

3.4 ネットワーク論区分

3.4.1 複雑ネットワークおよび総論

(1) 研究開発領域名

複雑ネットワークおよび総論

(2) 研究開発領域の簡潔な説明

統計物理学、確率論、シミュレーション工学などの理論的技術を用いた、現実中存在する生物、社会、工学などのネットワークの構造特徴抽出、ネットワークの構造やネットワーク上のダイナミクスの理解や予測、ネットワークの制御。

(3) 研究開発領域の詳細な説明と国内外の動向

本稿で述べる研究分野はネットワーク科学、ないし、複雑ネットワーク研究、と一般的に呼ばれ、本俯瞰区分の要に位置する。多数の要素が相互作用していることはシステム科学の対象であることの必要条件である。その中で、ネットワークとは、要素2つの相互作用の集合体として表されるようなシステムを指す。

人の組織というシステムについては、友人関係の織りなす網目、より具体的にはFacebookにあるような2者間の友人関係の集合体がネットワークの例である。環境というシステムにおいては、捕食—被食関係にある2つの生物種を結び、そのような捕食—被食関係とそれにかかわる生物種を集めたものがネットワークの例である。逆に、ネットワークという視点においては、要素2つの相互作用の集合体に分解できないものは扱わない。例えば、3人以上のグループでの友人関係は、グループ内の各2人の対人関係の集まりであると見なして解析を行う。3体以上の直接相互作用が本質的な役割を果たすシステムの取り扱い、他の方法論に委ねられる。

ネットワークとは離散数学でいうグラフと同値であり、グラフ理論は近代的研究に限っても、半世紀以上の歴史がある。これに対して、ネットワーク科学は、1998年を境に勃興した新しい分野である。グラフ理論と対象を同じくするにもかかわらず、異なる研究分野として認識され、2015年時点で多くの研究者が研究に参画するようになった理由は、主に2つある。

一つ目は、ネットワーク（すなわち、グラフ）を1998年頃までは数学的に扱うのが唯一の理論的手法であったのが、統計物理学者や数学をベースとしない工学者の研究への参画によって、厳密さに必ずしもこだわらずに、現実のデータを解析する手法が発達したことである。具体的には、マスター方程式、平均場近似、有限サイズスケーリングといった物理学的手法、ネットワークの生成やその上の感染症伝搬などのダイナミクスを解析するための数値計算の技術などが積極的に現実のネットワークの解析に用いられるようになった。それにより、多くのネットワークが頂点の次数が裾の長い分布に従うというスケールフリー性をもつこと、スケールフリー・ネットワークでは感染症伝搬が極度に起こりやすいこと、などが次々と明らかにされた。

二つ目は、1998年頃に、種々のネットワークのデータが一般の研究者に供用されるようになってきたことである。インターネットの発展や、計算機の飛躍的な性能向上がこ

の時期に起こった。このことは、分野横断的にネットワークの特性を比べることをも可能にした。例えば生態学における食物網、工学における電気回路や送配電網、社会学における人間関係ネットワークなどは、各個別分野においては 1998 年より何十年も前から研究が行われていたが、分野内の個論にとどまっていた。こういったデータをインターネット経由で黎明期のネットワーク科学研究者が使用できるようになりつつあったことに加えて、インターネット、ウェブグラフ、遺伝子発現制御ネットワーク、神経ネットワーク、電力網などのネットワークのデータが、インターネットや高速になった計算機を援用しながら、採取され、整理された。そして、分野横断的にネットワークを解析することによって、スケールフリー性に代表される普遍性や、分野別の個別性が明らかにされてきたのである。

2000～2005 年はネットワーク生成モデル、感染症伝播、中心性（頂点の重要性を定めて応用すること）、コミュニティ構造（計算機科学、離散数学などの分野ではグラフ分割とも呼ばれる）などが主要なテーマであった。それ以降は、本俯瞰区分の研究開発領域「機械学習・データマイニング分野におけるネットワーク構造解析」とも関係するネットワーク構造推定などの理論のさらなる発展とともに、感染症制御、生態系保全、人間の移動を考慮したネットワーク、電力網のコントロール、創薬などの研究が発展している。ネットワーク研究の理論発展に寄与したような確たる基礎理論を身につけた研究者が、これらの応用研究に参画していることが特徴的である。すなわち、理論と応用が異なる人材で担われている分野と事情が異なる。

国際比較としては、本領域を創成した 2 本の論文（1998、1999 年）の発祥でもある米国が圧倒的に他国をリードしている。その片方の著者である Barabási のチームを含む多くの強力な研究チームがボストンに存在し、シカゴも人材の蓄積が厚い。半世紀以上続く社会学のネットワーク研究も米国が中心であること、多くのインターネット企業が米国を本拠地にしていることも、米国が基礎研究から産業化フェーズまでの広きにわたって世界の研究フロンティアを牽引していることに寄与している。欧州出身の代表的な研究者も、その多くが米国で研究室を開設している。欧州はそれに次ぎ、2000 年代前半よりイタリア、スペイン、ハンガリー、フィンランドなどで研究が比較的盛んである。近年はイギリス、フランス、ドイツといった一般的に研究リソースが多い国に人材が移動している。アジアでは、基礎研究に限っては韓国が高いレベルを維持している。

（４）科学技術的・政策的課題

- ・ 基礎研究は 1998～2010 年頃の間非常に発展した。次のフェーズ、すなわち、今後 10 年程度での応用研究、産業化、あるいは、それに必要な種類の理論の構築の成否が、さらに長いスパンでネットワーク科学という研究分野がシステム科学技術分野の中において、さらには、科学技術全体の中において重要たり得るかどうかの分水嶺である。過去のカオスや複雑系の研究にも同様のフェーズの遷移があった。しかし、海外に比して、確かな基礎力をもちつつ応用や産業化に尽力する国内の人材は極めて少ない。
- ・ 特にインターネットや社会ネットワークの領域において、一般企業は多くのデータを有している。応用研究や産業化のもうひとつの障壁として、こういったデータや企業

側の問題意識に立脚するような産学連携が国内においては遅れていることがあげられる。多くの日本企業がかつてのような経済的・人的余裕をもたないこと、個人情報取り扱いについて匿名化などを施した上でも風当たりが強い日本の国民性などが主要な原因であると推測される。これに比して、海外では、携帯電話の通信データが研究用に供用されているなど、産学連携が進んでいる。

- ・ 日本学術振興会の分野区分などからもわかるように、日本の研究分野の予算配分、人的リソースの配分、研究教育制度は、既存の学問分野に基づく縦割りで行われる部分が多い。一方、ネットワーク科学は、扱うネットワークの対象が多岐にわたることからもわかるように、極めて学際的な分野である。学際的な分野を支援する枠組みは、存在するものの十分でない。審査側への外国人の登用などを含めた抜本的な対策が必要であると思われる。

（５）注目動向（新たな知見や新技術の創出、大規模プロジェクトの動向など）

- ・ 2011年11月に、Facebookと大学の研究者が協同して、Facebookのネットワーク構造解析結果を論文発表した。マーケティング、ユーザ利便性の向上などへの応用が見込まれる。
- ・ RFID（電波による個体識別）機器を人間に装着してもらい、人間同士の動的な社会ネットワーク（誰と誰が、いつ、どのくらいの長さ相互作用したか）を時空間的に大規模に計測するシステムが、2008年以降に、まずは基礎研究として調査され、病院への導入などの応用が模索されている。こういった型のネットワークはテンポラル・ネットワークと総称され、理論、応用の両面から現在の注目動向である。
- ・ 2010年にBoston大学Stanleyらのグループは、単独では障害耐性の高いネットワークが互いに接続しあうと脆弱になり、壊滅的な障害が起こりやすくなるという理論研究を論文発表した。これが契機となり、複数のネットワークが相互作用するネットワーク（マルチプレックス・ネットワーク）の研究が現在盛んである。
- ・ Northeastern大学のA.-L. Barabásiは、ネットワーク科学を創始する論文のひとつを1999年に著し、現在に至るまでネットワーク研究の第一人者であり続けている。その予算規模は公開資料からは明らかでないが、ネットワーク的に結合したシステムの制御性やネットワーク理論を使った創薬の研究などがBarabásiチームの現在のプロジェクトとされている。
- ・ 日本では2012年10月よりJST ERATO河原林巨大グラフプロジェクトが発足した。ネットワーク科学及び本研究俯瞰区分の研究開発領域「機械学習・データマイニング分野におけるネットワーク構造解析」、「ネットワークに関する離散数学」に関わる研究者が集まり、質の高い理論研究が発信されている。

