

組織移動情報と研究者の学術生産性を用いた 研究組織環境の定量評価手法の提案

Identifying Affiliation Impacts on Innovation Enhancement

三浦 崇寛*¹
Takahiro Miura

浅谷 公威*¹
Kimitaka Asatani

坂田 一郎*¹
Ichiro Sakata

*¹ 東京大学工学系研究科技術経営戦略学専攻

Department of Technology Management for Innovation, Graduate School of Engineering, The University of Tokyo

From the aspect of the efficient operation of a research organization, it is important to manage the research platform to promote the future publishing of scientists. However, the previous ways of organization evaluations cannot distinguish the evaluation based on scientists' community and that based on its research platform. We propose the methodology to extract scientist's movement and research productivity from bibliographic records on Computer Science between 2007 and 2018 on Scopus and make quantitative analyses about institutional impacts on scientific innovation. As a result, scientists moving to state-of-the-art institutions does not always improve scientist's productivity, rather collecting talented authors. Divided by its nationality, Chinese affiliations in low research productivity enhance the performance of scientists. Conversely, Japanese institutions give less contribution to scientists' productivity. This analysis contributes to a better understanding of scientists' incentive and what research organization can do to make scientific innovation.

1. はじめに

研究者が新しい知的発見を起こす具体的なメカニズムは未だ明らかになっていないが、研究組織が研究者の創造性に与える影響は非常に大きいことが指摘されている。適切な報奨設定や研究者同士の交流を活性化させるといった組織のインセンティブ設計によって研究者の創造性が向上することが分かっている[Verbee 2015]。特に、民間企業の参入によって近年急速に発展している Computer Science 分野では今後どの組織が成長していくかを予測することは研究者のキャリア形成のみならず、投資家や各国政府の戦略決定にも大きな影響を与えている。

しかし従来の組織評価指標である Nature Index や QS World University Rankings などの論文数や被引用数を軸とする評価は組織に所属する研究者の評価と組織環境の評価を区別していない。これは暗黙のうちに両者が相関していると考えられているためであるが、実際は両者が相関するとは限らない。例えば 1990 年代の日本の国立大学においては、優秀な研究者が劣悪な研究環境に置かれることで自身の研究能力を十分に発揮できない「頭脳の棺桶」と呼ばれる問題が古くから指摘されている[Asahi 1991]。こうした組織環境の悪化が 2000 年以降の日本の研究の国際競争力低下につながっていると考えられる[NISTEP 2017]。現状の組織環境評価手法は予算や学生の多様性などで測っているがこれらは環境の一側面に過ぎず、組織環境を定量的に評価できているものではない。

本研究では書誌情報から抽出される情報を元に組織環境を直接定量評価する手法を提案し、組織環境が研究者の学術生産性に与える影響を正しく評価できる指標となっているかを検証する。その結果、提案指標が既存の組織評価指標と比較してより研究組織環境の影響のみに焦点が当てられた指標となっていることが明らかになった。本研究によって研究者コミュニティとしての組織評価と研究プラットフォームとしての組織評価を区別

して行うことができることになることで、研究者の創造性を促す組織マネジメントを行うための意思決定補助になると考える。

2. 関連研究

研究環境に関する分析は大きく分けてチーム単位の研究者行動に着目した分析と地域単位の施策に関連づけた分析の 2 種類に分けられる。チームの研究環境に着目したものでは、各国の研究者の学術生産性が国際共著や国際間移動といった国の openness と相関関係にあることや[Wagner 2017]、論文投稿が活発で成熟した分野であるほど研究の team-size が大きいことが明らかになっている[Fortunato 2018]。また国や地域の施策と紐づけた研究では、近年科学的リーダーシップを獲得していると言われる中国が論文数やポストドクの数では欧米を上回りつつあるものの研究のインパクトはまだ追いついていないことや[Basu 2018]、アラブの春前後でアラブ諸国の研究者の論文数、被引用数ともに増加したこと[Ibrahim 2018]が計量書誌学上の研究から明らかになっている。しかし、研究環境の分析を組織評価において利用している研究は未だ行われていない。

