarrow}]} \quad (5)$$

$$C_{GF} = \langle C_{GF}(k) \rangle_k \quad (6)$$

ここで, d_i^{tot} はノード i から出るエッジと入るエッジの総数, d_i^{\leftrightarrow} は双方向のエッジの本数である。

2.2 提案手法

情報伝達の仕組みを考慮しつつ有向グラフのクラスタ係数を計算するうえで検討しなければならないことが2点ある。まず1点目として, 任意のクラスタにおいて, 親ノードが隣接ノードに影響を与える仕組みを検討する必要がある。考えてみると, ネットワーク上で各ノードが相互作用するタイプは, 情報などの媒体を単に伝送するケースと, パーコレート(浸透)するケースに大別できる。前者は, ニューラルネット上のパルスの伝送や Web を巡回する人口移動など, 媒体が送信ノードから受信ノードに移動するケースである。この場合, ノード間の結合強度を媒体の伝送率で定義する。後者は, うわさ, 伝染病, 火災などの伝搬といったパーコレーション現象を指し, 送信ノードの媒体を受信ノードに繁殖させるケースである。この場合では, ノード間の結合強度を媒体の増加率で定義する。

次に2点目として, 情報伝送(あるいは繁殖)に方向がある有向グラフにとって, 親ノードに対する隣接ノードとはいかなるものかを検討する必要がある。親ノードと隣接するパターンは, 結合を受ける, 結合を伸ばす, 相互結合するの3種がある。しかし, 1点目による定義を用いると, 隣接ノードは親ノードから結合を受けるものに限られることが導かれる。

以下, 具体的に提案法を示す。ただし, 4章において本提案手法を用いて線虫のニューラルネットワークや, ブログの相互リンクネットワークを解析するため, 本節では, 媒体を伝送するケースについて解説する。媒体を繁殖させるケースについては付録を参照されたい。しかし, いずれのケースでも, 後述の式(8)が導かれ, それ以下の議論はまったく同じになる。

まず手順1として, 親ノード k からの1次結合を考える(図1のSTEP1)。ノード i が有する親ノード k に関する情報量(媒体量)を $I_i(k)$ と書き, 親ノード k がこれから伝送

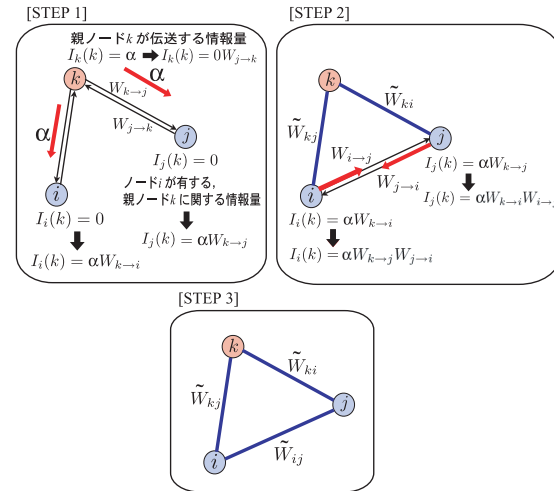


図1 提案手法の模式図。ただし, ノード間の相互作用により媒体を伝送するケース

Fig. 1 Diagrams of the proposed method in the case of transmitting information among nodes.

する情報量 $I_k(k)$ を α とする。最初, 親ノード k 以外のノードは, $I_k(k)$ に関する情報を得ていない。よって, 最初は $I_k(k) = \alpha$, $I_i(k) = 0$, $I_j(k) = 0$ である。

次に, 情報が伝送された後を考える。ノード i, j はそれぞれ伝送経路の結合強度(伝送率)に応じてノード k からの情報を受け, $I_i(k) = \alpha W_{k \rightarrow i}$, $I_j(k) = \alpha W_{k \rightarrow j}$ になる。そこで本研究では, ノード $k \rightleftharpoons i$ またはノード $k \rightleftharpoons j$ 間の双方向通信を1束にまとめて考えることにより, いったん, 有向グラフを無向グラフ化する。前述のように, 情報量の伝送率をネットワークの結合強度と見なすことより,

$$\tilde{W}_{jk}(= \tilde{W}_{kj}) = \frac{\alpha W_{k \rightarrow j} + 0 W_{j \rightarrow k}}{\alpha + 0} = W_{k \rightarrow j} \quad (7)$$