（６）キーワード

スケールフリー、スモールワールド、べき則、中心性、コミュニティ構造、感染症、生態系、脳、遺伝子発現、インターネット、ウェブ、ソーシャルネットワーク

（7）国際比較

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	○	↗	<ul style="list-style-type: none"> 理論研究や実データを用いた基礎研究が国内で行われるようになってきたのは、2004年頃であり、欧米に大きく遅れをとっている。そのひとつの証左として、ネットワーク研究の代表的な総説論文¹⁾に引用されている日本発の論文数はわずかである。 2007年頃以降、研究者の数は非常に増えており、日本物理学会、情報処理学会、電子情報通信学会などの中にネットワークを冠した学会のセッション、研究会組織などが存在し、定期的に研究会を開催している。 2012年10月よりJST ERATO 河原林巨大グラフプロジェクトがスタートし、コンピュータ科学の研究者の参入が進んでいる。
	応用研究・開発	△	→	<ul style="list-style-type: none"> 日立グループが、世界に先駆ける技術で、人の職場における動きや、対人相互作用を高精度で記録するシステムを開発した。国内外の大学にデータを供与したり、国内の企業への試験的な導入（一部は商品化）を行ったりしている²⁾。
	産業化	△	→	<ul style="list-style-type: none"> サイジニア社が、複雑ネットワーク理論に基づいて、アマゾンなどの購買サイトであるような、商品を買ったときに他の商品を薦めるレコメンデーションシステムを商品化し、国内外から高く評価されている³⁾。 職場などの人間関係ネットワーク解析を請け負うコンサルティングサービスは、過去10年程度の間いくつかの会社が提供した模様である。しかし、特筆すべき成功事例はない。顧客の収益化の見通しのないまま商品化を行ってしまったことが原因であると思われる。
米国	基礎研究	◎	→	<ul style="list-style-type: none"> 1998年と1999年にそれぞれ1本ずつ発表された論文⁴⁾がネットワーク科学の幕開けであると世界的に認識されている。これらの論文は米東海岸発であり、以来、米国は本分野の基礎研究を行う研究者の質量において、他国を圧倒している。 ネットワーク科学を代表する国際会議のシリーズが、米国を母体として毎年開催されていて、毎回数百人の研究者が集う⁵⁻⁷⁾。 Northeastern大学⁶⁾、Indiana大学⁷⁾などに、ネットワーク科学の研究センターがある。
	応用研究・開発	◎	→	<ul style="list-style-type: none"> ネットワークの基礎研究の礎をつくったような研究者が、遺伝子システム、感染症の制御などについて、大型プロジェクトを運営している^{6, 7)}。 Google社の検索エンジンのアルゴリズムに近いeigenfactorが、論文誌のランキングについて、既存の有名な手法（Impact Factor）の弱点を克服するものとして開発され、世界中の学術機関で使われているWeb of Scienceというデータベースにも実装されている⁸⁾。 シカゴにおいて、チーム作業の生産性などといった独創的な応用研究が進められている⁹⁾。
	産業化	◎	↗	<ul style="list-style-type: none"> ネットワークの理論を用いて、企業組織のネットワークをコンサルティングし、生産性の向上を目指すプロトコルが、商品化されている¹⁰⁾。 Googleの検索エンジンには、1998年に西海岸で発明されたページランクというアルゴリズムが搭載されている。ページランクはその後のGoogleの発展を牽引した。産業化における、ネットワークの基礎理論の最大の成功例として認知されている。 Facebookの中には、ネットワークの三角形を増やすようユーザーに推薦する機能など、ネットワーク科学の知見に基づく仕組みが搭載され、コミュニケーションの円滑化や拡大に間接的に貢献している。
欧州	基礎研究	◎	→	<ul style="list-style-type: none"> 特にスペインとイタリアにおいて、統計物理学の背景をもつ多くの研究者が基礎研究を行っている。ハンガリーとフィンランドも、他の分野の平均に比べて本分野における基礎研究が盛んである。近年は各国からイギリス、フランス、ドイツなどへ有力研究者の移動がある。 ネットワーク科学に関する（数百万ユーロから多いもので2000万ユーロを超える規模の）大規模プロジェクトが複数ある¹¹⁾。 スペインのBascompteのグループは、食物網の解析、その生態学的機能などの研究において、世界をリードしている¹²⁾。

	応用研究・開発	○	↗	<ul style="list-style-type: none"> ・米国の研究者と共同して、いくつかの欧州の国における携帯電話による社会ネットワークのデータの取得と解析が2007年以降行われている¹³⁾。災害時の移動、マーケティング、感染症制御などへの応用可能性がある。 ・フランスの携帯電話事業者が、自社の提供する携帯電話通話記録データを用いた研究のコンペティションを主催している¹⁴⁾。
	産業化	△	→	<ul style="list-style-type: none"> ・世界的に知られるような産業化の動きはなし。
中国	基礎研究	×	↗	<ul style="list-style-type: none"> ・国内の研究者評価システムが論文の数に偏っているためか、質の低い論文が大量に論文誌に現れている現状である。ただし、状況は少しずつ改善しているため、今後は上昇が見込まれる。
	応用研究・開発	×		<ul style="list-style-type: none"> ・応用研究・開発の動きなし。
	産業化	×		<ul style="list-style-type: none"> ・産業化の動きなし。
韓国	基礎研究	○	→	<ul style="list-style-type: none"> ・優秀な部分を含む統計物理学者のかなりの割合がネットワーク研究を行っている。そのため、上位の学術誌には、日本よりも多くの基礎研究成果が発表されている。 ・米国のボストンやシカゴにある世界的なネットワークの研究センターとの間に太いつながりが以前からあり、継続的に人員の交流が行われている。そのため、最先端の情報が入ってきやすい。 ・統計物理学関係以外のネットワークの基礎研究はあまり盛んではない。
	応用研究・開発	△	→	<ul style="list-style-type: none"> ・韓国最大のsocial networking serviceであるCyworldと大学が提携して、その構造解析などを行っている¹⁵⁾。ただし顕著な後続研究は見当たらない。
	産業化	×		<ul style="list-style-type: none"> ・産業化の動きなし。

(註1) フェーズ

基礎研究フェーズ：大学・国研などでの基礎研究のレベル

応用研究・開発フェーズ：研究・技術開発（プロトタイプの開発含む）のレベル

産業化フェーズ：量産技術・製品展開力のレベル

(註2) 現状

※我が国の現状を基準にした相対評価ではなく、絶対評価である。

◎：他国に比べて顕著な活動・成果が見えている、○：ある程度の活動・成果が見えている、

△：他国に比べて顕著な活動・成果が見えていない、×：特筆すべき活動・成果が見えていない

(註3) トレンド

↗：上昇傾向、→：現状維持、↘：下降傾向

(8) 引用資料

- 1) Albert, R.; Barabási, A.-L. Statistical mechanics of complex networks. Rev. Mod. Phys. 2002, vol. 74, p. 47-97.; Newman, M. E. J. The structure and function of complex networks. SIAM Rev. 2003, vol. 45, p. 167-256.; Boccaletti, S. et al. Complex networks: structure and dynamics. Phys. Rep. 2006, vol. 424, p. 175-308.
- 2) 日立ハイテク ビジネス顕微鏡.
<http://www.hitachi-hitec.com/jyouhou/business-microscope/>
- 3) サイジニア. <http://www.scigineer.co.jp/>
- 4) Watts, D. J.; Strogatz, S. H. Collective dynamics of 'small-world' networks. Nature. 1998, vol. 393, p. 440-442.; Barabási, A.-L.; Albert, R. Emergence of scaling in random networks. Science. 1999, vol. 286, p. 509-512.
- 5) Netsci. <http://netsci2014.net/>
- 6) Center for Complex Network Research, Northeastern University.
<http://www.barabasilab.com/>

