

論文の共著関係ネットワークの中心性分析

原田泰輔 山本学 寺野隆雄 (東京工業大学)

Central analysis of the collaboration relation network of a paper

*T. Harada, G. Yamamoto and T. Terano(Tokyo Institute of Technology)

Abstract— The relations between other researchers are very important like the Joint authorship of a paper and quotation of a paper. Creation of the researcher network about a certain area of research is performed, and the researcher who tows the field is specified. On the other hand, in this research, we analyzed about the researcher network between different areas of research. And we specified the researcher who tows the multi-area of research.

Key Words: Network of researchers, Complex networks

1 はじめに

論文の共著や文献の引用などに代表されるように、研究者にとって他の研究者との関わりは大変重要なものである。実際、これまでに、研究者ネットワークに関する多くの研究¹⁾が行われ、コミュニティマイニングのために研究者情報を可視化するなどの取り組みが行なわれている。²⁾また、近年では、異なる研究分野の連携が必要とされる局面も増え、効果的な共同研究行なうための Web からの研究者ネットワークの抽出などの取り組みも行なわれている。³⁾⁴⁾

上記のような現状のなかで、今後、研究者間、異なる研究分野間でのネットワークの重要性はさらに増していくものと考えられる。

そこで、本研究では、共著関係をもとにネットワークを作成し、その中心性を見ることで各研究分野で中心的な役割を果たしている人物の特定を行なう。

2 関連研究

研究者ネットワークについての研究では、特定の研究分野についての研究者ネットワークの作成が行なわれている。篠田は人工知能学会論文誌に掲載された論文の共著関係をもとに、ノードに時系列情報を持つ有向グラフを用いて人工知能研究の系譜を作成した。⁵⁾安田らは、ウェブマイニングの手法を用いて人工知能学会における研究者ネットワークを作成し、Web 上での共起関係ネットワークにおける研究者の媒介性により、将来の共著関係における研究者の位置特性がある程度予測できることを示した。⁶⁾内藤らは、日本におけるロボット工学の研究者ネットワークの分析を行ない、ロボット工学の研究者ネットワークが複雑ネットワークとして一般的な、スケールフリー性、スモールワールド性、クラスター性などの特徴を有していることを示した。⁷⁾

また、論文の引用関係から研究者ネットワークを芳鐘らは、これまでに行なわれてきた共著ネットワークの分析手法についてそれらの視点や用いられた指標、手法の整理を行なっている。¹⁾

3 対象データと分析手法

本研究では、オープンデータである JST 科学技術文献データをもとに、1981 年から 2012 年までの工学一般領域、エネルギー工学、熱機関学・応用熱力学、機械

工学、運輸工学、鉱山学の 6 分野を対象にした分析を行なった。6 分野全体では、論文数:627940 編、ユニーク著者数:734735 人、エネルギー工学分野では、論文数:58202 編、ユニーク著者数:93354 人であった。上記のデータを対象に、以下の手順で重み無し無効グラフを共著ネットワークを作成し、分析を行なった。

- (1) 各論文の共著者間に共起回数 1 を加える。ただし、三人以上の共著の場合は、著者リスト上で隣り合う著者同士のみ共起回数を加える。
- (2) 各著者をノード、共起をリンクとし、共起回数が 1 以上のノード間にリンクを張りネットワークの形成を行なう。ただし、共起回数が 2 以上の場合でも、重み付けは行なわず、単リンクとして扱う。
- (3)(2) で得られたネットワークの各ノードに対し、各中心性を求める。

(1) については、共著関係をその論文の著者数にとらわれず 1 とカウントする(一つの論文の共著者間で完全グラフが構成される)のが一般的である。しかし、共著者数が多い論文では、ひとつの論文で多くの共著関係を生み出してしまふ。⁵⁾そこで、本研究では、著者リスト上で隣り合う著者同士にのみリンクを張っていく。

4 分析結果

ここでは、上記の分析手法に基づいて行なった各分野の中心性分析の結果と、作成した共著関係ネットワークの評価を示す。前節で示した手順で共著関係ネットワークは、工学一般領域、エネルギー工学、熱機関学・応用熱力学、機械工学、運輸工学、鉱山学の 6 分野全体でノード数:411508、エッジ数:262278、エネルギー工学分野でノード数:39279、エッジ数:30445 のネットワークとなった。エネルギー工学分野のネットワークを図示すると、Fig 1, Fig 2 となった。

Fig 1, Fig 2 は、各ノードに、度数に応じた引力を持たせ配置を行なったものである。度数の多い著者を中心に、いくつかのグループに分かれていることがわかる。

