

輸送システムとコミュニケーションシステムにおける 顕著な重み付きネットワーク

青塚 恵^{†1} ホセ ナチエル^{†1}

近年の研究では非常に強く結びついた頂点が rich-club として知られる現象とその重要性同様に、ネットワークに存在する大部分の資産がどのように相互作用し交換しているかを取り上げている。ここで、私たちはコミュニケーションと輸送システムにおける rich-club の構造と効率を研究した。初めに、重みの付いた相互作用を用いることによって、eメール、バス、ニューラルネットワークに対する rich-club がもし共著論文や空港のネットワークのものと比較した場合になんらかの特性や独特なパターンを表すことを示す。次に、これらの顕著なサブネットワークがネットワークの資産を効率的に制御しているのかどうかを研究した。私たちの解析では、空港のネットワークが解析を行ったネットワークの中で最も効率の高い rich-club の核を持つことが示された。

Prominent Weighted Networks in Transportation and Communication Systems

MEGUMI AOTSUKA and JOSE NACHER

Recent studies have highlighted how highly connected nodes interact and exchange among themselves the majority of networks resources, a phenomenon known as rich-club, as well as its importance. Here, we investigate the rich-club structure and efficiency in communication and transportation weighted networks. First, by using weighted interactions, we show that the rich-club for email, bus and neural networks exhibits characteristic and unique patterns if compared to those from scientific collaborations and airport networks. Secondly, we investigate whether these prominent sub-networks efficiently control the network resources. Our analysis shows that the airport network has a rich-club core with highest efficiency among the analyzed networks.

1. はじめに

ネットワーク科学における近年の発達によって、実際の複雑なシステムの多くがネットワークによって表され、共通の構造的な性質を表わしていることが明らかになっている。ネットワーク研究は基本構成要素と同様にスケールフリーの性質や階層構造のみならず局所的な興味深い相互作用パターンとコミュニティの出現においても取り上げている¹⁾⁻³⁾。さらに、重み付き相互作用の存在を含めることによって、システムのお大半はネットワーク特性の綿密で正確な研究を可能にした⁴⁾。

近年の研究において、実際のシステムはランダムなものと比較して頂点間の接続がより強く結びついている傾向があると研究され、その特性は rich-club と呼ばれる。また、様々な研究により rich-club が影響した異なる性質の出現も示された。重みの付かない空港や共著論文ネットワークでは rich-club が観測され、タンパク質相互作用ネットワークでは rich-club サブグラフの頂点が強く結びついていないという興味深い結果となった⁵⁾。さらに、それらのネットワークに実際のシステムの相互作用の強さを与えることによって、研究者たちは重み付きネットワークの領域に rich-club の概念を広げた。重み付きネットワークにおいて多極化した特性⁶⁾の他に際立った制御特性⁷⁾もまた観察され、研究された。特に、ネットワーク全体において頂点同士が強く結びついて多くの情報を交換したり流れを提供したりする傾向があるということが研究された。この特性を表す要素は顕著な頂点の集合として定義された。

この研究では、私たちは初めに輸送や信号処理や社会コミュニケーションの分野での様々な実際のネットワークにおける重み付き rich-club の影響を計算した。次数と強度に対して豊かさを定義し、使用することによって重み付きネットワークを計算している。結果として共著論文ネットワークやアメリカ合衆国の空港ネットワークのような社会コミュニケーションと比較した場合、大学のメール送受信ネットワーク、函館市近郊のバスネットワーク、そして *C.elegans* ニューラルネットワークは特徴的なパターンを表している。この解析はある際立った頂点がシステムの資産や情報の流れの多くを交換し制御していることによってどのように集合を形成しているかを私たちに教える一方で、これらの突出した集合がネットワー

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

クの資産を効率的に制御できているのかどうかはまだ不明瞭なままである。この疑問に答えるために、確認された際立った頂点とそのネットワークの理想構造に一致するものの両方を計算する。その結果、全ての解析したネットワークの中で、アメリカ合衆国の空港ネットワークが圧倒的に資産を分かち合っている顕著な要素がより効率的に接続されたシステムである。逆に、C.elegans ニューラルネットワークの rich-club はネットワーク全体と比較してほんの少しだけ効率的な向上が見られる。さらに、eメールシステムと C.elegans システムは似たような効率の値を示した。それらの効率の割合はどちらかという完全ネットワークに対してほんの少しの向上が見られるかほぼ違いないように思える。さらに私たちの研究では重み付き rich-club のグラフ構造の影響に対する新しい見識を与え、高い効率性を持つネットワークサブグラフの実際の例を証明する。

