

イベント空間情報支援システムが抽出するヒューマンネットワークの構造

Structure of Human Network Extracted by the Event Space Information Support System

安田 雪

Yuki Yasuda

*1 東京大学大学院経済学研究科

Specially Appointed Faculty Member for the 21 COE, The University of Tokyo

My purpose here is (1) to investigate the structure of the Japanese AI researchers' network to identify the Invisible College of AI research and (2) to consider effects of their structural position within the network on their productivity. Findings suggest groups of researchers constitute an Invisible College, and that AI researchers' productivity highly correlates with their structural position within the network.

1. 目的

本報告の目的は、イベント空間情報支援システム[西村 03]が抽出した人工知能研究者のヒューマンネットワークの構造を分析し、人工知能学会における研究者のネットワーク内での位置特性と研究業績との対応関係を分析することである。

学問の各領域においては、先端的な研究を担う一部の優れた研究者の集団が存在し、「見えざる大学」(Invisible Colleges)を形成されていることが、科学社会学者によって指摘されて久しい[Crane72][Burt 82]。見えざる大学とは、大学や研究所などの制度的な組織に限定されることなく、空間的に拡散しながらも高い生産性を上げ続け、最先端の研究を担うことによって優れた研究者として認知されている人々の集団を指す。

現実には、研究動向や学会情報に詳しい者であれば、誰が見えざる大学の成員であるのかは、直感的に理解できることが多い。しかし、専門領域外の人間が見えざる大学の構成員を知るのは難しい。さらに、見えざる大学内部の者と外部の者を分ける境界は曖昧であり、学界事情に精通した者でもその境界線を引くことはきわめて困難である。

本報告では、イベント空間情報支援システムによって抽出された日本の人工知能研究者の協働関係の構造を分析し、抽出されたデータの制約下において、(1)人工知能研究者がもつネットワークの構造特性を明らかにし、(2)人工知能研究者のネットワーク内での位置特性と研究業績との関連性を検討し、我が国の人工知能研究者の「見えざる大学」を抽出することを試みる。

2. データと抽出方法

分析対象は、松尾ら[松尾 03]が開発したウェブマイニングの手法により、収集・構築された「人間関係ネットワーク」である。この人間関係ネットワークは、人工知能学会に属する研究者のうち298人の研究者間の協働関係をweb上の共起関係により抽出したものである。

基本的なデータ抽出手順は、以下のとおりである。まず、所属組織などで同定制限をつけた研究者の氏名を用い、単独の研究者名をキーワードにして検索エンジンが抽出するページ数を特性する。同様に二人の研究者のペアをキーワードにして検索エンジンが抽出するページ数を特定する。これらの共起数に一連の制約を加え、研究者のダイアドごとにシン普森係数を算出する。シン普森係数が、一定の閾値以上の場合に、研究

者間に関係があるとみなす。データ抽出の詳細については、[松尾 04]を参照して欲しい。

人間関係ネットワークは研究者が相互にとりもつ協働関係を複数の文脈でとらえている。Web上の氏名の共起関係に、共起するページの特徴を加味することにより、共著論文執筆、共同研究室所属、共同プロジェクト参加の3種のネットワークが抽出されている。マイニング技術の制約により、共同プロジェクト参加など再現性の低いネットワークと、共著論文執筆など再現性の高いネットワークの双方が含まれている。なお、再現性の確認にあたっては、パーソナルネットワークについてWeb上でアンケートを実施し、その回答との照合も現在進められている。

業績データは、国立情報学研究所の論文検索サービス[CiNii04]により検索された全文献数、2003年までの人工知能学会誌文献数、それぞれのページ数及び被引用数である。業績データ、共著者、所属組織データは松尾ら[松尾 04]による。

3. 研究者の関係構造のネットワーク分析

本論文では、人工知能研究者の「人間関係ネットワーク」に対して、中心的な研究者は誰か、ネットワーク内で似通った位置を占めている研究者は誰かをネットワーク分析の指標を用いて特定する。