4.1 中心性分析

ネットワークの中心性を評価する上で、本研究では以下の三つの指標を用いた。

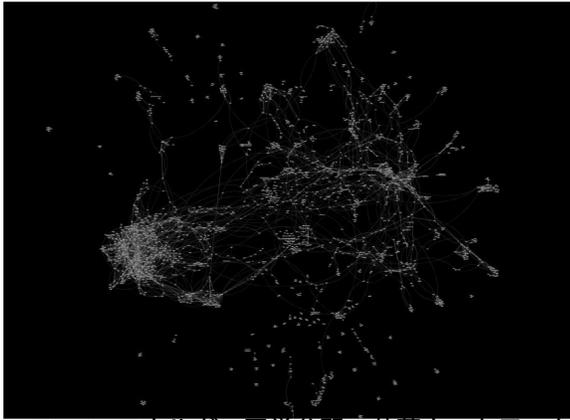


Fig. 1: エネルギー工学分野の共著ネットワーク

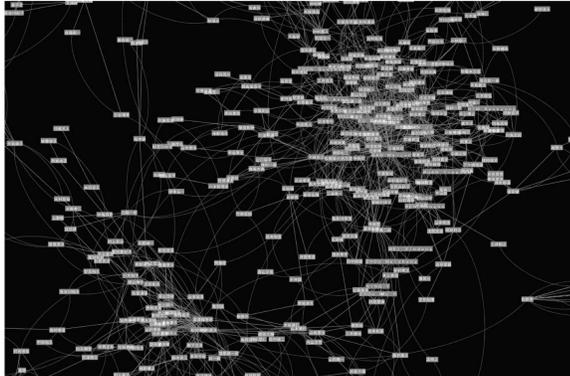


Fig. 2: エネルギー工学分野の共著ネットワーク 2

・ 次数中心性

次数中心性は、ネットワーク内でより多くのリンクを持つノードを高く評価する指標であり、グラフの隣接行列を a_{ij} とすると、次のように定式化される。

$$C_d(i) = \sum_{j=i}^n a_{ij} \quad (1)$$

ここで、 n はグラフに含まれるノード数である。

・ 媒介中心性

媒介中心性は、あるノードが他のノード間の最短経路上に位置する程度を示す指標であり、他のノード同士をつなぐ働きをするノードを示す。

$$C_b(i) = \sum_{i \neq j \neq k} \frac{g_{jk}(i)}{g_{jk}} \quad (2)$$

ここで、 g_{jk} は頂点 j と頂点 k の間の最短経路のうち頂点 i を通るものである。

・ 近接中心性

近接中心性は、他のノードへの最短距離の総和の逆数である。

$$C_c(i) = \frac{1}{\sum_{j=1}^n d_{ij}} \quad (3)$$

ここで、 d_{ij} は頂点 i から頂点 j への最短距離である。

工学一般領域、エネルギー工学、熱機関学・応用熱力学、機械工学、運輸工学、鉱山学の各分野での次数中心性、媒介中心性、近接中心性のランキングで上位にくる著者のリストを Table 1 から Table 3 に示す。

Table 1: 次数中心性

工学一般	エネルギー工学	熱機関・応用熱力学	機械工学	運輸交通工学	鉱山工学	6分野総合
田中啓介	美濃輪智朗	田辺新一	大森整	山本新	岡津弘明	田辺新一
江角務	松村幸彦	加藤信介	須田義大	中村英夫	山本晃司	大森整
久保司郎	吉野博	吉野博	永井正夫	野倫明	恒川昌美	加藤信介
秋庭義明	下田吉之	千田二郎	竹内芳美	泉隆	大和田秀二	吉野博
野口博司	赤司泰義	相良和伸	堤正臣	永井正夫	柴山敦	須田義大
北村春幸	丹羽英治	甲谷寿史	LIN Weimin	須田義大	広吉直樹	永井正夫
安藤柱	井上隆	山中俊夫	森田昇	水間毅	藤田豊久	村上周三
高橋学	渡辺俊行	秋元孝之	園枝正典	高橋聖	島田英樹	秋元孝之
寺田賢二郎	高口洋人	村上周三	毛利尚武	土居俊一	松井紀久男	下田吉之
菅田淳	村上周三	草鹿仁	青山崇一	山本修身	佐々木久郎	千田二郎