原稿は以下のように編成する。2章ではこの研究に用いたデータセットを紹介する。3章では方法とネットワークの指標を示す。そして4章で結果を、5章で研究結果の考察と結論を述べる。

2. データセット

私たちはコミュニケーションと輸送システムという異なる分野から様々な重み付きネットワークの解析を行った。ここで、各ネットワークに対する説明をする。アメリカ合衆国の空港ネットワークは空港を繋ぐ日に 5960 のフライト (辺) をもつ 500 の空港 (頂点) のネットワークである。ここでは、重み付きネットワークを考えているので、各 1 日のフライトは同じ出発地や目的地における異なる様々な便から成る。各辺の重みは 2 つの空港を行き来する乗客の合計数である。C.elegans ニューラルネットワークは 297 のニューロンが 3358 のシナプスやギャップの集合によって接続されている。相互作用の強さはニューロンの与えられたペア間のシナプスやギャップの数によって決定されている。キール大学のデータセットである eメールシステムは重み付きネットワークを明らかにする為にダウンロードし、この研究のために加工した。ユーザ (送信者) の数は 58339 人であり、生徒はもちろん、大学外にいる両親なども含んでいる。174200 通のメールのやり取りによって接続されている。重み付き相互作用は 2 人のユーザーがメールを送受信した回数を計算することによって構成されている。共著論文ネットワークではもし少なくとも 1 度はある研究を共同で行っていた場合に 95188 の双方向の辺が繋がれ、16726 の頂点 (研究者) から成る。研究者同士の相互作用の強さは共同で研究を行った回数を求めることによって形成されており、著者の研究チームや共著の行動性を浮き出している。最後に、函館市近郊のバスネットワークデー

タは 425 の頂点 (バス停) と 972 の辺を持ち、バス停間の連結で成り立っている。重みはバス停間の物理的な距離で表している。

これらの重み付きネットワークのネットワーク統計として平均次数や重み、次数、強度の最大値を表 1 でまとめている。

表 1 ネットワークのデータ

	平均次数	最大重み	最大次数	最大強度
空港ネットワーク	11.94188	2253992	145	98632722
C.elegans ニューロンネットワーク	11.3064	70	48	1700
e-mail ネットワーク	2.903382	9	6553	8353
共著論文ネットワーク	3.172986	22	214	190
函館市近郊のバスネットワーク	1.947896	10	6	19.2619

3. ネットワークの指標

3.1 次数や強度に対する顕著な頂点の豊かさ

重み付きグラフや重みのないグラフにおける rich-club を研究するために様々な手法が提案されているが、ここではシステムの多様な資産を分けあう顕著な頂点を分裂して求める手法を使用する⁷⁾。式は以下のように導かれる；

$$\phi^w(r) = \frac{W_{>r}}{\sum_{k=1}^{E>r} w_k^{rank}} \quad (1)$$

E はすべてのネットワークにおける辺の合計数であり、 $k = 1, \dots, E$ である w_k^{rank} は辺におけるランク付けした重みである。この式は分け合うことのできる重みの全ての量と比較した rich-club の頂点によって制御された重みの割合を得る。このような場合には、ネットワークの最も強い強度によってリンク付けされている。豊かさの値 r (次数 k や強度 s に対して考えることができる) を持つ頂点は、クラブメンバーに接続された辺の数 $E_{>r}$ と重み $W_{>r}$ の値を共に持ち、ある与えられた値 r よりも高い辺や重みのことである。重み付けされた rich-club 影響の存在を算定するために、同じネットワークのランダムバージョンと比較する必要がある。ランダムネットワークを作成する場合に、ヌルモデルは重みに対するスイッチングアルゴリズムを使うことを提案している。ランダムに選ばれた 2 辺が選ばれて重みが手順を入れ替えることによって交換されるアルゴリズムである。辺の重みのみを交換する方法なので次数自体に影響はなく、次数分布はこの過程では変化しないままである。

