

複雑な重み付きネットワークの解析とモデル化

青 塚 恵^{†1} ホセ ナチエル^{†1}

実際のシステムの多くは、重み付きネットワークによって表現されている。重み付きネットワークは、頂点間の相互作用がどれ程存在するかを示す強さの変化を生みだしている。ここで、私たちは様々なコミュニケーションシステムの統計的な解析を表わし、重み付きネットワークの進化に対する一般的なモデルを提案する。また、私たちの提案するモデルは、実際のシステムで観測されたネットワークの特性によく似た、統計的な特性を表す重み付きグラフを生み出す。

Analysis and modeling of complex weighted networks

MEGUMI AOTSUKA and JOSE NACHER

Most of real systems can be described by weighted networks, where the interaction between nodes generates a variation of their strength depending on the interaction intensity. Here, we show a statistical analysis of several communication systems and present a general model for the evolution of weighted networks. The model generates weighted graphs showing some similar statistical properties observed in real systems.

1. はじめに

複雑系システムの基本構成は、相互に作用しあいネットワーク化された構造を生み出している。ローカルな相互作用と自己編成のルールは、ネットワークの進化だけでなく、スモールワールドやベキ乗則、そしてクラスタリングのようなモジュラリティなどのグローバルな現象の出現を導く¹⁾⁻³⁾。分子生物学から社会におけるコミュニケーションシステムに至るま

で広範囲の分野において、ネットワーク科学が規則的格子モデルとランダムグラフ理論だけでは実際のネットワークを説明することができないことを示している^{4),5)}。グラフ理論で利用されている平均場近似とレート方程式アプローチに基づく新しいアプローチは複雑ネットワークへ初期の理論的な基礎を証明することができているし、実験に基づいたデータを比較するときに予測能力を示す⁶⁾⁻⁸⁾。さらに、辺によって接続された頂点を特徴づける単純なグラフは相互作用が働いているかどうか、つまり2つの頂点に何らかの関係があるのかどうかという2進状態のみで表わされているのに対して、実際のネットワークでは、関係性の他にも複雑な現象が存在する。ネットワークをより複雑で現実的に表現したい場合、相互作用の強さを考える。

例えば、社会システムにおける科学者 i と j 間の相互作用は、共著論文の数によって重み付けされた辺 w_{ij} によって表すことができる。他にも、輸送システムにおいては、空港間の相互作用は最大交通量、つまり空港から空港へと移動する場合に、直行便によって移動可能な数を座席数として視覚化できる。同様に、各頂点の強度 s_i は頂点 i に接続されている重み付けされたすべての辺の重みの和によって計算できる。同様に、研究者の共著論文ネットワークではある研究者によって書かれた論文の数を強度とし、空港ネットワークにおいては、ある空港を通る合計交通量によってそれを示せる⁹⁾。この変数は2進数グラフにおける頂点の民主的な視点をなくし、ローカルダイナミクスな相互作用に基づいた頂点の重要性を確かなものとして強調している。様々な研究者達が実際のネットワークを重み付きグラフとして考え、すでに解析が行われているが、まだ重み付きネットワークによる特性の発見は2進数グラフによって示された多くの特性に追いついていない。言い換えると、固有の複雑性と変数の数の増加によって、実際のシステムを調査するために提案された理論モデルの数は2進数グラフを研究するために提案したものよりも少ないのである¹⁰⁾⁻¹²⁾。

他の多くの実際システムにおいて、相互作用による弱さと強さの不安定な密度の存在の影には重要な意味が含まれている。生態系の安定性という点においては、種や代謝経路、流動速度による心臓血管や呼吸器官システムの間に相互作用の強さに関わっているということが論じられてきた。記憶のメカニズムと経路の認識問題では、辺の重みが同定プロセスに対してネットワークを最適化するため、大部分はダイナミックに調整する必要があるニューラルネットワークに基づいている¹³⁾⁻¹⁵⁾。さらに、進化する重み付きネットワークを理論的で実験的に理解することは、多方面の分野に渡って広く重要な意味を持つ。

ここで、私たちが現在行っている複雑な重み付きネットワークの研究について説明する。私たちはコミュニケーションと輸送システムを表す様々な重み付きネットワークの統計的な解