- 7) Center for Complex Networks and Systems Research, Indiana University.
<http://cnets.indiana.edu/>
- 8) Eigenfactor の HP. <http://mas.eigenfactor.org/>
- 9) Amaral 研究室の HP. <http://amaral.northwestern.edu/>
- 10) Cross, B. Driving Results through Social Networks – How Top Organization Leverage Networks for Performance and Growth. Jossey-Bass, 2009.
- 11) NETWORKS. <http://www.thenetworkcenter.nl/>
MULTIPLEX. <http://www.multiplexproject.eu/>
PLEXMATH. <http://www.plexmath.eu/>
- 12) Bascompte 研究室の HP. <http://bascompte.net/>
- 13) Onnela, J. P. et al. Structure and tie strengths in mobile communication networks. Proc. Natl. Acad. Sci. USA. 2007, vol. 104, p. 7332-7336.
- 14) data for development. <http://www.d4d.orange.com/en/home>
- 15) Ahn, Y. Y.; Han, S.; Kwak, H.; Moon, S.; Jeong, H. Analysis of topological characteristics of huge online social networking services. Proceedings of International World Wide Web Conference, 2007.

3.4.2 機械学習・データマイニング分野におけるネットワーク構造解析

（1）研究開発領域名

機械学習・データマイニング分野におけるネットワーク構造解析

（2）研究開発領域の簡潔な説明

機械学習やデータマイニングの技術を用いたネットワーク構造、あるいは、より一般的に関係データの解析。ネットワーク上の関係予測や分類、順位づけなど。

（3）研究開発領域の詳細な説明と国内外の動向

データ解析技術の一分野である機械学習、データマイニング分野では、近年、ネットワーク構造をもったデータの解析が盛んに行われている。複雑ネットワーク研究などの関連分野が「ある種のネットワークに共通して現れる性質を理解する」という科学的な動機を基礎としているのに対して、機械学習やデータマイニング分野におけるネットワークを対象とした研究は「与えられたネットワークデータにおける知識発見や予測を行う」という、より具体的で応用的な志向が強い。また、これらの技術と大規模な計算資源を用いて膨大なデータに対する計算を実際に行うためのアルゴリズムやシステム開発など、産業的な側面も強い。ただし、近年では様々なソーシャルネットワークサービス等の普及により、これを具体的な共通の興味の対象として、両者は急速に接近しつつある。

データマイニングとはデータベースに格納された膨大なデータの中から何らかの知見を発掘する（mining）ことを目的とする分野であり、1990年代の初頭に米IBMアルマデン研究所のRakesh Agrawal氏らの研究グループによって創始された。現在では、データマイニングはデータ解析技術における一大分野を成しており、産業界における注目度も非常に高い。初期の研究では、関係データベース中に繰り返し表れる特徴的なパターンを発見するという限定された問題設定から始まった¹⁾が、近年ではその守備範囲はデータ解析技術全般にわたっている。ネットワーク構造をもったデータも例外ではなく、これを対象としたデータマイニング分野はしばしばリンクマイニング²⁾と呼ばれ、特に米国を中心として盛んに研究が行われている。リンクマイニングにおける基本タスクとしては、ネットワーク上でのノードの分類や順序づけ（ランキング）など、従来の非ネットワーク型のデータ空間を対象とした解析における基本タスクをネットワーク上に拡張したものが挙げられる。また、ネットワーク型データに固有のタスクとしてリンク構造の予測、特徴的な部分構造パターンの発見、ネットワーク構造の分類なども挙げられる。特にソーシャルネットワークを対象としたネットワーク上の情報伝播（所謂「口コミ」）の解析も盛んである。

一方、ネットワーク構造は、より一般的にオブジェクト同士の関係を表現した関係データとしても理解することができる。人工知能から派生した機械学習分野では、関係データの解析は古くより一階述語論理を基礎とした知識獲得の枠組みである帰納論理プログラミング（関係学習）として、欧州を中心に研究が行われてきた。帰納論理プログラミングの枠組みは非常に汎用的であり、化合物内部の分子が形成するネットワーク構造の解析などにも用いられた。しかし、その強力さの代償として適用可能なデータ規模は

著しく限定される。そこで、近年ではより大規模でノイズをもった現実のネットワークデータへの適用を意識し、よりシンプルで統計的なモデル化に基づく方法の研究が盛んである。特に最近ではベイズ統計学を用いたモデル化が盛んに行われている。なお、統計的なモデル化は機械学習やデータマイニングで盛んになる以前からも、小規模な社会ネットワーク分析分野において、 p^* モデル、確率的ブロックモデルなどの名の下にすでに用いられてはいたが、機械学習やデータマイニングでは大規模で具体的なネットワークへの適用を意識して、モデルの複雑化や計算アルゴリズムの開発などが行われている。

関係データを対象とした解析の現実的課題への応用として最も成功しているもののひとつは、各々のユーザに合わせて適切な情報を提示する推薦システムである。推薦システムの研究自体は 1990 年代半ばから始まり、米 Amazon 社をはじめとするオンラインショッピングモールなどにおいて、顧客に対する適切な商品推薦に用いられてきた。推薦システムについて注目すべき近年の出来事は、米国最大のオンライン DVD レンタル企業である米 Netflix 社によって 2006 年から数年間にわたって開催された Netflix Prize³⁾である。これはユーザに対する映画の推薦精度を高めることを目指した予測モデリングのコンペティションであり、高額な優勝賞金が大きな話題をよび、多数の参加者を集め、結果としてこの分野を大きく進展させた。一方で、Facebook や Twitter といったソーシャルメディアが広く普及したことを背景に、これらにおける情報推薦、例えば新たな友人や話題の推薦といった機能を提供する目的にも、推薦システムの技術が用いられており、推薦システムは機械学習・データマイニング技術の「キラーアプリケーション」としての地位を確立している。

バイオ、創薬といった分野でもネットワーク構造は頻繁に現れ、機械学習やデータマイニングの技術が活用されている。例えばタンパク質の相互作用などの生体内ネットワークや、薬剤とその標的の間の相互作用ネットワーク、あるいはそれらが引き起こす副作用について、既知の関係から未知の関係を予測することができれば、実験のコストを大幅に削減することができる。一方、薬剤の候補となる化合物などもまた、原子間の共有結合などの作用を表したネットワーク構造を内部にもっており、これらを基に化合物の性質を予測することができれば、やはり薬剤設計の効率化につながる。

最近では、解析の対象となるネットワークが、静的で均一なものからより現実的かつ複雑な時間的に変化する異種混合的なネットワークへと広がりつつある。上記の推薦システムを例にあげれば、ユーザとコンテンツという 2 つのオブジェクトの間の関係の有無（もしくはその強さ）だけでなく、ユーザがコンテンツに対して取り得る行動の種類や、関係に付与されるタグといったより詳細な関係情報、あるいは関係の時間的変化といった動的特性にまで踏み込んだ解析が行われている。ここではより複雑な確率的な関係生成モデルや、テンソル分解と呼ばれるモデルや推定アルゴリズムが開発されており、ここ数年盛んに研究がなされている。同様の動きは複雑ネットワーク研究の分野でも起こっており、時間変化を伴う「時変ネットワーク」や異種混合的な「多重ネットワーク」として精力的に研究がおこなわれている。

情報科学分野の多分に漏れず、ネットワーク構造データの解析についても欧米、特に米国が主導的に研究を進めており、日本を含むアジアはその後追いとなっている感が否めない。しかしながら、データマイニングの分野にグラフ構造データを持ち込んだグラ