3. 提案手法

3.1 概要

本研究では、論文数や被引用数といった観測される学術生産性指標が研究者個人の研究力と研究組織の影響力の 2 つに分けられるとする Dundar の研究結果を用いる[Dundar 1998](図 1)。

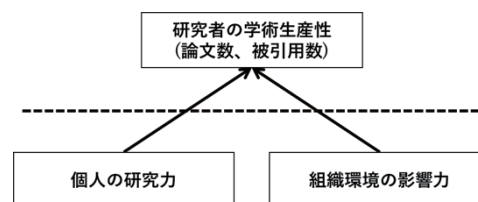


図 1. 学術生産性の捉え方

これまでの組織評価では書誌情報から直接観測できる論文数や被引用数のみを扱っていたため、その中で個人の研究力と

組織環境の影響力の占める割合が不透明であり組織環境を評価することができなかった。そこで研究者の組織移動に着目し、組織移動しても変わらない部分が研究者個人の研究力、変わる部分が組織環境の影響力であると仮定することによって本来観測することができない組織環境の影響を定量的に評価できるのではないかと考えた。提案手法の概要を図2に示す。以降で詳細を説明する。

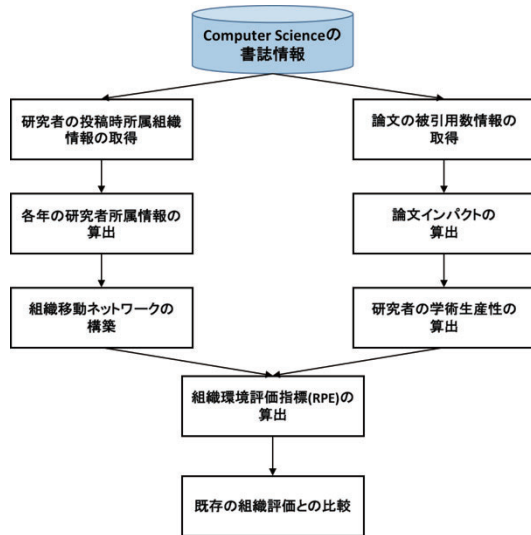


図2. 提案手法の概要

3.2 組織移動ネットワークの構築

各著者の論文投稿時における所属組織情報から著者をキーとしたスパースな所属組織辞書を作成する。次にその辞書から上山の提案した移動検出モデル[Kamiyama 2015]を用いて各著者の所属組織と流入組織を年ごとに特定する。上山はその組織で過去4年以内に投稿がなく新しく投稿された年を流入年、投稿以降4年間投稿がなかった年を流出年として研究者の所属情報を作成した。この時所属の精度を上げるために、データを通じて1回しか投稿のない組織には所属していないとする。具体例を図3に示す。例えば2010, 2011, 2013年に投稿があった場合は2010年を流入年, 2013年を流出年とする。

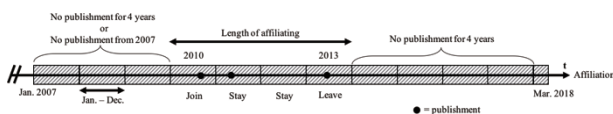


図3. 研究者の所属の検出方法

その後各年の流出流入の情報から、各研究者についてある組織を流出してから2年以内に流入している組織に対してリンクを結び、移動ネットワークを作成する。移動年は流出年と流入年の平均値をとって小数点を切り捨てる。

3.3 研究者の学術生産性の算出

本研究では、論文の2年以内の被引用数を論文のインパクトと考える。引用は多くないが優れた研究でもある sleeping beauty[Raan 2004]の存在も留意しなければならないが、Computer Science 分野では論文の投稿スパンが短いため2年内被引用数は本分野における学術生産性の一側面を示していると考えられる。各研究者の学術生産性は一般的に投稿論文の被

