

機関番号：12102
 研究種目：基盤研究(C)
 研究期間：2008～2010
 課題番号：20500225
 研究課題名(和文) 共著ネットワーク分析に基づく共同研究者間の影響関係に関する研究
 研究課題名(英文) An analysis of the influence among collaborators based on
 the observation of coauthorship networks
 研究代表者
 芳鐘 冬樹(YOSHIKANE FUYUKI)
 筑波大学・大学院図書館情報メディア研究科・准教授
 研究者番号：30353428

研究成果の概要(和文)：論文生産に関わる共同研究者間の影響関係についての知見を得ることを目的に、研究者の論文生産状況と、共著相手のその後の論文生産状況との関連について調査した結果、研究協力ネットワークの大域的構造を考慮に入れた指標が、共同研究者のその後のパフォーマンスとネットワーク構築状況の両方に強く関わっていること、そして、パフォーマンス指標に対する相関は、理論系分野の方が応用系分野よりも強い傾向にあることなどを明らかにした。

研究成果の概要(英文)：This study examines the relationship between researchers' performance, including the importance in collaboration networks, and the subsequent performance of their collaborators. The results show that the importance as the research leader or cooperater, reflecting the global structures of collaboration networks, is relatively highly correlated with collaborators' subsequent productivity, especially in a theoretical area.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	800,000	240,000	1,040,000
2009年度	800,000	240,000	1,040,000
2010年度	500,000	150,000	650,000
年度			
年度			
総計	2,100,000	630,000	2,730,000

研究分野：総合領域

科研費の分科・細目：情報学・図書館情報学・人文社会情報学

キーワード：図書館情報学，計量情報学，科学計量学，計量書誌学，研究評価

1. 研究開始当初の背景

近年、自然科学系の分野を中心に、研究者の専門分化や研究資金の削減といった、研究者、および研究者を取り巻く環境の変化から、専門知識や資金などを補い合うための研究協力の重要性が大きくなっている。それゆえ、研究者同士の共時的・社会認知的なつながりに基づく人脈形成、情報伝達、そして研究協

力が、研究活動全体に与える影響は、今日、非常に大きいものと考えられる。そうした考えのもと、本課題の研究開始までに、研究者の活動を見る観点として共時的なつながりの現れである「共著」に注目し、(1) 共著ネットワークの成長や、共著傾向の時系列変化といった点から、様々な研究分野について、研究協力に関する特徴を明らかにすると

もに、(2) 共同研究における役割を考慮したネットワーク指標を新たに提案し、特に、分野の中核を構成している研究者の特性を分析してきた。

特に、後者(2)の分析の中で、個人の生産性と、研究協力ネットワーク上の重要度との関連などを検証した結果、研究者の論文生産活動に対してネットワークが及ぼす影響の大きさが推測された。その影響をより直接観察するためには、研究者本人の特性間の関連性だけでなく、共同研究者の特性との関連性を分析する必要があるという考えに至った。共同研究者間の特性の関連を調査した研究は、論文生産性に注目したものなど多数存在する。例えば、相対共著頻度や相関係数などの指標を用いて、共著傾向(共著しやすさ)と、共著者の論文発表実績(共著者間の論文数の差)との関係を分析したものや、共著相手の所属機関や性別に注目したものなど様々な研究が存在する。

しかし、そこで見られているのは同一の時期の特性であり、時間が経過した後の状況との関連という観点も含めて、共同研究者間の特性の関連性に関する定量的な分析を行った研究は、これまでほとんどなされていない。共同研究者間の中・長期的な影響関係に関する示唆を与える研究は不十分であったと言える。

2. 研究の目的

本研究は、研究者の論文生産状況(生産性、共著ネットワーク上の重要度)と、共著相手のその後の論文生産状況との関連の有無を明らかにすることにより、論文生産に関わる共同研究者間の影響関係についての示唆を得ることを主たる目的とする。具体的には、計算機科学分野の研究者を対象に、

- (1) 本人の同時期の特性同士の相関(生産性と研究協力ネットワークの状況との相関など)
- (2) 本人の過去の特性と現在の特性との相関(ある時期に構築していたネットワークと、その後の生産性との相関など)
- (3) 過去の共著相手の特性と、本人の特性との相関(共著相手の生産性と、本人のその後の生産性との相関など)

を明らかにする。論文生産に関わる特性として、論文生産性と共著ネットワーク上の重要度を測定する。生産性の指標については、種々の論文計数法を用い、ネットワーク上の重要度については、次数中心性、媒介中心性などの伝統的なネットワーク尺度の他、研究協力における役割を考慮した指標も併せて用いる。それらの指標を、分析観点に応じ多

