

共著者ネットワークをもとにした 共同研究者探索支援システムの試作

吉田光男

豊橋技術科学大学

yoshida@cs.tut.ac.jp

概要 人工知能に関する旬な研究者を探索するために、人工知能学会や言語処理学会などの予稿集、論文誌のデータを収集し、ある著者の発表歴を追うためシステムを試作している。本稿では、その試作の為に収集した共著者ネットワークの変化について報告する。

キーワード 共著者ネットワーク, ネットワーク分析, 人工知能

1 はじめに

大学教員をはじめとする研究者は、その研究活動を一人で行なうわけではなく、様々な研究者と交流し、共同で実施している。このような共同研究は、大学間の研究者に限定されるわけではなく、大学と企業との産学連携による実施もある。文部科学省による平成 28 年度の調査では、全国の大学と企業との共同研究は、23,021 件にものぼる [1]。この調査によれば、企業による納入額が年間 300 万円未満のものが 84% (平成 28 年度納入分) を占めると報告されている。つまり、諸外国に比べて、比較的小規模に実施されているといえる。

共同研究に限らず、企業は製品開発などのために、外部の知識を求めている。科学技術・学術政策研究所は、2013 年から 2015 年度の研究活動において他組織との連携を実施した企業は 73.4% にのぼると報告している [2]。一方、情報サービス業においては 56.0% であり、比較的低いことが指摘されている。また、情報サービス業においては、現状、情報源として論文を重視するものの、学会での研究成果発表は重視されていない。つまり、情報サービス業においては、研究者との直接的な交流が十分に発展していない可能性が示唆される。

本研究では、企業が共同研究者を探すシステムについて検討する。特に、近年、社会的に注目を集めている人工知能研究を主な対象とし、ある旬な研究の中心的人物を探すこと、および、その研究者にアクセスするための周囲の研究者を探すことに着目する。共著者ネットワークの分析はこれまでも数多く行なわれており [3]、そのネットワークの有用性は明らかになっている [4, 5, 6, 7]。また、共著者ネットワークを利用し、日本の人工知能研究の系譜を作る試みも行なわれている [8]。本稿では、共同研究者を探すシステムの試作のために収集したデータについて報告する。

表 1 取得データの詳細 (人工知能学会全国大会の発表数)

| 年 | 論文数 | 著者数 | |
|------|------|-------|------|
| | | 重複あり | 重複なし |
| 2009 | 402 | 1318 | 879 |
| 2010 | 415 | 1280 | 891 |
| 2011 | 428 | 1338 | 943 |
| 2012 | 516 | 1634 | 1086 |
| 2013 | 672 | 2064 | 1356 |
| 2014 | 620 | 1908 | 1286 |
| 2015 | 651 | 2067 | 1352 |
| 2016 | 701 | 2200 | 1482 |
| 2017 | 748 | 2362 | 1552 |
| 合計 | 5153 | 16171 | 6157 |

2 データの収集

本稿では、2009 年から 2017 年までの人工知能学会全国大会の書誌情報を収集した。この書誌情報は kaigi.org¹ で公開されており、2018 年 5 月時点で、誰もが無料でアクセスすることができる。取得したデータについて、年別に表 1 に示す。著者数の重複ありは共著者全てを数え上げた延べ数であり、重複なしは対象年の著者のユニーク数を算出したものである。なお、重複なし著者数の合計については、全ての年における著者にユニーク数を算出したため、表中の合計とは一致しない。

2009 年から 2017 年まで使われていた人工知能学会全国大会の講演プログラムのシステムでは、それぞれに著者にユニークな ID が付与されている。しかしながら、筆者が確認したところ、同一著者が複数の ID に分かれている、同組織等における異なる著者で年別に ID を使い回しているケースがあったため、今回は、著者の集計においては、著者名をもとにした。そのため、姓名の表記が同じ著者がいる場合、その区別はされていない。

3 データの集計

まず、著者ごとに年別の発表件数を集計した。その結果を表 2 に示す。ここでは年別の上位 10 名の著者を求

Copyright is held by the author(s).

The article has been published without reviewing.

¹<https://kaigi.org/jsai/webprogram/>

表 2 年別の発表数上位 10 の著者一覧 (同一順位が 10 位をまたぐ場合は 10 位以内に切り捨て)

| 2009 | | 2011 | | 2013 | | 2015 | | 2017 | |
|--------|----|--------|----|--------|----|--------|----|--------|----|
| 著者 | 件数 | 著者 | 件数 | 著者 | 件数 | 著者 | 件数 | 著者 | 件数 |
| 本村 陽一 | 12 | 鷺尾 隆 | 7 | 高橋 達二 | 17 | 松尾 豊 | 15 | 松尾 豊 | 20 |
| 西田 佳史 | 11 | 西田 豊明 | 7 | 伊藤 孝行 | 16 | 長井 隆行 | 12 | 鳥海 不二夫 | 11 |
| 溝口 理一郎 | 11 | 小林 一郎 | 7 | 鳥海 不二夫 | 14 | 鳥海 不二夫 | 11 | 福田 直樹 | 10 |
| 大和田 勇人 | 9 | 山口 高平 | 7 | 松尾 豊 | 12 | 高橋 達二 | 11 | 白松 俊 | 10 |
| 松井 藤五郎 | 9 | 沼尾 正行 | 7 | 野田 五十樹 | 9 | 榊 剛史 | 11 | 池上 高志 | 10 |
| 廣安 知之 | 8 | 溝口 理一郎 | 7 | 三木 光範 | 9 | 福井 健一 | 10 | 長井 隆行 | 10 |
| 竹林 洋一 | 8 | | | 諏訪 正樹 | 8 | 中村 友昭 | 10 | 森 純一郎 | 10 |
| 三木 光範 | 8 | | | 小林 一郎 | 8 | 篠田 孝祐 | 10 | 吉川 大弘 | 10 |
| 石塚 満 | 8 | | | 溝口 理一郎 | 8 | 小林 一郎 | 10 | 伊藤 孝行 | 10 |
| | | | | 今井 倫太 | 8 | 栗原 聡 | 9 | | |