引用数の総和によって求められるが、より正確に研究のインパクトを測るために論文の分野と投稿年度によって標準化を行う。

分野が異なると引用構造が異なるため被引用数の比較では研究インパクトの比較ができないことが知られているが、Radicchiらによれば論文 p の被引用数 $w(p, s)$ をその分野の平均被引用数 $w_c(p, s)$ で標準化することで比較可能であることが指摘されており[Radicchi 2008],今回はこの手法を採用する。分野の特定には、論文の引用ネットワークを Louvain 法[Tang 2016]を用いてコミュニティに分割したものを用いる。

次に投稿年度についてである。Fortunatoらによれば知的発見あたりの論文数が増加しているため[Fortunato 2018],2007年の論文数を1とした各年の相対論文数 n_y で平均することによって論文の知的発見のみを計測することができる。また本手法では論文のインパクトを2年以内の被引用に限るため時系列情報が取得できるデータ内の論文からの被引用のみを考えるが、Computer Science は近年急速に成長した分野であり分野間引用から分野内引用へとシフトしていることが考えられるため、各年の2年内被引用数 $w_y(p, s)$ で標準化を行う。以上の2つから著者数 α_p の論文における1人あたりインパクトは、

$$i_{p,s} = \frac{w(p,s)}{(\alpha_p \times n_y \times (w_c(p,s) + w_y(p,s)))} \quad (1)$$

と求める。各年の研究者の学術生産性は $i_{p,s}$ をその研究者の投稿した論文数で足し合わせたものである。

3.4 提案指標の定義

組織移動情報と各年の研究者の学術生産性変化から今回の提案指標を算出する。今回提案する組織環境評価指標を Research Productivity Enhancement(RPE)と呼ぶ。RPEは流入した研究者の学術生産性が向上した場合、または流出した研究者の学術生産性が低下した場合に高い値をとる指標となる。提案手法と同様に、研究者の移動と学術生産性変化から各都市の傾向を分析した Verginer の計算手法[Verginer 2018]を用いて、各組織のRPEを以下のように求める。

$$RPE_{o,y} = \log \left(\frac{(\sum_{a \in A_{y,\sigma,o}} \forall \sigma \hat{I}_{a,y,s}^1) \times (\sum_{a \in A_{y,o,\tau}} \forall \tau \hat{I}_{a,y,s}^0)}{(\sum_{a \in A_{y,\sigma,o}} \forall \sigma \hat{I}_{a,y,s}^0) \times (\sum_{a \in A_{y,o,\tau}} \forall \tau \hat{I}_{a,y,s}^1)} \right) \quad (2)$$

表1. 各変数の定義

変数	定義
s	引用の判断基準(2年)
$A_{y,\sigma,\tau}$	y年に組織 σ から組織 τ に移動した著者群
$\hat{I}_{a,y,s}^0$	著者aの $\max(y-4, 2007)$ 年からy-1年までの平均インパクト
$\hat{I}_{a,y,s}^1$	著者aのy年から $\min(y+3, 2017-s)$ 年までの平均インパクト
$RPE_{o,y}$	y年の組織oのRPE

4. データ

本研究では Elsevier 社が運営する世界最大級の論文データベース Scopus から抽出した Computer Science カテゴリの論文のうち、2007年1月から2018年3月までの論文3,413,701件に絞って分析を行う。書誌情報は著者名、所属組織名、投稿年、引用文献がラベルづけされており、著者名と所属組織名はそれぞれユ

ユニークな ID で管理されている。今回は表現の揺らぎの影響を考慮して各 ID で最も頻度の多い名前と ID を紐づけた。名寄せの結果ユニークな著者数は 3,089,335 人、組織数は 237,179 件であった。各組織には所属国情報が付いており、その分布を表 2 に示す。

表 2. 各国の組織数 Top10

Chn	Usa	Ind	Fra	Jpn
37,245	37,070	16,656	13,104	12,280
Ita	Gbr	Esp	Bra	Kor
8,290	7,974	6,237	5,505	5,222