角的に組み合わせる。さらに、計算機科学の理論領域と応用領域とで結果の傾向を比較して、研究協力状況に関する領域間の差異についても明らかにする。

3. 研究の方法

計算機科学分野の理論領域と応用領域という2つの下位領域を分析の対象とした。計算機科学分野を含む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 *SCI (Science Citation Index)* 論文データベースを情報源として、1991年から2005年までの当該分野・領域の論文書誌データを抽出し分析に用いた。データの基本的な数量は、雑誌数68(うち理論領域35, 応用領域33), 論文数68,352(うち理論領域28,485, 応用領域39,867), 延べ著者数175,055(うち理論領域64,666, 応用領域110,389)である。また、名寄せ処理を施した後の異なり著者数は、理論領域で31,692, 応用領域で59,771である。本研究では、分野内の研究者の相対的關係(生産性の高低、共著傾向の強弱など)を内包する抽象的な論文生成源を潜在的母集団として想定する。観察される現象(生産性分布、共著ネットワークなど)は、生成源のポテンシャルティ(潜在性)が、その外側の条件(論文産出サイクルや発表媒体の数量など、そしてそれらによって規定される論文発表の機会、論文数)に応じて現実化したものと捉える。すなわち、用いるデータは、現実態としては全数調査に基づくが、母集団そのものを構成するものでなく標本と位置付ける。

研究者の活動状況を見る観点として、(A) 個人の論文生産性と、(B) 研究協力ネットワークを考慮した重要度を設定した。(B)については、(B)-1. 研究協力を行った相手の数だけ見る観点、(B)-2. 相手の重要度も含めて考慮する観点、(B)-3. 相手の所属の多様性、および、(B)-4. 間接的な経験・知識の媒介性に細分化した。(A)は、完全計数法、調整計数法、第1著者計数法、最終著者計数法による論文数(*CMP*, *ADJ*, *FST*, *LST*)と単著論文数(*SIM*)を指標に用いた(論文数指標族と呼ぶ)。(B)-1. は、支援を受けた支援者の異なり数と延べ数(*Cin*, *Cin'*)、支援した代表者の異なり数と延べ数(*Cout*, *Cout'*)、共著パートナーの異なり数と延べ数(*Cdeg*, *Cdeg'*)を指標に用いた(次数中心性指標族と呼ぶ)。(B)-2. は、HITSアルゴリズムを応用して算出する代表者としての協力実績・人脈の重要度(*CL*, *CL'*)、支援者としての協力実績・人脈の重要度(*CF*, *CF'*)、役割を区別しない協力実績・人脈の重要度(*CC*, *CC'*)を指標に用いた(HITS指標族と呼ぶ)。そして、(B)-3. は共著相手機関の異なり数(*Vorg*)と共著相

手国の異なり数 ($Vcnt$) を、(B)-4. は媒介する研究者ペア数 (CB) を指標に用いた (それぞれ、所属多様性指標族、媒介中心性指標と呼ぶ)。(B) 研究協力ネットワークを考慮した重要度のうち、 Cin' 、 $Cout'$ 、 $Cdeg'$ 、 CL 、 CF 、 CC は協力実績を重視するものであり、一方、 Cin 、 $Cout$ 、 $Cdeg$ 、 CL' 、 CF' 、 CC' 、 $Vorg$ 、 $Vcnt$ 、 CB は、重みなしグラフに基づき、実績よりも人脈の構成に注目するものである。また、(B)-1. 次数中心性指標族と (B)-2. HITS 指標族のうち、 $Cin(\cdot)$ 、 $CL(\cdot)$ は代表者としての重要度に、 $Cout(\cdot)$ 、 $CF(\cdot)$ は支援者としての重要度に、そして、 $Cdeg(\cdot)$ 、 $CC(\cdot)$ は役割を区別しない重要度に対応する。

研究者の論文生産に関わる特性間の関連性の分析では、現実には観察された論文数を重視し、現実の要因・制約のもとでの対象の特徴として、データそのものの特徴を指標で測る。標本の信頼性への配慮は、同じ論文生成源が同じ条件のもとで現実化する (同じ潜在的母集団から同じ量の標本が再抽出される) 場合に生じるランダムな誤差のみを想定した統計的有意性の確認にとどめる。ただし、特性間の関連性の分析結果をより適切に解釈できるようにするための下地として、各研究領域の研究者の生産性および共著傾向の基本的な状況については、その潜在的な特徴にも目を向け、標本をもとにした現象の再構築、論文数に応じた現象の変化の分析を行う。具体的には、各々の領域について、2,000 ずつに区切った標本量ごとに、ランダムサブサンプリングを 10,000 回繰り返すことで指標の平均値を求めるモンテカルロシミュレーションを行い、標本量の変化に伴う指標の挙動を観察する。