ここで、豊かさ r の式を導く；

$$\rho^w(r) = \frac{\phi^w(r)}{\phi_{switch}^w(r)} \quad (2)$$

この値が 1 より大きかった時ランダムモデルにおいて分け合っているものと比較して多くの情報の流れを制御する頂点の強い顕著な集合が存在していることを意味する。

3.2 ネットワークの効率の指標

重み付きグラフにおける最短経路の定義は複雑な重み付きネットワークにおけるネットワーク効率を正確に定義するために重要な役割を果たす^{8),9)}。最短経路の距離 d_{ij}^w は i から j における全ての可能な経路の合計は始めから終わりまでの距離の最も小さい和として定義することができる。また、 d_{ij}^w は l_{ij}^w を用いて求めることができる。さらに 2 点の頂点間のネットワークの効率性は $\epsilon_{ij} = 1/d_{ij}^w$ として定義される。繋がっていない頂点は $\epsilon_{ij} = 0$ という結果になる。すべてのグラフ G について考えてみると、重み付きネットワークの効率は以下のように平均化できる；

$$E^w(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in G} \frac{1}{d_{ij}^w} \quad (3)$$

このときの N はネットワークの全ての頂点数である。0 から 1 の間で効率的な値を持たせるためにこの式を正規化するために、グラフの理想効率を考えることができる。この理想的なグラフの接続形態はネットワークの完全グラフに等しく一致している。理想グラフの重みに対しては、全ての頂点の組み合わせ間の距離は重み付き辺が存在している辺の最短の長さに等しい。ネットワーク効率性に対する正規化した式は

$$E_{norm}^w = \frac{E^w(G)}{E^w(G_{ideal})} \quad (4)$$

と導かれる。重み付き距離の定義に対する具体的な設定は結果の節で説明する。

4. 結果と考察

この研究では、輸送システム（函館市のバスネットワークやアメリカの空港ネットワーク）、生物的信号処理システム（C.elegans ニューラルネットワーク）、そして社会コミュニケーションシステム（eメールシステムと共著論文）の分野から成るネットワークにおける重み付き rich-club の影響を解析し、ネットワーク全体との比較を行った。その結果を以下に述べていく。

共著論文ネットワークに対しては、 $r=k$ のとき、多くの共同研究者と研究を行っている活

動的な著者たちは、同じように共同研究者を多く持つ研究者と研究を行う傾向があるが、その割合はランダムに予想していたものと変化しない。強度 s では、科学的な生産性（論文の執筆数）に基づいていて、活動的な研究者はより活動的な研究者と共同執筆しがちである（図 1）。

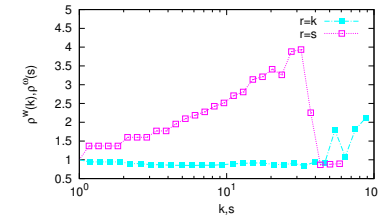


図 1 $r = k, s$ における共著論文ネットワークの rich-club

一方、 $r = k$ として考えたときの e メールコミュニケーションシステムは、互いにそれほど強くなく結び合った集合が見られる。更に、連絡をしあうとき、ランダムにむすびついたものよりも強い辺を発達させる。強度 s に対しては、非常に連絡を取り合うような e メールユーザは同じユーザと連絡をする機会はランダムにえられた時よりもより頻繁に互いに連絡をとる傾向がある（図 2, 図 3）。

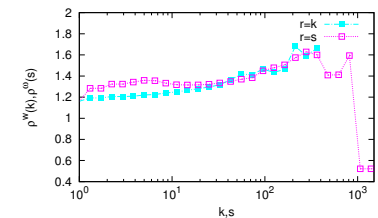


図 2 $r = k, s$ における e メールネットワークの rich-club の影響

重み付き rich-club に対する C.elegans ニューラルネットワークの解析もまた特有のパターンを示している。 $r = s, r = k$ の両方に対して、互いに非常に強く結びあった明らかなハブ



図 3 $r = k, s$ における email ネットワークのサブグラフ

頂点の集まりが見える。それらは密に集まって形成している。さらに、相互作用の強度はランダムによって予想されたものよりも高い。また、 $r = s$ の時、サブグラフの中に高いクラスタ性を持つ頂点の他に、強度の非常に大きい頂点が他の強度の弱い頂点と繋がっている集合も存在し、2つの性質を併せ持っている (図 4, 図 5)。

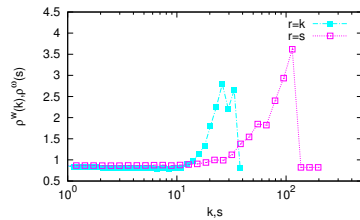


図 4 $r = k, s$ における C.elegans ネットワークの rich-club の影響



図 5 $r = k, s$ における C.elegans ネットワークのサブグラフ

空港ネットワークの重み付き rich-club の解析において、豊かさの変化は、頂点の次数が

k であるとき、ハブ空港間の接続が強い旅行の流れを示し、ほぼ完全なクラスタ構造をとるということを示している。同様に、強度 s に対してもまた、ランダムネットワークと比較した場合に、旅行者に多く利用される空港は大きな旅行者の流れを分けあっている。それらはまた高いクラスタ次数を示している (図 6, 図 7)。