3.1 ネットワーク分析とは何か

ネットワーク分析は、人々がもつ関係構造をネットワークとしてとらえ、ネットワーク全体の特徴とネットワーク内部において個々人が占める位置の特性を定量的に分析する手法である。関係を形成する要素をノード(Node)、要素間の関係を紐帯(Tie)としたグラフとすることで、関係を可視化して内部構造を分析する。人間間関係、企業間関係、産業間関係、国家間関係など、分析対象のミクロとマクロを問わず同一の手法で分析できる。グラフ理論的なアプローチを用いることにより、構成要素の属性的な要素を考慮せず、要素間の関係構造のみによって規定される情報を得ることが可能になるためである。

関係のパターンが規定するネットワーク指標は、グラフに描かれる関係を行列に変換したうえで計算され、定量的に算出される。中心性、結合関係、構造同値性などが代表的なものだが、それぞれから、権力構造、派閥の構造、競合関係を読みとることができる[安田 01]。

このような関係特性に注目するアプローチは、性別、年齢などの個人の生得的な要因だけでなく、むしろ個人がもつネットワークこそが、その人の行為、嗜好、パフォーマンスなどに大きな影響

連絡先 : 安田雪, 東京大学大学院経済学研究科,
yasuda@mmrc.e-u.tokyo.ac.jp

をあたえる要因となると考える構造社会学のものである[Burt82]。属性要因とネットワーク要因のいずれが、行為者の行為に大きな影響力をもつのかは二者択一の問題ではなく、現実には両者が複雑に作用しあって行為の決定要因となっていることが多い。しかし、説明変数に生得的属性のみを用いて、確率的に被説明変数に対する影響力を論じる従来のアプローチにも限界はある。個人が埋め込まれた関係構造を精査することによって初めて浮かびあがる関係的な要因もあり、行為の予測や行為者の理解には、これらの関係情報が不可欠なのである。

3.2 研究者のネットワークの重要性

研究者のネットワークに注目する理由は、研究者個人のミクロレベルと、学会全体のマクロレベルの双方にある。

研究者は長い研究生活において多くの研究者と出会う。誰を恩師として仰げるのか、誰の共同研究者になれるのか、研究会やプロジェクトでいかなる研究者と議論を交わせるのか。日常の研究生活や学会活動における出会いは、研究者相互に大きな影響を与える。とりわけ科学者としての方向性や規範が十分確立していない段階の若年研究者にとっては、誰と出会うかによってその後の研究者人生の方向が大きく異なる。個人レベルでも、協働関係のありかたは、研究者個人の方向性、生産性、ライフコースを決定づける重要な要因なのである。

学会という集団レベルで考えるならば、学界内部の研究者クラスターのありかたは、学術研究の動向や資源配分メカニズムと密接に関わりあっている。それぞれの学問領域においては、研究テーマ、投稿すべき学術雑誌、参加すべき学会や研究会活動、投入される研究予算額などについての規範が、その時点で最も権力のある一連の研究者集団によって規定される[Crane72]。極端に言えば、研究者ネットワークの権力構造は、科学における真実とは何か、研究の意義や適切性にまで影響を及ぼす。個人と集団双方のレベルにおいて、研究者のネットワークは看過できない影響力を科学研究に与える。

4. ネットワーク指標と研究業績

4.1 ネットワークにより異なる構造と中心的な研究者

人間関係ネットワークでは、共著ネットワーク、共同研究室所属、共同プロジェクトの種類のネットワークが抽出されている。この3種のネットワークは、いずれも同じ298人のメンバーとその相互関係によって形成されている。とはいえ、内部の関係構造は各ネットワークによって非常に異なる。

各ネットワークの紐帯数は、共著ネットワーク(1294本)、共同研究室所属(1660本)、共同プロジェクト(426本)であり、紐帯の分布はネットワークによって異なっている。ネットワーク密度は、共著ネットワーク(0.0146)、共同研究室ネットワーク(0.0188)、共同プロジェクトネットワーク(0.0274)である。共同プロジェクトネットワークの密度が最も高く、共著ネットワークの密度が最も低いのは、プロジェクトの共同関係から始まる関係が、共著論文に至る過程で多数淘汰されていくためであろう。研究者にとって共著関係を結ぶには、共同プロジェクト関係を形成するよりも強いコミットメントが必要である。共同プロジェクトは複数参加が比較的容易であるため、多数の重層的な関係が共同プロジェクトネットワーク内には存在し、共著ネットワークよりも多くの相互結合関係のサブグループが内包されている。