となる。ここで, 提案手法によって無向グラフ化された結合には \sim をつけている。同様に, $\tilde{W}_{ik}(= \tilde{W}_{ki}) = W_{k \rightarrow i}$ となる。逆に, $W_{k \rightarrow i} = 0$ ならば, $\tilde{W}_{ki} = \tilde{W}_{ik} = 0$ なので, 親ノード k の隣接ノードとなる必要十分条件は, $W_{k \rightarrow i} \neq 0$ である。

次に手順2として, 親ノード k からの2次結合を考える(図1のSTEP2)。手順1でのノード k に関する情報量がノード i, j 間の結合を経由し, $I_i(k) = \alpha W_{k \rightarrow j} W_{j \rightarrow i}$, $I_j(k) = \alpha W_{k \rightarrow i} W_{i \rightarrow j}$ に変化する。手順1と同様に, 双方向通信を1束にまとめて情報量

の伝送率を考えると,

$$\begin{aligned}\tilde{W}_{ij} (= \tilde{W}_{ji}) &= \frac{\alpha W_{k \rightarrow i} W_{i \rightarrow j} + \alpha W_{k \rightarrow j} W_{j \rightarrow i}}{\alpha W_{k \rightarrow j} + \alpha W_{k \rightarrow i}} \\ &= \frac{W_{k \rightarrow i} W_{i \rightarrow j} + W_{k \rightarrow j} W_{j \rightarrow i}}{W_{k \rightarrow j} + W_{k \rightarrow i}}\end{aligned}\quad (8)$$

が得られる.

以上により, 親ノード k を中心としたクラスタ内の全結合を, ノンパラメトリックに無向グラフ化することでできた (図 1 の STEP3). これにより, 文献 3) で提案されているような様々な重み付き無向グラフ用のクラスタ係数に代入することができる. ただし, 本提案法である無向グラフ化は, 有向グラフのクラスタ係数を算出することを目的として設計されているので, 平均最短経路長の算出には用いることはできない点に注意されたい. なお, 1 章で述べたように, 平均最短経路長は本来の有向グラフのまま算出できる.

本研究では, 上式を式 (3) の Fernandez のクラスタ係数に適用する. 式 (3) の W_{ij} に \tilde{W}_{ij} を代入すると,

$$C_{TS}(k) = \frac{1}{nP_2} \sum_{i \neq j \in N(k)} \frac{W_{k \rightarrow i} W_{i \rightarrow j} + W_{k \rightarrow j} W_{j \rightarrow i}}{W_{k \rightarrow j} + W_{k \rightarrow i}}\quad (9)$$

$$C_{TS} = \langle C_{TS}(k) \rangle_k\quad (10)$$

となり, まったく新しい有向グラフ用のクラスタ係数を定義することができる.

さらに, 提案手法と Fernandez のクラスタ係数を組み合わせる利点がある. もし, 解析対象のネットワークが無向グラフの場合に $C_{TS}(k)$ を適用すると, $W_{i \rightarrow j} = W_{j \rightarrow i}$ ($= W_{ij} = W_{ji}$) なので,

$$\begin{aligned}C_{TS}(k) &= \frac{1}{nP_2} \sum_{i \neq j \in N(k)} \frac{W_{k \rightarrow i} W_{i \rightarrow j} + W_{k \rightarrow j} W_{j \rightarrow i}}{W_{k \rightarrow i} + W_{k \rightarrow j}} \\ &= \frac{1}{nP_2} \sum_{i \neq j \in N(k)} \frac{W_{i \rightarrow j} (W_{k \rightarrow i} + W_{k \rightarrow j})}{W_{k \rightarrow i} + W_{k \rightarrow j}} \\ &= \frac{1}{nP_2} \sum_{i \neq j \in N(k)} W_{ij} = C_{LF}(k)\end{aligned}\quad (11)$$

つまり, 無向重みありグラフ用の Fernandez のクラスタ係数と等価になる. さらに, 解析対象のネットワークが無向かつ重みがなければ, 式 (11) において $W_{ij} = 0$ or 1 なので

$$C_{TS}(k) = \frac{t}{nC_2} = C_{WS}(k)$$

となり, Watts-Strogatz のクラスタ係数と等価になる. これにより, 式 (9) の提案指標 $C_{TS}(k)$ は, いかなるタイプのネットワークのクラスタ係数をも単一で算出することができ, 高い汎用性を有していることが分かる. なお, 解析対象ネットワークの結合強度の最大値を 1 に規格化しておけば, いずれのクラスタ係数も C_{TS} ($= C_{WS} = C_{LF} = C_{GF}$) $\in [0, 1]$ となる.