5. 結果

提案手法で移動を抽出した結果、ノード数 18,676、エッジ数 87,465 の移動ネットワークを構築した。流入は全年で観測されるが、流出は以降 4 年間の投稿情報が必要のため 2013 年までしか観測されない。そのため移動が観測されるのは 2007-2014 年の 8 年間で、各年の移動数は 2007 年と 2014 年が 5000 件弱、2008-2013 年では 10,000-15,000 件であった。

5.1 提案指標と既存指標を用いた組織評価

図 4 に主な結果を示す。横軸は既存指標で用いられる各研究組織から出た全ての論文のインパクトの総和、縦軸は今回の提案指標である RPE を各年で平均したものである。より精度を上げるために移動数が 250 回以上観測された 90 組織のみをプロットしている。点線はプロットされた組織の各平均線である。

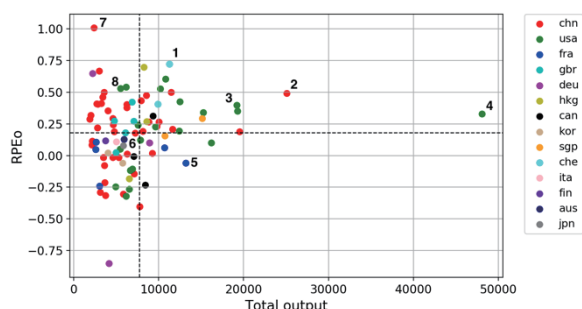


図 4. 各組織の学術生産性と RPE の関係

両者の相関係数は 0.166 程度で相関は見られず、背景で述べたように既存の組織評価指標が組織環境評価とは一致していないことがわかる。各軸の平均線によって区切られる 4 象限について特徴的な組織を挙げる。

(1) 第 1 象限

本象限は既存の組織評価でも提案指標でも高い評価を受けている組織群であり、優れた組織環境が研究者の学術生産性を高めたことで研究者コミュニティとしても活性化している正の循環の存在が推測される。OECD 内で最も 1 人あたりの教育費が高い 1: “ETH Zurich”や年間 1 兆円近い予算を有し 2016 年には科学技術の年間論文数がハーバード大を超えて世界 1 位となった 2: “Chinese Academy of Sciences”, Computer Science 分野を牽引する 3: “Microsoft Research”, 各分校が全て総和されているため既存指標で特に高い評価となっている 4: “California University”らがあった。

(2) 第 2 象限

本象限は既存指標では高い評価を受けているが、所属研究者の創造性を十分に引き出せていない「頭脳の棺桶」となっている可能性のある研究組織群である。フランスの 5: “CNRS”がこれにあたる。2000 年から CNRS の常勤研究員として勤めた小田によれば、CNRS はヨーロッパの優れた研究者がこぞって入る研究組織ではあるが政府からの資金が徐々に削られてしまい、資金を集めるために地方自治体や民間のプロジェクト用のテーマばかりにお金が割り振られてしまうため基礎研究が伸びにくい環境にあったという [Oda 2005]。また 2004 年には同組織内で研究者のストライキも発生している。この組織群は優れた研究者コミュニティを有しているがその能力を十分に活かすための組織環境に乏しく、現状優秀な研究者が集まることによって組織評価を高めているが将来の研究成果の低下が危ぶまれるともいえるべき組織群である。

(3) 第 3 象限

本象限は既存指標でも提案指標でもあまり評価されていない組織群であり、6: “University of Tokyo”があたる。東京大学は博士進学者が毎年減少してきており、研究環境の改善が叫ばれている。この象限に属する大学組織はこれから高い学術生産性を持つために組織マネジメントを見直す必要がある大学であると言える。