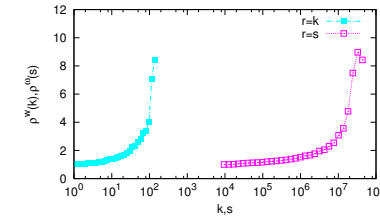


図 6 $r = k, s$ における空港ネットワークの rich-club の影響

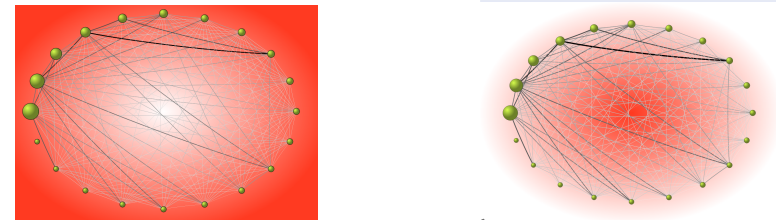


図 7 $r = k, s$ における空港ネットワークのサブグラフ

函館市近郊のバスネットワークについてもこの研究で解析を行った。このネットワークは、以前に解析したネットワークとは対照的に、スケールフリー分布に従っていない。物理学的なネットワークはバスネットワークのデザインを抑制している。その上、次数の範囲が表 1 で示したようにとても制限されている。従って $r = s$ に対する結果のみを示すことにする。強度 s に対する豊かさは減らされているが、ハブ頂点でのバス停の集合はバス停がランダムにおかれた場合よりも高い確率で最短距離によって強く接続されている (図 8, 図 9)。

私たちは上記で示したように輸送ネットワークとコミュニケーションネットワークを解析

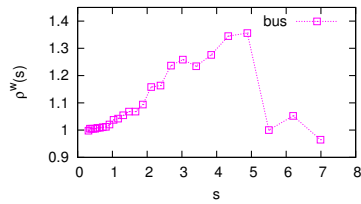


図 8 $r = s$ におけるバスネットワークの rich-club の影響



図 9 バスネットワークのサブグラフ

することで、ネットワークの特有のパターンを特定することが出来た。しかし、ハブ頂点で形成される集合のネットワーク効率におけるいくつかの疑問がまだ残ったままである。特に、これらのサブネットワークはシステム全体における性能と比較したときに高い効率性を示すのだろうか。重み付けされた rich-club の影響を持つ格サブグラフの存在に基づいてネットワークをデザインすべきなのだろうか。これらの疑問に答えるために、サブグラフとネットワーク全体におけるネットワーク効率をそれぞれ求めた。

これらの重み付きネットワークにおいて、 i から j へ向かうネットワークの全ての可能な経路の中から最短の距離の合計として定義した最短経路の長さ l_{ij}^s を計算した。各経路に割り当てる重みは各ネットワークの特性によって定義した。例えば、もし重み（可能なフライト数と座席数）の値が大きければ 2 点間の空港の相互作用は強いということが言える。移動量の流れの多さは高い効率性を示すので、 d_{ij}^s は i と j 間の重みの逆数 $l_{ij}^s = 1/w_{ij}$ として指定する。この考え方は e メールシステムやニューラルネットワークシステム、共著論文システムにも適用している。2 人のユーザ間での eメールの送受信の回数というのはユーザ同士が親しい場合や身近な存在である関係を示している。またニューロン間のマルチプルエッジ（シナプス）は強い相互作用を示している。活動的な研究者は科学的な関係によって論文の強さを分けあう。結果として、密接に接続された頂点は情報伝達において高い効率性を示すということになる。

全体のネットワークの結果は表 2 に、rich-club に対する結果は表 3 にそれぞれまとめた。全ての解析したシステムの中で、空港ネットワークは rich-club の頂点に対して最も高い効率性を持つように見える。 $r = k$ のときは 27%， $r = s$ のときは 30%を示している。つまりこれらのサブグラフは空港ネットワークの理想的な大規模グラフ構造による効率の約 27%と 30%に達しているということを意味している。しかしながら、ネットワーク全体の効率を計

