

# 研究協力ネットワークにおける役割から見た 研究活動の多様性：共著ネットワークの分析を通して

芳鐘 冬樹 (大学評価・学位授与機構 評価研究部, E-Mail: fuyuki@niad.ac.jp)  
辻 慶太 (国立情報学研究所 人間・社会情報研究系, E-Mail: keita@nii.ac.jp)

## Abstract

研究コミュニティのコミュニケーション構造の解明を目的として、共著ネットワークの分析を行った。計算機科学分野を調査対象とし、分野に属す研究者各々の研究協力ネットワークにおける中心性を、第1著者としての活性度と、第1著者以外の共著者としての活性度という2つの観点から計量し、研究活動の評価は観点によって異なってくることを、中心性が最も高い研究者がネットワーク上で占める位置の特徴は、研究主題によって異なっていることを示した。

## 1 はじめに

学術研究の世界では、通常、1人の研究者によって何の脈絡もなく唐突に成果が生み出されるということはない。研究分野のコンテキスト、すなわち、その分野における先行研究の蓄積や、研究コミュニティの中での協力関係の上に、新たな成果が積み上げられていくのが普通である。したがって、知識生産活動に関する分野の特徴の把握を目的に、研究者の活動を見るに際しては、研究者それぞれ単独の活動だけ見るのでは不十分で、何らかのつながり(知的紐帯の構造)の中での位置付けを考慮に入れる必要がある。

近年、研究の高度化・複雑化、研究者の専門分化、学際的分野の発達などを背景に (Bordons & Gómez, 2000)、研究成果の引用という経時的なつながりに加えて、研究協力という共時的なネットワークの重要性が大きくなっている。しかしながら、ごく局所的な研究協力のつながり(例えば、どういった属性を持つ研究者が、どういった属性の研究者と共同研究を行う傾向があるか)に着目した研究は、Kretschmer (1994; 1997)、Bahr & Zemon (2000)など、数多く存在するものの、分野全体を見渡した、より大域的なネットワークの分析は、未だ十分になされているとは言い難い。本研究は、研究協力ネットワークの中での役割を考慮したいいくつかの観点から、研究者の活動を分析することで、研究コミュニティのコミュニケーション構造に関する知見を得ることを目的とするものである。

研究分野のコンテキストの中に位置付けられるべきものという、研究活動に関する理念は、研究評価の場面においても反映されており、例えば、研究協力・研究連携が重要な評価の観点の1つとして取り

上げられている (大学評価・学位授与機構, 2003)。研究活動は、研究協力関係の中での役割も含む、様々な観点から評価する必要があり、また、観点が異なれば、評価の結果も異なりうると考えられる。研究活動の評価におけるいくつかの観点に対応する尺度の関連性について調べ、それらが必ずしも一致するものではないことを示すことも、本研究の目的の1つである。

上記の目的のため、本研究では、共著関係のネットワークの分析を行う。研究協力関係は、論文の形で成果が現れない、あるいは、謝辞に記されるだけの場合もあり、必ずしも共著者として論文のクレジットに記されるわけではないが、本研究では、共著で論文を著してはじめて専門的・実質的な協力関係が成立するものを見なし<sup>1</sup>、共著論文の発表状況をもって研究協力関係を計量する。また、そもそも、貢献なき「名誉」共著者の存在が指摘されるように (Cason, 1992)、論文の著者クレジットそのものの疑わしさもあるが、近年、各分野で、倫理ガイドラインが定められるようになってきていること、そして、少なくとも意識調査の結果では、ガイドラインの示すクレジット付与における倫理が研究者に浸透していること (Hoenら, 1998; Bartleら, 2000)を考慮し、本研究は、専門的・実質的貢献を表したものとしてクレジットが妥当であるという仮定を置く。

ところで、今回の研究は、現象の背後にある潜在的な構造(潜在態)というよりも、観察された現象そのものを対象としている。大域的なネットワークの特性を観察する場合、母集団確率を一定と置いてしまうと、同じ条件(同じ論文数)の下でも、潜在態の現実化のばらつきが非常に大きくなる。例えば、低頻度の研究者が仲介者として重要な位置を占めるとき、それが出現するかどうかでネットワークの特性が極端に変わりうる。母集団確率を一定と置くのではなく、それが出現しているからこそ、その周辺の構造がそのように形成されている、と捉えるべきと考えられる。本研究では、現実の共著の連鎖によって構築されたネットワークの特性を見ることを優先することとした。現象の背後にある構造を捉えるためには、仲介者の存在を条件とする、ある種の条件付き確率を組み入れた潜在態のモデルが必要になる。これについては今後の課題としたい。