1991 年～1995 年と 1996 年～2000 年、および 1996 年～2000 年と 2001 年～2005 年を、それぞれ対にして、対象研究者の状況と共著者の状況、そして特定の時期の状況とその後の状況という視点も含めて各指標の相関を調査した。相関の強さは、影響関係を直接示すものではないが、その可能性の有無、程度についての示唆が得られるものと考ええる。また、因子分析などの手法を採らず、相関係数に基づく分析を行う理由は、複数の変数の背後に潜む要因ではなく、設定した観点そのものが重要と考える立場を本研究はとり、それらの観点に対応する指標同士の関連性が見たいからである。

4. 研究成果

本研究では、論文生産性や共著ネットワークを考慮した重要度を測る指標間の相関分析を行った。計算機科学分野の傾向として、分析結果から明らかにした点のうち、主要なものについて以下にまとめる。

(1) 全般的な状況の動的傾向に関しては、

①論文生産数：

理論領域よりも応用領域の方が、同数の論文を生産するのに要する著者の数が多い。領域全体の論文の数 (論文を発表する場・機会) が同じという条件のもとでは、理論領域に属する研究者の方が、応用領域の研究者より、平均して多くの論文を生産しており (完全計数法に基づく論文数 CMP)、平均論文数の成長率を見ても、理論領域の方が一貫して高い値を保ち続けている。つまり、応用領域に比べて、理論領域は、領域の論文生産数が増えていく過程で、新たな周辺の著者が加わる傾向は弱く、同じ著者が繰り返し論文を生産する傾向が強い。また、領域全体の論文数を揃えたときの比較では、両領域とも、時期を遡るほど平均論文数が多く、時期が降るにつれて、同じ著者が繰り返し論文を生産する傾向は弱くなっている。

②共著パートナー数：

領域内の論文数が等しいという条件では、応用領域の方が、理論領域よりも、共著パートナー数 ($Cdeg$) の平均値が大きい。応用領域に属する研究者の方が、各々、多くの研究協力パートナーを必要とすることが窺える。また、どちらの領域においても、2001 年以降は、2000 年以前と比べて共著パートナーの数が多く、各パートナーとの共著頻度の偏りが小さい。2001 年以降は、多くのパートナーと同じような頻度で共著を行う傾向が強い (コアパートナーへの相対的な依存度が低い) と推測される。

平均共著パートナー数の成長率の変化では、1995 年以前の応用領域の減衰が遅いことが、最も顕著な点として観察された。1995 年以前の応用領域は、論文の生産数を増やすのに、多くの新たなパートナーを探さなければならない傾向が相対的に強いと推測される。

(2) 同時期の特性間の相関、指標の代替可能性に関しては、理論領域・応用領域ともに、

①完全計数法による論文数 (CMP) は、全般的に他の指標との相関がある程度は高かった (0.20 以上の有意な相関 ($p < 0.05$) を持つ相手の指標の数が最も多かった)。特に、調整計数法による論文数 (ADJ) との相関が比較的高かったが、それでも両指標の相関係数は 0.70 を下回っており、研究パフォーマンスの評価の場面などにおいて、完全計数法で調整計数法を代替することは妥当とは必ずしも言い難い。

②単著論文数 (SIN) は、ほとんどの指標に

対して相関が低く、負の相関を示すものが多かった。この観点から評価を行う場合、別の指標をもって測ることは難しい。

③第1著者計数法による論文数 (*FST*) と代表者としての協力実績 (*CL*) など、注目する共著の役割が同じ指標同士の相関は高かった。ただし、多少なりとも信頼に足る代替ができるような、極めて高い相関を示す組み合わせとなると限られていた。

研究協力を行った相手の重要度も考慮する協力実績 (*CC*) は、共著パートナー数 (*Cdeg*(\cdot)) との相関係数がそれほど高くなかった。共著ネットワークの大域的な構造を考慮に入れる重要度は、単なるパートナー数では代替し難い場合がある。

④第1著者計数法による論文数 (*FST*) と支援者としての協力実績 (*CF*) など、注目する共著の役割が異なる指標同士は、相関が低かった。特に、第1著者としての論文数 (*FST*) と最終著者としての論文数 (*LST*) は、数編しか論文を発表していない周辺的な研究者を含める含めないに関わりなく、ある程度強い負の相関を示していた。共同研究に基づく論文生産活動を評価するにあたって、焦点を当てる役割に応じた指標を採用する必要性が強い。

⑤共著の相手国の多様性 (*Vcnt*) は、相手機関の多様性 (*Vorg*) に比べて、概して別の指標との相関が低く、前者の方が、別の指標による代替可能性がより低い。