^{†1} 公立はこだて未来大学
Future University Hakodate

析に焦点を当てている。2進状態のネットワークと重み付きネットワークのそれぞれに対して、次数と重みと強度の確率分布、強度と次数間の結合、そして、次数相関とクラスタ係数を解析した。同時に、私たちは重み付きネットワークの進化と形成に関する一般的なモデルを表わす。提案するモデルのパラメータを一致させるため、実際に行った統計的な解析結果を使用した。モデルはパラメータをフィットする統計的な解析から導き出される情報を使用することができる。現在のモデルは、実際のシステムにおいて観測されたものとよく似た統計的な特性を表わす重み付きグラフを発生させる。

2. 重み付き進化グラフに対するモデル

これから、私たちは重み付きネットワークの進化を調査するためにモデルを提案する。私たちのモデルは重み付きネットワークの位相的な特性を解析的に予測するために、確率的に頂点と辺が結合する2つの基本的な過程に基づいている。(1) 頂点の強度において優先的な結合をもっている。これは新しく追加された頂点 j が強度に比例した確率をもつ元の頂点 i に接続することを意味する。(2) 現存している頂点間で新しい相互作用の追加を持つ。この過程は重みに基づいて優先的な辺の選択を用いることによって実行される。優先的な辺の選択は(16)によって調査されたが、頂点に接続することができるのは新しく追加された頂点のみであった。逆に、ここでは元の頂点間に新しい辺を追加することにする。この過程は、新しく追加される辺の基点としてランダムに新しい頂点を選択することで実行される。頂点を選択した後、重み付けされた辺は重みに比例した確率で優先的に選択され、その辺の両端にある頂点に対して2本の辺を、ランダムに選ばれた頂点と繋げる。私たちはこの過程を internal weighted edge selection メカニズムと呼ぶ。(1)と(2)の両方の過程が確率的に結合するが、過程(1)にあたる新たに追加された頂点が元の頂点への強度によって優先的な選択をされる確率は $(1-p)$ としている。その時、確率 p では過程(2)となる元々存在していたネットワークの中で重みづけされた辺の選択が起こる。後者の過程は論文(11)で示されているように local weight-driven dynamics を引き継いでいる。優先的に選択されたそれぞれの辺は、定数因子 ρ によってその重みを増加させる。その上、ランダムに選択された頂点は2つの新しい辺に接続することになるので、追加された2辺の重みによってその強度を増加していく。

2.1 アルゴリズム

私たちは以下のようにアルゴリズムの詳細を説明する：
 m_0 の頂点からなる過程を初期値にセットする。各タイムステップ毎に確率 $1-p$ で新しい

頂点がシステムに加えられていく。一般的に、新しい頂点はすでに存在している m 個の頂点に接続させることができる。新しい頂点が頂点 i を選択する確率は頂点 i の強度 s_i に比例し、その時の確率を

$$\prod(s_i) = \frac{s_i}{\sum_j s_j} \quad (1)$$

と表す。

確率 p では、辺の内部追加を行う。ランダムに頂点 l を選択し、辺は優先的な選択 w_{ij} によって行われる。選ばれた頂点 l が選択された辺 (i,j) の両端にある頂点に結合される。重みは $w_{ij} \rightarrow w_{ij} + \rho$ によって更に強い結合へと変化する。選ばれた頂点 l もまた接続した2つの頂点によってその強度が上がる。

この過程をネットワークの頂点の数が十分に大きな数になるまで繰り返していく。

2.2 連続理論に基づく解析

ここからは、連続理論を用いた式の項において、上で述べた基本的な過程を書く。このモデルでは、3つの基本的な変数を持つ。各頂点の結合性 k_i 、その強度 s_i 、そして相互作用している頂点間の重み付けされた辺 w_{ij} である。私たちは変数が連続して変化するという仮定をする。したがって、 $\prod(s_i)$ は s_i が時間ごとに変化することで変化率を表す。初めに、時間発展 t に依存する次数や強さ、そして重みや頂点の総数の値を求めていく。まず、頂点の総数は確率 $(1-p)$ によって決定され、 $N(t) = (1-p)t$ という値をとる。強度の総数 $S(t)$ は以下の式