<sup>1</sup> 謝辞は、専門的・実質的な協力関係というよりも、副次的な支援関係を示すものであることが、例えば、Croninら (2003)の調査で示されている。

## 2 分析対象およびデータ

本研究は、計算機科学分野を分析対象とした。共同研究が活発な分野であり、ネットワークを考慮する必要性が大きいこと、そして、理論的な研究だけでなく、学際的な応用も盛んな分野であり、研究主題(理論/応用)の違いに起因する傾向の差違から有用な知見が得られると予測されたことが、計算機科学を分析の対象に選んだ理由である。

本研究では、Thomson ISIが提供する *SCI* (*Science Citation Index*) データベースの CD-ROM 版、1999 年から 2003 年までの 5 年分を、共著ネットワークを観察するための情報源とする。*SCI* を情報源としたのは、自然科学系分野における最も包括的な書誌データベースのひとつであるため(英語圏の雑誌が中心という問題はあるものの)、そして、質的な基準<sup>2</sup>を満たすコアジャーナルのみを収載しているためである。*SCI* に収録されている文献には、論文だけでなくレビューやレターなど様々な種類が含まれているが、前節で述べたとおり、本研究の関心は、専門的・実質的な知的紐帯の構造にあるため、それが最もよく表れると考えられるオリジナルの論文(‘Document type’が‘Article’の文献)のみを観察の対象とした。

分野の区分、そして分野ごとのコアジャーナルの決定には、*SCI* の主題カテゴリ別収載誌一覧を利用した。計算機科学・理論(computer science, theory & methods)および計算機科学・学際的な応用(computer science, interdisciplinary applications)の2つのカテゴリを対象とし、各々について、*SCI* が収載するコアジャーナルに掲載された論文の書誌情報をデータベースから抽出した。本研究では、掲載雑誌に基づいて、研究者を分野・研究主題と対応付ける。つまり、*SCI* で「計算機科学・理論(あるいは応用)」というカテゴリが付与されている雑誌に掲載された雑誌の著者はすべて「計算機科学・理論(あるいは応用)」分野に属すと考える。

	応用	理論
雑誌数	22	21
論文数	11584	9663
異なり著者数	21801	14525
延べ著者数	32341	22485
平均著者数 (/1 論文)	2.55	2.33
平均論文数 (/1 著者)	1.48	1.45

表 1: データの基本的数量

表 1 に「応用」「理論」それぞれについて、データの基本的数量を示した。雑誌数や論文数、1 論文あたりの平均著者数、1 著者あたりの平均論文数では、両者にそれほど大きな差はない。著者数に関しては、異なり数で見ても、延べ数で見ても「応用」の方が多く、「理論」と比べて、多くの研究者の関与によって成り立っている領域であることが分かる。

<sup>2</sup> ピアレビューや引用分析に基づいて判断されている。また、雑誌の編集者や掲載論文の著者が属する国の多様性も考慮されている (<http://www.isinet.com/selection/>)。

## 3 分析手法

### 3.1 観点

本研究では、研究協力ネットワークにおける (A) 第 1 著者としての重要度と、(B) 第 1 著者以外の共著者としての重要度という 2 つの観点から、研究者の論文生産に関する活動状況を分析する。(A) と (B) を区別することは、第 1 著者はリーダーとして研究の設計を行う者であり、それ以外の共著者とは異なる特別な役割を担っている、という前提に基づく。ガイドラインにそのように明記されている分野もあり、また意識調査の結果からもそれが確認できることから (Bridgwater ら, 1981)、ある程度はその前提に妥当性があると考えられよう。論文生産におけるリーダーとしての役割、そして協力者としてそれを支えるフォロワの役割、どちらも研究協力ネットワークにおいて重要であり、かつ両者は本質的に異なるものと考え、(A)(B)2 つの観点から分析を行うことにした。