(3) 本人のその後の特性との相関、影響関係に関しては、理論領域・応用領域ともに、

①完全計数法による論文数 (*CMP* [本人・基点])、および媒介ペア数 (*CB* [本人・基点]) は、その後の状況を測った指標全般に対し、ある程度高い、有意な相関 ($p < 0.05$) を示していた。

②その後の論文数 (*CMP* [本人・その後]) に対して相関が最も高いのは、同じ指標 (*CMP* [本人・基点]) だったが、その後のネットワーク関係指標 (*Cdeg*, *Vorg* [本人・その後]) に対しては、それらの指標自身 (*Cdeg*, *Vorg* [本人・基点]) でなく、媒介ペア数 (*CB* [本人・基点]) の相関が高かった。媒介ペア数 (*CB*) は、パートナーを変えつつ論文の発表を繰り返す傾向に関わる指標であるが、そうした傾向が、単なるパートナー数の多寡よりも、その後のネットワーク構築に影響している可能性もあると考える。

(4) 共著者のその後の特性との相関、影響関

係に関しては、理論領域・応用領域ともに、

①その後の論文数に対しても、ネットワーク関係指標に対しても、相関が最も高い指標は、HITS 指標族 (*CC*(\cdot), *CL*(\cdot), *CF*(\cdot) [本人・基点]) だった。HITS 指標族によって測られるところの「隣接する研究者とのつながりの多さだけでなく、それら周囲の研究者が形成する関係をも含めた、より広範囲なつながりの構成」が、共著した相手のその後の状況に、ある程度強く影響している可能性が窺える。

②共著者のその後の論文数に対する相関係数 (HITS 指標族 (*CC* [本人・基点]) との相関が最大) は、本人のその後の論文数に対する相関係数 (論文数 (*CMP* [本人・基点]) との相関が最大) よりも 0.15 程度高かった。研究者の、ある時期における論文生産活動の状況は、自分自身のその後の生産性と同様に、あるいはそれ以上に、共著した相手のその後の生産性とも強く関わっている。

(5) (2)~(4)のいずれに関しても、理論領域の方が、応用領域よりも、指標間の相関が強い傾向が観察された。ただし、相関を見る指標の組合せによっては、比較する条件 (各々の領域内の論文数) が変わると傾向が変化することも推測される。両領域の比較結果を一般化するに際しては注意が必要である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計3件)

- ① Onodera, N., Iwasawa, M., Midorikawa, N., Yoshikane, F., Amano, K., Ootani, Y., Kodama, T., Kiyama, Y., Tsunoda, H. and Yamazaki, S. (2011). A method for eliminating articles by homonymous authors from the large number of articles retrieved by author search. *Journal of the American Society for Information Science and Technology*. Vol. 62, No. 4, p. 677-690. 査読有
- ② Yoshikane, F., Nozawa, T., Shibui, S. and Suzuki, T. (2009). An analysis of the connection between researchers' productivity and their co-authors' past attributions, including the importance in collaboration networks. *Scientometrics*. Vol. 79, No. 2, p. 435-449. 査読有
- ③ Yoshikane, F. (2008). An analysis of the correlation among research productivity and collaboration network

indices. *Research on Academic Degrees and University Evaluation*. No. 8, p. 43-56. 査読有

〔学会発表〕 (計 3 件)

- ① Kawamura, S., Suzuki, T. and Yoshikane, F. Classification of hyperlink networks by multivariate analysis of network indicators: the case of Japanese public libraries. Proceedings of ISSI 2009 (12th International Conference of the International Society for Scientometrics and Informetrics). Vol. 2, p. 648-659. (17 July, 2009 at Federal University of Rio de Janeiro, Rio de Janeiro, Brazil)
- ② Yoshikane, F., Suzuki, T. and Kawamura, S. Estimating the influence of researchers' and their collaborators' performance on their future productivity. Proceedings of ISSI 2009 (12th International Conference of the International Society for Scientometrics and Informetrics). Vol. 2, p. 1006-1007. (16 July, 2009 at Federal University of Rio de Janeiro, Rio de Janeiro, Brazil)
- ③ Yoshikane, F., Suzuki, T. and Tsuji, K. Investigation of the correlation between researchers' properties and productivity through analyzing co-authorship networks. 10th International Conference on Science and Technology Indicators: Book of Abstracts (Excellence and Emergence: A New Challenge for the Combination of Quantitative and Qualitative Approaches). p. 425-427. (18 September, 2008 at University of Vienna, Vienna, Austria)

6. 研究組織

(1) 研究代表者

芳鐘 冬樹 (YOSHIKANE FUYUKI)
筑波大学・大学院図書館情報メディア
研究科・准教授
研究者番号：30353428

(2) 研究分担者

(3) 連携研究者