さらに, 本提案手法におけるクラスタ係数と最短経路長の役割についてまとめる. 最短経路長は, 2 ノード間距離を $1/W_{i \rightarrow j}$ として計算すれば結合強度が大きいほど 2 ノード間距離は短いので, 情報をネットワーク全体に素早く伝達できるかを示す指標となる. 一方, クラスタ係数は, 局所的に結合する小ネットワークにおいて, 情報をいかに共有するかを示す指標である. 式 (7), (8) 導出の過程に鑑みれば, 手順 1 において親ノードからの 1 次結合を考慮することにより, 親ノードの情報が直接的に子ノード (1 次結合) に伝達される. さらに手順 2 において親ノードからの 2 次結合を考慮することにより, 親ノードの情報が子ノード (1 次結合) を仲介し, 間接的にも子ノード (2 次結合) に伝達され, 情報を共有する. このようにクラスタ係数が高いほど子ノード間で親ノードから得た情報を共有し合う. ニューラルネットワークを例にあげれば, 親ノードが発したパルスの子ノードらが受け取り, 互いにパルスを伝送し合うことで互いの内部状態を強化し合う. つまりクラスタ係数が高いほど, 同時発火が起こりやすい.

なお, 提案手法において親ノードと 1 次結合する子ノードを 2 つではなく, 同時にすべて考慮することで, より詳細に情報の伝達を分析できるが, 情報の合流や巡回など複雑なパターンが発生するため, 定式化は非常に困難である. そこで本論文では, 親ノードと 1 次結合する子ノードを 2 つに絞り, かつ 2 次結合までの情報伝達を考慮するにとどめた. 隣接する他の子ノードの効果は, 式 (9) のように平均化を施すことで取り込んでいる. 今後の課題として, より厳密に情報伝達を評価するために, 少しずつ複雑な状況を考慮し定式化することで, どのような新しい解析結果が得られるのかを分析しつつ, 提案手法の高度化を検討していく.

3. 提案手法の有用性の検証

次に本提案指標を適用するために, スモールワールド現象を示すシンプルな有向グラフのモデルを作成する. その際, Watts らが提案したスモールワールドモデル (以下, WS モデ

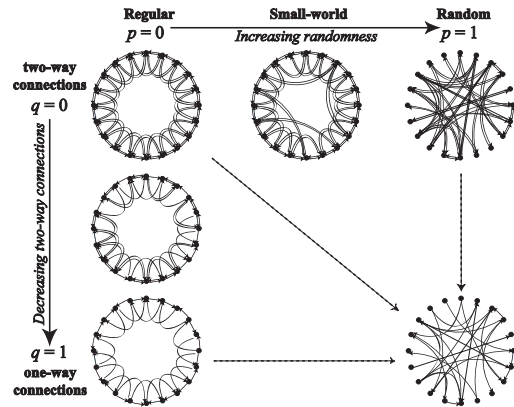


図 2 WS モデルの重み付き有向グラフ化 (拡張 WS モデル). 各ノードはノードから出て行く結合とノードに入る結合を持っている. 各パラメータ p と q は, それぞれ, 双方向結合をランダムに張り替える確率と, 一方の結合を削除して単方向結合にする確率を表している

Fig. 2 Extended WS model by making two-way networks according to two parameters p and q . The parameter p means the probability of randomly rewiring networks, and the parameter q means the probability of removing either incoming or outgoing connection.

ルと呼ぶ)¹⁾ を重み付き有向グラフ用に拡張する.

3.1 WS モデルの拡張

まず, 単純に WS モデル上の各エッジを双方向の結合に置き換え, さらに結合強度を $[0, 1]$ の一様乱数で分布させると, 図 2 上段にある重み付き双方向結合モデルになる. レギュラーグラフの双方向結合を保持したままランダムに再結合する確率を p すると, $p = 0$ でレギュラーグラフ, $p = 1$ でランダムグラフ, その中間でスモールワールドグラフとなる. 次に, 実際のネットワークは双方向結合と単方向結合が混在していることを考慮して, 双方向結合モデルからランダムに確率 q で双方向結合を選び, どちらか一方の結合を削除する. つまり $q = 1$ のとき, 図 2 下段のように完全な単方向結合モデルになる.

このモデルに対してノード数 N を 1,000, 各ノードから出力するエッジ数を 20 とし, 従来のように単純に無向グラフ化してクラスタ係数および平均最短経路長を計算する手法と, 本研究のように有向グラフのままクラスタ係数および平均最短経路長を計算する手法でどのような違いが生じるか確認する.