複数のネットワークの関係パターンの類似度を測る QAP 相関を算出すると、共著ネットワークと共同研究室所属ネットワークの相関計数は 0.534 ときわめて高い。共著ネットワークと共同プロジェクトネットワークの相関係数、共同研究室所属と共同プ

ロジェクトネットワークの相関係数は両方とも 0.099 であり、著しく低い。共同プロジェクトネットワークは、他とは異質な構造をもつ。

次に各ネットワークにおいて中心的な研究者(群)を特定するために中心性を算出する。中心性はネットワーク内で中心的な存在を特定する指標である。保持する紐帯数を数える次数、構成者全員への到達しやすさを計量する近接性、他者との関係を連結させるために不可欠な位置にいるかどうかを考慮する媒介性、他者との間接的結合関係の多寡を考慮する固有ベクトルによる中心性などの指標がある。

分析するネットワークと用いる指標により、研究者の中心性は微妙に変化する。共著ネットワークと共同研究室ネットワークの上位者はやや共通性が見られるが、前者で中心的な位置を占める研究者が必ずしも後者でも中心的であるとは限らない。また、共同プロジェクトネットワークの中心的な人物は、前二者のネットワークの中心的な人物とは異なっている。「名門大学の大御所」とみなしうる研究者が、全ネットワークにおいて中心的な位置を占めているわけでもない。一方、複数のネットワークの多くの指標において安定的に中心性の高い研究者も存在する。

中心性スコアと研究業績には高い相関が見られる。とりわけ共著ネットワークと共同研究室ネットワークの媒介性と全文の被引用数、人工知能学会誌における被引用数は、一貫して正の高い相関を示す。複数の研究室や共著関係を媒介させる役割をもつ研究者が、被引用頻度が高い。研究者がネットワーク内で占める位置と研究業績には密接な関連があるのだ。

4.2 見えざる大学の可視化

共著論文ネットワークに対して構造同値性に基づいてクラスター分析を行い、その関係パターンの類似性によって 298 人の研究者を 8 グループに分類した。グループの大きさは一定ではないが、グループごとに研究業績の量や所属などで特徴がある。また、グループ内及びグループ間の関係にも差があり、内部の相互関係が強いグループ、他者との関係を持たない研究者のグループなどが特定できる。

この一部のグループを構成する研究者こそが、「2003 年の Web 情報から抽出されたネットワークにおける見えざる大学」の構成者である。これらの一部の研究者によって、現在の人工知能学会の研究や論文生産の最先端が担われているものと考えられる。各研究者の中心性スコア、業績、クラスター分析の所属グループの詳細は当日に述べる。

参考文献

- [Burt 82] Burt, Ronald: *Toward a Structural Theory of Action*, Academic Press, 1982.
- [CiNii04] CiNii: NII 引用日期文献情報ナビゲータ, <http://ci.nii.ac.jp/cinii/servlet/CiNiiTop>.
- [Crane72] Crane, Diana: *Invisible Colleges*, University of Chicago Press, 1972.
- [松尾 03] 松尾豊, 友部博教, 橋田浩一, 石塚満: Web から人間関係ネットワークの抽出と情報支援, 第 17 回人工知能学会全国大会, 1F1-02 2003.
- [松尾 04] 松尾豊, 友部博教, 橋田浩一, 中島秀之, 石塚満: Web 上の情報から人間関係ネットワークの抽出, 人工知能学会誌 (投稿中).
- [西村 03] 西村拓一, 濱崎雅弘, 松尾豊, 大向一輝, 友部博教, 武田英明: 2003 年人工知能学会全国大会支援統合システム, 人工知能学会誌解説記事, Vol. 19 No. 2, 43-51, 2004.
- [安田 01] 安田雪: *実践ネットワーク分析*, 新曜社, 2001.