(4) 第 4 象限

本象限は既存指標では評価されていないが研究者の学術生産性を引き出しており今後の躍進が期待される組織群である。7: “Beijing University of Technology”, 8: “IBM Research”があたる。またこの象限に属する 24 組織のうち 15 組織が中国の大学であった。この傾向は現在の中国の社会情勢と比較しても理にかなっており、2019 年 1 月 11 日付の日本経済新聞の記事によればこれまで半導体や航空機のような戦略的製造業において技術流出を控えようとしてきたアメリカが、新たに AI やロボットを規制の対象としようとしている。この背景には今回明らかになった中国の研究機関の Computer Science における規模の拡大と実力の向上を予期したものであると考えられる。既存指標ではこれらの組織の優位性を説明できないが、RPE を用いることによって研究者の学術生産性を高めている組織環境を有していることが明らかになった。

5.2 移動数と提案指標の関係

各組織の移動数と RPE のばらつきを示したものが図 5 になる。移動数が多いほど RPE のばらつきは小さくなり、安定した値を取っていることがわかる。

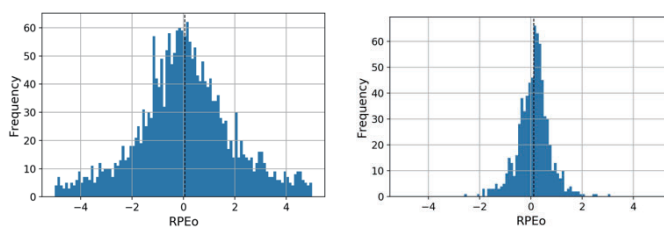


図 5. 移動数ごとの RPE の分布 (左が移動数 50 以下、右が移動数 50 以上)

提案手法では研究者個人の研究力の変動を全く考慮していなかったが、科学的発見というのはしばしば組織環境とは無関係に突発的に発生するものであり、研究キャリアの中でいつ大きな成功を収めるかは人によって様々で予測することは難しいことが分かっている [Fortunato 2018]。移動数が少ない組織では一研

研究者の学術生産性変化の影響が大きなウェイトを占めるため、突発的発見が起こった場合に組織環境が実際とは離れた値を取ってしまうことが考えられる。ある程度移動のサンプル数があれば突発的影響が平均化されより正確に組織環境のみを評価することができるため、RPE を用いた分析を行う際には一定以上のデータ数が必要であると考えられる。

6. 考察

本研究では組織環境を評価するための新しい指標とその算出方法の提案を行ったが、既存手法を用いても組織環境評価を行うことが可能であるかどうかをここで議論する。既存手法によって提案指標が代替不可能ならば、今回の提案手法が組織環境を評価する上で重要な役割を担うことができると考えられる。既存手法として、計量書誌学上で行われている移動ネットワークの構造的特徴に着目した分析を行う。例えば Life Science 分野の研究者に関しては PageRank の高い組織にインパクトの高い論文を出した研究者が移動しやすいことが明らかになっている [Verginer 2018]。

今回は抽出した移動ネットワークの各ノードの中心性と提案指標の相関係数を調べる。比較対象として各種中心性と既存指標である論文インパクトの総和との相関係数も調べる。中心性には代表的な次数中心性、近接中心性、媒介中心性、固有ベクトル中心性、PageRank [Page 1999] を用いる。PageRank は有向重みありネットワーク、それ以外は無向重みなしネットワークとして計算している。結果を表 3 に示す。

表 3. 各種中心性と既存指標、提案指標の相関係数

中心性	既存指標	提案指標
次数中心性	0.82	0.13
近接中心性	0.77	0.07
媒介中心性	0.81	0.07
固有ベクトル中心性	0.47	0.12
PageRank	0.72	0.05

ネットワークの各種中心性は既存指標と相関が強いものに対して提案指標との相関はほぼ見られないことがわかる。このことから研究者の移動は論文数や被引用数といった研究者コミュニティの評価を元に行われているものであり、組織移動に対して研究環境は大きな影響を与えていないことが示唆される。よって仮説で述べた通り既存手法である移動ネットワーク構造に着目した指標によっては今回の提案指標は代替されず、組織移動と研究者の学術生産性変化を用いた今回のアプローチは組織評価において重要な位置を占めると考えられる。