まず最初に, とともに従来手法であるが, 重みなし無向グラフに変換する方法 1 と, 重み付き無向グラフに変換する方法 2 で比較する.

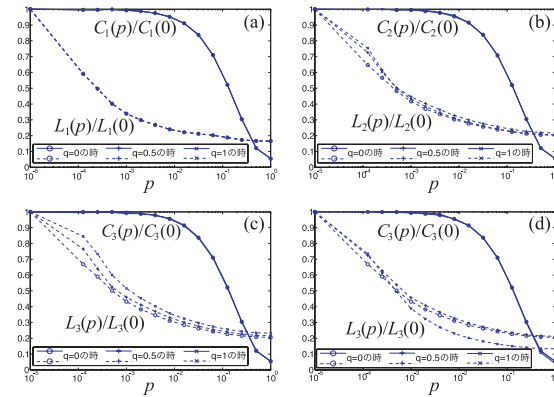


図 3 拡張 WS モデルにおけるスモールワールド現象の確認. (a) 方法 1, (b) 方法 2, (c) 方法 3 によって計算されたクラスタ係数と平均最短経路長を示している. (d) は 3.2 節で説明される非対称性を有する拡張 WS モデルに対して方法 3 を適用した結果である

Fig. 3 Estimated cluster coefficients and shortest path length of the extended WS models. Each result is calculated by (a) method 1, (b) method 2, and (c) method 3. Figure (d) shows the results in the case of asymmetric network models.

[方法 1] 各ノード間において単方向または双方向結合があれば, それを重みなし無向エッジで置き換えることで, 重みなし無向グラフに変換した. 変換後のグラフに対してクラスタ係数と平均最短経路長を算出した. 重みなしグラフなので, 各エッジの長さを 1 とした.

[方法 2] 各ノード間において結合強度の平均化 ($W_{ij} = (W_{i \rightarrow j} + W_{j \rightarrow i})/2$) を施すことにより有向グラフを重み付き無向グラフ化した. 変換後のグラフに対してクラスタ係数と平均最短経路長を算出した. 重み付きグラフなので, 各エッジの長さは結合強度 W_{ij} の逆数とした.

方法 1 で求めたクラスタ係数と平均最短経路長をそれぞれ C_1, L_1 と表記し, 方法 2 で求めた場合は C_2, L_2 と表記した. なお, 以後で用いるクラスタ係数および平均最短経路長は, グラフのエッジ数が異なる場合でも比較できるように, グラフを完全にランダム化したときの値で割って標準化している (ただし, 双方向結合の関係を保持したままグラフのランダム化を行っている).

まず, 図 2 のように拡張された WS モデルが, Watts らが考案したスモールワールドネットワークの定義を満たしているのかを確認する. 図 3 (a), (b) が示すように, 方法 1, 2 と

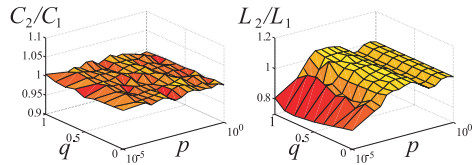


図4 方法1と方法2によるクラスタ係数および平均最短経路長の比較

Fig. 4 Comparing cluster coefficients and shortest path length by method 2 with those by method 1.

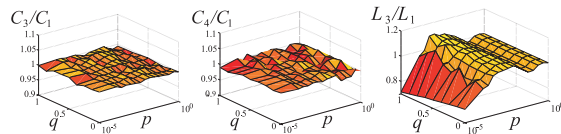


図5 図4と同様。ただし、方法1と方法3および方法1と方法4の比較。また、 L_3 は L_4 と等価である
Fig. 5 The same as Fig. 4, but comparisons between method 3 and method 1, and between method 4 and method 1. The L_3 is equivalent to L_4 .

