

ネットワーク調査の問題と展開可能性



社会ネットワークや社会関係資本論への関心が高まっている。また、それに伴い、社会ネットワークにかかわる調査を行いたいという需要も高まっている。しかしながら、社会ネットワークの調査は難しいのではないかという漠然とした不安から、社会ネットワークにかかわる調査にもう一步踏み出せずにいる研究者や学生も多いのではあるまいか。

そこで、社会ネットワーク調査にはどのような困難な点があるのかを明らかにし、可能な限りその対策を示そうと考えた。座談会では、社会ネットワーク調査にかかわってきた研究者が自らの経験を開示しつつ、社会ネットワーク調査の問題点と対応、そして今後の展開可能性を示した。また、個々の領域における研究者には、当該領域における社会ネットワーク調査の固有の問題と対応について論じていただいた。これらを参考にして、それぞれの関心をもつ方々が、自らの社会ネットワーク調査を展開していただければ幸いである。

(辻 竜平)

社会ネットワーク調査から見えるもの、 調査法上の問題点、そして展開可能性

辻 竜平 (信州大学人文学部准教授)

社会ネットワークについて調査したいと考える人が増えていると感じている。1つは、社会関係資本論が社会科学の多くの領域に浸透してきており、その大きな要素である社会ネットワークを調査する必要性を感じる人が増えているからだだろう。もう1つは、ネット社会やスモールワールド研究などにおいて、人々がインターネット上でどのようなつながりをもっていて、どのように情報が伝播するのかといったことを研究するにあたり、インターネット上の関係を可視化する必要性を感じる人が増えているからだだろう。しかし同時に、ネットワークの調査は、実施や分析が難しいのではないかと躊躇している人も多いと思う。本特集は、「ネットワーク調査の問題と展開可能性」と題されている。社会ネットワーク調査に潜む問題点を指摘し、その対応策を可能な限り提示し、そして、今後のネットワーク調査法の姿を考えるというのが趣旨である。これを通して、少しでも多くの人たちにネットワーク調査に参入してもらいたいと願っている。

ただ、すべての読者が社会ネットワークやその調査に関心をもっているわけではないだろう。そこで、ここでは、特集への導入として、社会ネットワーク調査によって見えてくるものは何か、その調査法上の問題、そして今後の展開可能性について、簡単に述べてお

きたい。より具体的には、この後の座談会や、個別領域における論稿を参考にしてもらいたい。

1 社会ネットワーク調査から見えるもの

社会ネットワークの調査法には、大雑把に言って、ホール・ネットワーク調査とパーソナル・ネットワーク（エゴセントリック・ネットワーク）調査の区別がある。

ホール・ネットワーク調査は、ある集団や社会の成員全員について、その間にある関係（たとえば、友人関係や取引関係）を測定するものである。そうして測定されたデータから、集団や社会の全体構造の特性や、全体構造における個々人の位置や役割、重要性などを明らかにしていくことができる。たとえば、集団の全体構造の特性がわかると、どんなメリットがあるのだろうか。辻（2000）は、集団全体における信頼関係の推移性（自分の信頼している人が信頼している人を、自分も信頼しているということ）の程度が高いほど、集団における社会的ジレンマ状況における協力的行動が促進されることを明らかにした。このように、集団の中での関係のあり方（この場合は信頼関係の推移性）に着目することによって、集団全体としての創発特性（この場合は協効率）を説明できるようになるわけである。

一方、パーソナル・ネットワーク調査は、調査対象者がどのような人々とつながっているかを測定するものである。測定法としては、特定の基準（たとえば、親しさ）から関係のある個人を挙げてもらい、それらの人々について詳細にたずねていくネーム・ジェネレータや、特定の社会的カテゴリ（たとえば、職業）に所属している友人・知人がいるかどうかをたずねていくポジション・ジェネレータといった方法がある。また、その調査対象者が、全体としてどのくらいの人々と知り合いなのかを測定する知人数推定法もある。ネーム・ジェネレータは、多くの場合「親しい人」を挙げてもらうので、その親密な他者との交際の仕方や他の側面での評価がわかるし、挙げてもらった人同士のつながりについて例えば、対象者が親密と思う他者同士の関係性が明らかになる。また、ポジション・ジェネレータは、調査対象者がどのような資源をもつ他者にアクセス可能かといったことがわかるし、職業カテゴリについて問う場合には、それに職業威信スコアを当てはめて、アクセス可能な威信スコアの高さといった指標を作り、交際する他者の階層と自己の階層との関係について明らかにすることもできる。

他にも、工夫次第で、ネットワークにかかわる調査は、さまざまなものがありうるし、その分析もさまざまなやり方があるだろう。何を知るために、調査にどんな工夫をするか、そしてどう分析するかをきちっと計画して調査し、問題解決を試みることは、一度ハマると抜け出しがたい面白さがある。しかし、それと同時に、ネットワーク調査特有の問題があるのも事実である。そこで、ホール・ネットワークとパーソナル・ネットワークの調査に絞って、簡単にその問題点を指摘することにする。

2 | それぞれの調査法の問題点

ホール・ネットワークの調査は、10年ほど前までは、もっぱら小集団を対象としてきた。小集団であれば、人々全員に自分以外の全員に対する関係性を問うことができるが、規模が大きくなるにつれ、全員からの調査協力を取りつけるのが難しいとか、その調査協力を取りつけるまでに手間や時間がかかったり合意が得られなかったりするとか、成員同士が互いのことを知らないケースが増える、といった実行上の問題が生じるからであった。

また、小集団で収集されたデータを用いて、個人の集団内のネットワーク上の位置と個人の行動との関係を分析しても、それは特定の集団で見られた固有の関係性なのであって、他の集団で同じような関係性が見られるかどうかはわからない、といった批判がありうる。一般的な結論を得るためには、多数の小集団からデータを収集しないとイケないのではないとか、データが取れた集団がどのくらい集団らしい特性をもっているのか（一般的な集団を代表していると言えるような特性を備えているか）という問題もある。これに対して、社会ネットワーク分析の研究者たちは、データを持ち寄って互いに提供しあうことによって少しでもデータの数を増やしたり、特殊な統計モデル（たとえば、 $p1$ モデルとか p^* モデル、並べ替え検定など）を工夫・開発したりしてきた。もちろん、それで問題が十分克服できたというわけではないのだが。

最近10年ほどの間に、インターネットから、ウェブサイト間のリンクや、BBSにおける書き込みの連鎖、SNSにおける「友だち」関係など、人々がインターネット上に築いている関係を抽出することが可能になった。

そしてそれは、小集団に属する少人数の人々の間だけでなく、何千、何万といった人々の間の関係についてのデータが入手できるようになったことを意味していた。

しかし、全国民（あるいは全世界の人々）がインターネットを利用しているわけではなく、また、SNSなどの各種サービスとなると、それらを利用しているのは、インターネット利用者の一部である。全国民を母集団とした場合、インターネット利用者は、何らかの特性をもっているのではないかと、また、インターネット利用者を母集団とした場合でも、各種サービスの利用者は、さらに何らかの特性をもっているのではないかと、という母集団とサンプルの代表性の問題は、まだ解決されたわけではない。あるサービス内における人々のネットワーク上の位置とその行動との関連性について何かがわかったとしても、それがどこまで一般化できる知見なのかは、なおよくわからないのである。なお、インターネット調査会社のもつ「モニター」についても、さまざまな工夫はなされているものの、根本的には同種の問題を抱えているものと考えられる。

この点について、パーソナル・ネットワーク調査には、メリットがある。ホール・ネットワーク調査は、どうしても特定の集団の中で行うほかないが、パーソナル・ネットワーク調査は、通常のランダムサンプリング調査の中で実施できるから、その分析結果も一般的だということができる。

しかし、ホール・ネットワーク調査に比べてパーソナル・ネットワーク調査は、調査対象者から見た他者たちとの関係と、やろうと思えば、その他者間の関係を調査対象者の主観的判断として測定することはできるが、調査対象者から直接つながっている人々以外の

人々が、どのようにつながっているのかについてはわからない。この意味でネットワーク構造から調査対象者が受ける影響については、部分的にしかわからないと言える。また、当然ながら、ランダムサンプリング調査の場合は、調査対象者がランダムに選ばれるので、調査対象者間の関係については、いくら工夫しても測定することは難しいだろう。この点においても、パーソナル・ネットワーク調査は、個人を中心とした限定的な関係しか扱えないと言える。

このようなパーソナル・ネットワーク調査の欠点の一部を補うために、東京大学の池田謙一や日本女子大学の石黒格らが行ってきたのが、ランダムサンプリングによって得た調査対象者（主対象者）がネーム・ジェネレータによって挙げた他者に調査を依頼するという方法である。いわゆるスノーボール・サンプリングと呼ばれる手法である。これによって、主対象者が直接に関係していない他者についてのデータを取得して、その影響について考えることができるわけである。ただし、スノーボール・サンプリングによって新たに生じる問題もある。それは、本特集号の座談会を参照してもらいたい。

3 社会ネットワーク調査の展開可能性

いずれにしても、現時点で、サンプリングを含めた調査法のさまざまな観点から見て完全な方法は存在しない。したがって、社会ネットワークの調査を試みようと考えたとき、自分の問題関心のために、完全ではないけれども、考えうる最善の調査法は何かを検討してみることが非常に重要だと言える。また、大きなくくりでは、ホール・ネットワーク調査かパーソナル・ネットワーク調査かという

ことにはなるのだろうが、その中で創意工夫をする余地はかなりある。そのヒントは、座談会記事や個別領域の論稿の中に散りばめられているのではないかと感じている。

2010年、平松闢らが『社会ネットワークの研究・メソッド——「つながり」を調査する』（平松ほか、2010）を出版した。これは、海外にもほとんど類例のない画期的なものであり、これによって、社会ネットワーク調査は、飛躍的に手の届きやすい調査となったのではないだろうか。この本に沿って研究を進めていけば、何とか一通り社会ネットワーク調査を行うことができるだろう。しかしながら、この本は、技法の紹介という側面が強く、その技法の背後にある考え方については、素通りされているようにも感じられる。そのため、そこで紹介されている技法を応用してみたいとなると、突破口が見出せないような事態に陥るのではないかという懸念もある。

本特集は、社会ネットワーク調査を実施するにあたって考えるべきことを基軸に編まれており、『社会ネットワークの研究・メソッド』とは独立した読み物になっていると思う。これはこれで、社会ネットワーク調査を志す人の参考になると思うが、『社会ネットワークの研究・メソッド』と併せて読まれると、なお効果的だろう。多くの研究者が社会ネットワーク調査に参入してもらえれば、と切に願っている。

なお、この特集は、主として社会科学の立場から社会科学の研究者や学生を対象として書かれているが、社会ネットワークに関心をもつ情報工学などの領域の研究者にも手にとってもらいたいと思っている。その理由は、座談会、とりわけその中の松尾豊氏の発言の中に表明されている。社会科学と情報工学の

優れたコラボレーションが生まれることも、また社会ネットワーク調査の新たな展開にとって重要である。

文献

- 平松闢・鶴飼孝造・宮垣元・星敦士、2010、『社会ネットワークの研究・メソッド——「つながり」を調査する』ミネルヴァ書房。
辻竜平、2000、「集団における信頼関係の構造化と集団内秩序の変化」『理論と方法』15（1）：197-208。



座談会

「ネットワーク調査の問題と 展開可能性」

日 時 2012年5月26日(土) 13:40～
場 所 有斐閣神保町ビル10階会議室

出席者 辻 竜平 (司会, 信州大学人文学部准教授)
石黒 格 (日本女子大学人間社会学部准教授)
野沢慎司 (明治学院大学社会学部教授)
松尾 豊 (東京大学大学院工学系研究科技術経営戦略学専
攻准教授)
安田 雪 (関西大学社会学部教授)
中里裕美 (オブザーバー, 明治大学情報コミュニケーション
学部専任講師)

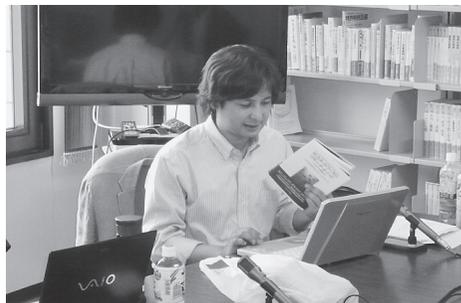


はじめに

辻 ただいまから、「ネットワーク調査の問題と展開可能性」という座談会を始めます。この会のために、本日は私以外に5人の先生にお集まりいただきました。どうもありがとうございます。まず、この座談会の目的を簡単にお話します。

ここ数年、ネットワークの調査に関心をもっている人が非常に増えてきているという印象があります。しかし一方で、ネットワークの調査というのは特有の難しさがあるとも感じています。そんな中、最近、平松闊先生たちが『社会ネットワークの研究・メソッド』（2010年、ミネルヴァ書房）という本を著されました。そこにはネットワーク調査のためのさまざまな方法が紹介されています。このような本は、英書を含めてほとんど類書がなく、たいへん画期的なものです。

しかしながら、これでネットワーク調査ができるようになるかという点、私はこれを読んでいるとやや難しいところがあるのではないかと感じています。なぜなら、ネットワーク調査の技術的な側面の紹介に力点が置かれていて、その背後にある考え方の部分が素通りされているように感じられるところがあるからです。ですから、この座談会では、この本には紹介されていない技法についてご紹介いただくということもありますが、そのような技法の背後にある考え方や問題点についても議論していきたいと思います。また、今後のネットワーク調査の展望についても考えたいと思います。私の気づかない点や考えが及んでいない点も多々あると思いますので、みなさんのお考えをどんどん出していただきたいです。そのことが、ネットワーク調査に関心のある人たちのためらいを払拭し、



辻 竜平氏

より多くの人々の参加とこの分野の発展につながっていくのではないかと考えています。そんなわけで、本日は調査法に関する座談会にお集まりいただいた次第です。

ではまず、お集まりいただいた皆さんの自己紹介をお願いいたします。最初に私は、信州大学の辻竜平です。ネットワーク分析との出会いは、学部の3年生のときに、平松闊先生が、私の母校である関西学院大学に非常勤で来られて、先生が1990年に出版された『社会ネットワーク』（平松闊編、福村出版）のゲラを教科書にして習ったことでした。修士課程修了後に、カリフォルニア大学アーバイン校に留学し、本格的にネットワーク分析を学びました。

私自身の関心は、人々がネットワークを通じてどのように協力を引き出すことができるかというところにあります。そういう関心から、現在は長野県内の地域社会において、人々がどのようにして協力関係を維持しているのかを研究しています。

あとは、ダンカン・ワッツの『スモールワールド・ネットワーク』（原題、Six Degrees）を翻訳（2004年、阪急コミュニケーションズ）したこともあって、さまざまな現象をスモールワールドの視点で考えると、どのようなことになるかを考えています。たとえば、流行

であるとか、世論であるとか、そういうことの形成にも関心をもっています。

安田 安田です。ネットワーク分析はアメリカ、コロンビア大学で出会ってそれ以来ネットワーク一筋で研究をしています。基本的には人のネットワークに限らず、松尾さんと組んでミクシィのSNSのネットワークの研究をしたり、あるいは経営の先生方と組んで生産工程とか、企業間のネットワークの研究をしたり、分野横断的にやっています。非常に多様なネットワークを見ながら、ずうっと25年さまよっています、いまだにわからないことのほうが多いです。

ネットワーク調査で今いちばん問題だと思っていることがあります。関係というのは、目には見えません。この絶対に見えないものを、われわれはあたかもあるかのごとく集めて、その状態を記述して、その効果、あるいはその影響力について語ります。ある意味、魔女みたいなことをやっているわけです。誰にも見えない、本当にあるのかないのかわからないものについて集めて形にして、無理やり絵を描いて可視化して、その効果を語る。今まで25年やってきて、私は本当に何を作ってきたのか、測ってきたのか、抽出してきたのか、本当にきちんと調査ができてきているのか。ものすごい嘘をついていたような気もするし、ものすごく重要なことを指摘してきた気もします。

ネットワークの社会調査に入りたい方が増えてきたように思いますが、それだけに、危険やリスクや、あるいは面白さや可能性も同時に感じているというのが現状です。

松尾 松尾です。呼んでいただいてありがとうございます。私自身は、もともと工学系、情報系で、2002年に東大の情報系で博士号を取りました。そのときの研究室は人工知能の研究室で、石塚満先生の所でした。私自身

はずっと人工知能の研究をしてきて、人工知能を作りたいなど、コンピューターに知能をもたせるといのはどういうことなのだろうかと思って研究をしてきました。

ちょうどWebがすごく広まってきた頃だったので、検索エンジンの研究が進んできたり、扱えるデータがたくさん増えてきていました。このデータが人工知能を作るうえで1つの重要な要素になるのではないかと、つまり、人々がWeb上でいろいろな行動をしたり、いろいろな考えを書いたりするので、それをうまく使えないだろうかと考えました。

人工知能の中では、昔から関係性がすごく大事なのです。知識イコール関係性みたいな面があって、意味ネットワークという研究分野があります。そこでは、世の中の概念間の関係を書き表すことによってコンピューターは知能をもてるのではないかと考えます。そして、人工知能における知識表現においては、何らかのエンティティとか、オブジェクト同士の関係性を記述します。そこでは、何かと何かがどうしたという関係の記述の総体が知識であると考えられているので、関係性ということにもともとすごく興味がありました。

ちょうどその頃、辻さんの訳したワッツの『スモールワールド・ネットワーク』が流行っていて、私もすごく興味をもちました。たとえばキーワード抽出法では、文章の中にある単語の関係をネットワークとして表すとスモールワールドになるので、ある単語を取り除くと、その世界がいちばん大きくなるような単語がキーワードになるのではないかという論文を書いたりしました。そんな感じでネットワークに興味をもちはじめました。本日はいろいろ勉強させていただきたいと思います。

野沢 明治学院大学の野沢です。私は、1980年代半ばに、大学院で都市社会学の研

研究室に所属して、団地などのコミュニティを研究する調査に加えてもらったのですが、当時のコミュニティ研究は「居住地域」を前提にしていました。しかし、バリー・ウェルマンやクロード・フィッシャーなどの研究に触れて、居住地域に限定されないネットワークという観点の有用性に気づかされました。同時に、家族とコミュニティがいかにかかわっているかに興味をもって、1990年代には夫婦関係や親子関係と、それを取り巻くパーソナル・ネットワークとの関連を、いくつかの場所で調査してきました。

1990年代半ばに、1年ほどトロント大学のウェルマンのもとで勉強する機会を得て、ネットワーク論についての知識を少し深めることができました。それが刺激になって、後にネットワーク論の主要な論文を翻訳して、リーディングスを編みました。今でも家族とネットワークの研究もしていますが、この10年ぐらいは、ステップファミリー（子どもがいる人が再婚して作る家族）についての研究をしています。家族の関係自体も複雑なネットワークになってくるという意味で、ネットワーク論の視点が活かせると思っています。が、その調査・分析は難しく、その点为本日のテーマにも重なってくるようなところもあります。

石黒 日本女子大学の石黒です。私は、東京大学の池田謙一先生の弟子でして、その池田先生が、私が学部の3年生か4年生の頃に、ネットワークの同質性と社会意識、池田先生の場合は、社会意識の中でも政党支持ですけども、その話を始められまして、それが非常に面白くて、巻き込まれるようにネーム・ジェネレータとか、スノーボール・サンプリングの調査に参加させていただくようになりました。

私がネーム・ジェネレータに最初にかかわ

ったのはわりと早くて、1996年です。スノーボール・サンプリングを最初にやったのが1998年。それ以降博士論文を書くまでは同じようにスノーボールの手法で調査をやってきました。

その後、青森県の弘前大学に2001年から2010年までいて、経済学の李永俊先生と一緒に仕事をするようになりました。青森の若者たちの就職とか、社会参加について、経済学者は経済的な側面を、私は階層的な側面や対人関係、社会ネットワークの側面を研究してきました。

その延長として、地方から大都市に移動をした人たちと、地元に残っている人たちの人間関係が切れるのかどうか。切れてしまったとしたらどうやって再生しているのだろうかというのを、フィッシャーの都市社会学的な視点を含めながらずっと研究をしてきました。

中里 明治大学の中里裕美です。2004年の秋に、京都大学で開催されたロナルド・バート教授の講演会への参加をきっかけに、社会ネットワークの視点からの地域通貨取引に関する研究を進めています。京都大学の若林直樹先生や、安田先生、金光先生に教えていただきながら、当初は地域通貨組織の30人ぐらいを対象とした取引ネットワークの中心性などの構造的特徴量を測るというようなことから始めました。

日本において地域通貨活動は、その経済的な面というよりは、地域コミュニティ内の人同士のつながりを取り戻すという社会福祉的な側面に重きを置いて展開されています。私自身は、そういう地域通貨組織内の会員（参加者）の取引ネットワークの構造であったり、それが地域社会に対してどのような効果をもたらしているのかというようなところに関心をもって研究を進めています。本日はオブザーバーとして参加させていただき、勉強させ