Table 2: 媒介中心性

工学一般	エネルギー工学	熱機関・応用熱力学	機械工学	運輸交通工学	鉱山工学	6分野総合
LIM C. w.	美濃輪智朗	田辺新一	竹内芳美	須田義大	柴山敦	大森整
酒井達雄	鳥田荘平	千田二郎	大森整	桑原雅夫	山富二郎	田辺新一
田中啓介	下田吉之	加藤信介	梅原徳次	水間毅	安達毅	ZHANG X.
高橋学	鈴木善三	森吉泰生	須田義大	赤松幹之	山口勉	WU Y.
大崎純	松村幸彦	後藤雄一	永井正夫	網島均	岡津弘明	吉野博
菅田淳	山本博巳	村上周三	中村隆	中村英夫	大和田秀二	ZHANG X.
野口裕久	坂西欣也	柳原隆司	辻本良信	永井正夫	山本晃司	須田義大
寺田賢二郎	三浦秀一	吉野博	佐々木信也	押野康夫	松岡俊文	村上周三
CHEN W. q.	伊香賀俊治	門出政則	森田昇	鎌田実	佐々木久郎	LIN W.
高橋正気	小木知子	池田耕一	LIN Weimin	緒方正剛	西山孝	竹内芳美

Table 3: 近接中心性

工学一般	エネルギー工学	熱機関・応用熱力学	機械工学	運輸交通工学	鉱山工学	6分野総合
酒井達雄	美濃輪智朗	田辺新一	大森整	須田義大	山口勉	田辺新一
菅田淳	鳥田荘平	加藤信介	竹内芳美	桑原雅夫	山富二郎	大森整
越智保雄	鈴木善三	村上周三	梅原徳次	網島均	佐々木久郎	村上周三
田中啓介	坂西欣也	柳原隆司	LIN Weimin	永井正夫	天満則夫	吉野博
小笠原俊夫	松村幸彦	吉野博	中村隆	片山硬	坂本靖英	LIN Weimin
野口裕久	花岡寿明	秋元孝之	佐々木信也	丸茂喜高	大貫光太郎	井上隆
伊藤隆基	山本博巳	池田耕一	森誠之	鎌田実	安達毅	須田義大
佐野雄二	三浦秀一	桑沢保夫	須田義大	水間毅	大和田秀二	竹内芳美
秋田真一	藤本真司	千田二郎	上原嘉宏	大野寛之	野井武	伊香賀俊治
大野信忠	井上貴至	後藤雄一	野口昭治	佐藤安弘	鳥田荘平	松村隆

Table 1 から Table 3 を見ると、順位の変動はあるものの、どの中心性においても大きな入れ替わりは見られず、各分野における各中心性においても特定の人物が上位にランキングされているのがわかる。しかし、6分野総合においては、単一分野よりも各中心性で変化が大きい。これは、各中心性に各研究分野の特徴（論文数が多い、共著人数が多い論文が多いなど）が現れたものだと考えられる。

また、エネルギー工学分野については、エネルギー工学、エネルギー資源及び開発、エネルギー変換及び貯蔵、エネルギー工学消費・省エネルギー、エネルギー利用と環境との関係さらに細かい分類に基づき、ランキングを行なった。結果を Table 4 に示す。

Table 4 を見ると、エネルギー工学分野内の各分類（エネルギー工学、エネルギー資源及び開発、エネルギー変換及び貯蔵、エネルギー工学消費・省エネルギー、エネルギー利用と環境との関係）においても、6分野（工学一般領域、エネルギー工学、熱機関学・応用熱力学、機械工学、運輸工学、鉱山学）と同様に、各中心性でのランキングの上位人物に大きな変動はないことがわかる。