フマイニングにおいて、日本は、研究初期の段階から重要な基礎研究を着実に積み上げてきている。また、情報科学全般の動向とも一致して、中国の存在感が急激に増している。

（４）科学技術的・政策的課題

近年、「ビッグデータ」の旗印のもと、多種多様な巨大データを活用することでさまざまなビジネス的あるいは社会的要請の高い実問題の解決を目指す機運が高まっている。ビッグデータの特徴としてしばしば挙げられるのが、量、スピード、多様性の3つである。このうちネットワーク構造データは、従来のデータ解析技術が対象としていた定型の表形式ではないという点で、多様性に深く関係している。そして、現在もさまざまな努力がなされてはいるものの、量とスピードへの挑戦は今後の技術的課題といえる。

また、これらの技術を真に社会に役立てていくためには、社会的意義のある課題設定と、なによりもデータが利用可能となることが必須である。近年オープンガバメントの流れが注目されているが、今後は公民連携のデータ共有が一層必要になっていくと思われる。一方で個人の行動履歴などの情報は、ともすると極めてデリケートな個人情報となり得る。データ解析によって得られる利便性とプライバシーの兼ね合いは、技術と政策の両面で考えていかなければならない重要課題である。

（５）注目動向（新たな知見や新技術の創出、大規模プロジェクトの動向など）

Carnegie Mellon 大学の Christos Faloutsos 氏を中心とするグラフマイニングプロジェクト⁴⁾は数十億ノードをもつ大規模ネットワーク処理を目指したプロジェクトであり、このような解析技術の大規模化は Facebook などの大規模ソーシャルネットワークの隆盛なども相まって当該分野におけるひとつのトレンドといえる。Stanford 大学の Jure Leskovec 氏を中心とした SNAP (Stanford Network Analysis Project) プロジェクト⁵⁾もまた様々な種類のネットワークデータを解析するためのツールを提供するだけでなく、ネットワークデータのアーカイブとしての役割も果たしており、強い求心力をもっている。

また（３）でも述べたように、近年では解析の対象となるネットワークが静的で均一なものから、より複雑な時間的に変化する異種混合的なネットワークへと発展しつつあり、モデルや手法の高度化が著しい。

（６）キーワード

グラフマイニング、リンクマイニング、協調フィルタリング、マーケティング、ソーシャルメディア、ソーシャルネットワーク、バイオインフォマティクス、ケモインフォマティクス、創薬

（7）国際比較

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	○	↗	・グラフ構造データを扱うデータマイニング技術については、その初期の段階から、大阪大学のチームをはじめとして日本が重要な貢献を行っている ⁶⁾ 。
	応用研究・開発	△	→	・基礎研究と産業化の結びつきが弱い。
	産業化	○	↗	・国内ソーシャルメディアにおける情報推薦、楽天、アマゾンなどのオンラインショッピングモールなどでの推薦システムなど、関係予測の産業利用が見られる。
米国	基礎研究	◎	↗	・当該分野を牽引する研究グループが多く存在する。Carnegie Mellon大学のChristos Faloutsos氏のグループやStanford大学のJure Leskovec氏らのグループが大規模なグラフ解析を積極的に進めている。Stanford大のDaphne Koller氏、Maryland大のLise Getoor氏のグループなどが、統計的な関係データ解析研究を積極的に推進している。
	応用研究・開発	◎	↗	・Netflixのチャレンジが特に推薦方式の進展を加速した。また、Twitter、Facebook、LinkedInなどの主要なソーシャルメディアの多くは米国発であることもあり、これらの研究者・エンジニアによる研究開発も盛んである。
	産業化	◎	↗	・Twitter、Facebook、LinkedInなどのソーシャルメディアにおける情報推薦、Amazon.comなどのオンラインショッピングモールなどにおける商品の推薦など、世界的に利用されているサービスにおける技術利用が盛んである。
欧州	基礎研究	○	↗	・米国に準ずる形で発展している。欧州は伝統的に論理に基づく関係データ解析に強く、帰納論理プログラミングを発展させてきた。Imperial College LondonのStephen H. Muggleton氏のグループなど多くの有力な研究グループが存在する。
	応用研究・開発	○	↗	・Mines ParisTechのJ.-P. Vert氏のグループをはじめとして、生体ネットワークや薬剤-標的間ネットワークなどをターゲットとしたネットワーク解析が盛んである。
	産業化	△	↗	・欧州独自の目立った動きはない。
中国	基礎研究	○	↗	・Microsoft Research Asiaや有力大学からコンスタントに論文が発表されている。
	応用研究・開発	△	↗	・特に目立った動きはない。
	産業化	○	↗	・海外ソーシャルメディアへのアクセスが制限されるなか、中国国内のソーシャルメディア、オンラインショッピングなどが独自に進化しており、推薦などの技術が用いられていると推測される。
韓国	基礎研究	△	→	・特に目立った動きはない。
	応用研究・開発	△	→	・特に目立った動きはない。
	産業化	△	→	・特に目立った動きはない。

（註1）フェーズ

基礎研究フェーズ：大学・国研などでの基礎研究のレベル

応用研究・開発フェーズ：研究・技術開発（プロトタイプの開発含む）のレベル

産業化フェーズ：量産技術・製品展開力のレベル

（註2）現状

※我が国の現状を基準にした相対評価ではなく、絶対評価である。

◎：他国に比べて顕著な活動・成果が見えている、○：ある程度の活動・成果が見えている、

△：他国に比べて顕著な活動・成果が見えていない、×：特筆すべき活動・成果が見えていない

（註3）トレンド

↗：上昇傾向、→：現状維持、↘：下降傾向

（8）引用資料

- 1) Agrawal, R.; Imielinski, T.; and Swami, A.N. Mining Association Rules between Sets of Items in Large Databases. Proc. SIGMOD Conference. 1993, p. 207-216.
- 2) Getoor, L.; Diehl, C. P. Link Mining: A Survey. SIGKDD Explorations. 2005, vol. 7, no. 2, p. 3-12.
- 3) Netflix Prize. <http://www.netflixprize.com/>
- 4) Pagasas project. <http://www.cs.cmu.edu/~pegasus/>
- 5) Stanford Network Analysis Project. <http://snap.stanford.edu/index.html>
- 6) Inokuchi, A.; Washio, T.; Motoda, H. An apriori-based algorithm for mining frequent substructures from graph data. Proc. Principles of Data Mining and Knowledge Discovery (PKDD). 2000, p. 13-23.

3.4.3 ネットワークに関する離散数学

(1) 研究開発領域名

ネットワークに関する離散数学

(2) 研究開発領域の簡潔な説明

グラフ理論は、その誕生以来、電気回路、分子構造などの具体的なグラフ構造に関する研究に動機づけられて発展してきた。その内容は伝統的に、直感的に理解しやすい問題設定を軸に展開され、多くの予備知識を必要としないといわれてきた。しかし、1980年代以降、グラフマイナー理論に代表される構造的グラフ理論の興隆によって、その数学的内容が充実するとともに、理論計算機科学にも大きな影響を与えている。同時に、計算機性能の向上に伴い、グラフ理論とその周辺が重要な役割を果たす応用分野も急速に拡大している。

(3) 研究開発領域の詳細な説明と国内外の動向

本開発領域は、グラフ理論の成果によって支えられている。現代のグラフ理論を牽引しているのは、グラフマイナー理論に代表される構造的グラフ理論である。実際、国際数理計画法学会 (MPS) とアメリカ数学会 (AMS) が共同で 3 年ごとに離散数学分野の優れた論文を表彰している Fulkerson 賞を見ると、過去 20 年間に受賞した 15 件のうち、グラフ理論に関するものは 6 件あり、そのうちで 5 件までが構造的グラフ理論の論文である。