いてはどんなことを意識されたのですか。

野沢 先ほども言いましたが、トロントでウェルマンの大学院の授業に出て勉強し直してみると、一通りこういうものを読めばネットワークについて問題の所在がわかる1冊のリーディングスがあると便利ではないかと思って作りました。

ネットワーク研究は、たしかに従属変数と独立変数の両面からネットワークを扱います。このテキストではその両方ありうるということを書いて、片方からだけしか扱えないものと考えさせないようにしていることは、重要なのではないかと思います。どちらを取るかは問いの立て方に依存するということをまずはわかってもらえればよいのではないのでしょうか。

石黒 因果という意味では、それを決定的に重要視している領域もあります。私が専門としている社会心理学もその1つです。たとえば、われわれが類似した相手を選択してネットワークを作っていくからホモフィリー、すなわち同類結合ができるのか、もともと結合があって、その後で人々が似ていくのかという研究には、30年、40年の積み重ねがあります。結局わかったのは、関係ができた人が、それから似るのではなくて、関係ができた相手を似ていると感じるということでした。

松尾 今石黒さんが言われた、似るから近くにいるのか、近くにいるから似るのかというホモフィリーにかんする点については、コンピューターサイエンスの分野にもそういった研究があります。たとえば、eコマースで商品を推薦するときに、近くの人の買ったものを推薦すると効果があるという研究がある一方で、自分と同じ物を買った人を推薦してあげると仲間や友だちになるということもあります。

私がやったのは、その両方の効果の研究で



野沢慎司氏（左）、石黒 格氏（右）

す。2人がどのぐらい同じ商品を買うと、推薦したときに効きやすいかとか、逆にどのぐらい共通の友人がいたりネットワーク上で近いと、推薦が効きやすいかを測るのです。これは、@cosmeという化粧品サイトのデータでやったのですが、わかったのは、その両方の効果の強さはブランドごとに違うということです。

とくに強いブランドというのは、両方の効果が強いのだということ。そうするとだんだん似てくるのです。似てきて、かつ集まってくるのです。

そうすると、どんどんその集まりが隔離されて結びつきの強いコミュニティができていく感じになるのです。これは、一方向だけでなく、両方の効果があるのではないかというのをデータで示したことになります。

安田 それは、リアリティが可視化したということですね。

松尾 そうです。

石黒 eコマースの場合は見えるのです。先ほど言った、似ていると思っているだけというのは、だいたいパーソナリティとか、価値観とか、見えないものを扱うという心理学の研究なので、バイアスというか思い違いがいっぱい出るのです。eコマースははっきり見えるところを扱いますから、当然、推薦とか共通の友人が見えれば効くわけです。私

の立場からすると、見えるのに、効かないブランドがあるということの方が面白いです。

松尾 効かないブランドの場合、人々は、ソーシャルグラフ的に見ると、いろいろな所に散らばったままです。私は、コミュニティグラビティと名づけたのですけれども、グラビティが働かないから、どんどん集まってこないのです。商品力があつたり、資本力がある会社がやっていると、グラビティが働いてマーケティングが自然にできてしまうのではないかと思うのです。

安田 山火事が発火しないのです。面白いですね。

松尾 発火しないのです。ともあれ、e-コマースでは、両方向の効果がダイナミックに起こることでネットワークが発生していると思うから、因果関係の特定はそもそも難しいと感じます。

辻 ネットワークの形成は、本当はダイナミックなプロセスなのです。しかし、社会学者がやっているような、1回限りの郵送調査の場合には、ダイナミクスというか、そういう部分は落とさざるをえなくて、とりあえず想定された原因となる変数、想定された結果となる変数を決めて分析していかざるをえないという難しさがあるのでしょうか。

もっと短いスパンで機動的にデータを時系列にうまく集められるような、そういう研究ができれば、松尾さんがやられているように、短いスパンでデータを集めて、それで因果あるいはその循環を扱うこともしやすくなるのかという気がしました。

石黒 もう1つは、そんなことは何も考えなくて、とにかく記述ができればいいという研究スタイルもありだと思のです。どこどこにつながりがありうるかという、それだけで十分な場合もあります。たとえば自治体がやってみるということもあってよいと思う

のですが、そういうときにあまり因果についてよく考えなさいということ強調する必要はないと思います。

野沢 私は、『社会ネットワークの研究・メソッド』の第3部のNPOがどう展開したか、そのダイナミクスを追いかけた質的な事例研究の紹介の価値が高いと思いました。因果関係などはあまり気にせずにデータを集めて、偶然がいろいろ働き、誰かと誰かが出会ったことで、次の展開が導かれたというような記述的な分析です。出会わなければこんなことは起きなかったという意味では因果なのだけれども、ネットワーク形成は偶然性に支配されているところもある。選択する以前に、たまたま近くに住んでいた、何かに所属していたことが決定的な分水嶺です。

そういう意味で、因果でもあり、因果を超えているような要因が働いて、物事が展開したというような質的な研究もここに入っているのが非常によいと思いました。厳密な測定の問題だけではなく、質的な調査をどうやるかについてもニーズがあるように思います。ここに書いてある以上に、もっと多くの調査事例を共有できるようにして、たとえば学部の学生が卒論研究のために自分1人で団体の変化を追いかけるのに役立つようなガイドブックがもっと出てきてもよいと思います。その意味で、これは1つのよい試みだったと思います。

辻 いろいろな論点を出していただいてありがとうございます。これはここまでとしましょう。

❧ プライバシーへの配慮

辻 次に扱いたいのは、プライバシーの問題です。『社会ネットワークの研究・メソッド』でもこの問題が取り上げられていますが、第2章第1節に、プライバシーの問題

をクリアして、可能であるならば想起法よりも認知法は確実性が高いと、さらっと書かれているだけです。ちょっと簡単すぎるかなという気がしています。ネットワーク調査の場合には、通常の郵送調査以上にプライバシーの問題が起りやすいと思うのです。その辺りについて、一般論としてではなく、たとえば、このような配慮を自分はしてきたというような経験をお聞かせいただけたらありがたいと思うのですが。

安田 私の場合は、配慮はそれほどせずよいデータは取ったのだが、逆に論文が公開できないというものがいっぱいあります。とくに企業間の取引データとか、売上データなどというのはお金もいただいてコンサルとしてお受けしますが、実データが生々しすぎる。10年、20年と寝かせても許可を貰えるかどうか分かりません。

また別のときには、農村に調査に入って、個々の農家の関係を取って、望ましい関係、今ある関係を対比して見せて、その差分について考察をさせていただきました。一応その土地の方へのフィードバックはしたものの、これも公開はできない、名前が出せない。

このように、実データを取ったのだけれども、個人情報がいかに多くて論文化できない研究を、私はものすごく抱えていて、墓場までもっていくべきかと迷っています。

辻 私も、ネットワーク調査をやっている人からそういう話をよく聞きます。本当にそこは難しいです。

石黒 人類学者は、アフリカとかニューギニアへ行ってきた、やはりジレンマに遭うみたいですね。あの人たちはいくつかのことをして、事実を一部改変する。たとえば調査地をずらすとか、内部の要素を、本質的なこと、自分の議論がずれない範囲でずらすとかいうことをやっています。

多数の個人のパーソナル・ネットワークを計量的に分析するときには、データはすべて匿名になってしまいますし、測定も限定的ですので、生々しい問題はあまりありません。しかし、スノーボール・サンプリングでデータを収集するときに、回答者から紹介してもらう人の住所をこっちに送ってもらうというやり方をしたときには、いわゆる個人情報の問題が出たように思います。

小規模な所で、回答者との間にラポールができた所でやるのならまだわかるのですが、あのときは何人だったか、とにかくランダム・サンプリングで、郵送でワッとやっちゃいました。

あの当時は、池田先生の調査に私も参加させていただいた初期の頃なのですが、暗中模索ということもあって、いろいろと調査法を試そうして使った方法の1つでした。結局、こっちに個人情報を譲渡させるやり方はよくないねということがわかって、それ以降やらなくなりました。

ただ、プライバシーというべきか、主観的な認識の問題というべきか難しいのですが、人間関係について聞かれただけで嫌がる人はかなり多いです。とくにネーム・ジェネレーターで配偶者のことなどを聞くと、うちの夫婦の関係が、お前に何の関係があるんだといった反発は当然あります。どこからがプライバシーかという基準を示すことができるのだろうかという難しいのですが、状況にもよりますが、唯一あるとしたら、通常の調査で回答者に聞けないことは、スノーボール調査を含めネットワーク調査で聞かないほうがいいのだろうというぐらいかと思います。

もちろん通常の調査で聞けることでも問題になることもあります。スノーボール形式の調査が来たから姑に渡した。しかし、姑が私について書いた内容がひどかった。それで夫

婦で大喧嘩だ。離婚の危機をどうしてくれると感想欄に書いてありました。これは自分たちの問題なのか、あちらの問題なのかまったくわからなくて困ったのです。お嫁さんが元のターゲットで、スノーボール・サンプリングで渡したようです。そこで、お互いをどう思っているかの評価があったのですけれども、姑さんの結果を見たら、自分の評価が低かったようです。

野沢 見せずに返す方法ではあったのですよね。

石黒 そうです。お嫁さんにもお姑さんにも、個別に返送用封筒を提供していました。

野沢 それでも問題が起きるのですね。

石黒 これはプライバシーなのかどうかかわからないのですけれども、そういう危険性がある。もう1つは、それぞれ村落でスノーボールをやると、普通は隠されているはずの、隠蔽されているはずの関係が顕在化してしまう。仲のよい人に回してくださいという、お前の所には来たの、俺には来ないよというのが見え見えになってしまうことがあります。

安田 今おっしゃった、潜在的な関係を顕在化させてしまうリスクと、その前におっしゃっていた、回答者本人の話ではなくて、回答者がその先に連なっている人の情報を間接的に取る側に回ることになってしまうリスクの2つのリスクがあります。この2つの問題は全然別ですが、両方とも怖いんです。だって、結局どちらも関係を壊すことになりえます。

石黒 今出した夫婦の例は、私がいちばん最初の頃にやったようなスノーボール・サンプリングですけれども、忘れられないのは、どうしようと思ったからです。モヤモヤしています。それは、もともと関係が悪かったのだろうと思いつつ、私たちが調査しなければ起きなかつたらうし、やはりずうっと覚えていますね。

安田 ただ、その頃から、制度も意識もだいぶ変わりましたからね。先生方はかなり早い頃からネットワーク・バッテリーとか言ってやっていたから、当時はまだゆるゆるで、その先の人の情報を取ってきて、住所も教えてね、というのがまだ許されるところが雰囲気としてあったのではないですか。その頃と比べると、だいぶ規範としてはよくなったのか、厳しくなったのか、変わったのは変わったなと思います。

石黒 当時は個人情報保護法もなく。

安田 10年以上前。やはり先駆者だけのご苦労があるのですね。

野沢 先ほど夫婦関係について苦情が出たということでしたけれども、家族研究はまさにそこに焦点をあてるわけです。別にネットワーク研究でなくても、夫婦関係とか親子関係についての評価をたずねるのは、家族についての大規模調査ではむしろ一般的です。ほかの家族に見られずに回答できるようになっていけばよいという前提で実施されていると思います。

私がかかわった調査では、夫と妻の両方に答えてもらう、さらには子どもにも回答してもらうことがありました。ネットワーク的な家族の捉え方だと言えるかもしれませんが、そこには同じような問題が生じる可能性があります。もちろん1人ずつ個別に調査票を入れる封筒があって、それに入れて送り返してもらうのですが。

ネットワークの研究では、ポジティブな関係にある人たちを引っ張り出してくることが多いのですが、対立関係や非常に不満が大きい関係も現実のネットワークを捉えるときには重要です。だとすると、(プライバシーの問題から逸脱してしまうかもしれませんが)家族というのはみな仲よくあるべきという社会規範がある一方で、現実にはそうではない実態を

ネットワーク的な観点から析出するのは非常に難しい。また質的なインタビュー調査などで、夫と妻の両方に話を聞くと、妻から聞いた話を夫には言えないことが多々出てきます。先ほど安田さんがおっしゃっていたように、墓場までもっていくしかないと思います。出せるところは結構限られてしまって、抽象的にぼやかして書くしかない。ジレンマを感じます。

安田 批判的すぎるのもいかがかと思うけれども、今おっしゃったようなネガティブな関係の調査とか、その影響力みたいなものは、たしかに分析して考えなければいけない。実態を知りたいと思うのだけれども、まず論文に出てこない。

ただ、研究者はみんなわかっていて、ネガティブなことがあって、ここは触ってはいけない。でも、それは機能としては動いている。それを人を傷つせず、対象を傷つせず、でも重要な問題であるから論じるかというところは難しく、あまり（研究が）ないです。

松尾先生のようなご研究では、Web上に公開されている情報を取ってきますが、人が進んで開示しているネットワークであれば、それをどう使ってもいいではないかという、ものすごく強いところがあります。

松尾 フェイスブックでは、情報を友だちまで開示するとか、友だちの友だちまで開示するとか、プライバシーの設定があります。その設定がうまくいっていないと、意に反して情報が出てしまうことがありますけれどもね。

SNSの研究では、たとえばプライバシーを保護しつつ、いかにマイニングするか、いかに匿名化するかといった研究がされていますが、私自身はあまり興味がありません。後からそういった情報をきちんと処理してから出せばよいのだと思います。

石黒 論文に公開される瞬間にどれだけ消せるか。

松尾 そうなのです。

石黒 調査の時点でプライベートな情報を取るなどと言う人が結構いますが、そんなのは質的研究者からすれば言えない、不可能なことなのです。彼らが気にするのは、公開されるときにどれだけ隠せるかなのです。そういう意味で言えば、データ採取のときに、一定程度個人的なデータがもし入ってしまったとしても、個人が特定できる形で公開しなければ問題はやや低い。ただ、これは調査側の言い分かもしれませんね、取られる側からすると。

個人情報保護法では、当人が同定できない限りそれは個人情報ではないのです。定義が、その人個人を同定できることなのです。メールアドレスは個人情報かという、かなり微妙になってしまいます。その辺はたぶん法律の問題はあるのですけれども、人々の認識の中には、法律とは別に、これは私の情報で、勝手に使うなというのが出てくる。何万というデータを集めて、1%でも怒る人がいたら大変だろうと思ってちょっと気になるのです。

松尾 私たちの研究では、公開されているデータを分析しますが、当然個人が出る形では出さないですし、統計的にどうかという全体の性質の話しかしないので、そこは大丈夫なのではないかと思います。実際にどういう処理がされてどうなるかはわからないですが、最後のところは研究者として正しいことをやっているのかどうかという研究者倫理みたいなところで縛るしかないのかなという感じはします。でも、難しいところだと思います。

少し関係するかもしれませんが、最近見つけた面白いことがあります。ツイッターユーザーはだいたい全体の10%ぐらいが自分の現住所を書いていて、残り90%は自分の住

所を出していません。だから、ある人と相互フォローしている人が100人いるとすると、その100人の中の10%、すなわち10人は自分の住所を書いていることとなります。そこで、その10人の住所の最頻値を取ると、じつはこの人の住所と95%ぐらい一致するのです。この人は全然思ってもみないと思いますが、自分の相互フォローの関係から住所が同定されてしまうのです。もちろん市町村レベルしかわからないので、そんなに問題視することもないとは思いますが。どのぐらいマイニングできるか、分析できるかという技術の裏返しとして、プライバシーというものがあるのかなと思います。だから、プライバシーを守るためには、技術としてはもっと先に進みながら、危ないところをちゃんと守っていくのが大事なのかなと思います。

野沢 自分とつながっている友だちの情報は、自分の個人情報に近い、重要な情報なのです。

松尾 自分の友だちが「僕は出していいよ」と言ってしまうと、そこから自分の情報が漏れてしまうわけです。

石黒 われわれが従来考えてきたプライバシーと呼ばれるものをもう少し広げて考えないとまずいかもしれないですね。

辻 私も、自分が想定していた話よりも随分いろいろな問題があるのだなということを再認識しました。ありがとうございました。

❖ 社会ネットワーク研究におけるサンプリング

辻 次は、ネットワークの調査のサンプリングにかかわる問題点で、ネットワーク・データの代表性と、特にスノーボールの代表性辺りの話が中心となります。『社会ネットワークのリサーチ・メソッド』では、サンプリングの問題については、ほとんど触れられて

いません。第3章で、パーソナル・ネットワークの調査例を扱いながら、その調査例においては、どのようにサンプリングをしたか、そのサンプリングは必ずしもランダム・サンプリングではなかったと述べられているだけです。また、調査目的によっては、ランダム・サンプリングにこだわらなくてはならない場合もあるのだとも述べられています。しかし、サンプリングについては、それ以上のことは述べられておらず、不完全燃焼な感じは否めません。どういうときに、ランダム・サンプリングが必要で、どういうときに必要ないか、また、どういったサンプリングがよいのかといったことはまったく書かれていないのです。

野沢 実際、家族社会学会の全国調査のように、夫婦関係やそれ以外の関係の実態を見るために、代表性のあるサンプルを取ってくる調査はあります。現代の家族とパーソナル・ネットワークを論じるうえでは、代表性のあるサンプルを取ってくるが必要になるからです。でも、違う目的でネットワークを調査しようとしたら、ランダム・サンプリングのデータにはまた別の限界があるので、どこかで全数調査をやるのが相応しいという場合もあるでしょう。そちらのほうが知りたことがわかる。サンプルの代表性が必要になる場合とそうでない場合がある、ということでしょうね。

辻 社会ネットワークの調査には、ある集団1つを取り上げてそのホール・ネットワークを全数調査するということがあります。その場合、その集団が母集団でもあり、その集団の各種のネットワーク特性を記述すればよいこととなります。この場合は、欠票が生じるということを除いては、サンプリングは問題になることはあまりありませんし、サンプルの代表性を気にすることもありません。

安田 でも、その場合は、回答してくれるかという問題よりも、回答はしてもらえても、その内容のほうの問題になることが多いです。結局、顔が見える人間関係の範囲で調査をするのが、関係データを取ったりするときにはいちばんいいのです。

たとえば、商店街の調査に入るのも、お酒を飲んだり、付き合ったり、お届け物をしたり、頭を下げたり、買物に回ったり、そういうことを普段からやっていないと、関係データを取ることはなかなか難しい。顔も名前も知らない人に、調査会社に頼んで調査をかけるだけで取れる社会意識の変数とか個人の性質の変数だと、それは楽でいいとは思いますが。ただ、いわゆる自分のプライベートにかかわる友人や人間関係のデータには、何か違いがある。調査を行う前の事前の準備とか根回しが必要な特殊なデータを、われわれはほしがっているのだなということがありますね。

辻 今の話で気づいたというか、認識を改めたところがありました。私は、後で扱うネーム・ジェネレータとかジオグラフィック・ジェネレータは、問題の構成が複雑であるうえに質問自体が難しすぎて理解されていないのではないかと思っていたのですが、今の話を聞いていると、これは答えたくないというか、答えられないというか、そういった理由で無回答が多くなっている可能性も相当高いのかなという気がしました。そのへんは私も認識を改めないといけないと思いました。本当は答えられる、答えようと思えば知っているのだけれども、答えを出すのを躊躇するとか、そういった側面もあるのでしょうか。

石黒 ネーム・ジェネレータだと、普段はそんなことはしないのに、身の周りにいる人を評価させたりするじゃないですか。「あなたはこの人に不満がありますか」とか、「どのぐらいの親しいですか」とか。

野沢 とくに「親しい人」を挙げるのは難しいかもしれませんね。

石黒 ああいうものに抵抗を感じる人がいるのはわかりますね。こちらからするとほしい変数なのですが。

野沢 先ほどの安田さんの商店街の話聞いて思い出したのですが、私の院生時代、1980年代に、東京の下町の一町内会で近所付き合いをたずねる全数世帯調査をやったことがありました。町内会長にまず話を通して、何度も通って、1軒1軒訪ねて面接していくスタイルでした。今は郵送調査が非常に多くなってしまいましたが、ややこしい人間関係をたずねるときに、直接あるいは電話で詳しく説明してからたずねれば、丁寧に対応できるかもしれません。もう一方で、直接訪問されるほうが面倒だから拒否されやすかったり、訪問のコストが大きかったり、さらにプライバシーの問題もあってジレンマなのですが。いずれにせよ、ベストな方法があるわけではありません。目的によって違いますが、数量的なデータを集めるときにも、丁寧に集める方法を少なくとも試してみてもいいのかなと思いました。

安田 ネットワーク・データに限った話ではないにしても、とりわけいちばん出にくい性質をもっているから、こういった方法を継承していくことは課題かなという気がします。

辻 ここまでのテーマは、ネットワーク調査におけるサンプリングについてでしたが、それよりも、調査を行ううえでの欠票や、ネットワーク項目への無回答を減らすための留意事項の話になりましたね。たとえランダム・サンプリングによる調査でも、ネットワーク項目に回答していただけないようでは、ネットワーク調査としてはまずいわけですから、このような問題をご指摘いただいたことは意味があったと思います。