表 3 年別の PageRank 上位 5 の著者一覧

| 2009 年 | 2013 年 | 2017 年 |
|--------|--------|--------|
| 小方 孝 | 伊藤 孝行 | 松尾 豊 |
| 本村 陽一 | 鳥海 不二夫 | 鳥海 不二夫 |
| 石塚 満 | 高橋 達二 | 荒井 幸代 |
| 溝口 理一郎 | 松尾 豊 | 白松 俊 |
| 竹林 洋一 | 三木 光範 | 栗原 聡 |

め、同一順位によって 10 名をまたぐ場合、10 名以下になるように切り捨て、また、紙面の制限により奇数年のみを記載している。表の通り、発表数の多い著者は年ごとに入れ替わっており、旬な研究者を抽出する場合には、使用するデータの期間を適切に調整する必要があることが分かる。なお、対象 9 年間の総発表件数の上位 5 名は、松尾豊 (84 件)、鳥海不二夫 (75 件)、小林一郎 (69 件)、溝口理一郎 (62 件)、栗原聡 (57 件) であった。

次に、共著者ネットワークを構築し、著者の PageRank スコアを算出した。その結果を表 3 に示す。ここでは年別の上位 5 名の著者を求め、また、紙面の制限により一部の年のみを記載している。共著者ネットワークは、共著者間に完全グラフを作り、複数の共著関係があったとしてもエッジに重みを付けずに構築している。多くの場合、上位の著者は表 2 と一致するものの、一部のユーザは PageRank スコアで高く評価されており、発表件数で漏れている著者を抽出することができている。ただし、年別の共著者ネットワークは図 3 のように疎なネットワークであることに留意する必要がある。なお、対象 9 年間で共著者ネットワークを構築した場合、PageRank スコア上位 5 名は、鳥海不二夫、松尾豊、山口高平、小林一郎、伊藤孝行であった。

4 おわりに

本稿では、人工知能関係の共同研究者を探索するシステムの構築を目指し、そのためのデータについて報告した。人工知能学会全国大会における 2009 年から 2017 年の書誌情報を取得し、発表数の多い著者、共著者ネットワークにおける PageRank スコアの高い著者を確認したところ、年ごとに変化が大きいことが分かった。

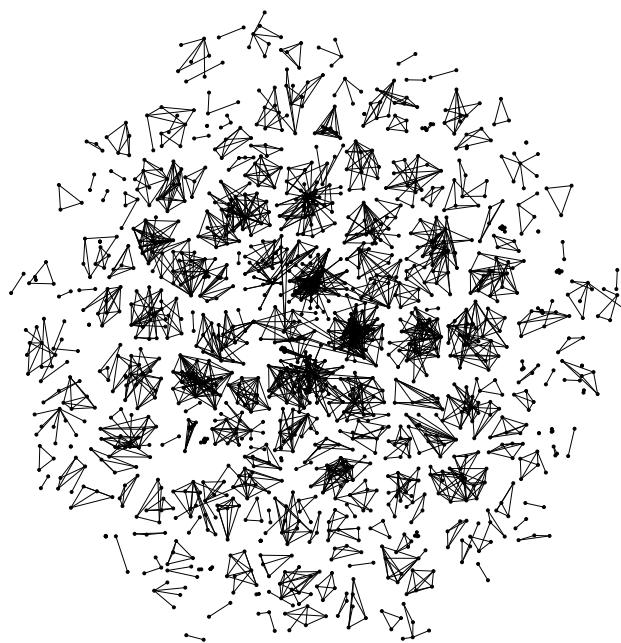


図 1 2017 年の共著者ネットワーク

参考文献

- [1] 文部科学省：平成 28 年度 大学等における産学連携等実施状況について (2018).
- [2] 科学技術・学術政策研究所：民間企業の研究活動に関する調査報告 2016 (2017).
- [3] Kumar, S.: Co-authorship networks: a review of the literature, *Aslib Journal of Information Management*, Vol. 67, No. 1, pp. 55–73 (2015).
- [4] Biscaro, C. and Giupponi, C.: Co-Authorship and Bibliographic Coupling Network Effects on Citations, *PLoS ONE*, Vol. 9, No. 6, p. e99502 (2014).
- [5] Mena-Chalco, J. P., Digiampietri, L. A., Lopes, F. M. and Cesar, R. M.: Brazilian bibliometric coauthorship networks, *Journal of the Association for Information Science and Technology*, Vol. 65, No. 7, pp. 1424–1445 (2014).
- [6] Uddin, S., Hossain, L. and Rasmussen, K.: Network Effects on Scientific Collaborations, *PLoS ONE*, Vol. 8, No. 2, p. e57546 (2013).
- [7] Uddin, S., Hossain, L., Abbasi, A. and Rasmussen, K.: Trend and efficiency analysis of co-authorship network, *Scientometrics*, Vol. 90, No. 2, pp. 687–699 (2012).
- [8] 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， Vol. 26, No. 6, pp. 584–589 (2011).