7. 結論

本研究では、書誌情報から抽出した研究者の組織移動ネットワークと学術生産性変化から研究組織環境を定量評価する指標 RPE を提案し、実際に RPE を用いた環境評価を行なった。その結果、Computer Science 分野を牽引する代表的な研究組織や既存指標ではまだ評価されていない中国の研究組織が高い RPE を保持していることが明らかになった。一方でフランスの CNRS に代表されるように、優れた研究者コミュニティを有しているが組織環境がその能力を十分に引き出せていない「頭脳の棺桶」となっている組織の存在も明らかになった。このことから RPE は研究組織が研究者の学術生産性に与える影響を表していると考えられるため、研究環境評価指標としてだけでなく研究者の将来性を推測するためにも利用できると考えられる。また既存手法である研究者の移動構造から組織の特徴を抽出する手

法によっては組織環境を表現することができず、今回の提案指標が組織環境評価に対して有意な結果を示していることが考えられる。今後の展開としては、現在の組織環境評価指標で用いられている予算や留学生数、所属研究者へのアンケートを用いた現状の組織環境評価と比べて提案手法が妥当な結果を示しているのかを調査する。

参考文献

- [Verbree 2015] Maaik Verbree, Edwin Horlings, Peter Groenewgen, Inge Vander Weijden and Peter van den Besselaar: Organizational factors influencing scholarly performance, *Scientometrics*, Vol. 102, No.1, pp.25-49, Jan, 2015
- [Asahi 1991] 朝日新聞 Weekly. 「頭脳の棺桶 国立大学」, AERA, 朝日新聞社, p9-14, 1991
- [NISTEP 2017] 村上昭義, 伊神正貴: 科学研究のベンチマーキング 2017. Technical Report, 文部科学省科学技術・学術政策研究所, August 2017.
- [Wagner 2017] Caroline S Wagner and Koen Jonkers. Open countries have strong science. *Nature News*, Vol. 550, No. 7674, p. 32, 2017.
- [Fortunato 2018] Fortunato Santo, et al.: Science of science, *Science* 359.6379: eaao0185, 2018
- [Basu 2018] A. Basu, P. Foland, G. Holdridge, and R. D. Shelton: China's rising leadership in science and technology: quantitative and qualitative indicators. *Scientometrics*, Vol. 117, No. 1, pp. 249-269, Oct 2018.
- [Ibrahim 2018] Bahaa Ibrahim. Arab spring's effect on scientific productivity and research performance in arab countries. *Scientometrics*, Vol. 117, No. 3, pp. 1555-1586, Dec 2018.
- [Dundar 1998] Halil Dundar and Darrell R Lewis: Determinants of research productivity in higher education, *Research in higher education*, Vol. 39, No. 6, pp. 607-631, 1998
- [Kamiyama 2015] 上山隆大: 研究者養成と研究のマネジメント - アメリカの経験から学ぶ. *日本労働研究雑誌*, No.660, pp.87-102, Jul 2015.
- [Raaijmakers 2004] Anthony F. J. van Raaijmakers: Sleeping beauties in science. *Scientometrics*, Vol. 59, No. 3, pp. 467-472, Mar 2004.
- [Radicchi 2008] Filippo Radicchi, Santo Fortunato, and Claudio Castellano: Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences*, Vol. 105, No. 45, pp. 17268-17272, 2008.
- [Tang 2016] Jian Tang, Jingzhou Liu, Ming Zhang, and Qiaozhu Mei: Visualizing large-scale and high-dimensional data. In *Proceedings of the 25th International Conference on World Wide Web*, pp.287-297, 2016
- [Verginer 2018] Luca Verginer and Massimo Riccaboni: Brain-circulation network: The global mobility of the life scientists. Working Papers 10/2018, IMT Institute for Advanced Studies Lucca, October 2018
- [Oda 2005] 小田玲子: フランスの研究者の教育・就職状況と研究費, 2005. <http://mathsoc.jp/publication/tushin/1004/oda.pdf> (1/27 アクセス)
- [Page 1999] Page, Lawrence, et al.: The PageRank citation ranking Bringing order to the web. *Stanford InfoLab*, 1999.