Table 4: エネルギー工学分野の中心性

エネルギー工学			エネルギー資源及び開発			エネルギー変換及び貯蔵			エネルギー工学消費・省エネルギー			エネルギー利用と環境との関係		
近接	次数	媒介	近接	次数	媒介	近接	次数	媒介	近接	次数	媒介	近接	次数	媒介
未広茂	未広茂	未広茂	美濃輪智朗	美濃輪智朗	美濃輪智朗	長屋重夫	長屋重夫	長屋重夫	下田吉之	吉野博	下田吉之	伊香賀俊治	市川徹	伊香賀俊治
小宮山涼一	伊藤浩吉	坂田興	花岡寿明	松村幸彦	小木知子	平野直樹	金子宏	平野直樹	村上周三	下田吉之	井上隆	近本智行	下田吉之	西村郁夫
藤井康正	山地憲治	山本博巳	鈴木善三	近藤昭彦	花岡寿明	式町浩二	玉浦裕	津田理	宮田征門	井上隆	伊香賀俊治	永瀬修	外岡豊	遠藤康之
山本博巳	山本博巳	松村幸彦	柳下立夫	中村真人	天野寿二	後村直紀	平野直樹	長嶋賢	吉野博	村上周三	柳原隆司	柳原隆司	丹羽英治	戸田貴光
坂村幸彦	前川忠	藤井康正	松村幸彦	井田民男	横山伸也	大崎博之	秋山友宏	大崎博之	井上隆	赤司泰義	渡辺健一郎	渡辺健一郎	西村郁夫	伊香賀俊治
坂田興	中田俊彦	小宮山涼一	澤山茂樹	井上貴至	長嶋賢	相良和伸	野村新一	奥宮正哉	奥宮正哉	渡辺俊行	吉野博	村上周三	村上周三	永田勝也
前川忠	藤井康正	坂西欣也	多田千佳	吉川邦夫	近藤昭彦	玉田勉	松本昭	新富孝和	坂本雄三	宮田征門	村上周三	村上周三	戸田貴光	佐土原聡
永富悠	近藤康彦	近藤康彦	柳田高志	藤本真司	岩田光夫	福垣淳二	青田浩幸	村上雅人	高口洋人	高口洋人	宮崎隆彦	宮崎隆彦	坂本雄三	岡建雄
柳沢明	小宮山涼一	柳沢明	藤本真司	坂志朗	袖山義人	中村武恒	式町浩二	嶋田隆一	坊垣和明	奥宮正哉	丹羽英治	丹羽英治	外岡豊	吉田聡
坂西欣也	松村幸彦	美濃輪智朗	井上誠一	笹内謙一	高橋徹	大久保仁	熊野寛之	広瀬誠	伊香賀俊治	伊香賀俊治	坂本雄三	坂本雄三	水石仁	松嶋聖

4.2 ネットワークの評価

現実のネットワークでは、スケールフリー性、スモールワールド性、クラスター性を有することが知られている。8) ここでは、作成した共著関係ネットワークについて、次数分布、平均最短距離、クラスター係数を求め、その評価を行なった。

4.2.1 スケールフリー性

あるネットワークで各ノードの次数 k の分布 $P(k)$ が、べき指数 γ のマイナスのべき乗に比例しているとき、このネットワークはスケールフリーであるという。

$$P(k) \propto k^{-\gamma} \quad (4)$$

エネルギー工学分野の次数分布を Fig. 3, 工学一般領域, エネルギー工学, 熱機関学・応用熱力学, 機械工学, 運輸工学, 鉱山学の6分野全体での次数分布を Fig. 4 に示す。

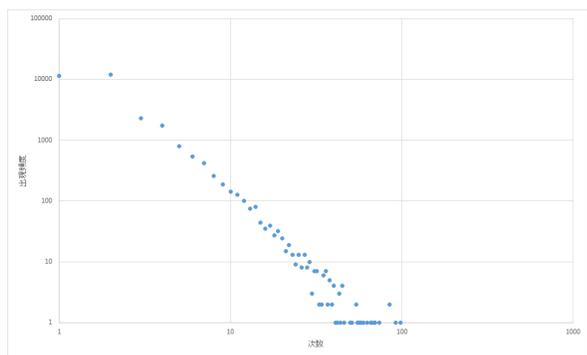


Fig. 3: エネルギー工学分野での次数分布

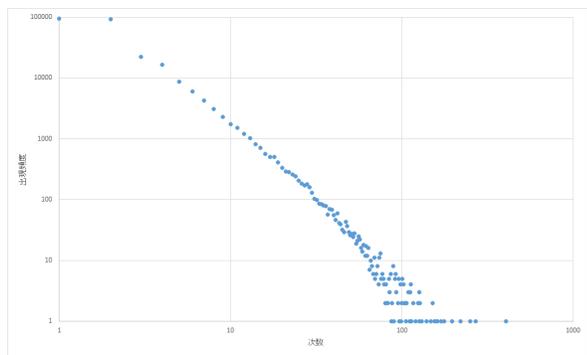


Fig. 4: 6分野全体での次数分布

Fig. 3, Fig. 4 では縦軸を出現頻度、横軸を次数としており、ともに対数をとっている。どちらもほぼ直線を描いており、べき乗則に沿っているといえる。エネルギー工学分野, 6分野全体のどちらの共著関係ネットワークもスケールフリー性を有していることがわかる。