アメリカ合衆国の Robertson、Seymour は、1983 年から 2012 年にかけて、全部で 23 編からなる連作論文を発表してグラフマイナー理論を確立した。グラフのマイナーとは、元のグラフから辺の開放除去 (両端点をそのままにして辺を消去する操作) や縮約除去 (両端点が一致するまで辺を縮める操作) と孤立点の除去を繰り返し適用して得られるグラフのことである。グラフが平面に埋め込み可能であることを意味する平面性のように、マイナーを取るという操作に関して閉じている性質は多い。そのような性質を満たさないグラフがひとつ見つかり、それをマイナーとして含むグラフもその性質を満たさない。このようなグラフのうちで極小なものを禁止マイナーと呼ぶ。グラフの平面性の禁止マイナーは、完全グラフ K_5 と完全 2 部グラフ $K_{3,3}$ のみであることが知られている。このようにマイナーに関して閉じた性質に関しては常に有限個の禁止マイナーが存在するというのが Wagner 予想である。Robertson、Seymour の一連の論文における最終目標は、この Wagner 予想の証明であった。これによって、与えられたグラフにおいて、マイナーに関して閉じた性質の有無を調べることが、禁止マイナーの存在を調べることに帰着された。

グラフマイナー理論は、Wagner 予想を証明しただけでなく、さまざまな副産物をもたらしている。理論計算機科学の観点で特に重要なのは、木幅の概念の導入である。木幅とは、グラフがどの程度木に近い構造を有しているかを表すパラメータである。木幅 1 のグラフは木であり、木幅が 2 以下のグラフは直並列グラフになる。一般のグラフ上で NP 困難な問題に対しても木幅が限定されたグラフのクラスに対しては、多項式時間アルゴリズムが設計できることが多い。この例が示すように、パラメータを固定したグ

ラフに対して計算量を議論する研究分野が生まれ、今日に至っている。

グラフマイナー理論の応用として、大規模集積回路の設計にも関連の深い、指定された端子対間を結ぶ点素道の有無を判定する問題に対して、端子対数を固定した場合の多項式時間アルゴリズムが設計されている。ただし、得られたアルゴリズムは多項式時間であるものの、実際の計算時間は端子対数が少し大きくなると爆発的に大きくなる。この難点を克服して、現実的なアルゴリズムを実現することを目指したアルゴリズム的グラフマイナー理論が河原林らによって始められ、成果が次々とあがっている。

グラフマイナー理論の鍵は、禁止マイナーを含まない任意のグラフが、小さな部品を特定の方法でつなぎ合わせることで構成できることである。同様のパラダイムが、マイナーとは異なる操作に関して閉じている性質に注目した場合にも適用可能であることが期待される。こうしてグラフマイナー理論と同様の方法論による研究が進められ、構造的グラフ理論と総称されている。その中でも最も有名な成功例は、グラフの彩色に関する強パーフェクトグラフ予想の Chudnovsky、Robertson、Seymour、Thomas による証明であろう。ここでは、マイナーを取る操作の代わりに、誘導部分グラフを取る操作が単位操作となる。

グラフマイナー理論に代表される構造的グラフ理論の研究は、アメリカ合衆国を中心に展開されている。そこから派生した木幅などのパラメータ固定計算量の研究は、ヨーロッパで盛んに行われている。日本では、河原林らの活躍によって、アルゴリズム的グラフマイナー理論が推進されているという状況にある。

一方、構造的グラフ理論以外でも、以下にあげる重要な展開がなされている。

既存のグラフにできるだけ少ない辺を加えることによって所望の連結度を達成することを目的とするグラフの連結度増大問題は、通信網の設計との関連も深く、注目を集めてきた。特に 1980 年代後半の渡辺、中村による結果を契機に、多くの研究者が参入した。中でも、Frank を中心としたハンガリーの研究者集団の活躍は特筆に値する。わが国では、グラフの合法的順序づけの手法を編み出した永持、茨木が、辺連結度増大問題に対する最も効率的な決定性アルゴリズムを設計した。近年では、ハンガリー出身の Végő が、長年の課題であった点連結度増大問題に関して、画期的な成果をあげている。

ネットワーク中に容量制約を満たしつつ多種類のフロー（水、物流、交通流などに対応）を効率的に流すことを目的とした多品種流問題は、単品種の場合と異なり、整数性を課すと NP 困難になる。現実では、流したい対象は自由に分割できないことも多いため、フローの量などが整数であるという制限を課すことは、応用目的のためにも重要である。整数性を課さない問題は線形計画法で簡単に解けるため、どのような条件下で線形計画緩和の最適解の整数性が保証できるかが重要となる。実際には、整数性が保証される場合は非常に限られているが、ある種の条件下では、分母が 2 の分数最適解をもつことが保証される。さらに進んで、どのような需要設定ならば、ネットワークの構造や容量によらずに、任意の整数容量の問題に対して、最適解の分数性が保証されるかという課題が Karzanov によって 1990 年に提起された。この問題も長い間未解決であったが、最近の平井による一連の研究の結果、完全に解決された。

ネットワークは、通信や輸送の媒体であるだけでなく、建造物の骨組みとしても現れる。骨組みの組合せ的な構造から、その剛性を判定することを目的とした組合せ剛性理

論もグラフ理論と同様に古くから研究が行われてきた。特に、2次元骨組みの一般剛性を特徴づけた Laman の定理以来、多くの研究者が参入している。近年では、組合せ剛性理論の応用分野が、建築構造物だけでなく、ネットワークセンシングや分子構造解析へも広がってきている。ネットワークセンシングに関連しては、点対間の距離情報に基づいて各点の位置を決定する問題において、解が一意的に決定できる条件として、大域剛性の概念が登場した。大域剛性に関する最も基本的な問題であった Hendrickson 予想は、最近、Connely と Jackson、Jordán が解決した。分子の剛性を判定する際には、グラフの「2乗」として得られるグラフを扱うことが理論的な筋道として自然である。このクラスのグラフの3次元における一般剛性が効率的なアルゴリズムで判定できる全域木の詰め込みによって特徴づけられる、という予想が1980年代に Tay と Whiteley によって提示されていた(分子剛性予想)。この予想は、最近、加藤と谷川によって肯定的に解決された。今日では、組合せ剛性理論の研究者の多くが、有用な物質の設計への応用を念頭において、結晶構造のように周期性のある構造の剛性の特徴づけに取り組んでいる。

以上のように、ネットワークに関する離散数学は、急速な進展を見せており、長年の未解決問題が次々と解決されてきた様子には眼を見張るものがある。その中で、わが国の若手研究者も本質的な貢献を果たしている。

(4) 科学技術的・政策的課題

近年の離散数学の進展は、理論の深化がもたらしたものであり、対象の直感的な把握のしやすさと予備知識を必要としない点が離散数学の特徴といわれていた時代からは隔世の観がある。それゆえに、当該研究開発領域の進展は研究者個人の資質と努力に大きく依存している。同時に、その成果は、分野横断的にさまざまな応用技術の基礎となっている。日本が本領域の優位性を獲得するためには、優れた研究者が研究に没頭できる環境を維持することが必要不可欠であろう。

一方で、実用にかかわる問題が新たな研究対象を生み出して、分野の発展を刺激してきたことも事実である。したがって、本分野の研究者が異分野と協働することを奨励する施策には、長期的に見ても意義があると思われる。

(5) 注目動向(新たな知見や新技術の創出、大規模プロジェクトの動向など)

従来のグラフ理論・グラフアルゴリズムの研究では、対象となるグラフの基本的な構造が把握できること、言い換えると、グラフが計算機に格納できることが暗黙の前提とされていた。しかし、ネットワーク科学の進展に伴って、爆発的な大きさのネットワーク、あるいは進化し続けるネットワークも興味の対象となってきている。このようなネットワークを扱うには、計算モデルの設定から再検討する必要がある、計算機科学における新たな研究テーマともなっている。2010年に京都賞を受賞した Lovász は、このように巨大なネットワークを調べるアプローチとして、グラフの無限極限に当たるグラフオンの概念を導入し、研究を進めている。今後の展開が注目される。