算した場合、効率は 2%減ってしまっている。逆に残りのネットワークは rich-club のサブグラフに対して微妙な効率の向上を示している。ニューラルネットワークに対するネットワーク効率は $r = k$ と $r = s$ の各々、サブグラフの効率は全体のネットワークに比べて 2 倍高くなり、3%と 4%に達している。メールコミュニケーションシステムの効率は $r = k, s$ の時共に約 5%になっている。全てのシステムの中で、共著論文ネットワークが rich-club において 1%よりも遥かに低い効率の値を持ち、最も効率性が悪いように見える。研究者たちは論文執筆に参加するけれども、大部分の割合が全体のネットワークの効率性を下げて分裂するかのように研究を行う傾向がある。一般的に、ネットワーク全体とサブグラフの結果を比較した場合には、私たちの研究では表 3 で表したようにクラブメンバーの数が増加したときに効率性が下がることを示している。

5. 結 論

この研究では、コミュニケーションと輸送システムの分野からネットワークを解析し、重み付き rich-club 影響もその効率性も計算した。結果は近年解析された eメール、ニューラルネットワーク、そしてバスネットワークは共著論文や空港ネットワークと比較した場合に特有の性質を示す。コアサブグラフの理想構造は情報の流れとシステムの資産の制御における新しい見識を導く。その一方で、効率測定の計算は空港ネットワークがとても高い効率性（完全グラフの効率と比較して 15 倍）をもつ rich-club の核を持っていることを示している。逆に、C.elegans サブネットワークは完全ネットワークに対してほんの少し効率の向上が見られたり、同じような効率である。興味深いことに、eメールシステムと C.elegans ニューラルネットワークに対する rich-club 効率は 3-5%間という似かよった結果を示している。しかしながら、空港システムの効率は 30%にも達している。更に、競争メカニズムが存在することによって現在と異なる新たな効率というものが生み出されるかもしれない。なぜなら、空港のデータセットを考えてみると、便を提供している全ての会社が他社との競争するために毎年旅行客の需要や移動の流れに基づいてよりよい運行状況へとアップロードを行うからである。また、高いクラスタ次数の出現といったものも残りのネットワークで観察されていないサブグラフの特有な性質を表す。空港ネットワークが高いクラスタ性を持つことから、高いクラスタ次数（三角形など密に繋がっていること）を持つ重み付き rich-club の組み合わせは大部分のネットワークの流れをコントロールし、全てのネットワークパフォーマンスを向上させる高い効率性を持つ核のサブグラフを導くかもしれない。将来的な研究として実際のネットワークを考える上で、観測された効率性とネットワークコストについての

関係性もまた研究してみる必要がある。

表 2 ネットワーク全体の効率

	E	E_{id}	E/E_{id}
空港ネットワーク	0.040659	2.253994	0.018039
C.elegans ニューロンネットワーク	0.974814	69.9987	0.013926

表 3 サブグラフの効率

	次数 20	次数 50	強度 20	強度 50
空港ネットワーク	0.27292	0.184414	0.304797	0.196788
C.elegans ニューロンネットワーク	0.02794	0.023359	0.038829	0.017717
e-mail ネットワーク	0.046757	0.035521	0.052581	0.036015
共著論文ネットワーク	0.005621	0.004257	0.003461	0.003163

6. 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 22 年度）の助成を受けている。ご支援して下さっている関係者の方々に深く感謝の意を述べる。

参 考 文 献

- 1) R. Albert and A-L. Barabasi, Rev. Mod. Phys. 74, 47 (2002).
- 2) G. Caldarelli, Scale-free networks: Complex webs in Nature and Technology (Oxford University, New York, 2007).
- 3) A. Barrat, M. Barthelemy and A. Vespignani, Dynamical processes on complex networks (Cambridge University, Cambridge, 2008).
- 4) A. Barrat, M. Barthelemy, R. Pastor-Satorras and V. Vespignani, Proc. Natl. Acad. Sci. USA, 101, 3747.
- 5) V. Colizza, A. Flammini, M.A. Serrano and A. Vespignani, Nature Phys. 2, 110 (2006).
- 6) M. A. Serrano, Phys. Rev. E, 78, 026101 (2008).
- 7) T. Opsal, V. Colizza, P. Panzarasa and J.J. Ramasco, Phys. Rev. Lett, 101, 168702 (2008).
- 8) V. Latora and M. Marchiori, Phys. Rev. Lett., 87, 198701 (2001).
- 9) I. Vragovic, E. Louis and A. Diaz-Guilera, Phys. Rev.E 71, 037122 (2005).