$$\sum_i s_i(t) = 2m(1-p)t + 2p(2+\rho)t \quad (2)$$

によって与えられる。そして合計の重みに対しては

$$\sum_{i>j} w_{ij}(t) = m(1-p)t + p(2+\rho)t. \quad (3)$$

と書き表すことができる。また、合計次数は

$$\sum_i k_i(t) = 2m(1-p)t + 4pt \quad (4)$$

のように計算できる。この式を用いて、連続理論における各変数の変化を表すことが可能である。初めに、 w_{ij} が変化する変化率を求める。私たちのアルゴリズムに従うことによって、確率 p で起こる過程(2)を通して元々存在していた頂点の重みに変化が起きていく。したがって、ネットワーク内の重みの変化率を以下のように、重み付けされた辺の優先的な選択

メカニズムを通して得ることができる：

$$\frac{dw_{ij}}{dt} = \frac{p\rho w_{ij}}{\sum_{i>j} w_{ij}} \quad (5)$$

この式は初期条件 $w_{ij}(t_{ij}) = 1$ を使って解くことができる．簡単にするため，今から w_{ij} の代わりに w と書くことにする．基本的な形である $w(t)$ を

$$w(t) = w(t_{ij}) \left(\frac{t}{t_i} \right)^{\frac{p\rho}{[q(2+\rho)+(1-p)m]}} \quad (6)$$

とすることにより，重み分布を計算することが可能である．計算をしてみると，

$$P(w) = w^{-\gamma_w} \quad (7)$$

が導かれる．この時の γ_w は

$$\gamma_w = \frac{p(2+\rho) + (1-p)m}{p\rho} + 1. \quad (8)$$

となっている．この結果は，重みの指数が確率 p と相互作用の強度の増加をしめす ρ と位相パラメータ m によって成長する過程へと結びつく 2つのパラメータに依存することを示している．次に，以下のように頂点 i の強さにおける変化率の値を求める：

$$\frac{ds_i}{dt} = \frac{m(1-p)s_i}{\sum_j s_j} + p \left[\frac{2}{N(t)} + \frac{(1+\rho)s_i}{\sum_{i>j} w_{ij}} \right] \quad (9)$$

第二項では $s_i = \sum_j w_{ij}$ を用いた．

上記のように求めることによって，初期条件 $s(t_i) = m$ を用い， $s_i(t)$ に対する式を決定する． $s_i(t)$ は

$$s_i(t) = \frac{1}{((1-p)B(p,m))} \left[\left(\frac{t}{t_i} \right)^{\frac{B(p,m)}{A(p,m)}} [(m(1-p)B(p,m) + 2pA(p,m))] - 2pA(p,m) \right] \quad (10)$$

と導ける．この時， $A(p,m)$ と $B(p,m)$ の値は

$$A(p,m) = 2m(1-p) + 2p(2+\rho) \quad (11)$$

$$B(p,m) = m(1-p) + 2p(1+\rho) \quad (12)$$

とする．

頂点の強度分布は

$$P(s) \propto \left[s + \frac{2pA(p,m)}{(1-p)B(p,m)} \right]^{-\gamma_s} \quad (13)$$

と導く．そして，

$$\gamma_s = \frac{A(p,m)}{B(p,m)} + 1 = \frac{2m(1-p) + 2(2+\rho)p}{m(1-p) + 2p(1+\rho)} + 1 \quad (14)$$

のように表す．

言い換えると，頂点 i の次数が

$$\frac{dk_i}{dt} = \frac{m(1-p)s_i}{\sum_j s_j} + p \left[\frac{2}{N(t)} + \frac{s_i}{\sum_{i>j} w_{ij}} \right] \quad (15)$$

のように変化していく様子を計算することができる．

同様の方法で，次数分布 $P(k)$ を計算できる．また，コンピュータシミュレーションによっても実行する予定である．

2.3 パラメータの推定

今までは，分布に対して基本形の主な解析を導き出してきた．しかしながら，私たちが解析した各実際ネットワークの確率パラメータ p の値をもた決定することができる．モデルから，

$$\langle k \rangle = \frac{1}{N} \sum_i k_i \quad (16)$$

として平均次数を計算できる．モデルにおいて，頂点の合計数は $N(t) = (1-p)t$ と $\sum_i k_i = 2m(1-p)t + 4pt$ に従い，そしてそれは以下のような式

$$p = \frac{\langle k \rangle - 2m}{\langle k \rangle - 2m + 4} \quad (17)$$

を導く．この式から実験的なデータから計算したネットワークの平均次数 $\langle k \rangle$ を用いて各ネットワークに対して各確率 p の識別をする．現在行っている研究の次の段階として実際のネットワークにおいてこれら 2つの過程の発現率を特定するのにこの式を使うことができると期待している．