左より安田 雪氏, 松尾 豊氏, 中里裕美氏

辻 では次に、スノーボール・サンプリングの問題点に移りましょう。これについては、この中では石黒さんがお詳しいかなと。

石黒 スノーボール・サンプリングで問題になるとしたら、最初の回答者（主回答者）からもう一段スノーボール・サンプリングするときで、そこでどんな問題が発生するかはきっと誰も全体的にはわかっていないのではないかと思います。でも、主回答者から次の回答者に回してもらうときに、そこがランダムであると仮定するのはばかげています。配偶者が出やすかったり、家族が出やすかったりするものです。スノーボールでわかるのは、それがその人たちの想起しやすいネットワークだということなのです。問題は、方法論によって結果が変わるときでしょうか。たとえば私たちが2004年に行った調査のデータだと、主回答者のネーム・ジェネレータでは半分ぐらいは非親族で占められていたのに、そこからさらにスノーボール・サンプリングを依頼することによって、非親族が回答される割合が3割まで減ってしまうのです。こうなると、これはバイアスと呼ばざるをえない。

では、このネーム・ジェネレータで測られた親族半分というのがベースラインだと言っていいのかというと、わかりません。ここから先はもはや形而上学ですね。

安田 補正をするなり前提条件なりをしっ

かり論じて、今おっしゃったようなスノーボールのときはここの先はこういう偏りがあって、このぐらいの割合を補正して割り引いて考えなければいけないとか、ここは元の基礎数字とは言っているけれども、実際この前提条件を変えるとこのぐらい変わってしまうみたいな、その細かいところがまだ詰められていない感じがするのです。スノーボール・サンプリングという呼び方には、サンプリングという言葉が入っていて、そのために、そうやって得られたサンプルには代表性がありそうだなみたいな気持ちにさせられているところがあります。私がそれを見てそう思うってしまうのが悪いのかもしれないのですが。

石黒 もう1つの問題は、主回答者にネーム・ジェネレータで次の回答者となる人の名前を出してもらって、そこからサンプリングするわけですが、スノーボールとネーム・ジェネレータで出てきた人の対応が取れないということがよく起こることです。

普通の調査でも、回答すべき人と回答した人が違うということはあるんですが、そういうことがもっとずっと起こりやすくなっている気がします。何でわかるかという、主回答者にこの人は男性ですか、女性ですかと聞いて、次に出てきた本人に、あなたは男性ですか、女性ですかと聞くと、こんなレベルでさえ、もう食い違いが数%も出てきたりする

のです。関係はどうですかと聞くと、片方は家族と言っているのに、こちらは友だちと言っていたり。先ほど話題にした、本当はやるべきではない主回答者に他者の住所と氏名を聞くというのは、そのためにやったことなのです。当時は調査票にIDを入れて本人確認をするということを普通にやっていました。それと同じことを、スノーボール・サンプリングでもやろうとしていたのです。

他にも、スノーボールをやったときに、全部1人で書いたという人がいました。「人に迷惑はかけられませんから、全部書きました。あしからず」と言って、全部同じ回答をして、1つの封筒に突っ込んで返してくる人がいるのです。そう言うのであればわかりますが、言うてくれなければ調査者にはわかりませんので、問題が発見されずに素通りされてしまうことになります。

野沢 ほかにはどういう可能性があるのですか。非親族だったものが親族になったりというのは、誰か違う人を介して頼んだとか。

石黒 おそらく、ネーム・ジェネレータの時点では、思いついた人をバツと書くわけですね。質問紙では、だいたいその後ろに「では、この人たちに封筒を送ってください」と書いてあるわけです。そここのところがきちんと行われぬ。

野沢 それで違う人に送ってしまうと。

石黒 そうだと思います。

安田 それに加えて、そこまで親しくはない、それは頼めないというバイアスがそこで出てくるということもありますか。

石黒 それも、たぶん、あると思います。

野沢 それで、頼める人に頼むということになってしまったと。

石黒 だと思います。たとえば、自分が主回答者の場合、職場の何とかさんはよく話すということを書いたとしても、上司であるこ

の人に次の回答を頼めない。

安田 躊躇しますよね。

石黒 多くの人は、調査者が、たんに協力を増やしたいだけだと思っているのだと思うのです。それなら、手近な人に頼めばいいということになるのでしょうか。

安田 話は少し違いますが、企業や会社の社長さんとか、商店街の店主さんなどに調査をお願いすると、他の従業員やアルバイトが答えてしまったということも起こります。中堅の従業員と社長が答えたのでは、全然違った答えが返って来たりします。ネットワーク調査に固有ではない話ですが。

野沢 先ほど触れた家族調査でも似たような問題が生じます。回答者の配偶者にも回答をお願いして調査票が戻ってきても、本当にその配偶者が書いたかの確認が難しい。とくに夫婦両方の回答が揃ったら謝礼を多くする場合だと、代わりに配偶者の分も書いてしまうことが起きやすくなります。

安田 関係データを集めようとする、おっしゃったようにいちばん頼みやすい、いちばん公的に認められている夫婦とか、表に出してもいいところが出てきて、実際に調査で回答しやすい関係と実態で結びついている関係とは差があるのだらうと思います。

石黒 そうですね、かなり大きい差があるのだらうと思います。今日は前向きに対応を検討するという会の趣旨もあるので、どうすればいいのだらうということなのですが、1つは人数を増やすのがバイアスを下げる方法ですね。回答してもらう他者の数を増やしていく。家族は核家族化が進んでいてせいぜい4、5人しかいないので、極端な話、10人聞けば家族だけで占められるということとはほとんどない。その辺は、パートがGSSのデータを1980年代に分析していますが、回答他者数を増やすと親しさのバイアスも減るし、

接触頻度の高さのバイアスも急激に減っていくという分析をしています。たぶんそれは日本でも同じです。でも、人数を増やすのは回答者の負担が重いとすると、そういった調査項目を設けた理由を正直に言うしかないと思うのです。このデータは人間関係を調べたくてやっていて、あなたのデータとあちらの人のデータを対にしてデータを分析するつもりですよ、と言うしかないのではないのでしょうか。

このようなやり方はたぶん回収率を下げることになってしまうのですが、もう一方で、あの人でいいやとか、数を集めればいいのだから頼りやすい人にするとか、私が全部書いてしまうとかということとはしなくなると思うのです。

安田 精度は高いデータが返ってくる。

石黒 精度は高いけれども、回収率が下がるデータになる。でも、精度を上げるほうが大事な気がするの、それがいいのではないかと。このやり方は、調査票にIDを振ることの理由付けになるのかなと思っています。

これは弘前大学在職時に青森でやったからうまくいっただけかもしれませんが、若者の就業の話をするときに、親子のペアデータを取ろうとして、親子でスノーボール・サンプリングをしました。このとき、IDを振って、このIDは親子をペアのデータとして分析するから付けましたと回答者に伝えました。すると、IDを付けた調査でよく遭うID欄が切り取られてきたといったことがいっさいなかったのです。他のネットワーク調査の場合でも、この調査ではあなたの回答と挙げてもらった人たち（1番、2番、3番……）の回答を組み合わせて使いますから、ここのネーム・ジェネレータで主回答者が挙げた人とスノーボールを転がした先の回答者は同じ人にしてくださいと正直に頼み、封筒に1番、2番、3

番……と書いて、その人たちは一致させてくださいとお願いするのが、たぶん1つの手かなと。

野沢 回収率は下がったのではないですか。

石黒 いや、そんなに下がった印象はないですね。もともと若者ターゲットだから低いというのもあるのですが。若者にやると、その年齢層でダイアド（ペアデータ）が取れるのはせいぜい1割5分から2割とかなので。ただ、そこから劇的に下がった感はないのです。

ほかにも、主回答者のネーム・ジェネレータの問いに、挙げてもらう人に応じて、回答欄の色が赤い縁取りをしてある質問と青い縁取りをしてある質問を並べておき、また、封筒に赤と青の縁取りをして、赤の人には赤の封筒を送ってください、青の人には青の封筒を送ってくださいとしてわかりやすくするという工夫をしたこともあります。

安田 今の色のノウハウは、すごく面白いことをお伺いできたと思います。

石黒 結局、正直に言うことが信頼を育む部分もあって、自分たちの手で自分たちの大学や研究所の近辺でやったときにはスノーボールの回収率が高いのです。でも調査会社に頼んで広域にやろうとすると、ガクッと落ちるというのは経験しましたね。

弘前に移って1年目か2年目にスノーボールをやって、びっくりしました。77%の回収者から少なくともダイアドが取れたのです。40%程度の調査回収率で、77%のダイアドというのは高い値です。辻さんと2004年にやった全国調査では、ダイアドがとれた回答者は15~20%ですよ。回答者を分母にしても30%ぐらいでした。

辻 他に、ネットワーク調査のサンプリングや代表性にかかわる問題はありますか。

松尾 私がいつも抱えている問題に、ネッ

トワーク・データが逆に取れてしまっただろうとすればよいか困るということがあります。でも、取れてしまったときに、その取れてしまったネットワークにそもそも意味があるのかどうかよくわからないのです。たとえば、お買物アプリを使った人が5万人いて、その5万人のネットワーク・データが取れたときに、これは何なのかみたいなのがあるわけです。そのときに、たとえばお買物アプリを使っているユーザー層の年齢、性別の比率が、インターネットユーザーの年齢、性別の比率と一致していれば、インターネットユーザーを代表するよいサンプルなのかもしれないと思うのですが、そこがずれていることがある。

ずれているとわかったら、ではこのネットワーク・データはどういうデータなのかというところはやはりわからなくて、どうしたらいいのかなと。ここで実数を出したり中心性を出したりクラスタリングしたりしてどうこう言うのですが、この結果が何なのかが言えないのです。たとえば、日本の主婦層なり対象としたいところとある程度近い分析結果になっているはずだというようなことが言えればよいのですが。

石黒 一般化可能性ですよ。代表性というのは、母集団をどこに決めるかによります。母集団の中で結果がわかれば一般化できるのですけれども。

松尾 データがある中で、逆にこれを代表している母集団を知りたいという。

安田 逆問題ですね。

野沢 それは、ネットワークをどうやって引き出すかというときの問題ともかかわっていると思うのです。ポジション・ジェネレータだけではなくて、たとえば年賀状という方法を使っている人もいます。年賀状を見せてくださいと言って見せてくれれば、こういう付き合いがあるということはわかるのだけ

ども、これが一体何を表しているかという問題と同じですね。

そういえば、笹森秀雄さんの1956年の論文「都市における社会関係に関する実証的研究」(『社会学評論』6(2):58-83)もありました。これは香典帳を使って社会関係を分析した論文です。香典帳にはお葬式に来た人の名前が載っているんで、この人はどのような関係の人ですかと全部聞いていくという方法です。これも年賀状と同じで、亡くなった人とその家族とつながっている人が一目瞭然でわかるのです。

いろいろな方法でそういうものは集められるので、逆に言うとデータはあるのだけれど、それが何を意味するかを後から確認するという共通の問題を抱えているのかなと思いました。

ネットワークを研究するのは難しいと思われているので、参入の敷居は低くしたいと、私も同じ思いはあるのですが、なかなかそうはなりにくいところもありそうです。

❖ パーソナル・ネットワークの測定(1) 知人数推定

辻 次はパーソナル・ネットワークの測定ですね。パーソナル・ネットワークを測定するための測定法がいろいろあります。測定法にはそれぞれの固有な問題点があるので、それについて指摘して改善策を考えたいと思います。まず、電話帳法や携帯電話の登録件数をはじめとする知人数推定について扱いたいと思います。

従来のネーム・ジェネレータという測定法が、きわめて近い人たちについて詳しく聞くという方法であるのに対して、電話帳法や携帯の登録件数の場合は少し話したことがある程度、顔を覚えている程度、名前を覚えている程度という、そういった弱い関係、弱い

紐帯の範囲を定めていくことが目的です。なぜこういったことが必要になるかというと、マーク・グラノヴェターの弱い紐帯の強さの指摘をはじめとして、その後も、パートの構造的空隙、ワッツのスマールワールドといった弱い関係やちょっとした関係の重要性が理論的に指摘されているほか、被災地におけるボランティアとのつながりなど実践的にもその重要性が明らかになってきているためです。

これまでは知人数推定のために、とくに日本では年賀状がよく用いられていたわけなのですが、最近は年賀状を書かない人が、とくに若い人たちを中心にしてすごく多くなってきていて、もうこれは使えないのではないかと思っています。学生などに聞いてみると、「書かないです」という人が大半なので、調査対象に20代、30代の半ばぐらいまでが含まれるようだったら、年賀状の枚数を知人数の推定値として用いるのはもう駄目だろうと思います。

携帯電話の登録件数については石黒さんのほうが詳しいので、詳しいことは後ほど石黒さんに聞きたいと思います。なので、電話帳法について、私のほうからお話します。

電話帳法の測定の仕方は10個ないし20個程度の苗字を示して、そういった苗字の人を何人知っているか聞いておきます。そして、10個ないし20個の苗字の件数が電話帳全体の中に占める割合を基にして、全体として何人知っているか、すなわち知人数を推定していくのです。私が2004年度版写録宝夢巢（シャーロックホームズ）を利用して作成した苗字リスト（次ページの表を参照）の場合、その正確な推定式は、苗字数10個版の場合、 $\text{回答された知人数の合計} \times 28191237 / 572340$ 、20個版の場合、 $\text{回答された知人数の合計} \times 28191237 / 1181786$ です。

この問題点はいろいろとあるのですが、まずは、尺度の感度が低いことです。たとえば10個版として私が使っているものだと、いずれかの苗字をもつ人1人を知っていると、それが知人数としてはだいたい50人ぐらいに相当します。20個版でやっても1人知っていると、だいたい半分の25人ぐらいに相当するのです。数人単位の細かい推定ができないのです。

平均値を取ると、いろいろな所で調査をやっていますが、だいたい200~350人ぐらいの平均になります。ただし、この中でいちばん大きい問題が知人数ゼロという人がいちばん多いことです。そこで、全体が無回答なのか、リストに載っている苗字の人を誰も知らないのかを区別するため、最後に「誰も知らない」という選択肢を設けています。「誰も知らない」を含めてリスト全体が無回答の場合は0人とはせず、全体を欠損値扱いにしています。そのうえでの知人数が200~350人ぐらいになります。

分布はゼロの辺りが最頻値で、だんだんと減っていくという感じの歪んだ分布になってしまっています。知人数がゼロというのはありえないことですが、0人~50人の間の場合に、尺度の感度が低いために、0人になってしまう人がたくさん出てくるのです。

石黒 ゼロの多い問題は、この間論文を書かせてもらったのですが（石黒格「分位点回帰分析を用いた知人数の分析——分布の差異を予測する」『理論と方法』26(2):389-403,2011）、平均値ではなくて中央値を使う分析法を使うと、たぶん解消できると思います。

私は辻さんに教えてもらってこの10人のリストを使わせてもらっています。辻さんの推定法だと50人から下の人をゼロ人としてしまいますが、実際には、それらの人たちの知人数はだいたいゼロから50人の範囲に収

表 電話帳法で使用している苗字リスト

苗字数 10 個版の苗字	岩崎, 平野, 上野, 高田, 石田, 和田, 田村, 中野, 松田, 青木
苗字数 20 個版の苗字	(10 個版の苗字に加えて) 松井, 小山, 村田, 大野, 横山, 柴田, 中山, 岡田, 小川, 藤田

まっているというように考えます。辻さんの推定法で知人数の平均を出すと、ゼロから 50 人の範囲の人たちをすべて 0 として加算するので、低いほうに推定値が引っ張られてしましますが、中央値だったら、生の数字ではなくて順位を使いますので、それがゼロだろうが、ゼロから 50 の間のどんな数だろうが、推定値が引っ張られる心配をしなくてよくなるのです。

辻 なるほど、そういう手もありますね。

ところで、苗字のリストは写録宝夢巢という電話帳ソフトを使って作りました。そして、ここがポイントなのですが、全国における苗字の密度というものが、あまり大きな偏りのないものから選んでいます。写録宝夢巢は非常によくできたソフトで、県レベルや市町村レベルで、各苗字の軒数や密度などを出してくれます。それを見ると、本当に苗字というのは地理的に偏っているのだなということがよくわかります。そういう偏りのある苗字を使わないようにしながら、10 個版では、上位 40 位から 100 位、20 個版では 30 位から 100 位に入る苗字の中から選んでいます。

全国調査とか、各地から人が集まってくる都市部の調査をするときには、その 10 個ないし 20 個のリストでいいような気がするのですが、とくに地方に行くと同じ苗字の人ばかりが集まっている集落とかがあるわけです。また、ある特定の地方に多い苗字みたいなのがあったりします。そうすると、その人たちの付き合いの範囲が、そういった同じような苗字をもつ人たちの間にかなり限られてしまうというようなことになります。だから、全

国に満遍なく分布しているからといっても、地方の町村部で調査するときには、ちょっと問題があるのではないかと考えています。しかし、補正をかけるというのも難しいです。写録宝夢巢は、市町村ごとに苗字の軒数や密度が出せますが、人々は、同一市町村の中でしか付き合わないということはないですから、補正のためにどの範囲まで付き合いの範囲を考えればよいかは難しいです。まだ残っている課題です。

辻 では次に、石黒さんから、携帯電話のアドレス帳の登録数を知人数の推定値とする方法について説明いただけますか。

石黒 携帯電話のほうはだいぶ普及率は上がったと思うのですが、難しいのは利用のスタイルが間違いなく高齢者と若い人とで違っていることです。お年の方は知り合った人をかたっぱしから入れていくとか、会ったらまずアドレスの交換をして、あとは何も連絡していないといった使い方をしていないと思うので、そのところで、この方法は若い人限定だと思います。

もう 1 つはハードウェア上の制限があって、今何人ぐらいでしょうか、私が調べた 2004 年の頃は 700 件がたぶんトップだったのですが、当時は 200 とか 300 しか入らない携帯電話も多くて、もし 200 だとしたら、辻さんなどが明らかにした平均値でも上限に達してしまうことになって、かなり低いほうにバイアスがかかるという問題点が間違いなくあっただろうと思います。

これは知人数の使い方の問題なのですが、

多い人と少ない人を順序尺度的に並べられればいいというのであれば、それほど問題にはならないのでしょうか。しかし、知人数をピンポイントで推定するということに意図がある場合には、おそらく致命的になります。ネットワーク・ハブみたいな人の知人数に興味があるのだというような場合にも、すごくたくさんさんの知人がいることを順序では表現できないので、致命的になるだろうと思います。

ただ、携帯電話の登録数は、非常にお手軽です。これに答えられない人はあまりいません。

辻 もう数年するともうメモリーの上限みたいなことは心配しなくてよくなりますかね。

石黒 と思います。

松尾 でも、最近は、スカイプ上で電話をかけられたりしますし、フェイスブックだけしか知らない人とかもいますし、そういう人によってコミュニケーションのチャンネルが変わってくるということがあります。

辻 はあ、これからは、その問題が出てくるのですね。うーん。

松尾 私は携帯を落としまくるので、そのたびにもう入れ直すのが面倒くさくなります。そうすると、ここ1年とか2年で連絡した人しか入っていない。

辻 そうすると、知人数が過小評価されてしまうかもしれませんね。携帯の別の問題は、一度登録した人をわざわざ消さないということがありますね。この場合は、過大評価されてしまうかもしれません。

石黒 いや、問題にもよりますが、それは弱い絆なんだというのであれば、むしろポジティブに評価できる可能性もあるかなと。

❖ パーソナル・ネットワークの測定(2) 3 つのジェネレータ

辻 では次にいきましょう。ネーム・ジェネレータですね。本当はこのところがいち

ばんニーズがありそうですね。

野沢 ネーム・ジェネレータでは、かなり狭い範囲の相手との関係についてしか明らかにできないといわれています。けれども、まさにそこを知りたいというときには、やはり使える方法なのではないでしょうか。それをもとにして、たとえばネットワーク密度などの構造的な変数を引っ張り出すということを私もしてきました。

たしかに、それは本当に十分信頼できるデータなのかという疑問が付きまわっていると思います。ただ、この方法で結構見えてきたこともある。そこにはやはり無意味ではなくて、意味がある。一概に、使えないと言い切ってしまうのは惜しいかなと思っています。ウェルマンやフィッシャーなども、常にというわけではありませんが、パーソナル・ネットワークを使ってうまくやっているように見えます。

日本でも、たとえば松田茂樹さんの育児ネットワークの研究はネットワーク密度が中ぐらいのときにより効果をもたらすという知見を出していて、なるほどと思わされます。同様の知見が常に得られるわけでもないし、いろいろ文脈的条件に依存しているようですが、郊外の、外から転入してきた母親たちが多いような地域で、なおかつ子どもとくに何か手がかかる要因があるようなお母さんたちにとっては、そのネットワーク構造が、1つのグループにドブプリ納まってしまっているような高密度ネットワークでもなく、完全にバラバラでもない場合に母親のウェルビーイングにより効果をもたらすという知見は、了解可能な知見です。