もに q の値によらず高いクラスタ係数と小さい平均最短経路長を同時に満たしている p が存在する。よって、拡張された WS モデルはスモールワールドネットワークの定義を満たしている。

次に、方法1, 2における各指標の推定結果の違いを調べ、結果を図4に示す。平均最短経路に関しては違いを確認できるが、クラスタ係数に関してはグラフの重みを考慮しても、結果に影響しないことが分かる。

さらに、重み付き有向グラフを本提案指標で評価するために、下記の方法3を導入する。
[方法3] 重み付き有向グラフのまま平均最短経路長を計算し、式(9), (10)の提案指標を用いることでクラスタ係数を算出した。方法2と同様に重み付きグラフなので、平均最短経路を計算する際に、各エッジの長さは結合強度の逆数とした。

図3(c)より、提案手法を用いて拡張 WS モデルを解析してもスモールワールドネットワークの定義を損なわないことが確認できる。

さらに、方法1, 3における各指標の推定結果の違いを調べ、結果を図5に示す。この場合も同様に、平均最短経路長に関する違いは確認できるが、最も基礎的な方法1と比較したにも関わらず、クラスタ係数に関する違いは確認できない。なお、方法2と方法3を比較しても違いは現れなかった。

次に、Fagioloのクラスタ係数の性質を調べるために下記の方法4を導入する。

[方法4] 方法3と同様に平均最短経路を計算し、式(5), (6)の指標を用いることでクラスタ係数を算出した。

図5に示すように、この場合も先の結果と同様であり、結局、方法1~4において違いは確認されなかった。

以上の結果をふまえると、有向グラフのように複雑な形状のまま解析しても違いが見られないのなら、WSモデルのようにシンプルな重みなし無向グラフをベースにして解析した方が、ネットワークの本質を議論するうえで有効かもしれない。しかし次節では、有向グラフ解析が本質的に必要となる1要因として、ノード間の最短経路長の非対称性について紹介する。

3.2 ノード間距離の非対称性の導入

前節で提案した有向グラフのモデル(拡張 WS モデル)は、いずれもバイアスのない確率によって生成されていたので、各ノード間の行きと帰りの最短経路長がほぼ等しいという特徴を有している。つまり、マクロ的に見れば無向グラフのような性質を持つ。本研究では、この性質を最短経路長の対称性と呼び、非対称度:

$$\delta = \langle |l_{i \rightarrow j} - l_{j \rightarrow i}| \rangle \quad (12)$$

が、十分に小さいと定義する。ここで、 $l_{i \rightarrow j}$ はノード i からノード j への最短経路長を意味する。

次に、この最短経路の対称性を破るネットワークモデルとして、前節の拡張 WS モデル(図2)において、双方向結合の削除の仕方にバイアスを設ける。前節では、確率 q に応じて単方向結合にするかを決定し、さらに双方向結合のうち片方をランダムに選んで削除した。しかし、最短経路長の非対称性を持たせるために、以下のルールで双方向結合の片方を削除する。

$$\begin{cases} W_{i \rightarrow j} \rightarrow 0 & \text{if } j \geq i + \frac{N}{2}, \text{ and} \\ W_{j \rightarrow i} \rightarrow 0 & \text{if } j < i + \frac{N}{2}. \end{cases}$$

ここで結合強度が0とは結合がないことを意味し、 N は全ノード数である。つまり結合する相手が、インデックスの観点においてネットワーク全体の半分より近ければ、相手から受ける結合を削除し、逆にネットワーク全体の半分より遠ければ、相手へ送る結合を削除する。この操作により、円環状の有向グラフモデルにおいて反時計回りの結合が残りやすくなり、ノード間の最短経路長に非対称性を持たせることができる。その様子を図6に示す。 δ

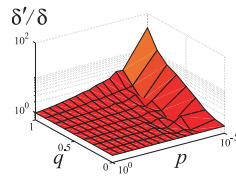


図 6 非対称化処理によって変化する最短経路長の対称性

Fig. 6 Asymmetric shortest path length between two nodes.

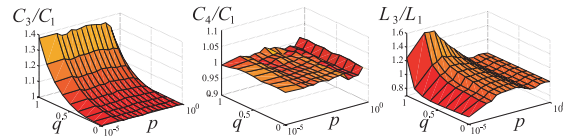


図 7 図 5 と同様。ただし、非対称性を有する拡張 WS モデルの結果

Fig. 7 The same as Fig. 5, but the case of asymmetric network models.