3. 統計的な分析

3.1 ネットワークデータ

ここで，いくつかの実際の重み付きネットワークに対する解析結果を示す．私たちはアメリカ合衆国の空港ネットワーク，Condensed matter に提出された原稿を書いた研究者たちから成るネットワーク，そして Kiel 大学のメールサーバーのログファイルから構成された

email ネットワークを解析している。また、これ以降は研究者間のネットワークは、共著論文ネットワークとする。各ネットワークの頂点、次数、そして重みは表 1 に示した。

世界中の空港ネットワークが研究者たちによって研究されている^{9),12),17)}。その中で私たちが使用したデータセットは、直行便が通っている空港間でのその座席数によって重みを付けている。一方で、共著論文ネットワークは頂点を研究者とし、4 年間の間に共著論文を何度執筆したかによって重み付けしている。このネットワークはすでに解析されているが⁹⁾、図を描画し、私たちのモデルに必要なパラメータを得るために解析を行った。更にコミュニケーションネットワークにあたる email ネットワークでは、112 日間の間に大学の生徒と外部の家族や友達などによって送受信された全ての発信元と宛先を記録したものを使用している¹⁸⁾。次数分布はすでに研究されているが、今回は重み付きネットワークの特性を調べるために、アドレスや送受信回数などを抜き出し、データを加工した。個々の間でメールを何度送りあったかということを email ネットワークの重みとした。

3.2 ネットワーク解析結果

この節では、上記のネットワークにおける解析結果を示す。図 1 から、ネットワークが一般的にべき乗則か、少なくとも 2 桁かそれ以上の桁数伸びている尾部の重いものによく近似している。その結果は email ネットワークのようなコミュニケーションネットワークが頂点の次数から相互作用の強度や重みに至るまで、他のコミュニケーションネットワークや大規模輸送システムと同様の特性をもつことを示している図 2。更に、重みが付いているデータとそうでないデータに対してクラスタ係数と次数相関を計算した。図 3 では、email ネットワークの重みの有無によって別々にデータの解析結果を示した。クラスタ係数は共に k^{-1} に近似している。また、次数相関においては共著論文ネットワークが assortative mixing の振る舞いが見られたのに対して⁹⁾、正反対である disassortative mixing が観測された。更に、文献 9) の共著論文ネットワークに関して述べれば、重み付きネットワークでのその振る舞いは更に強まっていた。最後に、私たちは図 4 に示したように、強度と次数間における結合を調べた。email ネットワークは線形結合 $s(k) \propto k$ に特徴付けられ、その結果は共

著論文ネットワークに対しても一致している。輸送ネットワークである世界中の空港ネットワークとアメリカ合衆国の空港ネットワークは非線形結合 $s(k) \propto k^{1.5}$ によって特徴付けられている。

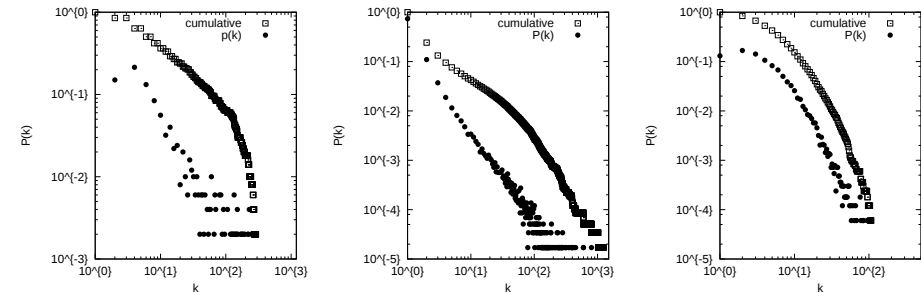


図 1 ネットワークの次数分布
Fig. 1 probability distribution of network

4. 考察と結論

この研究では、コミュニケーションと輸送システムに相当する様々な重み付きネットワークを解析した。その解析において、多くのネットワークが広範囲のべき乗則の分布を示し、相関関係と同様、高いクラスタ性を持っている。しかしながら、これらのネットワーク間で目立った違いを示している。強度と頂点の次数の相関関係を見てみると、アメリカ合衆国の空港ネットワークは非線形に連結していることによって特徴づけられている。逆に、コミュニケーションシステムは線形な結合を表わしている。分析的な解析をモデル化し導き出す場合、強度における非線形結合を示すのは困難である。