そんなインスピレーションをもたらすような研究もあるので、パーソナル・ネットワークの中核部分の構造を捉える方法の1つとしての意義は残されていると思っています。私

自身の研究例だと、夫と妻の両方に、それぞれがもっている親密なネットワークについてたずねました。親密な相手が、夫婦間で分離されているのか、共有されているのかを分析すると、共有されているほど夫婦間の情緒的結合が強まるというポジティブな効果が見られました。

安田 いちばん問題なのは、やはり調査対象者である本人と挙げられた他者の関係は実測だけど、その他者同士の関係は、もう調査対象者の推測にしかすぎない、全然違った異質のものが交じってしまっている点です。もうちょっとわれわれが大きな声で言っていかなければいけないのですが、そこはあまり伝わってなくて、ネーム・ジェネレータで出されるネットワークは異質な紐帯が交じって、一部は認知で一部は実際である、そのところではかなりの不正確な他者間の関係を含んでいる可能性があることは、認識していただきたいところです。

辻 これはネーム・ジェネレータの研究ではないけれども、1987年のデイヴィッド・クラックハートの集団内の認知ネットワークの話(D. Krackhardt, "Cognitive Social Structures", *Social Networks*, 9: 109-34)があったでしょう。あれも他者同士の関係がどのようになっているかということも聞いているわけです。あの研究だと、会社とか、そういう普段から付き合っている人たちの間のことなので、ある程度のことは観察していて、ある程度正確にわかるころはあると思うのです。しかし、他者同士の関係というのは、状況によっては必ずしもわかるわけではないことも多いです。

野沢 つまり、分析対象は、本人が認知してつながっていると思っている範囲になる。認知していないけれど、他者同士がつながっている可能性はもちろんある。けれども、つながっていないと思って相談したのだとする

と、本人が認知している構造自体が、何らかの効果をもたらすことがありそうに思うのです。

先ほどの育児ネットワークの効果の例に戻れば、この人はあちらの人とつながっていないから、こういう話をしても大丈夫と思い込んで愚痴を言えるとしたら、そしてそれが母親のウェルビーイングによい結果をもたらすとしたら、本人の認知ネットワークの構造はそれなりに有用な変数だということになります。

安田 実際あそこはつながっているかどうかということよりも、つながっていると思って私は行動するのか、つながっていないと思っているから行動するのかということはあると思う。

辻 ある側面ではそうですね。ただまあ、その後、他者同士がつながっていることによって、とんでもないことが起こってしまうということもあるかとは思いますが。

野沢 ところで、辻さんが言っていた、「ジオグラフィカル・ジェネレータ」という方法は、特別新しいものというわけではないのですか。

辻 最初にちょっと説明しておきますと、ジオグラフィカル・ジェネレータという言い方は、私がここ10年ほど使ってきたものに対して最近付けた名前です。欧米の人は知りません(笑)。これは、人々が地理的にどのくらい隔たったところに住んでいるか、たとえば、同居、町内、市内、県内、県外といったカテゴリに何人の人が住んでいるかを聞く方法です。私の場合は、家族・親戚、仕事関係、友人、ただの近所の人といったカテゴリと地理的なカテゴリを掛け合わせ、たとえば、町内に住む家族とか、市内に住む友人といった人たちが何人いるかをたずねました。また、場合によっては、町内に住む家族の人数をた

ずねた後に、その中でよく手伝いをしてくれる人は何人いますか、といったサブクエスチョンを設けることもできます。ただし、ただでさえ分量が多くなるので、あまりサブクエスチョンを設けすぎると、質問紙で見た感じも圧迫感があって無回答も多くなるので、注意しないとはいけません。

石黒 ポジション・ジェネレータもジオグラフィカル・ジェネレータも、ネーム・ジェネレータなどにあるようなどういふ人を挙げてもらうかという曖昧さを減らし、想起の面倒くささを減らし、さらに言うなら身近な人以外の情報がとれないという制約減らしのための技法ですよ。

野沢 そうですね。必要に応じて組み合わせていけばよいのではないかという気がします。

石黒 問題点も同じですよ。不適切に使うと何のために取ったのかよくわからなくなる。

辻 私がジオグラフィカル・ジェネレータを使い始めたのは、都市と村落のネットワークの比較をしたと思ったときに、その違いは、町内あるいは集落という身近なところにあるネットワークのボリュームに違いがあるのではないかと考えたからです。実際、そういうことが確認されました。また、中越地震の被災地研究(辻竜平・針原素子「新潟県中越地震におけるパーソナル・ネットワークと一般的信頼の変化—震災前後のパネル調査を用いて」『社会学研究』84: 69-102, 2008)では、どういったところにあるどういふ関係が増えるとか減るとかいったことを示すのにも利用しました。

今、石黒さんが挙げられた「ポジション・ジェネレータ」というのは、社会的なポジション、よく使われるのは職業ですが、いくつかの職業を挙げておいて、そういった職業の人がいるかどうかを問う方法です。そして、

SSM 調査などで明らかにされた職業威信スコアを対応づければ、各回答者の付き合いのある人の威信スコアの最大値、威信スコアの上下の幅といった指標も計算することもできます。ナン・リンがこういった方法を用いているのは、よく知られていますね。私も何回か使ったことがあります。とくに階層関係絡みの問題だとポジション・ジェネレータを使う価値はすごくあると思いました。

安田 量なのか、内容なのか、つながりの型なのか、ネットワークとか絆といて、いったい何を測っているのかというのを自覚的に考えてほしい。日本社会全体もつながり、つながりなどと言っているだけで、今自分は量を増やそうとしているのか、関係を強めようとしているのか、数を増やそうとしているのか、どういふ質のものを増やそうとしているのかはどうもよくわかっていない。ネットワーク質問とかつながり調査も、そこがいちばん問題。つながり、絆とか、ネットワーク調査・測定というだけの言葉で、うやむやにしないでほしいというのはありますね。

🔗 パーソナル・ネットワークの測定(3) ネットワーク時代のネットワーク・データ収集法

安田 ところで、今までは想起とわれわれの認知で済んでいたのですが、それこそフェイスブックだったりスパイシーだったり、どんどん関係が可視化されて、認知部分の補正をするものがいっぱい、日々周りに現れてきたわけですよ。じつはここはちゃんとつながっているということは、少なくともいろいろなところからわかるようになり始めている。そういう中でわれわれがデータを取っていると、ネーム・ジェネレータ自体も、今まで認知があやふやだった部分は多かったけれども、少しずつ精度が上がりつつあって、さらに上がる可能性はある。あの人はどういふ

人間関係やどういう人間関係地図を頭の中にもって動いているのかということが、わりと予測が付けやすくなったし、実態としても取れる部分は取れるようになってきている。そのときに、ネーム・ジェネレータだけの質問調査と推測部分に頼っていた時代のもの、これからの人間関係、パーソナル・ネットワーク・データの取り方というのは、やはりちょっと変わっていく可能性もある気がします。

野沢 本人の認知を補う方法の1つはスノーボールの併用、名前が挙がった人たちに聞きに行くという方法ですね。家族研究では、「あなたにとって誰が(重要な)家族ですか」というネーム・ジェネレータを使用したスイスの家族社会学者、エリック・ウイドマーの研究がそれです。自分が家族だと思っている人たちは、自分のことを家族だと思っていると思ったら、違う人を家族と思っていることもありえますよね。ウイドマーが1人の人物から始めて、スノーボール・サンプリングでインタビューをしたところ、その人の家族は密度の高いネットワークではなくて、かなりスカスカでゆるやかなネットワークで、境界がなく外に広がっていました。

対象が1ケースだけの分析ですから、サンプリングの代表性の問題があって、そこから何が言えるのか難しいのですが、そういう方法もある。とくに家族のような親密な小集団と思われている関係についてたずねるときには、その方法が非常に難しい。家族のメンバー全員に協力してもらえとは限らないという問題もあります。そこに対立関係があったりする場合には、描き切ったときにどのように提示し、フィードバックするかというやっかいな問題も出てくる。難しいですよね。本人だけに聞いていれば、本人の認知、本人の世界観で完結しているからOKなのですが、本人の認知を補正する情報を複眼的に追加し

ていくと、倫理的なものを含む、いくつかの問題にぶつかるところがまたジレンマです。

聞いてみたら誰が家族かすぐわかるかと思ったら、じつはそうでもない。いろいろなタイプの家族があるから難しいのです。血縁があっても、一緒に暮らしていても、それだけでは家族とまらない場合もある。

石黒 ネーム・ジェネレータの範囲だったら問題はないのですよね。

野沢 本人の認知の範囲、つまり第1次ゾーンで止まっていればわかりやすい。

石黒 どういうときにもう一歩知りたくなるのか。紐帯の先に行きたくなるのかなのですよね。私なんかはホモフィリーの研究者だったので、ネーム・ジェネレータで「この人はどういう意見をもっていると思いますか」と聞いて、回答者の意見と比べるというやり方をずっとしてきたのですが、こうなるとネーム・ジェネレータで出てきた人当人にも質問したくなります。なぜかという、回答者の推測が歪むからですね。ネーム・ジェネレータについて私が出た教訓は、本当のところはこの人が見たことしか聞けないということです。

「向こうの人はどうだと思おう」と聞いても、単純に無回答がバツと増えて20%くらいになったことがありますし、回答が出てきたとしても相当誤差がある。これに意味を見出すかどうかは人次第だとは思いますが。だから、問題は、紐帯をたどらなくてはわからないような問題を抱えているかどうかだと思います。

ところで、社会脳仮説で有名なロビン・ダンバーですが、あの人たちが最近ネットワーク研究に入ってきています。やり方はネーム・ジェネレータで、調査によって少しずつ値は違いますが、最大150人程度まで挙げさせているのです。平均値にしても50人くら

い、ネーム・ジェネレータで人を挙げさせている。

なぜこれができるかという、彼らが聞きたいのは、たとえばその人が親族かどうか。その人と親しいかどうか、直近で何日前に会ったか、連絡したか、これだけなのです。こうやって質問をものすごく区切り、かつ客観的で答えやすいようにし、しかも自分の評価でよい。そういうものに限って聞いていくと、1人当たり2時間かかるとは書いてありますが、できるというのです。私はネーム・ジェネレータの最大の制約はたぶん人数だと思っております。だけど、それだってやり方を制限すればできてしまうのです。

私も4人までしか取ったことがありません。相乗り調査で、私はこれが聞きたい、私はこれが聞きたいと詰め込んでしまった。私たちのネットワーク・バッテリーは4頁もあるものがあります。もう少し冷静になって削れば人数は増やせたかもしれない。密度はちょっと厳しいですけどもね。

野沢 150人になったら密度は厳しいね。

石黒 けれども、構造的な側面はやはり取れないのですが、認知ネットワークの諸々の側面、たとえばサポーターがどのぐらいいるかとか、そういったことはかなりできるという気があらためてしています。だから、じつはネーム・ジェネレータの本質的な問題ではなくて、よく調査屋がやってしまう詰め込みすぎの問題というのがじつは結構あったのではないかという気がします。これは面接だからできるのですね。

野沢 面接ですか。

石黒 はい。方法を読むと、「アドレス帳とかどンドン使っていていいですから、できるだけ思い出してください」といって。

野沢 何か手がかりを使わないと、いきなり思い出せというのは難しいだろうなと思

います。

石黒 150人答えた人、すごいですよね。

野沢 日本だったら年賀状を見てもいいですと言ってやればどうでしょう。

辻 私はじつは似たようなことを電通総研の人たちとやりました。最初、電話帳法を応用してやりましょうということだったのですが、郵送調査というのは紙幅の関係もあるので何人かしか聞けないでしょう。だから、ネット調査にしましょうと。ネット調査だったら、その苗字をもつ知人数だけではなくて、もっといろいろ聞けるのではないかと思ったのです。それで、岩崎という苗字の知人がたとえば2人いたら、その2人のニックネームとかイニシャルを書いてもらう。次に、平野という苗字を提示して、同じように聞くというふうに聞いていきます。そして、ここからがネット調査のうまみですが、こうして挙げてもらったニックネームについて、その人たち同士が知り合いかどうかすべてのペアについて1つずつ聞くということをやったのです。こんなこと、郵送調査じゃできませんよ。最大で137人挙げた人がいて、かわいそうでした。

石黒 答えさせたのですか。

辻 出るんだもん(汗)。

石黒 出るんだもんじゃないですよ(苦笑)。

辻 まさかそんなことは起こりえないだろうと思っていたのだけれども、起こってしまったと。本当に申し訳なかったな。知人数推定だと、優に6,500人は超えていますね。

石黒 50×50でも、地獄のようなものではないですか。

安田 それで正確なのですかね。どこまで精度が出るのか。

石黒 知っているかどうかぐらいなら。

安田 いやいや、途中で嫌になってしまう

的に電子ネットワークに人間関係が依存しているわけではなくて、外側に依存している部分がすごく多い。そして、そういう親密な関係の中で、たとえばフェイスブックでつながっているということは、何の機能を果たすのか、つながっている関係とつながっていない関係は、どう違うのかといったことをやりたがっているのです。そうすると、どれだけデータが取りやすくても、全面的に電子ネットワークのデータに依存した話し方がしにくいのです。

だから、聞いていて面白いと思うことはいっぱいある反面、それをどうやって取り込んでもいけばいいのかは、非常にわかりにくい、アイデアがないところです。

安田 全面的に移行ということではありません。今申し上げたいのは、インターネットや大規模ネットワークの研究を行うためには、デジタルのものを使ったり、工学の方と組まなければできないことがあるということです。そういったこととは親和性の低い領域やテーマの研究をなさる方たちに、変わってくれというつもりはない。ただ、そういったことをやるのだったら、組まないと損ですよ。ポーンとしていると領空侵犯されて、われわれが語るべきことが全部、経済物理学や工学系研究者に語られてしまうという危機意識があります。

松尾 研究の興味がおありのところとかを、工学系のエンジニアとかにお話いただければと思います。

辻 少なくとも、現状では、社会学系の人と情報工学系の人ほとんど対話をしていないと思うのですよ。おそらく最初のとっかかりは億劫だと思うし、実際問題として、私も最初、情報工学系の人たちに呼ばれて会合に行ったときに、いったいここはどういう世界なのだろう、この人たちは何を考えているの

だろうという感じでした。まずはお互いになりみ合うまで時間やコストがかかります。そのハードルというのか、壁というのか、そういうものがあるような気がするのです。5年くらい前と比べると、現在はもうちょっと対話ができそうな状況になりつつあるように思います。社会学者の中に、「やっぱりインターネットを利用して、あるいはインターネット上のネットワーク・データを取ることを、もうちょっとちゃんとしないとイケないんじゃない」という意識をもつ人がずいぶん増えてきた気がします。うまく乗りさえすれば、たとえば家族研究だって、これまでは家族関係のデータをネットワーク上で取ろうなんてあまり考えられてこなかったのかもしれませんが、家族関係と、情報工学系の人々がこれまで扱ったことのある関係が、どこかしら似ている側面があって、お互いの関心をすり合わせていくことによって、うまくコラボレートできたりするようなことが、もしかするとあるのではないかという気がするのです。

松尾 たぶんわれわれのやり方でいちばん違うのが、データがたくさんあることでしょうか。そうすると、まさにマイニングですが、たくさんあるところから何かわかるかもしれないというような調べ方ができます。われわれのやり方は、そこが特徴なのかという感じがしました。

野沢 マイニングの仕方ですが、こういうものを使ったらできますというのはありますか。たとえば「どの範囲を家族と思うか」に関する先ほどの話では、家族と指名された人がさらに同じ質問をされると、同じ人が指名されるとは限らないということですね。それをもっと簡単に調査する方法はあるでしょうか。ウイドマーがやったのは、1人ずつ訪ねて聞くのを繰り返していったわけですが、何か工学的なうまい方法を使ったら、紹介した

りインタビューしたりするのをネット上でやる方法、スノーボール・サンプリングをアシストするテクノロジーという感じかもしれませんが、そういうことはありうるのですか。

安田 言葉の共起関係などを使っていけば、こういう言葉をよく使うような人同士は、お互いを家族と認識し合っているとか、こういう言葉を使わない人同士は家族と認識していないとか、そういうことだったら、会話ログをテキストマイニングに放り込んでいけばできそうですよね。

松尾 面白いですね。人工知能の中でもフォーマルなロジックからいく人と、逆に世の中に何も無いところからボトムアップに形成されているという人がいて、私は後者なのです。家族についても、トップダウンに定義された家族というのはそれはそれであるかもしれないけれども、それが本来指すべき先の何らかの家族の概念があって、それはボトムアップに形成されているものだと思うのです。そこはデータを通じて見える可能性があるものではないかと思っていて、その関係性が私のいちばん興味があるところなのです。

野沢 家族社会学における家族研究にもそういう転換に近いものがあったように思います。それまでは家族というのはこういうものだという定義があって研究が始まるものだった。そうではなくて、誰が家族かということ自体が研究されるべき問題に、1980年代の終わり頃からなってきた、先ほど言ったようなネットワークとしての家族調査に似たかたちで、上野千鶴子さんなども「ファミリー・アイデンティティ」という概念を提唱して、「あなたにとっての家族は誰ですか」を問う研究を日本でもやってみせたのは、今おっしゃったようなゼロからこの人に聞いてみようという流れですね。テキストマイニングでさ

らに何かできるかもしれないし、そういうことまではまだやられていない。であれば、何か可能性を考えてもいいかもしれないですね。

辻 繰り返し同じことになりますが、本当に対話をすればできる可能性はある気がします。

❖ どのように社会科学の研究者と情報工学の研究者は関係を構築すればよいか

辻 さて、ここまで、ネットワーク時代のネットワーク収集法というようなものについて話をしてきましたが、われわれは、ネットワークを扱うことをそれぞれの領域で専門にしてきたので、会うべくして会ったという側面もありますが、それでも、社会科学系と情報工学系の研究者がコラボレートしていくのはたいへんです。しかし、インターネット上の人々の関係に関心があっても、技術がないからということ諦めてしまっている人たちもいっぱいいると思うんです。これは、社会科学側にとっては、自ら研究領域に縛りをかけてしまうことになっていると思うんです。この点について、どんなふうにして研究者間につながりを作っていけばよいか、ご提案いただきたいと思います。

安田 わかりやすさというのは大事です。知人数の推定なら辻先生だし、ミスター・スノーボールだったら池田組だったり、松尾さんにはAIを任せておけばいいとか。もう少し外から見てネットワーク研究者としてわかりやすくすることは、あってもいいと思います。

野沢 外からは、われわれの中での違いについて判別がつかないということなのでしょうね。

安田 「社会学者」みたいにしか思っていないでしょう。

松尾 やはり私は1つ学会の機能として、

そういうディレクトリ機能みたいなものは必要なのではないかと思うこともあります。

たぶん、いろいろ教えていただきたい人、一緒にやりたい人はいるとは思いますが、そこに辿り着けないと思うのです。やはりどこか大きな旗を立ててもらって、そこから入っていけるようになるといいなと思います。

野沢 社会科学の中の領域横断的な研究者ネットワークが未発達という部分もありそうですね。ネットワーク分析の学会はないし、横のつながりは少なく、それぞれテーマごとにやっているという感じになっているという構造の問題もあるのかと。

辻 そうですね。あと、社会学者はとくにそうですが、単著にこだわるところがあるでしょう。それが一緒にやることを邪魔している気もします。

安田 工学系のほうはいいですよ。ピシピシして、早く作って、さっさと終わらしましょうみたいな。手続とか、準備とか、根回しとか言いに来て、何をいつまでに完成させると言ってくればバツとできるのですが、先が見えないと、それだけで萎えてしまう。社会科学では、さっと作って終わらせるという文化があまりない。試行して哲学して頑張って、果てしなく時間をかけてローカルな穴を掘り続けるみたいところは、やはり怖いですね。アウトプットがでない共同作業は嫌です。

松尾 それはそうで、アウトプットがないと嫌だから作るのですが、アウトプットを作るときに本当によく考えているかという、やはり自分たちだけでは考えきれないのです。だから、ずっと研究されている方の意見とかを聞きたいし、そこに何か知見があるのだったら、知りたいというのはあるのです。でも、そこまでやっている時間がないので、しょうがないからバツとやってしまいます。だけど、そこを知りたいというのはすごくあります。

辻 情報系の学会に行くと、最近、ずいぶんましになったかと思うのですが、とにかく2004、2005年ぐらいの頃の話だと、社会ネットワークに関するモデルを作りましたとか言っているいろいろなモデルの発表がされるのですが、「何で俺に一言聞いてくれない。そうしたら、こんな仮定をぶち込むのなんか絶対許さなかったよ」とかいうものが、たくさんありました。だから、そういうのを消していくと、やはりお互いにとって納得のできるよいものができてくると違うかなという気はしますね。