（6）キーワード

グラフ、ネットワーク、グラフマイナー理論、木幅、固定パラメータ計算量、アルゴリズム的グラフマイナー理論、構造的グラフ理論、パーフェクトグラフ、連結度増大問題、多品種流問題、組合せ剛性理論、分子剛性予想、グラフオン

（7）国際比較

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	◎	↗	・アルゴリズム的グラフマイナー理論、多品種流問題、組合せ剛性理論において、画期的な結果が、若手研究者によって発表されている。 ・世界水準の競争に参加している研究者層の厚さという点では、改善の余地がある。
	応用研究・開発	○	↗	・ERATO河原林巨大グラフプロジェクトが発足し、多くの若手研究者が応用研究に積極的に参画して成果をあげている ¹⁾ 。
	産業化	○	→	・2部グラフのDM分解を用いて半導体メモリの不具合を救済する方式が開発され、実用化された（東芝） ²⁾ 。
米国	基礎研究	◎	↗	・構造的グラフ理論の研究で、他を圧倒している。 ・ネットワークゲーム理論、評判の伝搬モデルなど、新たな問題設定で野心的な研究テーマが追究されている。
	応用研究・開発	◎	→	・IBM、Microsoft、Yahoo!、Googleといった企業が、離散数学・計算機科学の分野の学位取得者の主要な就職先になっている。その中には、世界的に広く知られた研究者も少なくない。 ・分子剛性予想が証明される以前から、その正当性を前提に、分子運動の自由度を解析するソフトウェアが開発されている ³⁾ 。
	産業化	○	→	・木幅の変種を用いたルーティング方式の開発された（Bell 研究所） ⁴⁾ 。
欧州	基礎研究	◎	→	・ドイツ、フランス、オランダにおいては、グラフアルゴリズム、特にパラメータ固定計算量の研究が充実している。 ・グラフ連結度に関する研究では、ハンガリーの研究者の活躍が顕著。
	応用研究・開発	○	↗	・ドイツでは、数学の工学的応用を推進するプロジェクトMATHEONが推進され、産業界の諸問題に数学者が取り組む体制が整ってきた ⁵⁾ 。
	産業化	○	→	・グラフアルゴリズムのパッケージソフトLEDAが販売されている ⁶⁾ 。
中国	基礎研究	△	→	・論文数は多いのかもしれないが、質の高いものは少ない。
	応用研究・開発	△	→	・応用研究・開発の動きは不明。
	産業化	△	→	・産業化の動きは不明。
韓国	基礎研究	○	↗	・KAISTでは、Sang-il Oumが中心になって、外国人若手研究者を多数含んだ活動的なグループを形成している ⁷⁾ 。
	応用研究・開発	△	→	・応用研究・開発の動きは不明。
	産業化	△	→	・産業化の動きは不明。

（註1）フェーズ

基礎研究フェーズ：大学・国研などでの基礎研究のレベル
 応用研究・開発フェーズ：研究・技術開発（プロトタイプの開発含む）のレベル
 産業化フェーズ：量産技術・製品展開力のレベル

（註2）現状

※我が国の現状を基準にした相対評価ではなく、絶対評価である。
 ◎：他国に比べて顕著な活動・成果が見えている、○：ある程度の活動・成果が見えている、△：他国に比べて顕著な活動・成果が見えていない、×：特筆すべき活動・成果が見えていない

（註3）トレンド

↗：上昇傾向、→：現状維持、↘：下降傾向

（8）引用資料

- 1) <http://www.jst.go.jp/erato/kawarabayashi/>
- 2) <http://www.j-tokkyo.com/1999/G11C/JP11238396.shtml>
- 3) FIRST. <http://flexweb.asu.edu/software/first/>
- 4) Communications Network Ring Router, US patent number 5533016.
- 5) MATHEON. <http://www.matheon.de/>
- 6) LEDA. <http://www.algorithmic-solutions.com/>

3.4.4 ネットワークを扱うソフトウェア

(1) 研究開発領域名

ネットワークを扱うソフトウェア

(2) 研究開発領域の簡潔な説明

社会・生物・人工物システムに見られる依存関係や共起関係、相関などの接続関係を記述したデータ集合からネットワーク構造を構成し、ネットワークの表現・操作・分析における適切な形式でデータを管理し、ネットワーク構造を特徴付ける諸統計量を計算し、ネットワークを低次元空間に埋め込み視覚化するための計算機ソフトウェア。

(3) 研究開発領域の詳細な説明と国内外の動向

本稿が想定するネットワークを扱うソフトウェアとは主に以下の2つの機能のいずれか、あるいは両方を併せ持つものである。一つには①ネットワークという構造を持つ位相的性質を統計的・数値的に評価するための計算機能。もう一つには②本質的に高次元の空間でのみ距離的な矛盾なくレイアウト可能なネットワーク構造を、次元圧縮や最適化の技法を用いて矛盾を最小化しつつ低次元空間（主に2次元）に埋め込むための視覚化機能である。ソフトウェアが主たる対象とする学術領域、実装された分析・視覚化アルゴリズムの種類、スケーラビリティのような計算パフォーマンス、商用／非商用、ライセンス形態、開発言語、動作するプラットフォーム、他の統計分析ソフトウェアとの連携、ユーザインターフェイス、インタラクティブ性など、多様なニーズに応じて数多くの開発プロジェクトが歴史的に存在し、全体を俯瞰するのは容易ではない。またソフトウェアと一言で言っても、利用者にとってプログラミングの経験を必要としないパッケージ化されたソフトウェアから特定のプログラミング言語向けのライブラリまで、その使い方や利用者に要求されるスキルレベルは様々である。

①の構造分析機能についてみると、技術的に高度なものとしては主に以下の3つを挙げることができる：1) ネットワークの持つ行列的性質を求める固有値解析機能、2) 最短経路や媒介性といったネットワーク上の経路または距離に関する特徴量を求める経路探索機能、3) 相対的に密に接続されたノード集合—“クラスター構造”や“コミュニティ構造”と呼ばれる—を発見するクラスタリング機能。その他にも次数（ノードにつながるエッジ数）やモチーフ（あらかじめ定義された単位構造）の分布、さらには隣接関係に対するネットワーク特徴量間の相関といった指標もネットワークに固有の計算問題として重要である。固有値解析機能を例にとれば、ネットワーク構造の数学的表現は行列を用いて自然に与えられるため、その行列の持つ数学的／統計的性質を求めることはネットワークの機能を推定する上で重要な意味を持つ。例えば1998年にGoogle社が発表した“ページランク”¹⁾と呼ばれるウェブサイトの評価手法は、ウェブが持つネットワーク構造を遷移確率行列で表現したときの固有ベクトルを求める問題に帰着される。行列表現に対する数値計算手法は工学的にも長い歴史を持っており、商用／非商用を問わず様々な数値計算ライブラリが存在する。経路探索やクラスタリングについても同様、いずれも応用分野は制御や最適化といった工学の多岐にわたり、数多くの計算アルゴリズムとその実装例が存在する。したがってそれらは統合型の数値計算パッケージ

ソフトウェアが提供する定番機能ともなっている。つまり、比較的新しい学問領域であるネットワーク科学においても、その根幹の部分では従来の多くのソフトウェア資産が利用可能である。また近年では R²⁾や Python³⁾といった主に学術・教育分野で急速に普及が進んでいるプログラミング言語・環境上で動作するネットワーク分析パッケージの開発も盛んである。ユーザが増えることで様々なニーズに応じたチュートリアルが整備されるとともに最新の研究成果もいち早く取り込まれ、その結果益々ユーザが増える、という一極集中的な循環が予想される。