非線形結合は $\rho - \rho_w$ に置き換えることによってコンピュータシミュレーションを取り入れて実行することができた。近年の研究では、この結合が確率的な過程の結果になるかもしれないと提案され、この問題に取り組んでいる。特に、乗法過程ではこのように観測された非線形結合を形成することができたと論じられている。言い換えれば、私たちはある確率 q に従ってリンクを繋ぎ直すこともできる。ただ、輸送システムでは確かに可能である一方

表 1 各ネットワークの頂点と辺と重みの関係
Table 1 Relation of values in each network

	頂点	辺	重み
空港ネットワーク	空港	航路	座席数
e-mail ネットワーク	ユーザー	メールの送受信	送受信回数
共著論文ネットワーク	研究者	共著論文	共著論文数

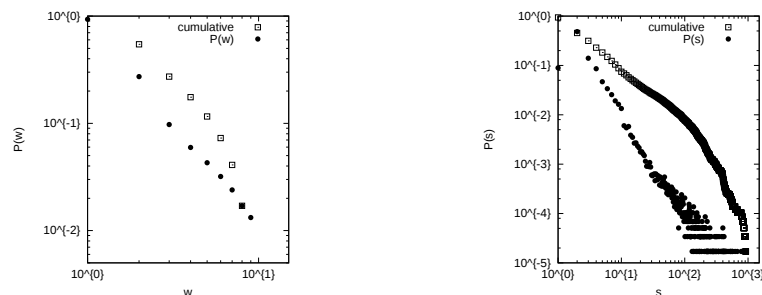


図 2 email ネットワークの重みと強度における分布
Fig.2 Weight and strength distribution of email network

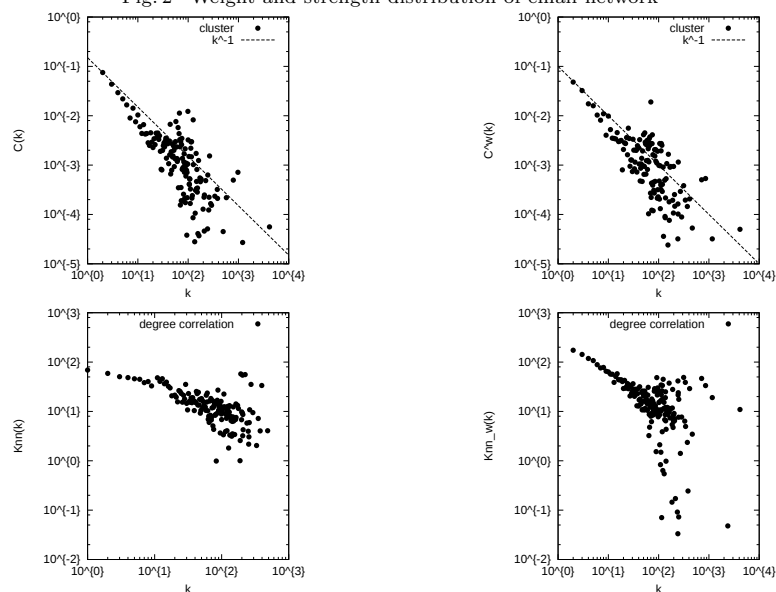


図 3 email ネットワークのクラスタ係数と次数相関 (左: 重みの無いネットワーク, 右: 重み付きネットワーク)
Fig.3 Clustering coefficient and degree correlation of email network (left: non-weighted network, right: weighted network)