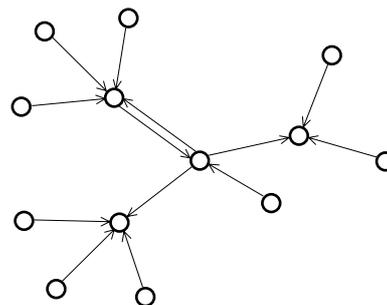


図 1: 共著ネットワークの例

それぞれの操作的定義・指標を設定するにあたり、次のモデルを想定する。(i) 第 1 著者を到達点、それ以外の共著者を出発点とする有向グラフを想定する(図 1 参照)。(ii) 共著関係の強度を考慮に入れた重み付きグラフを想定する。共著関係の強度の尺度としては、Narin ら (1991) の指標や Arunachalam ら (1994) の指標など、いくつかの指標が存在するが、本研究では、共著関係の強度は、単純に共著回数に比例して強くなると仮定し、共著で論文を著した頻度そのものをグラフの重みとした。このような重み付き有向グラフを想定した上で、次に述べる HITS アルゴリズムを応用し、ネットワークの大域的な構造を考慮した重要度の計算を行った。

### 3.2 HITS アルゴリズム

共著関係を結んだ相手の数と、関係の強さ、さらに相手の重要度にも注目して、研究協力ネットワークにおける重要度を計算する。研究者  $n_i$  のリーダーと

しての重要度  $C_l(n_i)$  と、フォローとしての重要度  $C_f(n_i)$  は、それぞれ次の式を尺度にして求める。

$$C_l(n_i) = \sum_{j=1}^g a_{ji} C_f(n_j) \quad (1)$$

$$C_f(n_i) = \sum_{j=1}^g a_{ij} C_l(n_j) \quad (2)$$

ここで、 $g$  は研究者の数を指す。また、 $a_{ij}$  はネットワークの隣接行列  $A$  の成分を指し、研究者  $n_j$  から研究者  $n_i$  に向けた結合の強度、つまり  $n_i$  が第1著者になり  $n_j$  と共著した頻度を値としてとる。ただし、対角成分  $a_{ii}$  は0とする。ここで置いている仮定は、「重要なリーダーを支えている研究者は、フォローとして重要な役割を担っており、重要なフォローをまとめている研究者は、リーダーとして重要な役割を担っている」という相互の依存関係である。式(1)(2)で、再帰的な代入を繰り返すことにより、ネットワークの大域的な構造が各々の研究者の重要度に反映する。この再帰的な繰り返しは、隣接行列  $A$  の固有ベクトル問題に帰着するものである。

「より重要なノードとの関係の方が、そうでないノードとの関係よりも、重要度への寄与が大きい」というアイデアは、Bonacich (1987) の中心性や、HITS アルゴリズム (Kleinberg, 1998)、PageRank アルゴリズム (Brin & Page, 1998) に共通している。<sup>3</sup> 関係の方向性を考慮した2つの役割を設定しているという点で HITS に最も近いが、本研究の手法では、さらに結合の強度(重み)も反映させている。本研究が扱う共著ネットワークは、ノード数の多さから、隣接行列の固有ベクトルを求める方法は現実的でないため、HITS と同じステップで、代入とベクトルの正規化を再帰的に繰り返すことにより(10回ループ)、 $C_l(n_i)$  と  $C_f(n_i)$  を計算した。

## 4 分析結果

計算機科学分野の応用領域と理論領域のそれぞれについて、各々の研究者の重要度  $C_l(n_i)$  と  $C_f(n_i)$  を求めた。2つの重要度、すなわちリーダーとしての重要度とフォローとしての重要度の相関を調べた結果が表2である。両指標ともに比率尺度であるが、外れ値が存在しているため、ここではスピアマンの順位相関係数  $\rho$  を用いている。また、併せて、1著者あたりの平均入出次数も表に示しておいた。