野沢 橋渡しですね。

辻 今まで何々ジェネレータというものがいくつか出てきましたね。ただ、たとえばウェルマンがやった、フィッシャーがやった、そういうものをただ横流し的に翻訳して、それをそのままこの研究で使いましたというのでは、たぶんどうにもならないことも多いと思います。同じような関心をもった研究だったらいいかもしれないけれども、研究者も人それぞれ違う関心をもっているわけですね。だから、研究関心に合わせて、何々ジェネレータみたいなものがもっと開発されているし、どれだけ先のことかわからないですが、情報工学などの人と一緒にコラボレートして、ネットワーク・データを取得する新しい方法を開発していくというような、結局そういう流れなのかと、話を聞いていて思いました。

✿ ネットワーク研究への第一歩：ネットワークのデータ・アーカイブについて

辻 ところで、ネットワークに関心はあるけれど、すぐに自力で研究を始めることにはためらいを感じたり、まして他の研究者に声をかけるのはこわいなと思う人もいると思います。ともかく一度始めてみたいという人に対して、よい方法はありますか。

石黒 JGSSにはネットワーク・バッテリーが入っています。池田先生がJGSSのネットワーク・バッテリーの作成にかかわっていらっしやいました。まずは、それを使ってみてはどうでしょうか。また、あれがスタンダードな形になっていけば、みんなもっとやりやすくなります。あのデータそのものを使うという手もあるし、あれに合わせた形で自分の調査でも使ってみることができるようになります。公開データが出てきたというのは、今後のためにもものすごく大きな意義がありますね。

これは、パーソナル・ネットワークの場合ですが、ホール・ネットワークも、すごく興味深いデータがポコッと公開されたりすると、それを使って練習したり、それを共通の土台にして議論したりできうるのかもしれないです。公開されているものは、何かないですか。

辻 UCINETには、使い古されていますけど、ホール・ネットワークのデータがたくさん付属していますよ。また、国際社会ネットワーク学会(INSNA)のウェブサイトにも、データへのリンクがあります(<http://www.insna.org/software/data.html>)。

安田 最近のでは、エンロン・データはありますよね。

松尾 そうですね。

ところで、ワールド・ワイド・ウェブというわれわれの分野だといちばんよい国際会議があって、そこで今年話し合われていたことは、Web上のネットワーク分析を研究者たちはいろいろやっているのですが、独占所有権のあるデータを分析して、それを論文にするというのは認めない方向にしようと言っていました。パブリックなデータにするか、論文が通ったらパブリックにしないと駄目ですよと。そうでないと検証できないからです。最近では、パブリックなデータにして共有しよ

うという流れになってきている。

石黒 そうすると、データを取るところでまずハードルがあった文系の社会学者にも自分なりに何ができるかということがわかってきて、それはそれで単純にやりやすくなるし、そこからも進みますよね。このデータを取ったこの人と一緒にこういうことが新しくできるな、とか。

安田 協議はしやすくなりますよね。

石黒 やはり共有データというのは、プラットフォームとしてすごく大事ですよ。本場にJGSSにネーム・ジェネレータが入っているのはありがたいです。

安田 そこで最初の話に戻ってしまうのですが、調査をするときに、このデータはいずれ公開して共有しますなどということを書いたときに、また回答者の方々がどう思われるか、どう説得的にそこでモチベーションを上げて回答してもらおうかというのは、課題になります。

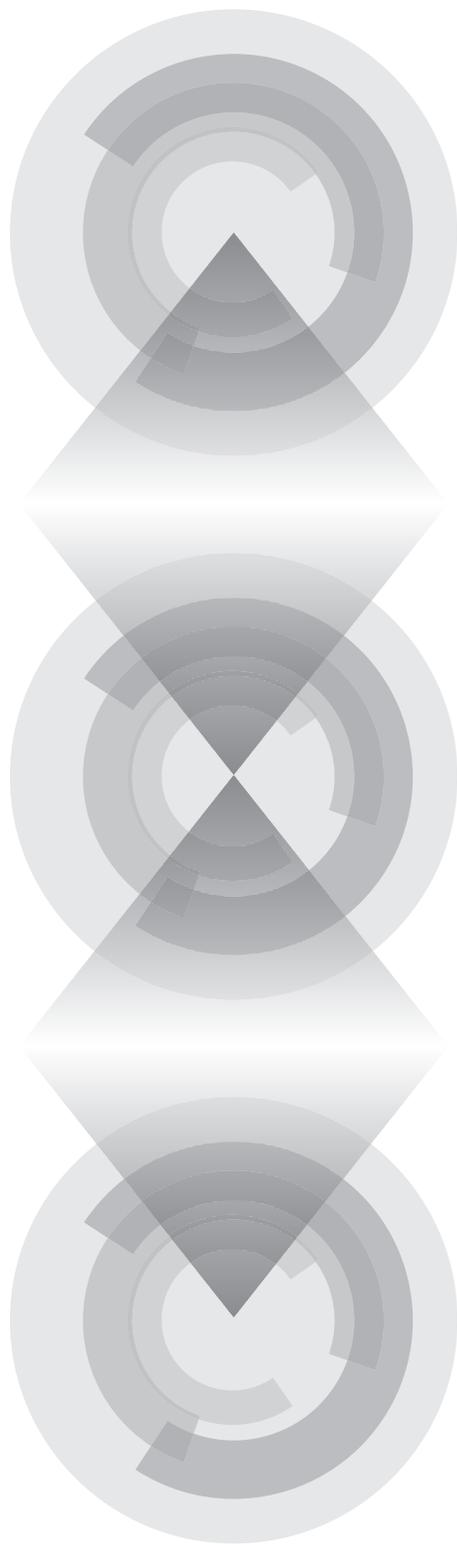
石黒 2004年に辻さんととったあのデータも、いい加減に公開しなければいけないですよ。

辻 あれは石黒さんの書いた教科書『Stataによる社会調査データの分析——入門から応用まで』(北大路書房、2008年)の中に、一部公開されていますが。

石黒 ネットワーク・データが公開されていないので、スノーボールを丸ごと公開するというのをそろそろやらないと。全国でやられたスノーボールは5本の指以下しかないので。

辻 では、そろそろ最後にしたいと思えます。みなさんには、事前にかくさんの準備をしていただいたのですが、それらについて、まんべんなく取り上げることができませんでした。この座談会で扱った題材はその一部だ

った気がしますが、それでも、何となく最後には方向性が見えてきたのかと思います。社会ネットワークの調査は、固有の難しさもありますが、既存のネットワーク・バッテリーを標準的で変更できないものと決めつけず、自分の研究課題に合わせて試行錯誤しつつ尺度の開発などをしていく面白さもあります。場合によっては、情報工学系の人々とのコラボレーションも有効だと思います。この座談会では、われわれがこれまでどのようにネットワーク調査の課題に取り組んできたのかを、バッテリーの開発の裏側にある考え方や、成功や失敗の経験なども含めて開示してきました。万能な方法というものはありませんが、新たにネットワーク調査に参入しようとする方に参考となる素材を提供することはできたのではないかと考えています。その意味で、とてもよい座談会だったと思います。これまで、社会ネットワークの調査法に関して、あまりにも何もなさすぎました。平松先生たちの『社会ネットワークのリサーチ・メソッド』にきっかけを与えていただいて、それについて私たち社会ネットワークを専門として研究している者が、社会ネットワークの調査法についてさらに議論を深められたのではないかと思います。『社会と調査』のネットワーク特集号では、この座談会をもとにしてさらに個別領域にかかわる論稿が何本か出てくることになっています。新たに社会ネットワークにかかわる調査をやってみようかという人には、これらの論稿も含めて大いに参考にしてもらえるのではないかと思います。私はこういう座談会是不慣れで、うまくまとめられなかった感じもしまして、非常に申し訳なかったのですが、全体としては私にとってたいへん勉強になる座談会になりました。みなさんのご協力に感謝したいと思います。今日はどうもありがとうございました。



地域通貨取引のネットワーク分析

中里裕美 (明治大学情報コミュニケーション学部専任講師)

平本 毅 (京大経営管理研究部付属経営研究センター研究員)

1 地域通貨

本稿では、ホールネットワーク分析を用いた社会調査の一事例として地域通貨 community currency 組織の調査を取り上げる。まずは地域通貨とは何かを簡単に説明しておく必要があるだろう。地域通貨とは、いくつかのやり方でそのはたらきに制限と変更を加えた貨幣である。まず地域通貨は、「地域」や「コミュニティ」にその流通範囲を制限している。またその機能にかんしても、地域通貨は法定通貨と比して一定の制限が加えられている。地域通貨は利子をつけないため、「資本」として自己増殖する機能をもたないし、流通期限が決まっているものの場合、価値の貯蔵機能も弱い。このような制限を加えるのは、地域通貨が特定の目的に供するためである。ここでその目的のすべてに触れることはできないが、その代表的なものを挙げれば地域通貨は、価値の貯蔵機能を弱体化させ、利子を設けず、流通範囲を限定することによって、資本の地域外流出を防ぎ、地域内の経済循環を活性化する。加えて近年の地域通貨は、社会関係資本の形成、社会福祉サービスの供給、地域への愛着の高揚、環境保全などの、非経済的な効果に照準したものが主流になってきている¹ (Nakazato and Hiramoto, 2012)。

地域通貨の取引は互いの顔の見える範囲に限定されるので、地域内でのコミュニケーションが活性化され、また市場に出回らない財やサービスが交換対象になるため、リユースが促進され、非市場的な労働力が生かされる。こうした種々の非経済的な効果を総称して、便宜的に地域通貨活動の「社会的」効果と呼ぶことにしよう。中里がスウェーデンと日本で地域通貨組織の調査を行ったのは、国内外において地域通貨への関心が高まり、とくにその社会的効果が注目されたという背景を受けてのことだった。つまり、通貨の流通量や流通速度を調べるだけではなく、取引という社会的行為が、どんな社会的効果をもたらすかということに目を向ける必要があった。

中里が本格的に調べた組織は、ストックホルムで活動する LETS (Local Exchange Trading System/Scheme)³ 組織 (Bytesring Stockholm (以下 <BYTS>)) と兵庫県旧村岡町 (現香美町) で活動していたエコマネー組織⁴ (1 むらおか) である。<BYTS> はわが国でいうところの NPO 組織の形態で活動しており、他方で <1 むらおか> は町の社会福祉協議会が活動の一環として運営する事業である。スウェーデンの首都であるストックホルムが、6519.30 km² の面積に 209 万 1,473 人 (2011 年; Statistiska centralbyrån) が住み、交通網も整備された都市圏であるのに対し、旧村岡

町は山間地域に位置し、165.66 km²の面積に6,448人（2006年度国勢調査）が暮らし、高齢化率が非常に高い。土地の大部分が山林に覆われており、周囲にはスキー場や植物園などの観光資源が豊富である。このような意味で両組織の活動が置かれる社会文化的文脈には大きな差がある。両組織ともに、当時主流だった個対個の取引方式をとっている。

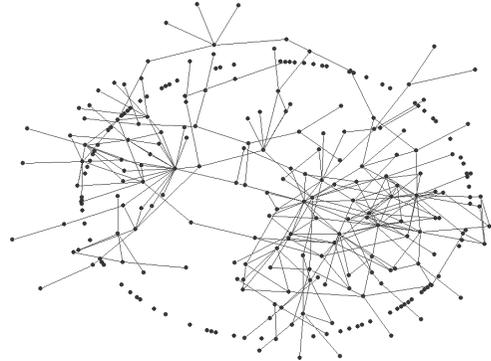
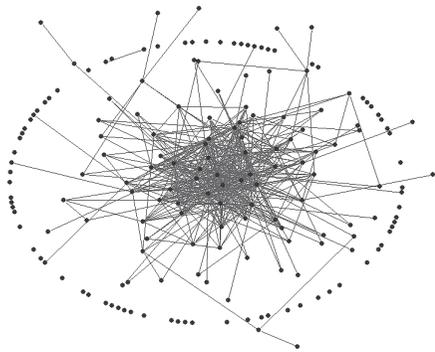
取引ネットワークを調べようと考えたのは、比較的単純な理由と、それに端を発して抱いた1つの疑問とに突き動かされてのことである。単純な理由とは、地域通貨をめぐる言説の中に、「ネットワーク（あるいはネットワークキング）」や「社会関係資本」という概念が繰り返し現れてくるというものである。こうした概念群に彩られつつも、当時はまだ地域通貨取引ネットワークの経験的な研究がほとんど存在しなかった。⁵これに端を発した1つの疑問とは、そのような「ネットワーク」や「社会関係資本」を形作るはずの人の「行為」に、従来の地域通貨研究が正しい位置づけを与えてこなかったように思うことである。取引の仕組みが法定通貨のそれと大きく異なり、そしてそれが単純な経済的交換とも社会的交換ともいえないようなものであるとすれば、地域通貨を取引する「行為」はいったい、どのような原理にしたがって組織化されるのであろうか。これがわからないと、「ネットワークキング」や「社会関係資本」の形成がどう行われるかもわからないだろうし、地域通貨活動を活性化したり、その問題点を指摘して改善することにも限界があるはずである。これらの問題意識に支えられて、私たちは新しい経済社会学（Granovetter, 1985）を背景とした地域通貨取引「行為」の社会ネットワーク分析（中里, 2007）を始めた。

2

ネットワーク分析用に作成したわけではないデータの扱いとその「客観性」

取引ネットワークのデータは比較的簡単に手に入った。なぜなら地域通貨の多くはもとも、法定通貨と異なり貨幣ごとの取引記録を残しておくので、誰と誰の間でどのような取引が行われたかが記録されているのである。〈BYTS〉は取引情報が電子的に管理されており、〈1むらおか〉は紙券の裏書に誰と誰がいつ取引したかが記載されていた。〈BYTS〉については2002年1月～05年12月、〈1むらおか〉については2002年12月～06年3月の期間の取引記録を収集した。そこから取引行列を書くうえでの苦労は、一般的に「客観的」とみなされるであろう取引記録を、ネットワーク分析用のデータに整えることにあった。とくに、データの「境界」をどこに定めるかが問題になった。両組織ともに会員（利用者）名簿が存在し、これが公的な意味での「コミュニティ」の「境界」となる。だが実際に取引記録をみみると、名簿に記載されていない人物の名が出てくることがある。交換が相対取引で行われる以上、取引の場に居合わせた友人や家族が取引に参加するなどして、実際の取引の現場ではこうしたことが生じる。この2つの「境界」のどちらを調査対象となっている地域通貨組織の活動の記録として認めるかという判断は、調査者に委ねられている。この調査では、地域住民間の相互行為を調べるという目的上、公的な組織の境界ではなく、実際に取引に参加している人々を外縁とした。

考えみると、ネットワークデータを調査設計に基づいて収集する場合、得られるデータはネットワーク分析用に整っている。他方で取引記録などの「客観的」なデータは、あく



注) Pajekの描画コマンドを用いて筆者作成。

図1 〈BYTS〉(左)と〈1 むらおか〉(右)の取引ネットワークのソシオグラム(全期間)

まで「取引を記録する」といった活動のために存在するのであって、ネットワーク分析用のデータではない。「境界」判断のもう1つの例を挙げると、地域通貨の取引行列をそのままグラフ化すると、孤立点が多くなる。図1は、〈BYTS〉〈1 むらおか〉の取引ネットワークのグラフである。

外周に環状に配置した点が、孤立点である。地域通貨取引は多くの場合、住民や会員間で満遍なく行われるのではなく、特定の者に集中して行われ、他の多くの者はあまり取引に参加しない(中里, 2007)。しかも集中して取引を行う層の人々はその層の内部で取引する傾向にあるので、ネットワーク分析的にはこれは、コア-周辺構造をとることがある(中里・平本, 2011)。周辺に位置する者の中には、まったく取引に参加しない「幽霊会員」が一定数存在する。彼らを取引行列に含めると孤立点が増えるわけである。もちろん孤立点の多さが分析を本質的に妨げるわけではないが、問題になるのは、孤立点を分析に含めるか否かである。当然のことながら孤立点を母数に含めると、密度をはじめとしたホールネットワークのさまざまなネットワーク指標が低い値を示すようになり、指標の差に与えられる

重みが軽くなる。この場合、まったく取引に参加しない人々を、地域通貨活動に参加していない層とみなして除外することが1つの解決策になるだろう。だがこの判断は、もはや「客観的」な判断ではありえない。むしろそれは、地域通貨活動というものが、参加している者にとってどのような意味をもつものかということにかんする調査と観察の「結果」である。私たちは取引記録の分析だけではなく、聞き取り調査、参与観察、質問紙調査を組み合わせる両組織の「活動」の意味と、そこで行われるメンバーの「行為」を理解することにつとめた(中里・平本, 2011)。たとえば地域通貨活動をソーシャル・サポート供給の仕組み(Nakazato and Hiramoto, 2012)と考えたとき、次のことがいえるだろう。「幽霊会員」であろうとも、会員の「できること」はカタログに掲載され、その会員が「幽霊会員」であるかどうかは他の者からみて多くの場合わからないのだから、カタログを閲覧する者にとって、その「幽霊会員」は財・サービスの提供者として実在している。いまずぐにはその者と取引が行われずとも、「幽霊会員」がカタログに載せた「できること」(あるいは、これにより増すことになる受領可能な

財・サービスの多様性)は、閲覧者にとっての精神的な心の拠り所になったり、一種の保険としてはたらく可能性がある。このような調査の「結果」を考慮しながら、孤立点を分析に含めるか否かを調査者は判断することになる。この調査では、ソーシャル・サポートの供給などの、孤立点が一定の意味をもつと考えられる事柄にかんして、孤立点を含めた分析を行った。

ネットワーク分析が不可避的に直面する問題の1つは、調査対象とするネットワークの性質によって、どのような分析が可能かが大きく左右される点にある。この点も、調査設計に基づいて収集したネットワークデータを分析する場合と、すでに存在するネットワークデータを分析する場合とで事情が異なる。次節では、具体的な分析の紹介を通して、どのような分析が取引記録により可能になったかを述べたい。

3 分析例

📍 地域通貨活動の分析

ネットワーク分析に用いたデータはまさに取引の活動の記録なので、地域通貨活動にまつわる言説が実態を表したものであるかどうかを、かなり精確に調べることができる。たとえば地域通貨は「助け合い」の仕組みであるといわれるが、これは具体的には、①ネットワークの互酬性、②ネットワークバランスといった指標で検証することができる。

まず、①互酬性の指標は、ネットワーク内での相互対(双方向の紐帯をもつ二者)の占める割合を示す。〈BYTS〉の互酬性は0.18、〈1むらおか〉の互酬性は0.09なので、〈BYTS〉の方が直接的な互酬関係への志向

が強いといえる。次に、②ネットワークバランスは特定の二者間ではなくネットワーク全体としてみたときの互酬性の指標であり、各ノードの入次数と出次数の相関係数として定義される。個対個型地域通貨は、互酬関係を特定の二者間に閉じるのではなく会員全体に拡大する仕組みである。〈BYTS〉のネットワークバランスは0.7、〈1むらおか〉は0.16であり、〈BYTS〉で間接的な互酬関係がある程度形成されている一方で、〈1むらおか〉はむしろ、財・サービスを「与える側」と「与えられる側」に会員が二極化されている様子が伺えた。このような二極化は、〈1むらおか〉が「助け合い」の道具としてというよりも、社会福祉協議会が提供する福祉サービスの一部、あるいはその補完手段として捉えられていることを示唆している。

📍 地域の分析

取引はそれが行われる行為主体間の関係性や社会ネットワーク、社会文化的文脈などに「埋め込まれて (embedded; Granovetter, 1985)」行われるので、取引行為のあり方を調べるのが、その地域がどのような地域かを明らかにすることにも繋がる(中里, 2007)。ホールネットワークに特徴的な指標の1つとして、ネットワークの連結度(connectivity)を挙げることができよう。連結度とは、ネットワークの中での到達可能なペアの割合を表す。つまり、ネットワークが1つに結合されていてすべてのペアが到達可能な状態にあるとき、「連結度」は1の値をとる。この「連結度」を比べると、〈BYTS〉が0.93に対して〈1むらおか〉が0.28と、〈1むらおか〉の取引ネットワークの「連結度」はかなり低い。簡単にいうと、〈1むらおか〉の取引ネットワークは全体としてのまとまりがなく、

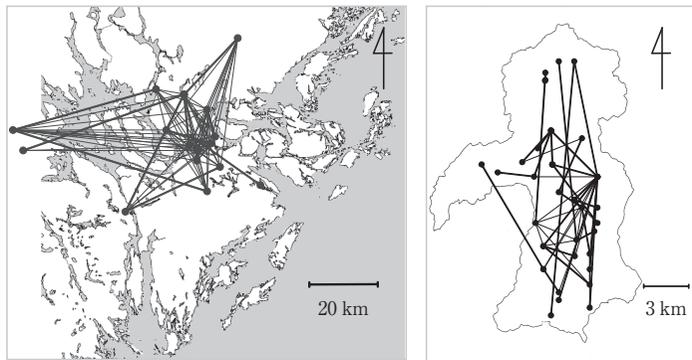


図2 〈BYTS〉(左)と〈1 むらおか〉(右)の地区間の取引ネットワークのソシオグラム(全期間)