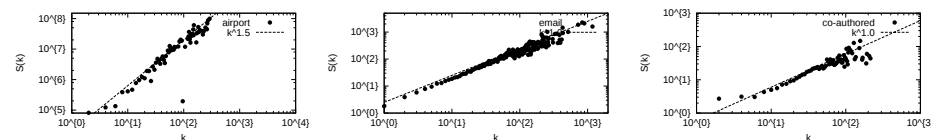


図 4 強度と次数間の結合
Fig.4 The coupling between the strength and degree

で、送受信する email のようなコミュニケーションシステムにおいて繋ぎ換えは必ずしも行えるとは言えない。なぜなら、メールを送受信した瞬間のみ辺が繋がるだけで、形跡しか残らないからである。非線形結合と繋ぎ換え過程は輸送システムをより現実的に表すことができ、モデルの将来的な拡張が期待される。ここで、私たちは確率分布に焦点を当てる。しかしながら、同様のアプローチを用いることで私たちはクラスタリング次数に対する式もまた表現できる。確率 $(1-p)$ によって進化していく最初の過程が高いクラスタ係数を発生させる可能性が低い一方で、2 番目のメカニズムはネットワークのクラスタ係数を増加させるような頂点の内部優先的選択に基づいている。したがって、このクラスタ係数の提案はネットワークを特徴づけるパラメータ p によって左右される可能性がある。私たちは更なるコミュニケーションシステムにおいてその基本形の解析的な導出を含める予定である。モデルを解析結果と比較するために、モデルのシミュレーションを実行し、より洗練されたメカニズムを持つモデルとして拡張させる。また、日本の空港と鉄道システムの時間発展に基づくデータ解析を行い、このシステムにおける非線形結合の存在を確かめるつもりである。

5. 謝 辞

著者（青塚恵）は、公立はこだて未来大学特別研究費重要区分（平成 21 年度）の助成を受けている。ご支援して下さい関係者の方々に深く感謝の意を述べる。

参 考 文 献

- 1) D.J. Watts and D.H. Strogats, Nature (London) 393, 440(1998).
- 2) A.-L. Barabási and R. Albert, Science 286, 509(1999).
- 3) E. Ravasz et al., Hierarchical Organization of Modularity in Metabolic Networks, Science, Vol. 297, no. 5586, pp. 1551-1555
- 4) A. Barrat, M. Barthélemy and A. Vespignani, Dynamical processes on complex networks, (University Press, Cambridge 2008).

- 5) M. E. J. Newman, A.-L. Barabási and D.J. Watts, The structure and dynamics of complex networks, (University Press, Princeton 2007).
- 6) G. Caldarelli, Scale-free networks: Complex webs in nature and technology, (University Press, Oxford 2007).
- 7) S.N. Dorogovtsev and J.F.F. Mendes, Evolusion of networks: From biological nets to the Internet and WWW. (University Press, Oxford 2003)
- 8) A.-L. Barabási and Z.N. Oltvai, Nat. Rev. Genetics, 5, 101 (2004).
- 9) A. Barrat, M. Barthelemy, R. Pastor-Satorras and V. Vespigani, The architecture of complex weighted networks, PNAS 101, 3747(2004).
- 10) A. Barrat, M. Barthelemy, and V. Vespigani, Weighted evolving networks: Coupling topology and dynamics, Phy. Rev. Lett. 227701(2004).
- 11) Z. Pan and S. Wang, Generalized local-world model for weighted networks, Phys. Rev. E. 056109(2006).
- 12) A. Gatreaua, Al. Barrat and M. Barthelemy, Microdynamics in stationary complex networks, PNAS 106, 8847(2009).
- 13) E.L. Berlow, Nature (London) 398, 330 (1999).
- 14) J.R. Banavar, A. Maritan and A. Rinaldo, Nature (London) 299, 130 (1999).
- 15) B. Muller and J. Reinhardt, Neural Networks: An introduction (springer-verlag, Berlin, 1991).
- 16) S.N. Dorgovtsev and J.F.F. Mendes. Evolved weighted scale-free networks. In Proc. Science of Complex Networks: From Biology to the Internet and WWW CNET 2003, aveiro, Portugal. AIP Conf. Proc. 776.
- 17) G. Bagler, Analysis of the airport network of India as a complex weighted network, Physica A, 387, 2972(2008).
- 18) H. Ebel, et al., Scale-free topology of email networks, Phys. Rev. E 66, 935103(2002).
- 19) Stanley, M. H. R. et al., Nature 379, 804(1996).