4.2.2 平均最短距離

平均最短距離は、全てのノード間の最短距離の平均を取ったもので、次のように定式化される。

$$L = \frac{1}{N(N-1)/2} \sum_{i>j} d_{ij} \quad (5)$$

ここで、 d_{ij} は頂点 i から頂点 j への最短距離である。

4.2.3 クラスター性

クラスター性とは、友達の友達がやはり友達であるという場合の確率のことで、次のように定式化される。

$$C = \frac{1}{n} \sum_{i=1}^n C_i \quad (6)$$

ただし、

$$C_i = \frac{2m_i}{k_i(k_i - 1)} \quad (7)$$

ここで、 n はノード数、 m_i はノード i に隣接する頂点間のリンクの数、 k_i はノード i の次数である。

4.2.4 ネットワーク密度

ネットワーク密度は、実際のエッジ数を全てのノードが完全グラフを構成した場合のエッジ数で除したものである。

エネルギー工学分野, 6分野全体の共著関係ネットワークの平均最短距離, クラスター係数を Table 5 にまとめた。Table 5 を見ると、エネルギー工学分野, 6

Table 5: 共著関係ネットワークの評価

	ノード数	エッジ数	ネットワーク密度	平均最短距離	クラスター係数
エネルギー工学分野	39279	30445	1.19642E-5	10.74694	0.11736
6分野全体	411508	262278	8.47564E-5	9.91911	0.14241

分野全体で共に、クラスター係数が低いことがわかる。一方で、ネットワーク密度では、エネルギー工学分野のネットワーク密度が6分野全体と比べて約7倍の密度となっている。このことから、異分野間よりも単一分野内の方が共著関係が多く結ばれていることがわかる。

5 まとめと今後の課題

本研究では、JST 科学技術文献データをもとに、工学一般領域, エネルギー工学, 熱機関学・応用熱力学, 機械工学, 運輸工学, 鉱山学の6分野に対して論文の

共著関係ネットワークを作成し、各分野でのネットワーク中心性の高い研究者の特定を行なった。

また、作成した共著関係ネットワークについて、複雑ネットワークとしての評価を行なった。本研究では、共著者が複数存在する場合、共著リスト上で隣り合う著者のみにリンクを張り、ネットワークを作成したためだと考えられる。

今後の課題として、以下のことが挙げられる。

- 多重リンクの扱い。
- 単著の多い研究者の評価。
- 論文数は少ないが重要な論文を出した研究者の評価。
- 時系列情報。

より高精度の分析を行なっていくために、これらの課題に対し今後は共著者ネットワークを評価する指標として、論文の被引用数などを用いた評価を取り入れていく。また、共起回数に応じた重み付けや、時系列情報を取り入れることで、共著関係ネットワークの成長過程についての分析も行なっていく。

本研究では、工学一般領域、エネルギー工学、熱機関学・応用熱力学、機械工学、運輸工学、鉱山学の6分野に限定した分析を行なったが、今後は、さらに多くの分野で分析を行ない、特定分野内での中心人物の特定のみでなく、分野と分野をつなぐ、ハブとしての役割を果たしている研究者の特定を行なっていく。

参考文献

- 1) 芳鐘冬樹, 影浦峽: 共著ネットワークの分析: 視点と手法, 日本図書館情報学会研究大会発表要綱, 第49回, 55/58 (2001)
- 2) 市瀬龍太郎, 武田英明, 植山浩介: コミュニティマイニングのための研究者情報の視覚化, 信学技報, 104.587, 1/6 (2005)
- 3) 浅田 洋平, 松尾 豊, 石塚 満: Web からの研究者ネットワーク抽出の大規模化, 人工知能学会誌, 第20巻6号 (2006)
- 4) 松尾 豊, 石黒 周, 松原 仁, 橋田 浩一, 中島 秀之: 効果的な共同研究を支援するため Web からの研究者ネットワーク抽出, 情報学シンポジウム講演論文集 (2005)
- 5) 篠田 孝祐: 日本における人工知能研究の系譜, 人工知能学会誌, 26巻6号 (2011)
- 6) 安田 雪, 松尾 豊: 人工知能学会における研究者ネットワークの分析, 第19回人工知能学会全国大会, 2A3-02 (2005)
- 7) 内藤理, 佐藤啓宏, 工藤俊亮, 池内克史: 日本におけるロボット工学の研究者ネットワークの分析, 日本ロボット学会誌, 第30巻-6号, 629/638 (2012)
- 8) 増田直紀, 今野紀雄: 複雑ネットワークの科学, 126/127, 産業図書 (2005)