は非対称化処理を施していない拡張 WS モデルの非対称度 δ' は本節で紹介した非対称化処理を施した拡張 WS モデルの非対称度である。図 6 の縦軸 δ'/δ は、マクロ的な有向グラフ度と見なすことができる。結果のとおり $p \rightarrow$ 小かつ $q \rightarrow$ 大において $\delta'/\delta \gg 1$ なので、新しいネットワークモデルは確かに非対称性を有し、マクロ的にも有向グラフ的性質を有していることが分かる。

さて、この非対称性を有するネットワークモデルに対して、前節と同様に従来手法と提案手法で比較を行う。ただし、従来手法である方法 1 と方法 2 を用いた場合、無向グラフに変換する際にこの非対称性は消滅してしまうので、 C_1, C_2, L_1, L_2 の値は変化しない。よって、方法 1 と方法 2 の比較は省略する。

まず、図 3(d) に示すように、非対称性を有する拡張 WS においても、スモールワールドネットワークの定義を満たしていることが確認できる。さらに、図 7 において、方法 1 と方法 3, 4 の解析結果を比較する。解析結果から分かるように、平均最短経路長に関しては方法 1 との差異を確認できるが、方法 4 においてはクラスタ係数に関しては差異を確認できない。しかし、提案手法である方法 3 を用いた場合、明らかに異なるクラスタ係数を得ている。つまり無向グラフとは異なる特徴を抽出するには、提案手法を用いることで初めて可能になる。そこで次章では、実際のネットワークデータに対して本提案手法を適用し、これまで調査されてきた実ネットワークのスモールワールド現象について再考する。

表 1 実ネットワークデータでの解析結果。C.elegans と Political blogs のノード数は、それぞれ 297, 1,224 である。また、 L_3 は L_4 と等価である

Table 1 Results of real networks: C.elegans and Political blogs, whose numbers of nodes are 297 and 1,224, respectively.

	δ'/δ	$C_3/C_4/C_1$	L_3/L_1
C.elegans	140.9/95.4	17.9/7.1/6.4	2.5/1.0
Political blogs	0.74/0.67	21.3/19.4/16.1	1.2/1.1

4. 実ネットワークのスモールワールド現象の再考

現在、様々な実際のネットワークの結合の様子を記録したデータが Web 上で公開されている^{8),9)}。本研究は有向グラフ解析に着眼しているため、実データとして C.elegans (線虫の神経回路網) と Political blogs (アメリカ政治に関するブログの相互リンク網) を解析対象とした。前者は重み付き有向グラフで、後者は重みなし有向グラフで表現されている。まず、有向グラフの非対称性について検証する。実ネットワークを式 (12) に適用した結果を δ' とする。比較のために対称性を有するネットワークを用意する必要があるため、実ネットワークの双方向結合を保ったまま完全ランダム化することで作成した。そのネットワークを式 (12) に適用した結果を δ とする。 δ はランダム化に関する 30 回の試行平均値とした。表 1 に示すように、いずれのネットワークも $\delta' > \delta$ なので非対称性を有することが分かり、本提案手法がネットワークの特徴解析に必要であることが分かる。

そこで、従来の枠組みである方法 1, 4 と本提案手法である方法 3 を適用して解析結果を比較しつつ、実ネットワークのスモールワールド現象について再検討する。文献 1) では C.elegans に関する Watts らの解析結果は $C_1 = 5.6$ および $L_1 = 1.2$ であるが、おおむね表 1 の結果と同じである(若干の誤差は解析したデータが異なるためと思われる)。文献 1) より、スモールワールド現象は、クラスタ性が高い割にはノード間の最短経路長が小さいことによって定義されているので、スモールワールド現象の判断基準として、単純にクラスタ係数を平均最短経路長で割って方法 1 と方法 3 を比較すると、C.elegans では $6.4 \rightarrow 7.3$ に増大、Political blogs では $14.6 \rightarrow 17.8$ に増大することが分かる。しかし、方法 4 を用いた場合、前者は $6.4 \rightarrow 2.8$ 、後者は $14.6 \rightarrow 16.2$ となり減少することもありうる。スモールワールド現象の判断基準について議論の余地はあるが、本提案手法を用いることで、スモールワールド現象の見落としや過小評価を低減させる効果があるといえる。

5. 結論と今後の課題

本研究では、有向重み付きネットワーク解析においてグラフを簡略化せざるをえないという問題点をあげ、情報伝達を考慮することで新しい方法でクラスタ係数を計算する手法を提案した。さらに、その妥当性を検証するために、Wattsらが提案したWSモデルを重み付き有向グラフ用に拡張し、本提案手法を適用した。その際、拡張されたWSモデルが従来のスモールワールドネットワークの定義に反していないことを確認し、提案手法を用いてもスモールワールド現象を分析できることを確かめた。