	応用	理論
平均入出次数 (/1 著者)	1.72	1.60
$C_l \cdot C_f$ 間の相関係数 $\rho$	-0.488	-0.438

表2: ネットワークの特性

平均入出次数は、平均何人のリーダーあるいはフォローと結合しているかを表す。1論文あたりの平均

<sup>3</sup> HITS と PageRank は、ウェブページの検索結果のスコア付けを目的に考案されたものである。

著者数の比較と同様、「応用」の方が若干高く、より多くの研究者が協力関係を結んでいることが窺える。2つの重要度の相関は、「応用」「理論」どちらも負の相関を示している。リーダーとしての役割と、フォローとしての役割は、兼ねられるというよりも、別の研究者によって担われる傾向があることが確認できる。「応用」の方が、負の相関が強く表れており、より明確に役割が分かれていると推測できる。

重要度が上位の著者の特徴を、リーダーについては表3に、フォローについては表4に示した。 $D_{in}$  と  $D_{out}$  は入次数と出次数を、 $P$  は発表論文数を、 $P_{sin}$ 、 $P_1$ 、 $P_{2-}$ 、 $P_{last}$  は、それぞれ単著論文数、第1著者としての共著論文数、第2著者以降での共著論文数、最終著者としての共著論文数を表す。「応用」ではリーダーの10位とフォローの9位が、「理論」ではリーダーの3位とフォローの2位が同じ研究者である以外は、10位以内にリーダーとフォローの重なりはない。

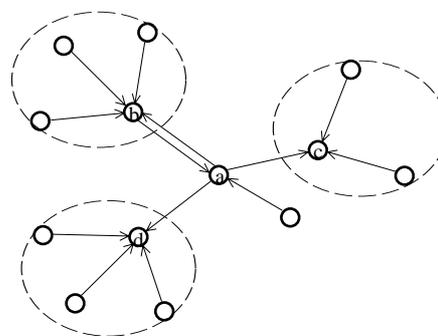


図2: コーディネータとリーダー

表4、フォローの特徴では「応用」と「理論」の差が観察できる。「理論」では、上述のフォロー2位のように、特定のリーダーを支えつつ ( $D_{out} = 1$ ,  $P_{2-} = 4$ )、自身もリーダーとして活発に論文を発表している ( $P_1 = 9$ ) 研究者がいる一方で、フォロー1・6・9位のように、多数のリーダーを ( $D_{out}$  が大)、最終著者という監督者的な立場で支える ( $P_{last}/P$  が大) 研究者も、フォローとして重要な役割を演じている。後者は、ある研究の遂行のために不足していると思われる専門知識などを、自らの人脈の中から見つけてきてつなぎ合わせるコーディネータの役割を果たしていると推測できる。知識伝達のネットワークにおいて、リーダーがハブとして機能しつつ、そのようなコーディネータはバブとハブを仲介するブリッジとして機能するものと考えられる(図2のa)。一方「応用」では、フォローの上位にはそのような研究者は現れていない。「応用」は、社会一般におけるアプリケーションのコンテキストから問題設定が行われるモード2(ギボンズ, 1997)に近く、コーディネータが分野の外に存在して、リーダーとの間に専門的・実質的な協力関係がないこともしばしばあり、それがネットワークに反映しているという

順位	応用							理論						
	$D_{in}$	$D_{out}$	$P$	$P_{sin}$	$P_1$	$P_{2-}$	$P_{last}$	$D_{in}$	$D_{out}$	$P$	$P_{sin}$	$P_1$	$P_{2-}$	$P_{last}$
1	69	0	2	0	2	0	0	41	1	20	1	18	1	1
2	20	2	14	0	12	2	1	6	2	12	1	8	3	1
3	5	2	16	1	9	6	0	18	1	13	0	9	4	0
4	6	2	18	0	8	10	0	1	0	3	1	2	0	0
5	2	0	1	0	1	0	0	5	5	7	0	2	5	1
6	12	8	33	0	23	10	9	3	0	2	0	2	0	0
7	3	2	10	0	4	6	0	21	0	9	0	9	0	0
8	3	0	2	0	2	0	0	8	1	10	0	9	1	0
9	3	0	1	0	1	0	0	2	1	2	0	1	1	0
10	3	1	3	0	1	2	0	2	2	3	0	1	2	1