②の視覚化機能についてみると、古くには階層構造を表すツリーの描画のような単純なものから始まり、電子回路図や人間関係といったより複雑な構造に対する工学的・社会科学的応用の面から 1990 年前後にネットワークの自動描画の研究が本格的に行われ始めた。初期の代表的な研究はエッジをバネに見立てた力学エネルギー極小化に基づくノードのレイアウト手法であり、1989 年の鎌田・川合らの研究は現在でも標準的に用いられる手法の一つである⁴⁾。国際的な研究集会として、例えばネットワーク可視化の国際会議 Graph Drawing⁵⁾は 1992 年、ネットワークを含む情報可視化一般の国際会議 InfoVis⁶⁾は 1995 年が第 1 回の開催年である。時期を同じくして Pajek⁷⁾や Graphviz⁸⁾といったネットワーク可視化ソフトウェアがそれらの研究成果を取り込んでユーザを集めた。初期の可視化ソフトウェアの多くはノード間の隣接関係を記述したデータファイルを読み込み、あらかじめ指定したアルゴリズムを用いてレイアウトを行った結果を画像ファイルとして出力するものであった。したがって、クラスタリングや経路探索の結果をレイアウトや視覚属性に反映させたい場合には、予め分析ソフトウェアを用いてネットワークの構築や属性付けを行い、それを一般的なネットワーク記述形式でファイルにエクスポートし、ネットワーク可視化ソフトで画像を生成するという手順になる。そのためソフトウェア間の入出力データの形式互換性も重要となる。多くのソフトウェアがサポートする形式としては例えば GML(Graph Modeling Language)⁹⁾、GraphML¹⁰⁾、dot¹¹⁾などがある。あるいは単純なカンマ/タブ区切りテキスト(csv、tsv)を用いるケースも一般的である。近年ではファイル入出力を基本とする従来の手順から、一つのソフトウェアの中で様々な分析結果をインタラクティブにレイアウトや視覚表現に反映する機能を備えたものも増えてきている。扱うネットワークが大規模になればレイアウト計算の負荷は高くなるため、可視化とその前処理としての構造分析の融合は自然な流れである。また「探索的ネットワーク分析」と呼ばれるような、ネットワークを様々な視点から眺めながら分析の糸口を掴んでいく研究アプローチでは、上述した分析と視覚化の統合的活用は必須である。

このような構造分析と可視化の機能を総合的に備えたソフトウェアの開発が加速するのはネットワーク科学が大きく注目を集めた 2000 年代初頭からである。開発主体はソフトウェア開発能力を持った研究者個人によるものから、ネットワーク科学を牽引する研究者グループによるもの、あるいはそのスピンオフ企業のような高い専門性を持った企業である。個人開発から始まりオープンソースプロジェクトに移行していくものも多くみられる。主に学術利用において現在使われているソフトウェアのリストを表 1 に示す。ここで挙げた例は汎用性や開発の規模、ユーザ数、将来性(実装の一般性、オープン性、開発主体)の面で主観的に特に重要と認めたものである。国際的に見た場合、量

的にも歴史的にも米国が圧倒的に先行している。欧州ではドイツが先行していると考えられ、他の国でも特徴的な活動は認められる。日本では大学の研究室等の小規模なレベルでの開発例は過去にいくつか見つけられるが、国際的にユーザを集めることはなく活動も継続的ではない。同様に中国、韓国でも目立った活動は見られない。

表 1：ネットワークを扱うソフトウェアの例

プロジェクト名	開発主体	用途	開発開始年
Pajek	スロベニア：Vladimir Batagelj	可視化	1996
Graphviz	米：AT&T 研究所	可視化	2000
Siena/RSiena (with R) ¹²⁾	英：オクスフォード大学	分析・可視化	2000
BGL (Boost Graph Library) (with C++) ¹³⁾	米：インディアナ大学	分析	2000
GTL (Graph Template Library) (with C++) ¹⁴⁾	独：パッサウ大学	分析	2001
visone ¹⁵⁾	独：コンスタンツ大学	分析・可視化	2002
UCINET for Windows (商用) ¹⁶⁾	米：Analytic Technologies 社	分析・可視化	2002
Cytoscape ¹⁷⁾	米：Institute for Systems Biology	分析・可視化	2002
NetDraw ¹⁸⁾	米：Analytic Technologies 社	可視化	2002
statnet (with R) ¹⁹⁾	米：ワシントン大学	分析・可視化	2003
Tulip ²⁰⁾	仏：ボルドー大学	分析・可視化	2003
LGL: Large Graph Layout (with Java) ²¹⁾	米：テキサス大学オースティン校	可視化	2003
JUNG (with Java) ²²⁾	米：カリフォルニア大学アーバイン校	分析・可視化	2003
SNAP (with C++/Python) ²³⁾	米：スタンフォード大学	分析	2004
igraph (with R) ²⁴⁾	洪：科学アカデミー	分析・可視化	2005
CFinder ²⁵⁾	洪：エトヴェシュ・ロラード大学	分析・可視化	2005
Network Workbench (with Java) ²⁶⁾	米：インディアナ大学	分析・可視化	2006
GraphStream (with Java) ²⁷⁾	仏：ル・アーヴル大学	分析・可視化	2007
NetworkX (with Python) ²⁸⁾	米：LANL 研究所	分析・可視化	2008
NodeXL (Excel プラグイン) ²⁹⁾	米：Microsoft Research	分析・可視化	2008
Gephi ³⁰⁾	仏：Gephi コンソーシアム	分析・可視化	2008
Cuttlefish ³¹⁾	瑞：チューリヒ工科大学	分析・可視化	2009
graph-tool (with Python) ³²⁾	独：ダルムシュタット工科大学	分析・可視化	2009

(4) 科学技術的・政策的課題

- ・近技術的には大規模データに対する適用性が伝統的かつ現代的な課題である。組み合わせ爆発に由来する計算量的困難を克服するためのアルゴリズムや、また扱うデータ量の増加に対応する分散処理フレームワーク³³⁾といった新しい研究成果を積極的に取り込む開発体制が重要である。
- ・一方で、例えばソーシャルネットワーク研究の実際においては扱うデータ量は必ずしも大規模でない場合も多い。探索的なデータ分析においてはむしろユーザインターフェイスの作りこみとソフトウェアとしての完成度が重要になる。具体的な研究課題を抱えたアカデミック内部ではこういったソフトウェア開発は難しく、開発力のある集団(企業やコミュニティ)をどのように味方につけていくかが重要である。
- ・アクティブな開発体制を維持するためには、高いコミットメントを持ったコアメンバーはもちろん、オープンソース開発に見られるような一定規模のコミュニティを形成することが重要である。また、広く開発への貢献やユーザを集めるにはソフトウェア開発の国際化(英文ドキュメントの充実など)が重要である。
- ・ソフトウェア開発は科学的な知見を得るための活動とは必ずしもベクトルを同じくせず、アカデミズムの中では過去には積極的には評価されにくい風潮があった。現在ではいくつかの国内論文誌にシステム開発論文というカテゴリが設けられ、また未踏IT人材発掘・育成事業(年齢制限あり)³⁴⁾のような資金提供も存在し、一定の評価がなされるようになってきている。天才発掘的な個人サポートに止まらず、科学研究を支援するソフトウェア開発を中・長期的プロジェクトとして奨励する資金的枠組みも重要である。
- ・一般にソフトウェア開発は科学・技術的な探究活動とは異なるスキルとメンタリティを必要とするものであり、先進的なソフトウェアの発展に必要なのは最新の研究動向と技術動向に一定の関心と理解を持つスキルを持った研究者である。学部から大学院につながる一貫性を持ったスキル教育が重要である。

(5) 注目動向(新たな知見や新技術の創出、大規模プロジェクトの動向など)

- ・最近のトレンドとして、一般的なブラウザ上で動作する JavaScript ベースのネットワーク可視化フレームワークの開発が進んでいる³⁵⁾。プラットフォームに依存せず、他者へのプレゼンテーションやブラウザの機能を用いた外部データの利用が容易であるため、そういったブラウザベースの可視化手法はウェブデザインやマーケティング、ジャーナリズムに浸透しつつある。
- ・2012年度の報告では「クラスタリングのような計算負荷の高い分析機能を提供する外部のクラウド型計算サービスと融合した、ウェブベースの高度なネットワーク可視化サービスが登場する可能性は高いと考えられる」と予測したが、現時点ではまだそのような動向は見られない。
- ・Gephi³⁰⁾(高い審美表現を目指すインタラクティブ性の高い汎用のネットワーク可視化ソフトウェア)、Cytoscape¹⁷⁾(システム生物学のための専門的ソフトウェアとして開発がスタートしたものの、その後開発コミュニティを充実させながら汎用の可視化ソフトウェアとして多くのユーザを集める)、NetworkX²⁸⁾(教育用プログラミン