分断化されている。その理由は、地区をノードとした地区間の取引ネットワークを描くことによって明らかになる。

この地区間取引ネットワークのソシオグラムは、どの地区に住む個人の間で取引が行われたかを調べることにより描くことができる。図2の2つのネットワークの「密度」(すなわち「地区間取引密度」)を比べればわかるように、〈1 むらおか〉の取引の多くは地区内、あるいは特定の少数の地区間で行われており、全体的な地区間の交流は盛んでない(〈BYTS〉の密度=0.43に対して〈1 むらおか〉の密度=0.05)。〈BYTS〉の年平均取引量が256.25回、〈1 むらおか〉の年平均取引量が283.43回でさほど変わらないにもかかわらず、地区間取引ネットワークの「密度」が大きく異なるということは、〈1 むらおか〉の取引が同一地区内の者同士で行われているか、あるいはネットワーク上に存在する少数の紐帯の間(特定の少数の地区に住む者の間)で行われる傾向にあることを意味する。兵庫県旧村岡町は地縁が強く残る土地であり、また交通網も整備されていない。さらに山間の高齢化地域に住む〈1 むらおか〉会員の年齢は

〈BYTS〉会員のそれより高めである。つまり、地縁の存在や移動能力の低さといった要因が、取引ネットワークの地区間での分断を生じさせている。

縦断的分析

取引記録には時間情報(取引を行った日付)が残されているので、縦断的な観点からの分析を行うことができる。縦断的なネットワーク分析において重要な点は前後のネットワークの境界をどこに設けるかということにあるが、ここで注目すべきは、メンバー自身の活動の記録である取引記録の場合、その境界がメンバー自身によって、境界づけることに志向した形で設けられている場合があるという点である。〈1 むらおか〉では当時地域通貨活動に適用されていた「前払式証券の規制等に関する法律」⁷のため、6ヵ月の流通実験を繰り返すことにより活動が行われていた。この法律は、地域通貨が第三者発行型前払式証券(発行主体以外の第三者に対して使用できる前払式証券)とみなせる場合には、発行主体の法人が、資本または出資額が1,000万円を超える等の要件を満たして事前登録を行わねば

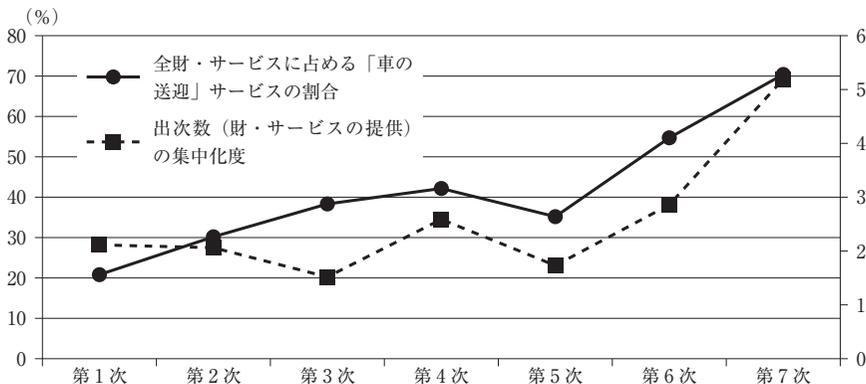


図3 〈1むらおか〉の各流通期間における全財・サービスに占める「車の送迎」サービスの割合と、各流通期間における出次数の集中化度

ならないことを定めるものである。このような活動の「区切り」は、地域通貨活動に対して与える成員の意味付けの変化と対応しうる。

図3に、〈1むらおか〉の各流通期間における全財・サービスに占める「車の送迎」サービスの割合と出次数（ここでの出次は通貨の支払いではなく、財・サービスの提供を表す）の集中化度を示す。

「地域通貨活動の分析」と「地域の分析」でみてきたように、旧村岡町では主に高齢者が、福祉サービスの一環として地域通貨を利用していた。流通実験が回数を重ね、使い方がわかってくると、地域住民は地域通貨の使いみちを固定するようになる。第7次の流通実験に至っては、じつに全体の7割が「車の送迎」サービスである。山間に位置する旧村岡町においては高齢者の移動が大きな問題になるため、地域通貨がここで使われる。このとき、車を出せる者は限定されているため、財・サービスの提供における、特定の行為者への集中が生じることになる。この、取引ネットワーク内での特定の行為者への「提供」の偏りは、出次の次数集中化度により算出することができる。図3は、全財・サービスに

占める「車の送迎」サービスの割合が流通実験を重ねて高まってくるにつれ、出次数（財・サービスの提供）の集中化度も高まる——つまり特定の者からの財・サービスの提供が多くなる——様子を表している。

4 まとめ

本稿では地域通貨組織の調査例の紹介を通じて、ネットワークデータが調査対象の活動の記録として入手できる場合のホールネットワーク調査と分析の一事例を示した。本稿は明確な結論の提出を目指す類のものではないが、以下のことは指摘できるように思う。

- (1) 組織や集団の公的な「境界」は、分析上のホールネットワークの「境界」の客観的な基準ではない可能性がある。何をデータに含めるかということにかんして、ネットワーク分析の調査対象として何が適切かを、ネットワーク分析以外の調査の「結果」として理解したうえで定めることになる場合もある。
- (2) 地域通貨の取引記録は、①地域通貨活動の実態を調べる、②活動が埋め込まれる

「地域」の特徴を調べる、③縦断的な活動の展開を調べる、といった分析を可能にする。

ネットワークデータが調査に先立って存在するホールネットワーク分析の場合、少なくとも調査を始めた当初は、データは調査者の手を離れている。そのデータを整える方法や、そのデータを用いてどんな分析ができるかを考えることを通じて、調査者はデータに接近していく。それはデータの「客観性」を離れて調査者の「主観性」に近づいていくことではおそらくなく、むしろデータを生み出した社会成員の活動や行為の意味を理解するという、調査の一環なのである。

[付記] 本稿は、平成 23~25 年度科学研究費補助金(基盤研究 C)「NPO における社会的行為の組織化とそれが地域社会に及ぼす影響に関する実証的研究」(研究代表者・中里裕美)の研究成果の一部である。

注

- ・1 本稿では詳述しないが、どの目的に供するかによって地域通貨の種類が異なる場合がある。
- ・2 本稿で取り上げるものもそうだが、地域通貨の仕組みのいくつかはメンバー間の相対取引を基本とし、また取引に市場が介入するとしても、その主体は地域の小規模商店などが多いので、取引の匿名性は低いものとなる。
- ・3 北米や西欧を中心に普及する個対個型取引の地域通貨方式であり、最も一般的かつ汎用的な流通の仕組みをもつといわれる。
- ・4 市場で流通する財・サービスを交換対象から除外すること等に特徴をもつ、日本独自の個対個型取引の地域通貨方式。
- ・5 現在はこの状況が変わりつつある。地域通貨取引のネットワーク分析の代表的な研究例として、西部(2008)、Collom(2012)を挙げておこう。
- ・6 ストックホルムには 26 の地方自治区があり、なかでも中心となるストックホルム市はさらに 14 の区に細分されている。一方で旧村岡町は 52 の行政単位に分かれている。二組織について各会員の居住地を調べ、上記の地区単位毎の分布をみたくて、最終的にストックホルムについて 20、旧村岡町について 34 の地区を点として選んだ。

- ・7 現在この法律は廃止されている。
- ・8 集中化度は、各行為者の中心性がネットワーク内部でどの程度特定行為者に偏っているのかということ比率で表す。すなわち、0(中心性が完全に分散したネットワーク)から 1(中心性が完全に集中化したネットワーク)に近づくにつれて、ネットワークの集中化傾向が強くなる。この集中化度はさまざまな中心性について求めることができるが、ここでは次数の、出次について求めている。次数集中度の算出式を以下に示す。

$$C_D = \frac{\sum_{i=1}^g [C_D(n^*) - C_D(n_i)]}{[(g-1)(g-2)]}$$

ここで、 C_D は次数の集中度、 $C_D(n^*)$ はネットワーク内に存在する最大の次数中心性をもつ行為者 n の次数中心性の値、 $C_D(n_i)$ は各行為者の次数中心性の値、 g はネットワーク内の行為者の数を表す。

文献

- Collom, E., 2012, "Key Indicators of Time Bank Participation: Using Transaction Data for Evaluation," *International Journal of Community Currency Research*, 16 (A): 18-29.
- Granovetter, M., 1985, "Economic Action and Social Structure: The Problem of Embeddedness," *American Journal of Sociology*, 91 (3): 481-580.
- 中里裕美, 2007, 「地域通貨の取引行為にみられる経済-社会の相互関係に関する一考察——社会ネットワーク論の視点から」『経済社会学会年報』29: 74-85。
- 中里裕美・平本毅, 2011, 「実践コミュニティとしての地域通貨組織への会員の参加の構造と組織のパフォーマンスの関係——スウェーデンの LETS 組織を事例として」『ノンプロフィット・レビュー』11 (1): 1-10。
- Nakazato, H. and T. Hiramoto, 2012, "An Empirical Study of the Social Effects of Community Currencies," *International Journal of Community Currency Research*, 16 (D): 124-35.
- 西部忠, 2008, 「地域通貨の流通ネットワーク分析——経済活性化とコミュニティ構築のための制度設計に向けて」『情報処理』49 (3): 290-97。

育児ネットワークに関する調査

1 育児へのネットワークアプローチ

本稿では、私が行った育児ネットワークに関する調査（松田，2008）を取り上げて、その方法とえられた知見を紹介したうえで、調査方法の課題を論じたい。

この調査は、当時社会的な問題になっていた育児をする母親の孤立や育児不安等に対して、解決策を導き出すことを目的として実施したものである。背景には、育児はもっぱら家族の中で行われるもののように思われがちだが、世帯外の人々の支えを受けながら行えば、そうした子育ての問題が起るこゝろが減るのではないかという問題意識がある。

仕事や暮らしが高度に複雑になった今日、人々が接する相手は、仕事、余暇、育児など用途に応じてある程度専門特化している。人々は仕事や余暇などの目的ごとに異なるネットワークの束をもっている。育児ネットワークとは、その1つである。

育児研究の分野では、「親」「きょうだい」などどのような人たちが育児を支援しているかを調べた既存調査はなされていた。これに対して、ネットワーク分析のアプローチを用いて育児の実態を把握したところに、本調査の特徴がある。

この研究では、①育児ネットワークの構造、

松田茂樹（株式会社第一生命経済研究所主席研究員）

②育児ネットワークを規定する要因、③育児ネットワークが親子にもたらす効果、を調査・分析している。本稿では、このうち育児ネットワークの構造と効果に関する部分を紹介する。

2 調査方法

❖ 育児ネットワークの範囲

わが国では、母親が育児を中心的に担っている。その母親に対して育児の援助を行う人々の紐帯の集合を育児ネットワークとする。これは母親からみた育児支援のパーソナル・ネットワークである。それを、①父親、②同居親族、③世帯外にいて育児の援助を行う人々（＝世帯外ネットワーク）に分ける。このうち、ネットワーク分析の方法を用いて調べたのは、世帯外ネットワークである。育児をされる子どもは、小学校就学前である。

育児をされる側の子どもの視点から、ネットワークを捉える方法もありうる。しかし、本調査の目的は母親の孤立や育児不安等の問題を解決する方策を探ることにあつたため、母親を中心としたネットワークを取り上げることにした。育児ネットワークを、①②③全体について測定する方法もありうる。しかし、育児の現状を調べると、世帯内と世帯外では、

子どもの世話をしている内容・頻度、世話をすることの意味が大きく異なる。本調査は世帯外の人々の支えに注目しようとしたこともあり、世帯内にいる父親・同居親族と、世帯外の人々の関係を分けて調べた。

❖ 調査概要

調査は、2002年笹川科学研究費補助金を受けて、2002年9～10月に東京都の府中市、国立市、立川市、多摩市、日野市の計5市において行った。対象者は、住民基本台帳から無作為抽出した、未就学児がいる世帯の母親である。対象地域は、東京都西部に位置する、都心に通勤するサラリーマンが多いベッドタウンである。調査は郵送で行い、発送数は2,000人、有効回収数は481人（有効回収率24.1%）であった。

❖ 測定方法

世帯外ネットワークの測定方法は、この特集で紹介されている、ネームジェネレータ方式である。具体的な方法は次のとおりである。

第一に、回答者に、世帯外の人のうち、何らかの育児の支援をしている人を4人まであげてもらった。これを世帯外ネットワークの中核とした。ここであげる人はインフォーマルな人たちに限定し、保育園や幼稚園等の育児の専門機関およびそこにいる保育者は除外した。

この種の調査において、ネットワークの中核を何名分測定するかは、先行研究間で一致していない。それは研究対象、研究目的、調査方法に依存している。たとえば、面接者が質問をして筆記する場合よりも、回答者自身が質問を読んで調査票に回答する形式の場合、回答者の負担や回答の正確性を考えて、測定人数が少なくなる傾向がある。本調査は郵送

法であるという事情を考慮して、回答の混乱を避けるために、上位4人までの構造を測定した。

第二に、それらの人が行っている具体的な育児支援の内容をたずねる。その内容としては、世話、遊び相手、（母親に対する）助言、相談、努力の評価である。

第三に、その人たちがどのような属性の人かをたずねた。自分との関係（「自分の親」など）、性別、年齢、未既婚、仕事の有無、居住地域（自宅からの距離）である。これが、中核の構成を測る変数である。その4人の相互の交流関係（親しい交流あり／交流あり／交流なし）もたずねた。これは中核の密度を測るために用いる。

第四に、世帯外ネットワークには、中核となる4人よりも多くの人間が関わっている。そこで、中核となる人以外に育児支援者がいれば、親族と非親族別にその人数もたずねた。育児支援は、中核としてあげられた人たちからのものが多く、かつ頻度も高い。それに比べて、中核以外にあげられた人からのものは、相対的に育児支援が少なく、頻度も低いとみられる。そこで、中核としてあげられた人の数を「強い紐帯」の規模、中核以外としてあげられた人を「弱い紐帯」の規模とした。なお、ここでいう強い紐帯と弱い紐帯という表現は、あくまでも育児ネットワークについてのものであり、グラノベッターなどの既存研究の指標とは異なる。

以上の測定方法でえられた情報から、世帯外ネットワークの規模（強い紐帯／全紐帯／全紐帯のうちの非親族）、構成（親族割合、女性割合、無職者割合、近隣者割合）、密度、平均年齢の各変数を作成した。

3 調査結果

❖ 世帯外ネットワークの構造

この測定方法によって、えられた世帯外ネットワークの構造が表1である。

育児支援をする強い紐帯は約3人、弱い紐帯を含めた全紐帯は約7人である。差し引くと弱い紐帯は約4人。全紐帯のうち非親族は約5人である。ここから、現代の母親たちは多くの人々から支援を受けながら育児をしていることがわかる。

核家族化や都市化によって、母親たちは親族や近隣から育児支援を受けられず、孤立しているという指摘がなされたこともある。だが、この結果をみると、そうした指摘はあてはまらない。

ただし、標準偏差をみると、人によってばらつきが大きいことがわかる。強い紐帯が0~1人しかいない人は8%、非親族の支援者がまったくいない人は15%いる。世帯外ネットワークが脆弱な家庭が一定程度いる。

中核の構成を表す変数は、強い紐帯が3~4人の人について作成した。規模が0~1人の場合、ネットワーク分析でいう構成の概念に合う指標を作成できない。規模が2人の場合には、親族割合は0%、50%、100%ときわめて粗い変数になる。こうした理由から、本調査では、規模が3人以上に限定して構成の変数を作成した。

親族の割合は49%。約半分が親族、残り半分が非親族である。なお、この表では「自分との関係」を省略しているが、内訳として最も多いのは、非親族では「子どもを通じた友人知人」（いわゆる子育て仲間）、親族では「自分の母親」であった。

表1 世帯外ネットワークの記述統計量

		平均	標準偏差
規模	強い紐帯	3.35	1.08
	全紐帯	7.29	3.59
	全紐帯・非親族	4.75	3.25
構成	親族割合	0.49	0.32
	女性割合	0.88	0.17
	無職者割合	0.58	0.32
	30分圏居住割合	0.48	0.35
重みなし密度		0.44	0.30
平均年齢		44.65	8.93

注) 構成と密度は、規模が3以上の者が対象。
(出所) 松田(2008)。

世帯外ネットワークは女性が88%と大半を占める。無職者の割合が58%と比較的高いのは、専業主婦同士のネットワークが多いためである。30分圏居住者が48%であることから、約半数は同一町内など近いところに住んでいる。

密度は、この中で最も指標化が難しいものである。密度の代表的な計算方法には、重みなし密度と重みつき密度がある。ネットワークのメンバー同士の関係を、親しい交流あり/交流あり/交流なしの3段階で測定している場合には、両者の計算方法は下記のようになる。

$$\text{重みなし密度} : D = (S + A) / (N \times (N - 1) / 2)$$

$$\text{重みつき密度} : D = (S + A \times 0.5) / (N \times (N - 1) / 2)$$

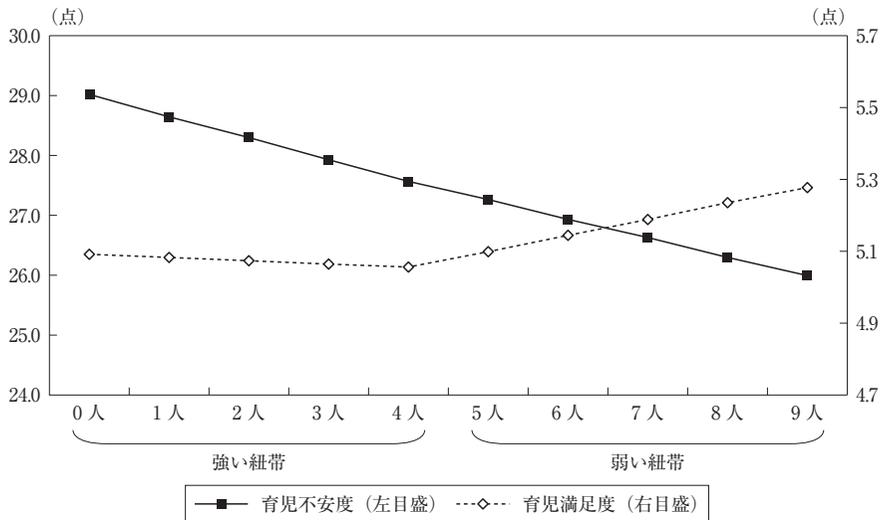
S : 構成員同士が「親しく交流している」紐帯の数

A : 構成員同士が「交流がある」紐帯の数

N : ネットワーク規模

このうち、表中には重みなし密度の結果を掲載した。重みなし密度0.44ということは、中核の人たちのうち、約半数のメンバー同士に交流があることになる。

本調査では、重みなし密度と重みつき密度の両方の変数を用いて分析し、世帯外ネット



(出所) 松田 (2008)。

図1 強い紐帯と弱い紐帯の規模別にみた母親の育児不安度と育児満足度

ワークの密度が親子に及ぼす効果は、いずれの結果もほぼ変わらないことを確認している。しかし、以下にあげる理由から、測定の妥当性という点では重みなし密度の方が高いとみられる。まず、調査の回答者が中核のメンバー同士に交流があるか否かは知りえても、それが親しい交流があるか否かまでは判断がつかない可能性がある。また、既存調査で用いられる「親しく交流している」を1点、「交流がある」を0.5点、「交流がない」を0点とすることの恣意性の問題がある。それぞれを何点に重みづけするかによって、微妙に結果は異なってしまう。

世帯外ネットワークの効果：中庸なネットワークの強さ

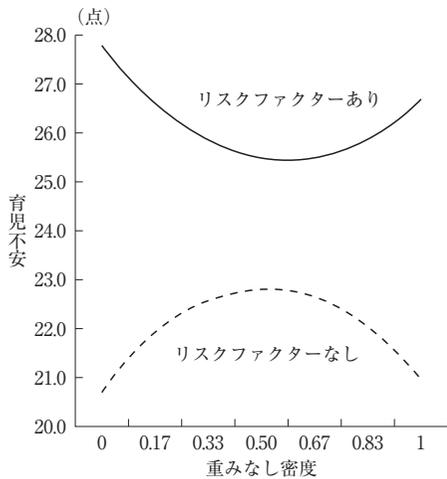
次に、世帯外ネットワークが母親のウェルビーイングに与える主な効果を紹介しよう。ウェルビーイングにあたる指標は、育児不安(牧野・中西, 1985)の10項目からなる尺度と育児満足度(7段階の尺度)を用いた。

規模と育児不安度・育児満足度の関係が図

1である。これは、被説明変数に育児不安度または育児満足度を取り、説明変数に育児ネットワークの各変数と統制変数などを用いた重回帰分析を行い、その結果をもとに規模だけの効果を推計したものである。