表 3: リーダとしての重要度が上位の著者の特徴

順位	応用							理論						
	$D_{in}$	$D_{out}$	$P$	$P_{sin}$	$P_1$	$P_{2-}$	$P_{last}$	$D_{in}$	$D_{out}$	$P$	$P_{sin}$	$P_1$	$P_{2-}$	$P_{last}$
1	0	3	4	0	0	4	2	2	9	18	0	1	17	16
2	0	2	3	0	0	3	0	18	1	13	0	9	4	0
3	0	2	3	0	0	3	1	0	6	8	0	0	8	0
4	1	2	4	0	1	3	0	1	4	8	0	1	7	2
5	0	2	3	0	0	3	0	1	3	9	3	1	5	2
6	0	2	3	0	0	3	0	0	7	10	0	0	10	6
7	0	2	3	0	0	3	1	0	1	8	0	0	8	2
8	0	2	3	0	0	3	1	0	3	4	0	0	4	0
9	3	1	3	0	1	2	0	0	5	7	0	0	7	6
10	1	1	3	0	1	2	0	1	3	9	4	1	4	1

表 4: フォロワとしての重要度が上位の著者の特徴

推測もありえるだろう。

## References

Arunachalam, S., Srinivasan, R. and Raman, V. (1994) "International collaboration in science - participation by the Asian giants," *Scientometrics*, vol. 30, no. 1, p. 7-22.

Bahr, A. H. and Zemon, M. (2000) "Collaborative authorship in the journal literature: perspectives for academic librarians who wish to publish," *College & Research Libraries*, vol. 61, no. 5, p. 410-419.

Bartle, S. A., Fink, A. A. and Hayes, B. C. (2000) "Psychology of the scientist: LXXX. attitudes regarding authorship issues in psychological publications," *Psychological Reports*, vol. 86, no. 3, part 1, p. 771-788.

Bonacich, P. (1987) "Power and centrality: a family of measures," *The American Journal of Sociology*, vol. 92, no. 5, p. 1170-1182.

Bordons, M. and Gómez, I. (2000) "Collaboration networks in science," *Web of Knowledge: A Festschrift in Honor of Eugene Garfield*, p. 197-213.

Bridgwater, C. A., Bornstein, P. H. and Walkenbach, J. (1981) "Ethical issues and the assignment of publication credit," *American Psychologist*, vol. 36, no. 5, p. 524-525.

Brin, S. and Page, L. (1998) "The anatomy of a large-scale hypertextual Web search engine," *Proceedings of 7th World Wide Web Conference*, p. 101-117.

Cason, J. A. (1992) "Authorship trends in poultry science, 1981 through 1990," *Poultry Science*, vol. 71, no. 8,

p. 1283-1291.

Cronin, B., Shaw, D. and Barre, K. L. (2003) "A cast of thousands: coauthorship and subauthorship collaboration in the 20th century as manifested in the scholarly journal literature of psychology and philosophy," *Journal of the American Society for Information Science and Technology*, vol. 54, no. 9, p. 855-871.

大学評価・学位授与機構 [編] (2003) 『自己評価実施要項：全学テーマ別評価「国際的な連携及び交流活動」(平成 14 年度着手分)』, 改訂版, 大学評価・学位授与機構: 東京, 56p.

ギボンズ, マイケル [編著] 小林 信一 [監訳] (1997) 『現代社会と知の創造：モード論とは何か (丸善ライブラリー, 241)』, 丸善: 東京, 293p.

Hoen, W. P., Walvoort, H. C. and Overbeke, A. J. P. M. (1998) "What are the factors determining authorship and the order of the authors' names?: a study among authors of the Netherlands Tijdschrift voor Geneeskunde (Dutch Journal of Medicine)," *Journal of the American Medical Association*, vol. 280, no. 3, p. 217-218.

Kleinberg, J. M. (1998) "Authoritative sources in a hyperlinked environment," *Proc. 9th ACM-SIAM Symposium on Discrete Algorithms*, p. 668-677.

Kretschmer, H. (1994) "Coauthorship networks of invisible-colleges and institutionalized communities," *Scientometrics*, vol. 30, no. 1, p. 363-369.

Kretschmer, H. (1997) "Patterns of behaviour in coauthorship networks of invisible colleges," *Scientometrics*, vol. 40, no. 3, p. 579-591.

Narin, F., Stevens, K. and Whitlow E. S. (1991) "Scientific cooperation in Europe and the citation of multinationally authored papers," *Scientometrics*, vol. 21, no. 3, p. 313-323.