グ言語として欧米の大学で積極的に導入されている Python のライブラリとして開発が進む）、igraph²⁴⁾（分野を問わず多くの研究者や学生が利用する非商用の統計分析ソフトウェア“R”のパッケージとして提供される）等はオープンソースプロジェクトとしてコミュニティを形成し積極的に開発が進められており、近年目にする機会の多いソフトウェアの例である。Network Workbench²⁶⁾は完成度が高いソフトウェアの一つだが 2011 年で更新が止まっている。ネットワーク分析のツール群をライブラリとして提供する SNAP²³⁾はベンチマーク用のデータセットも同時に多数提供しており、研究論文での引用も多くみられる。

- ・ ネットワーク構造をデータベース内に自然な形で保持することができる“グラフデータベース”³⁶⁾の開発も進んでいる。収集・蓄積されたデータに分析・可視化という付加価値を与えるビジネス分野は近年活発であり、データベースや分散コンピューティングとその出口としての分析・可視化のスムーズな連携は更に進んでいくだろう。

（6）キーワード

ソフトウェア、フレームワーク、統計分析、固有値分析、経路探索、クラスタリング、可視化、レイアウト、ユーザインターフェイス、プラットフォーム、開発言語、オープンソースプロジェクト、ライセンス、データベース

（7）国際比較

国・地域	フェーズ	現状	トレンド	各国の状況、評価の際に参考にした根拠など
日本	基礎研究	△	→	・ソフトウェア開発において基礎研究とプロトタイピングの差は現状小さくなっていると考えられるため、「応用研究・開発」にまとめるものとする。
	応用研究・開発	△	→	・国外で開発された定評のあるソフトウェアの日本語化といったプロジェクトは存在するものの、ソフトウェア開発に関する組織的で継続性を持った独自の活動は見られない。
	産業化	△	→	・科学・技術系ソフトウェアベンダーが商用の統計分析パッケージの一機能として一般的なネットワーク分析機能を提供するケースや、国外で開発された定評のある統計分析パッケージの輸入代理業のみ。 ・Big Dataへの取り組みの一環としてネットワーク分析をソリューションとして提供する企業は一定数存在するが、パッケージソフトウェアとして提供している例は見られない。
米国	基礎研究	◎	↗	・同下。
	応用研究・開発	◎	↗	・歴史的にも著名な開発プロジェクトが米国の大学や研究機関を中心に多く存在し、研究者にとっての標準ソフトウェアとしてよく利用されている。現在アクティブに開発が進行しているプロジェクトも多い。
	産業化	◎	→	・ネットワーク分析に特化した商用ソフトウェアを提供しているベンダーが多く存在する。
欧州	基礎研究	◎	↗	・同下。
	応用研究・開発	◎	↗	・個別ではドイツ発祥の開発プロジェクトが多いものの、国際的な人材交流も活発であり、ハンガリー、イギリス、フランス、イタリア、スペイン、北欧など、大学機関を中心に特徴ある開発プロジェクトが存在する。
	産業化	○	→	・ネットワーク分析に特化した商用ソフトウェアを提供しているベンダーが存在する。

中国	基礎研究	×		・ソフトウェア開発に関する特筆すべき成果なし。
	応用研究・開発	×		・ソフトウェア開発に関する特筆すべき成果なし。
	産業化	×		・ソフトウェア開発に関する特筆すべき成果なし。
韓国	基礎研究	△	→	・日本と状況に差はないと考えられる。
	応用研究・開発	△	→	・日本と状況に差はないと考えられる。
	産業化	△	→	・日本と状況に差はないと考えられる。

(註1) フェーズ

基礎研究フェーズ：大学・国研などでの基礎研究のレベル

応用研究・開発フェーズ：研究・技術開発（プロトタイプの開発含む）のレベル

産業化フェーズ：量産技術・製品展開力のレベル

(註2) 現状

※我が国の現状を基準にした相対評価ではなく、絶対評価である。

◎：他国に比べて顕著な活動・成果が見えている、○：ある程度の活動・成果が見えている、

△：他国に比べて顕著な活動・成果が見えていない、×：特筆すべき活動・成果が見えていない

(註3) トレンド

↗：上昇傾向、→：現状維持、↘：下降傾向

(8) 引用資料

- 1) Page, L., Method for node ranking in a linked database. U. S. Patent 6,285,999 B1. 1998.
- 2) The R Project for Statistical Computing. <http://www.r-project.org/>
- 3) Python. <https://www.python.org/>
- 4) 例えば Fruchterman, T. M. J. et al. Graph Drawing by Force-Directed Placement. Software—Practice and Experience. 1991, vol. 21 no. 11, p. 1129-1164.; Kamada, T. et al. An Algorithm for Drawing General Undirected Graphs. Information Processing Letters. 1989, vol. 31 p. 7-15.
- 5) Graph Drawing. <http://graphdrawing.org/symposia.html>
- 6) IEEE Information Visualization (InfoVis). <http://conferences.computer.org/infovis/>
- 7) Pajek. <http://pajek.imfm.si/doku.php>
- 8) Graphviz. <http://www.graphviz.org/>
- 9) GML (Graph Modeling Language). <http://www.fim.uni-passau.de/index.php?id=17297&L=1>
- 10) GraphML. <http://graphml.graphdrawing.org/>
- 11) The DOT Language. <http://www.graphviz.org/content/dot-language>
- 12) Siena/RSiena. <http://www.stats.ox.ac.uk/~sniijders/siena/>
- 13) BGL (Boost Graph Library). <http://www.boost.org/libs/graph/doc>
- 14) GTL (Graph Template Library). <http://www.fim.uni-passau.de/index.php?id=17297&L=1>
- 15) Visone. <http://visone.info/>
- 16) UCINET. <http://www.analytictech.com/ucinet/>
- 17) Cytoscape. <http://www.cytoscape.org/>
- 18) NetDraw. <https://sites.google.com/site/netdrawsoftware/>

- 19) Statnet. <http://statnet.org/>
- 20) Tulip. <http://tulip.labri.fr/TulipDrupal/>
- 21) LGL (Large Graph Layout). <http://lgl.sourceforge.net/>
- 22) JUNG. <http://jung.sourceforge.net/>
- 23) SNAP (Stanford Network Analysis Project). <http://snap.stanford.edu/>
- 24) Igraph. <http://igraph.sourceforge.net/>
- 25) CFinder. <http://cfinder.org/wiki/>
- 26) Network Workbench. <http://nwb.cns.iu.edu/>
- 27) GraphStream. <http://graphstream-project.org/>
- 28) NetworkX. <http://networkx.lanl.gov/>
- 29) NodeXL. <http://nodexl.codeplex.com/>
- 30) Gephi. <https://gephi.org/>
- 31) CuttleFish. <http://cuttlefish.sourceforge.net/>
- 32) graph-tool. <http://graph-tool.skewed.de/>
- 33) 例えば Ghemawat, G.; Gobioff, H.; Leung, S.-T. The Google File System, 2004.; Dean, J.; Ghemawat, S. MapReduce: Simplified Data Processing on Large Clusters, 2004.
- 34) (独) 情報処理推進機構 未踏 IT 人材発掘・育成事業.
<http://www.ipa.go.jp/jinzai/mitou/>
- 35) 例えば d3 (Data Driven Document). <http://d3js.org/>; sigma.js. <http://sigmajavascript.org/>
- 36) 例えば neo4j. <http://www.neo4j.org/>