図からわかるように、強い紐帯の規模が増えても、弱い紐帯の規模が増えても、母親の育児不安度は同じように低下する。しかし、育児満足度については、紐帯の数が5人以上になってはじめて、紐帯の規模が増えることが満足度の向上につながる。これは、弱い紐帯が増えることがもたらす効果または強い紐帯の数が5人以上になることがもたらす効果である。

この結果から、母親の育児を支えるには、規模を増やすことが大切であることがわかる。育児不安度と育児満足度をともに改善するには、とくに弱い紐帯を広げる必要がある。頻繁に育児支援をする強い紐帯と異なり、弱い紐帯は母親にいつもと異なる情報や空気を与えるので、このような結果になったとみられ



(出所) 松田 (2008)。
 図2 リスクファクターの有無と重みなし密度別にみた母親の育児不安度

る。

方法論的な観点からみれば、世帯外ネットワークの規模を単一の指標とするのではなく、その支援の強さなどを基準に、強い紐帯と弱い紐帯というように分けて指標化することも有効である。

次に、重みなし密度と育児不安度の関係が図2である。ここでは、育児において母親のストレス源となりうるリスクファクター²という変数を用いて、それがあつた場合とない場合というように条件を分けたうえで図1と同様の分析を行っている。

分析結果をみると、リスクファクターがある場合、密度が低すぎても高すぎても育児不安度は高く、密度が中程度のときに育児不安度が低い。密度にこうしたカーブの効果があることは、筆者が同じテーマについて別に調査した2つのデータでもみられたことである。

従来の研究では、ネットワーク構造が生み出す力について、大きく分けて「ソーシャル・サポート仮説」と「社会的資源仮説」という2つの異なる仮説が提示されてきた(た

とえば Lin and Ensel, 1989)。密度については、ソーシャル・サポート仮説では密度が高いほどサポート効果が発揮されると考え、社会的資源仮説では密度が低いほどサポート効果が発揮されると考えられてきた。ところが、本調査では、いずれでもない結果がえられた。育児においては、ネットワークが中程度のときに、最もサポート効果が発揮されるのである。これを、「中庸なネットワーク (Moderate Network)」と名づけた。

日本の育児は中庸なネットワークの中で行うのがよいのである。具体的なイメージは、世帯外ネットワークの中核が4人の場合、そのうち2人が幼稚園の仲間であり、残り2人が児童館や公園で一緒に子どもを遊ばせる仲間であるというように、異なる集まりに属しているような状態である。あるいは、公園で知り合った仲間が4人いて、それらの間の関係が緩やかな状態でもよい。中核4人が顔見知りでないような関係あるいは中核4人全員が緊密であるような関係であつてはいけない。

中庸なネットワークは、特定の種類のサポート力を強く発揮するように専門特化していないが、さまざまな種類のサポート力を発揮することで、人を支える総合力が高いといえる。これが、「中庸なネットワークの強さ」である。

なお、リスクファクターがない場合、密度が逆のカーブを描く。この理由は、現在のところ理論的に十分説明できていない。

4 ネットワーク分析の可能性と課題

以上が育児ネットワーク調査の概要である。最後に、本調査からえられるネットワーク分析の可能性や課題等を述べたい。

まず、可能性である。本調査では、ネット

ワーク分析の方法を育児の調査に用いることによって、従来にない知見をえた。たとえば、世帯外ネットワークから母親たちの多くは孤立しておらず、むしろ豊かなネットワークに支えられていること。ただし、一部に孤立している層がいること。強い紐帯と弱い紐帯では、サポート効果が異なること。中庸なネットワークという新しい知見がえられたことなどである。育児以外においても、中庸なネットワークというものが効果を発揮している領域はあるだろう。ネットワーク分析を用いることで、本調査でえられた以外の発見ができるかもしれない。

次に、課題である。第一は、パーソナル・ネットワーク研究に共通することだが、育児におけるネットワークを測定する範囲についてである。本調査では、世帯外ネットワークのうち上位4人までの中核の構造を測定した。ネットワークが5人以上いたとしても、中核の構造が変わるものではないので、この方法による研究は有効である。しかし、中核だけをみては、世帯外ネットワークの全体像を見誤る可能性がある。この調査では、限界はあるものの、中核以外に広がるネットワークの規模を測定することで、この問題を補おうとした。しかし、中核以外の部分の構成や密度はわからないままである。パーソナル・ネットワークを調べる際、中核以外の部分の構造を、中核と同じように調査することは、対象者に負担をかけすぎため現実的ではない。であれば、せめて中核以外の部分の構造については、規模や構成のいくつかなど主要な指標のみを測定するようにした方がよいと考えられる。

また、測定する中核の数は4人以上が望ましい。3人では、作成する指標が粗いものになる。

第二は、密度の操作化についてである。重みつき密度を用いてなされた研究は多いが、先述したようにその重みのつけ方がはたして妥当なものかという疑問が残る。その原因の1つが、重みをつける際に使う、ネットワークのメンバー同士の関係を測定する方法が粗いことにある。既存研究では、「親しく交流している」1点、「交流がある」0.5点、「交流がない」0点などの3段階の変数から重みを指標化したりしている。親しい交流かただの交流かを分ける基準が各人で異なる可能性があり、そうなれば1点、0.5点、0点という等間隔が必ずなりたつのかどうかわからない。この点の改善策としては、複数の変数によって交流の程度を測定し、それをもとに合成尺度を作成する方法が考えられる。

第三は、ネットワーク分析の質問量の〈ダウンサイジング〉である。従来のパーソナル・ネットワークを測定する方法を用いると、ネットワークを調べるだけで、相当な質問量になる。ちなみに、筆者の調査では、当該部分のみで約4ページの質問量がある。回答にも時間がかかる。このため、ネットワーク分析の質問を厳密に行おうとすると、ことに郵送による調査では、ネットワークの構造ばかりを把握する調査になってしまう。筆者の育児ネットワークの調査では、中核となる人の属性等を1人ひとりたずねていく方法を用いたため、その部分のボリュームが多いものになっている。

だが、通常調査の目的は、ネットワーク構造を知るだけでなく、それがどのような問題を引き起こしているか、どのような要因によってネットワークが作られているか、ネットワーク以外の実態はどうかなどを明らかにすることだろう。となれば、ネットワーク構造の質問量を必要最低限に小さくすることが望

ましい。たとえば、質問（横）×人間（縦、4人まで）というマトリックス形式が用いられる場合もある。ただし、質問を簡略にしようとするほど、対象者は回答しにくいという問題もある。

現在、各研究者が独自の経験と勘でダウンサイジングした質問を作成しているが、これがベストというものはまだないよう見受けられる。この分野の研究者が協力して、質問量のコンパクトさと回答のしやすさを天秤にかけた、最も効率的な質問の〈モデル〉を作成することが、この手法の普及には必要ではないだろうか。

注

- ・1 本データを使用した分析において、0.3~0.7点の範囲で重みを変更した場合には、世帯外ネットワークの構造が親子に与える効果の分析において結果が異なる場合があった。
- ・2 子どもが病気がち、反抗期、いうことを聞かない、すぐにぐずる、双子、未熟児、発達・発育に遅れの各項目に該当するか否かをたずねたものである。これらのうち1つでも該当するものがあればリスクファクターあり、まったくなければリスクファクターなしである。なお、これがその子ども自身の発達のリスクになるという意味ではない。ここでの視点は、それが親の育児不安やストレスを高める可能性があるという点にある。

文献

- 牧野カッコ・中西雪夫、1985、「乳幼児をもつ母親の育児不安——父親の生活および意識との関連」『家庭教育研究所紀要』6: 11-24。
- 松田茂樹、2008、『何が育児を支えるのか——中庸なネットワークの強さ』勁草書房。
- Lin, Nan and Walter M. Ensel, 1989, "Life Stress and Health : Stressors and Resources," *American Sociological Review*, 54: 382-99.

健康格差とネットワークをめぐる 研究上の諸問題とその克服

——大規模社会疫学調査研究の経験を踏まえて——

中川雅貴 (国立社会保障・人口問題研究所国際関係部研究員)

近藤克則 (日本福祉大学社会福祉学部教授)

鈴木佳代 (日本福祉大学健康社会研究センター主任研究員)

1 はじめに

2012年7月に厚生労働省が発表した『健康日本21(第2次)』では、目標の中に「健康格差の縮小」が加えられるとともに、「健康を支え、守るための社会環境の整備」を促進するために「ソーシャル・キャピタルの向上」に取り組むことが明記された¹。この健康格差や、それを規定する社会的要因および社会環境としてのソーシャル・キャピタルについては、格差社会に関する関心の高まりの中にあって、社会疫学や公衆衛生学分野においても重要な研究対象となりつつある。社会疫学とは「健康状態の社会的分布と社会的決定要因を研究する疫学の一分野」(Berkman and Kawachi, 2000: 6)であるが、ここで注目を集めているテーマの1つが、社会的ネットワークおよびソーシャル・キャピタルの概念の導入である。

本稿は、社会疫学的研究において、とりわけ社会的ネットワークやソーシャル・キャピタルをめぐる実証研究が直面している調査および分析上の諸問題を整理し、その克服に向けた取り組みを紹介することを目的とする。次節では、この研究領域において、社会的ネットワークおよびソーシャル・キャピタルの概念がどのように導入されて、いかなる実証

分析上の課題に直面してきたのかについて、具体的な事例をもとに概説する。つづく第3節では、前節で整理した調査および分析上の諸課題の克服に向けた日本国内における先駆的な社会疫学的調査研究プロジェクトの取り組みについて紹介する。終節では、これまでの成果と到達点を確認したうえで、今後の可能性と課題について述べる。

2 社会疫学におけるネットワークの計測——発展と課題

家族や友人、近隣関係などを通じた人的ネットワークによって形成される社会関係が個人の健康状態に与える影響について、初めて体系的な整理を行ったのは、1988年に*Science*誌に掲載されたJ.S. ハウスらのレビュー論文である。House et al. (1988)は、一連のコホート研究および実験研究における知見を整理し、保持するネットワークの量が少なく質の低い社会的に孤立した人は死亡リスクが有意に高いことを指摘した。このレビュー論文は、従来、遺伝子や生活習慣および行動といった個人的要因によって規定されるとみなされてきた健康に関する分析に、社会関係要因についての視座を導入することの重要性および可能性を提起した。一方、その経路(メカニズム)と促進要因あるいは制約要因については、理論的にも実証的にも今後解明

されるべき課題であると結論づけられた (House et al., 1988)。

ハウスらのレビュー論文が発表された1980年代以降、死亡が増える理由あるいはプロセスの解明に向けて感染症の発症率やうつ等の精神障害の頻度を含むさまざまな健康指標と社会的ネットワークとの関連について欧米を中心に研究が蓄積されてきた²。一方、日本においては、社会的ネットワークと健康に関する旧東京都老人総合研究所による一連の研究があるものの、健康格差研究や社会疫学については研究の蓄積が始まったばかりである³。たとえば高齢者約3万3,000人を対象としたAGES (Aichi Gerontological Evaluation Study, 愛知老年学的評価研究) プロジェクト (近藤編, 2007) では、町内会や老人クラブといった組織への参加やボランティア活動に示される社会参加と、主観的健康観やうつ状態といった健康指標との関連を観察している。また、ネットワークの機能的側面に着目し、悩みごとの相談や世話の支援といった種々のソーシャル・サポートの授受と、前述の健康指標との有意な関係を指摘している。さらに、社会参加の貧しさが、転倒や認知症リスクといった要介護状態の発生リスクと関連していることも確認されている (近藤, 2012)。

日本国内での研究は、一時点の横断データ分析に依拠したものが多く、いわゆる実験が困難な社会科学が共通して直面する「逆の因果関係の可能性」や「内生性」を排除できないという問題を抱えている。すなわち、社会関係が豊かなために心身の健康状態が良好なのか、逆に、良好な健康状態が積極的な社会参加に寄与しているのかといった因果関係については、明示することができない。「逆の因果関係」を排除するためには、縦断データの収集・蓄積およびコホート研究やパネル分

析の導入が課題として指摘されてきた。

ネットワークと健康の関係をめぐる分析が直面するもう1つの課題として、個人レベルと社会・環境レベルの2つのレベルをめぐる問題が存在する。この問題は、ネットワークは個人レベルと、社会・環境レベルの2つの異なるレベルで捉えられるという概念上の問題のみならず、個別的分析において、「集団レベルで観察されることが、個人レベルでは該当しない」といういわゆる ecological fallacy (生態学的誤謬) の問題と関連している⁴。概念上の問題について言えば、社会疫学分野においては、ソーシャル・キャピタルについても、個人レベルの変数と社会・環境変数に区別すべきとされることが多い。とりわけ、前者を「構成効果」、後者を「文脈効果」として分類し、それぞれが健康アウトカムに与える独立的な影響を峻別する際に不可欠な分析手法となっているのが、マルチレベル分析である (Kawachi et al., 2007)。この高度な統計分析の方法を用いるには、各個人が属するコミュニティが特定可能な階層構造をもつ「ネスト化」されたデータの構築が必要であることが、方法論上の大きなハードルの1つとなっている。

3 課題の克服に向けた取り組み — AGES プロジェクトの経験

ここでは、前節で紹介した社会疫学分野におけるネットワークの分析をめぐる方法論的課題ならびにそれらの克服に求められるデータの収集および計測について、AGES プロジェクトがいかなる取り組みをしてきたのかについて紹介する。1999年に愛知県の2自治体で始まったAGES プロジェクトは、介護保険政策の評価を目的に、調査対象地域における保険者 (自治体) の協力を得て、要介護

認定を受けていない高齢者（65歳以上）を対象とし、生活習慣や心理的・社会的側面に加え、世帯構造や社会関係など社会疫学的な項目を含む調査票を用いて、継続的に調査を実施してきた。⁵この調査研究プロジェクトでは、高齢者ケア政策の科学的知見を得るという基本的な目的に加えて、欧米における最先端の社会疫学的研究の潮流を見据え、健康格差の実態およびその社会的決定要因を解明するために、社会的ネットワークおよびソーシャル・キャピタルが健康アウトカムにおよぼす影響を解明できる研究デザインを導入することを設計思想の1つにしてきた。それに求められる大規模データ収集のため、2010～11年度で3回目となる調査では、その対象地域を12都道府県、31市町村、577校区、調査対象人数を約11万人に拡大し、名称もJAGES（Japan Gerontological Evaluation Study：日本老年学的評価研究）とした。

社会的ネットワークの計測については、「心配や愚痴を聞いてくれる／聞いてあげる人」といった情緒的サポート、「病気の看病や世話をしてくれる／してあげる人」といった手段的サポートの「受領」と「提供」それぞれについて、具体的に「配偶者」「同居の子ども」「別居の子どもや親戚」「近隣」「友人」といった選択肢からあてはまるものすべてについて調査対象者に選択してもらうという方式で計測を試みている。また、ネットワークの構造（質）や量を計測する目的で、「あなたの友人関係についておうかがいします」として、「1）友人・知人と会う頻度はどれくらいですか」「2）この1か月間、何人の友人・知人と会いましたか。同じ人には何度会っても1人と数えることとします」「3）よく会う友人・知人はどんな関係の人ですか。あてはまる番号すべてに○をつけてくださ

い」という設問を用いている。さらに「ボランティアのグループ」「老人クラブ」「スポーツ関係のグループやクラブ」「町内会」といった各種の会やグループへの参加状況について、それぞれその有無や頻度を調査してきた。これらの設問は、社会参加を通じた個人レベルで保持されるネットワークの質や量を計測するとともに、個人レベルでの社会参加に関する指標をもってソーシャル・キャピタルを計測しようという近年の社会疫学研究における指摘を反映したものである。⁶加えて認知的なソーシャル・キャピタルについては、「あなたの地域の人々は、一般的に信用できると思いますか」「あなたの地域の人々は、多くの場合、他の人の役に立とうとしますか」などの質問項目を用い、ソーシャル・キャピタルが集団的特性というだけでなく個人的特性ともみなしようという考えのもと、それぞれの質問項目によって計測された地域信頼性指標や地域扶助規範指標を、個人レベルの変数としてだけではなく、地域単位で集計することで社会・環境変数としてマルチレベル分析などに用いてきた。

さらに、前節で取り上げた社会的ネットワークやソーシャル・キャピタルと健康の関連をめぐる「逆の因果関係」問題の克服に向けた取り組みとして、(J)AGESプロジェクトでは縦断調査法を導入し、それによって得られたパネルデータおよびコホートデータを用いた検証を蓄積してきた。たとえば、吉井ほか（2005）では、社会的ネットワークやソーシャル・サポートの受領および提供といった高齢者の社会関係の特徴と、その後2年間の要介護状態発生の関連性を明らかにした。この研究では、質問票を用いたベースライン調査から得られたデータと、協力保険者（自治体）から提供された調査対象者の「要介護認定デ

ータ」をもとに要介護状態の新規発生の有無および死亡の時期も含め追跡したデータベースを用いたコホート分析により、高齢者の社会関係と要介護状態発生の因果関係には、性別や世帯類型による違いがあることが実証された。同様に、平井ほか(2009)は、観察期間を3年間に延長したコホート研究で、ベースライン調査時点での「友人と会う頻度月1回未満」「会への自主的な参加なし」といった社会的孤立状況が、男女共通して、その後の要支援以上の要介護認定発生リスクと関連していることを示した。

因果関係の検出に迫る分析を可能にする縦断調査法の導入に加えて、JAGESデータにおけるもう1つの大きな特徴が、マルチレベル分析に対応するためのネスト化された階層的データの収集である。全国代表サンプルでは、都道府県レベルでの分析が限度となるのに対し、JAGESデータでは、全国代表サンプルではないものの、31市町村において小学校区など小地域単位でのソーシャル・キャピタル分析ができるサンプリング方法をとっている。これは、プロジェクト発足当初より、各調査地域(市町村)における小学校区などを地域・環境単位とする階層的なデータの構築を念頭に置き、学区ごとの個人標本数(通常50~100サンプル)を確保するようマルチレベル・データの継続的な入手に努力してきたことの蓄積のうえに成り立っている。ソーシャル・キャピタルの計測に際しては、上述の「地域信頼性指標」や「地域扶助規範指標」について、個人レベルでの計測結果を校区ごとに集計するという方法で地域・環境変数としてのソーシャル・キャピタル変数を構築してきた。

JAGESデータの階層性を利用したマルチレベル分析では、たとえば、Ichida et al.