さらに、本提案手法による解析結果と、従来のようなシンプルなネットワーク簡略化手法による解析結果を比較した。その結果、各ノード間の最短経路長に非対称性を有するネットワークにおいて本提案手法（方法3）が有効であることが分かり、従来の手法（方法1, 2, 4）では十分にネットワーク構造を特徴づけられないことを示した。また、線虫の神経網などの実ネットワークの解析を行ったところ、従来よりも顕著にスモールワールド現象を確認することができた。つまり、本提案手法はスモールワールド現象を検出するための感度を向上させ、スモールワールド現象の見落としや過小評価を低減させる効果が示唆される。さらに、実データは強い非対称性を有する場合が少なくない。また、本研究で提案したクラスタ係数 C_{TS} は、いずれのタイプのネットワークに対してもノンパラメトリックに適用可能である。よって、本提案手法（方法3）を用いることはつねに有用である。

今後の課題として、以下について検討を行う。

- 最短経路長の非対称性がどのようなメカニズムで高いクラスタ係数に起因するのかを明らかにする。
- この非対称性以外にも、提案手法の効果が現れる例が存在する可能性があるので、検討を行う。
- 非対称性など、特殊な性質を持つネットワーク上で起こる現象や、そのようなネットワークで構成されるシステムが見せる振舞いについて解析する。
- 実データへの応用例を増やし、各分野における新たな知見を模索する。
- C_{GF} と C_{TS} の違いは、グラフ理論的な一般性を持ってネットワークポロジを評価するか、情報伝達の仕方にも着目するかにあると述べたが、この着眼点の違いによって解析結果として差が生じるメカニズムについても詳しく分析する。
- 式(7), (8)の導出過程において、より厳密に情報伝達を定式化するために、親ノードと隣接する子ノードを同時にたくさん考慮することが期待される。その際、少しずつ複

雑な状況を考慮し定式化することで、どのような新しい知見が得られるのかを分析しつつ、提案手法の高度化を検討していく。

謝辞 本研究の着想は、これまでの埼玉大学池口徹教授との共同研究の中で芽生えました。本研究の重要性に気付かせていただき、また本研究をまとめられたことを深く感謝いたします。なお、本研究の一部は、文科省科研費若手研究（スタートアップ）(No.18800061)のご支援により行われました。

参考文献

- 1) Watts, D.J. and Strogatz, S.H.: *Nature*, Vol.393, pp.440–442 (1998).
- 2) Barabási, A.L. and Albert, R.: *Science*, Vol.286, pp.509–512 (1999).
- 3) Kalna, G. and Higham, D.J.: Submitted to *Adaptation in Artificial and Biological Systems (ALSB'06)*. http://www.maths.strath.ac.uk/~aas96106/rep03_2006.pdf (参照日 2007.12.15).
- 4) Fagiolo, G.: *Phys. Rev., E* Vol.76, 026107 (2007).
- 5) White, J.G., et al.: *Phil. Trans. R. Soc. London*, Vol.314, pp.1–340 (1986).
- 6) Adamic, L.A. and Glance, N.: *Proc. WWW-2005 Workshop on the Weblogging Ecosystem* (2005).
- 7) Fernandez, L.L., et al.: *Proc. 1st Intl. Workshop on Mining Software Repositories (MSR2004)*, pp.101–105 (2004). <http://opensource.mit.edu/papers/llopez-sna-short.pdf> (参照日 2007.12.15).
- 8) Newman, M.: *Network data*. <http://www-personal.umich.edu/~mejn/netdata/> (参照日 2007.12.15).
- 9) Collective Dynamics Group: *Network Resources – Data Sets*. <http://cdg.columbia.edu/cdg/datasets> (参照日 2007.12.15).

付 録

2.2節の提案手法の解説において、ノード間の相互作用が媒体を浸透させるケース、つまり送信ノードの媒体を受信ノードに繁殖させるケースについての説明を省略したが、本付録において詳細を述べる。この場合では、ノード間の結合強度を媒体の増加率で定義する。

手順1として、親ノード k からの1次結合を考える（図8のSTEP1）。ノード i が有する親ノード k に関する情報量（媒体量）を $I_i(k)$ と書くと、親ノード k がこれから繁殖させる情報量 $I_k(k)$ を α とする。最初、親ノード k 以外のノードは、 $I_k(k)$ に関する情報を得ていない。よって、最初は $I_k(k) = \alpha$, $I_i(k) = 0$, $I_j(k) = 0$ である。

次に、情報（媒体）が繁殖した後を考える。ノード i, j はそれぞれ浸透経路の結合強度

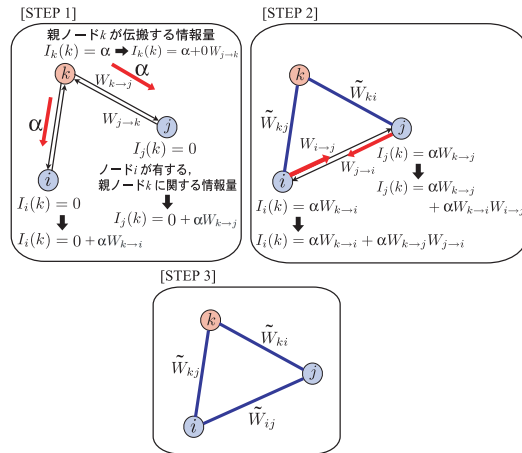


図 8 提案手法の模式図．ただし，ノード間の相互作用が媒体を繁殖させるケース

Fig. 8 Diagrams of the proposed method in the case of propagating information among nodes.

(増加率)に応じてノード k からの情報を受け, $I_i(k) = \alpha W_{k \rightarrow i}$, $I_j(k) = \alpha W_{k \rightarrow j}$ になる．そこで本研究では, ノード $k \rightleftharpoons i$ 間またはノード $k \rightleftharpoons j$ 間の双方向浸透を1束にまとめて考えることにより, いったん, 有向グラフを無向グラフ化する．本ケースでは, 情報量の増加率をネットワークの結合強度と見なすことより,

$$\tilde{W}_{jk}(= \tilde{W}_{kj}) = \frac{\{(0 + \alpha W_{k \rightarrow j}) + (\alpha + 0W_{j \rightarrow k})\} - (\alpha + 0)}{\alpha + 0} = W_{k \rightarrow j} \quad (13)$$

となる．同様に, $\tilde{W}_{ik}(= \tilde{W}_{ki}) = W_{k \rightarrow i}$ となる．逆に, $W_{k \rightarrow i} = 0$ ならば, $\tilde{W}_{ki} = \tilde{W}_{ik} = 0$ なので, 中心ノード k の近傍ノードとなる必要十分条件は, $W_{k \rightarrow i} \neq 0$ である．つまり, 2.2 節で紹介したケースと同じ条件が導かれる．

次に手順 2 として, 中心ノード k からの 2 次結合を考える (図 8 の STEP2)．ノード i, j 間の結合を経由し, ノード k に関する情報量は, $I_i(k) = \alpha W_{k \rightarrow i} + \alpha W_{k \rightarrow j} W_{j \rightarrow i}$, $I_j(k) = \alpha W_{k \rightarrow j} + \alpha W_{k \rightarrow i} W_{i \rightarrow j}$ に増加する．手順 1 と同様に, ノード k に関する情報量の増加率を考えると,

$$\begin{aligned} \tilde{W}_{ij}(= \tilde{W}_{ji}) &= \frac{1}{\alpha W_{k \rightarrow j} + \alpha W_{k \rightarrow i}} [(\alpha W_{k \rightarrow j} + \alpha W_{k \rightarrow i} W_{i \rightarrow j}) \\ &\quad + (\alpha W_{k \rightarrow i} + \alpha W_{k \rightarrow j} W_{j \rightarrow i}) - (\alpha W_{k \rightarrow j} + \alpha W_{k \rightarrow i})] \\ &= \frac{W_{k \rightarrow i} W_{i \rightarrow j} + W_{k \rightarrow j} W_{j \rightarrow i}}{W_{k \rightarrow j} + W_{k \rightarrow i}} \end{aligned} \quad (14)$$

となり, 式 (8) と同じ式が得られる．さらに, 以下の議論も 2.2 節とまったく同様である．

以上により, ネットワーク上での各ノードの相互作用を議論する際, 情報などの媒体を単に伝送するケースと, パーコレート (浸透) するケースに大別できる．しかし, 導出する過程は異なるものの, 最終的に同じ提案指標 (式 (9), (10)) に帰結する．

(平成 20 年 1 月 4 日受付)

(平成 20 年 6 月 17 日再受付)

(平成 20 年 10 月 14 日採録)



鈴木 智也 (正会員)

昭和 51 年生．平成 17 年東京理科大学大学院理学研究科物理学専攻博士課程修了．理学博士．同年東京電機大学工学部電子工学科助手, 平成 18 年より同志社大学理工学部情報システムデザイン学科専任講師, 現在に至る．複雑系, 時系列解析, 非線形現象等に関する研究に従事．電子情報通信学会, 日本物理学会, 人工知能学会各会員．