(2009)が、25地区から抽出された1万5,225人のマルチレベル・データを分析し、上述の地域信頼性指標を用いて計測された地域要因としてのソーシャル・キャピタル(地区ごとの集計値)が、個人レベルの要因を考慮した場合でも、地域在住高齢者の良好な健康と関連することを検証した。また、Aida et al.(2013)は、1万4,589人を4年間追跡した縦断データを用いたマルチレベル分析によって、ソーシャル・キャピタルが弱い地域に住む女性では、強い地域に住む女性と比べ要介護状態になるリスクが68%高いことを報告している。こうした「地域信頼性指標」や「地域扶助規範指標」を用いたソーシャル・キャピタルの分析は、一方で、「地域」の範囲についての認識が人によって異なるという計測上の問題を抱えている。そのため、(J)AGESプロジェクトでは、「(地域に関する)上の問いに回答するときあなたが地域と考えたのは、次の中のどれに近いですか」という設問によって、その範囲の特定化にも取り組んでいる。

4 おわりに

以上、本稿では、健康格差の実態やそれを規定する社会的要因および社会環境を主たる研究課題の1つとして発展してきた社会疫学の領域において、社会的ネットワークやソーシャル・キャピタルの分析上の課題およびその克服に向けた取り組みについて、具体的な事例を紹介しながら概観した。社会的ネットワークやソーシャル・キャピタル概念の導入は、社会的・環境的要因に健康の決定要因を求める社会疫学研究の視座を拡大させてきたが、その計測をめぐるデータ収集および分析の手法に関しては、最先端の社会疫学研究においても、模索が続けられており、多くの課

題が指摘されている。本稿で取り上げた「逆因果関係の排除」「生態的誤謬」「個人レベルと社会・環境レベルの要因の峻別」といった分析上の主要な課題の克服を目指して、(J) AGES プロジェクトは、その発足当初より、概念に関する理論研究、パネルデータ分析やマルチレベル分析といった分析技法の発達に加えて、それが可能となる縦断データおよび階層（マルチレベル）を蓄積してきた。その成果については一部しか紹介できなかったが、パネルデータ分析やマルチレベル分析によって、社会的ネットワークおよびソーシャル・キャピタルと健康のあいだには、単に「逆の因果」や「生体的誤謬」だけ、あるいは「個人レベルのみで説明可能」な関連に留まらない関連が観察されることを明らかにしつつある。

今後、これらの関連の解明にとどまらず『健康日本 21（第 2 次）』で示された「健康格差の縮小」に寄与する研究を展望すると、測定や分析のレベルを超える課題も多く指摘できる。たとえば、これまで蓄積されてきた知見は、人々の「つながり」や「きずな」を豊かにすることによって健康格差が緩和される可能性を示唆している。しかし、それが因果関係であることが解明されたとしても、原因に介入すれば期待通りに結果を変えられる（操作可能）とは限らない。本稿で述べた計測や分析上の課題だけではなく、社会的ネットワークやソーシャル・キャピタルの操作可能性の検証や、予期せぬ副作用のモニタリング、介入に伴う費用と効果からみた効率についても検証が必要である。ユネスコが提唱するように、「知識のための科学」ととどまらず「社会のための科学」とするために、社会調査が取り組むべきことは多い。

【謝辞】 本稿で紹介した日本老年学的評価研究（the Japan Gerontological Evaluation Study: JAGES）プロジェクトにおける調査研究は、私立大学戦略的研究基盤形成支援事業（文部科学省）、ならびに、厚生労働科学研究費補助金（長寿科学総合研究事業、H22-長寿-指定-008）による助成を受けて実施されました。記して深謝します。

注

- 1 『健康日本 21（第 2 次）』の詳細については、厚生労働省による『健康日本 21（第 2 次）の推進に関する参考資料』（http://www.mhlw.go.jp/bunya/kenkou/dl/kenkounippon21_02.pdf）を参照。
- 2 さまざまな健康アウトカム変数と社会的ネットワークの関連を取り扱った分析については、Berkman and Glass (2000) で詳細にレビューされている。
- 3 たとえば、日本学術会議が 2011 年に発表した『わが国の健康の社会格差の現状理解とその改善に向けて』と題する提言では、日本における社会格差問題に関する課題の 1 つとして「健康の社会格差に関する研究の不足」が指摘され、とりわけ学際的な視点からの研究をいっそう推進する必要性が提言された（日本学術会議、2011）。
- 4 「ecological fallacy」については、市田（2007）にわかりやすく解説されている。
- 5 JAGES プロジェクトの詳細については、近藤（2007）、Kondo et al. (2010) および以下のウェブサイトを参照。<http://square.umin.ac.jp/ages/>
- 6 たとえば、Harpham et al. (2002) など。
- 7 小学校区などを「地域・環境単位」とするマルチレベル・データの収集については、協力保険者（自治体）における高齢者ケア政策および事業の立案・評価に関連する地区診断への知見提供の要請に応えるという調査実施上の目的にも合致したものであった。

文献

- Aida, J., K. Kondo, I. Kawachi, S. V. Subramanian, Y. Ichida, H. Hirai, N. Kondo, K. Osaka, A. Sheiham, G. Tsakos, and R. G. Watt, 2013, "Does Social Capital Affect the Incidence of Functional Disability in Older Japanese? A Prospective Population-based Cohort Study," *Journal of Epidemiology and Community Health*, 67: 42-47.
- Berkman, L. F. and T. Glass, 2000, "Social Integration, Social Networks, Social Support, and Health," L. F. Berkman and I. Kawachi eds., *Social Epidemiology*, New York: Oxford University Press.

- Berkman, L. F. and I. Kawachi, 2000, "A Historical Framework for Social Epidemiology," L. F. Berkman and I. Kawachi eds., 2000, *Social Epidemiology*, New York: Oxford University Press.
- Harpham, T., E. Grant and E. Thomas, 2002, "Measuring Social Capital within Health Surveys: Key Issues," *Health Policy Planning*, 17(1): 106-11.
- 平井寛・近藤克則・尾島俊之・村田千代栄, 2009, 「地域在住高齢者の要介護認定のリスク要因の検討——AGES プロジェクト3年間の追跡研究」『日本公衆衛生雑誌』56(8): 501-12。
- House, J. S., K. R. Landis and D. Umberson, 1988, "Social Relationships and Health," *Science*, 241(4865): 540-45.
- 市田行信, 2007, 「4つの錯誤 (fallacy)」近藤克則編『検証「健康格差社会」——介護予防に向けた社会疫学的大規模調査』医学書院。
- Ichida, Y., K. Kondo, H. Hirai, T. Hanibuchi, G. Yoshikawa and C. Murata, 2009, "Social Capital, Income Inequality and Self-rated Health in Chita Peninsula, Japan: A Multilevel Analysis of Older People in 25 Communities," *Social Science & Medicine*, 69(4): 489-99.
- JAGES project 日本老年学的評価研究ウェブサイト, 2012, 「日本老年学的評価研究: Japan Gerontological Evaluation Study」(2012年11月9日取得, <http://square.umin.ac.jp/ages/>)
- Kawachi, I., S. V. Subramanian and D. Kim, 2007, "Social Capital and Health: A Decade of Progress and Beyond," I. Kawachi, S. V. Subramanian and D. Kim eds., *Social Capital and Health*, New York: Springer.
- 近藤克則, 2012, 「『医療クライシス』を超えて——イギリスと日本の医療・介護のゆくえ』医学書院。
- 編, 2007, 『検証「健康格差社会」——介護予防に向けた社会疫学的大規模調査』医学書院。
- Kondo, K. ed., 2010, *Health Inequalities in Japan: An Empirical Study of Older People*, Melbourne: Trans Pacific Press.
- 日本学術会議基礎医学委員会・健康・生活科学委員会合同パブリックヘルス科学分科会, 2011, 「わが国の健康の社会格差の現状理解とその改善に向けて」(2012年11月1日取得, <http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-21-t133-7.pdf>)。
- 吉井清子・近藤克則・久世淳子・樋口京子, 2005, 「地域在住高齢者の社会関係の特徴とその後2年間の要介護状態発生との関連性」『日本公衆衛生雑誌』52(6): 456-67。



ウェブを用いた社会ネットワークの取得

飯塚修平 (東京大学大学院工学系研究科技術経営戦略学専攻修士課程)

大澤昇平 (東京大学大学院工学系研究科技術経営戦略学専攻博士課程)

松尾 豊 (東京大学大学院工学系研究科技術経営戦略学専攻准教授)

1 はじめに

近年、社会におけるさまざまな活動にウェブが使われるようになり、そのデータがウェブ上に蓄積されるようになった。また、mixiやFacebook、Twitterなどのソーシャルメディアの登場によって、個人間の関係性をウェブで観察するための新しい方法が可能になった。私たちは、計算機科学の分野、特にウェブマイニングの研究として、こういったデータの収集と分析を行っている。ウェブ上のデータから観察される人と人との関係性が、いったい現実の何を示唆しているのかというのは難しい問題ではあるものの、そのデータは、今までになく大規模かつ多様であり、研究対象として興味深い。

本稿では電子メール、検索エンジン、メッセージ、SNS、ブログそしてTwitterを対象として、ウェブから社会ネットワークのデータを収集する方法を概観する。特にTwitterはプログラミングによるデータの収集が容易であるため、サンプルコードを交えてより詳細に解説する。なお、本稿のサンプルコードはすべてPHPで記述されており、プログラムの詳細については専門誌(小川ほか, 2010)などを参照されたい。

2 ウェブ上のデータを収集する

私たちはこれまで多くのデータを分析してきたが、データの入手方法としては、ウェブサービス運営者に交渉してデータを提供してもらうか、もしくは自分たちで収集するかのいずれかである。本稿では後者に焦点をあてる。

データをプログラムで自動的に収集することはクローリングと呼ばれ、クローラと呼ばれるプログラムに自動的にウェブページにアクセスさせてデータを収集する。なかでも、ウェブサイト上から得られたHTMLファイルを解析し、必要な該当部分だけを抽出することをスクレイピングと呼ぶ。ウェブサイトによってはスクレイピングを禁止しているものがあり注意が必要だが、学術研究としては以前から広範囲に行われている。一部のウェブサービスはApplication Programming Interface (API) という仕組みを提供しており、スクレイピングを行わなくても、プログラムから自動的にデータを収集することができる。以下、計算機科学の分野で多く研究が行われているデータについて解説する。

✉ 電子メール

電子メールの分析は比較的古くから行われ

ている。たとえば、ある組織の電子メールを分析したいときには、その組織の電子メールはすべて、メールサーバを通過しているので、メールサーバの管理者に依頼して、サーバに記録されたログをもらうことになる。

一方、こうした方法をとらなくとも、被験者から許可を得てデータを収集することもできる。これは、OAuth という仕組み¹によって実現され、被験者の権限を「借りて」プログラムからメールサーバにアクセスするものである。OAuth による IMAP 接続²に対応している GMail や Yahoo! Mail などのメールサービスに蓄積されたデータであれば可能である。

電子メールに関する研究として、メールのヘッダからその人物にまつわるキーワードを抽出する研究 (Culotta et al., 2004) や、メールのやりとりを用いてコミュニティを抽出する研究 (Tyler et al., 2003) などがある。

🔍 検索エンジン

一方、私たちが多く研究を行ってきた方法の1つに、ウェブ上に出現する名前の共起関係 (同時に出現する関係) から、2人の関係性を推定しようというものがある。Google や Yahoo! などの検索エンジンに対して2人の人物名を並べた検索クエリを投げると、両方の名前が出現するウェブページの一覧が検索結果として表示される。

こうした研究の例として、研究者の氏名から共著関係のネットワークを再現した研究 (松尾ほか, 2005) がある。検索エンジンは人物名以外にも応用することが可能であり、企業名から企業間の提携関係や訴訟関係のネットワークを再現する研究 (金ほか, 2007) も行われている。自然言語処理の分野では、検索エンジン (もしくは検索エンジンに蓄えられ



図1 Facebook アプリケーションの認可画面

たデータ) を利用して概念や単語同士の関係性を抽出することは、今では基本的な処理になっている。

🌐 SNS

Mixi や Facebook などのソーシャルネットワークサービス (SNS) からユーザー間のネットワークを取得することができる。データ分析の対象としては、非常に興味深く、多くの研究者がこのデータの入手を試みているが、なかなかハードルが高い。Facebook を例にとると、現時点で可能なのは、ユーザーに許可を得て、そのデータを取得する方法である。したがって、Facebook の運営者、もしくは非常に大規模なアプリケーションの運用者でなければ、Facebook ユーザーの大規模なネットワークデータを取得することは不可能である。

データ収集の認可は OAuth による認証によって行われる。被験者に実験用の Facebook アプリケーションにアクセスしてもらい、図1のような認可画面で認可をもらう。図1の例では被験者の基本データ (被験者の友人一覧も含まれる) の他に、被験者自身のメールアドレス、被験者自身のプロフィール情報 (自己紹介文、生年月日、学歴)、そして被験者の友人のプロフィール情報を収集するための認可を求めている。

SNS のデータの研究例として、mixi を

分析した研究（松尾ほか，2005），大学生の Facebook ユーザの友達関係のネットワークおよび写真をタグづけしあうネットワークの特性の研究（Lewisa et al., 2008）や，友達関係にあるユーザ数と情報発信力の関連性の研究（Tong et al., 2008）が行われている。

📖 ブログ

ブログは，すべてのデータがオープンになっているため，収集しやすいデータの1つである。基本的に1つのブログは1人のユーザが書いていると考えることができ，ブログ間の関係性はユーザ間の何らかの関係性を表すと解釈することができる。たとえば，ウェブ上に蓄積されたブログデータから，支持政党のネットワーク（Adamic and Glance, 2005）や現実の日常生活における交流のネットワーク（Ali-Hasan and Adamic, 2007）を再現する研究が行われている。

ブログから得られる関係性としては，引用（ブログ記事中の他ブログへのリンク），お気に入り（トップページに含まれる他のブログへのリンク），コメント（他のブログへの書き込み），トラックバック（引用時の逆方向のリンク）であり，その総合的な分析も行われている（Furukawa et al., 2007）。

🐦 Twitter

ブログの新しい形態としてマイクロブログがある。投稿が非常に短いテキストである点，リアルタイム性が重視される点が特徴である。マイクロブログの代表格である Twitter は 2012 年 8 月現在アクティブユーザ 1 億 4,000 万人を超え，大きなウェブサービスである。³ 投稿が一般公開されているうえに，API が充実しているのでデータを収集しやすい。そのため，ウェブマイニングの分野では，研究対

象として頻繁に用いられている。

Twitter は主に以下の5つのデータを手がかりにしてユーザ間の関係性を把握する。

1. フォロー ユーザは他のユーザをフォローすることで，他のユーザの投稿（ツイートと呼ぶ）を自分のタイムラインでチェックすることができるようになる。

2. リプライ あるツイートへの返事としてツイートを投稿すること。本文の先頭に“@（ユーザ名）”をつけて投稿する。メールの To と同様に用いられる。

3. メンション 本文の先頭ではなく，本文中に“@（ユーザ名）”をつけて投稿すること。メールの CC と同様に用いられる。

4. リツイート 他のユーザによるツイートを再投稿すること。

5. ダイレクトメッセージ 他のユーザに対してメッセージを送ること。ツイートとは異なり，当事者のみがメッセージを閲覧することができる。

リプライおよびメンションのデータから Twitter 上での会話のネットワークを再現し，そのコミュニティ形成について調査する研究（Java et al., 2007）やツイートで情報が伝播する様子から地震や台風などのイベントを発見する研究（Sakaki et al., 2010）が行われている。

こうしたデータを共有する試みもあり，ウェブマイニングに関する研究コミュニティの流れとしては，研究論文に用いるデータはできるだけ共有しようという方向に向かいつつある。たとえば，KONECT（the Koblenz Network Collection）には，Institute for Web Science and Technologies, University of Koblenz-Landau によって収集された各種ウェブサービスのネットワークのデータが公開されている。

3 Twitter ネットワークの分析

今回取り上げたメディアの中でも特に Twitter は容易に大規模なデータを収集することができる。本節では Twitter を対象にした具体的なクロール手法およびネットワークの描画方法を解説する。

🔗 API を用いてデータを収集する

お使いのウェブブラウザを立ち上げて URL バーに `https://api.twitter.com/1/statuses/user_timeline.json?screen_name=tushuhei` と入力していただきたい。するとソースコード 1 が表示される。これは特定のユーザのツイートを収集するための API によって返された、@tushuhei (私の Twitter アカウント) の最新ツイート一覧である。これが、API を用いたデータ取得の基本である。この処理をプログラムを用いて、自動的に次々を行うことによってデータを収集していく。

```
[{
  created_at: "Wed Jul 11 18:43:01 +0000 2012",
  id: 223125258679496700,
  text: "最近は論文書いてる.",
  source: "web",
  in_reply_to_status_id: null,
  user: {
    id: 109813658,
    name: "飯塚修平",
    screen_name: "tushuhei",
    location: "Tsukuba, Japan",
    followers_count: 703,
    friends_count: 634,
  },
  geo: null,
  (一部属性を省略している)
},
{
  created_at: "Tue Jul 10 05:14:09 +0000 2012",
  id: 222559313187713020,
  ... (略) ...
}]
```

ソースコード 1 Twitter API のレスポンス例

```
<?php
function getReplies($screen_name){
    # $screen_name の最新ツイート 200 件を収集
    $tweets=json_decode(@file_get_contents
("http://api.twitter.com/1/statuses/user_time
line.json?count=200&screen_name=$screen_name"
));
    $replies=array();
    # 各ツイートの返信相手のアカウント名を配列に追加
    foreach($tweets as $tweet){
        if($tweet->in_reply_to_screen_name){
            $replies[]=$tweet->in_reply_to_screen_name;
        }
    }
    return array_unique($replies);
}
```

ソースコード 2 get Replies. php

API で収集できるデータには、ツイートの本文だけではなく返信 (in_reply_to_status_id) や位置情報 (geo) などの、ツイートに付随する情報も含まれていることがわかる。ユーザのツイートを収集する以外の API や screen_name 以外のパラメータの詳細については Twitter の REST API Documentation を参照されたい。

🔗 ネットワークの可視化

次に、前項で紹介した API を用いてツイートを収集し、それらに含まれるリプライから会話のネットワークを再現する。その後、グラフ解析ソフトウェアを用いて可視化する。

まずは、特定のユーザが投稿した最新 200 件のツイートからリプライが含まれるものをすべて抽出する。その後リプライ相手のユーザ、すなわち会話相手の一覧をつくる。ソースコード 2 はその一覧を収集するためのプログラムである。

次にクローラを実装する (ソースコード 3)。このクローラはまずジョブリスト \$jobs に含まれるジョブ対象ユーザの会話相手一覧を収集する。そして収集した会話相手をジョブに投入すると同時に、(ジョブ対象ユーザ、会話

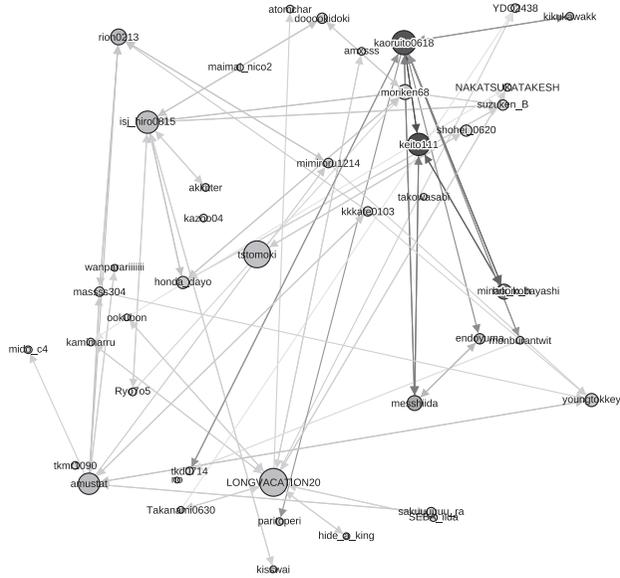


図2 Twitterの会話のネットワーク

相手のユーザ」というペアを出力する。会話相手の収集が完了したジョブ対象ユーザは、完了ジョブリスト \$done に投入される。以上を3回繰り返して、初期にジョブリストに投入されていたユーザの3ホップ先までの会

話のネットワークのデータを収集する。APIの1時間あたりの使用回数には制限があるため、今回はAPIにリクエストを送るたびに1秒の遅延処理を施している。

今回は初期にユーザ @tushuhei のみがジョブリストに入っている状態でクロウリングを開始し、クローラの標準出力結果を network.csv に書き込む。2012年7月14日時点でクロールした結果6,211ノード、7,932エッジからなる有向グラフを得ることができた。

次に、出力した csv ファイルをオープンソースのグラフ可視化ソフト gephi を用いて可視化する(図2)。可視性のために次数の和が50以上のものに絞って表示する。ノードの大きさが次数を、色の濃さが固有ベクトル中心性の高さを、各ノードのラベルはユーザ名を示す。エッジの矢印はツイートの投稿者からリプライの相手に向かっている。

```
<?php
include("getReplies.php");
$jobs=array("tushuhei");
$done=array();
# 3ホップ先まで収集
for($i=0; $i<3; $i++){
    foreach($jobs as $job){
        # ジョブリストに含まれているユーザの会話相手を取得
        foreach(getReplies($job) as $reply){
            # 会話相手がジョブリストにも完了ジョブリストにも含まれて
            # いなかったらジョブリストに追加
            if(!in_array($reply, array_merge($jobs,
            $done))){
                $jobs[]=$reply;
            }
            echo"$job, $reply\n";
        }
    }
    # 処理が完了したら、ジョブ対象ユーザをジョブリストから削除し、完了ジョブリストに追加
    $done[]=array_shift($jobs);
    # API 使用制限を超えないよう1秒遅延処理
    sleep(1);
}
}
```

ソースコード3 reply Network Crawler.php

4 おわりに

今後、ますます多くのデータがウェブ上で

利用可能になるだろう。データ収集のプログラミングはハードルが高いが、いったんできるようになると一気に入手可能なデータの量と多様性が広がる。本稿が何らかの形で、調査、研究の一助となれば幸いである。

注

- ・ 1 認証のプロセスにおいて、ユーザの認可をやりとりするための仕様。
- ・ 2 電子メールの通信規約の一種。
- ・ 3 Six-Year-Old Twitter Now Has 140M Active Users Sending 340M Tweets Per Day (<http://techcrunch.com/2012/03/21/six-year-old-twitter-now-has-140m-active-users-sending-340m-tweets-per-day>).
- ・ 4 Twitter のトップページに相当するページ。フォローしているユーザによる投稿が時系列で表示される。
- ・ 5 実際の表示には改行コードが含まれていないため、ソースコード1のように整形されたデータを見ることはできない。あらかじめ json データを整形して表示する拡張機能をブラウザにインストールすることを薦める。使用しているブラウザが Google Chrome の場合は JSONView という拡張機能がある (<https://chrome.google.com/webstore/detail/hkllaanhfefbnpoihckbnefhakgolnmc>)。

文献

- Adamic, L. A. and N. Glance, 2005, "The Political Blogosphere and the 2004 US Election: Divided They Blog," *In Proceedings of the 3rd International Workshop on Link Discovery*, 36-43.
- Ali-Hasan Noor and L. A. Adamic, 2007, "Expressing Social Relationships on the Blog through Links and Comments," *In Proceedings of the 1st Annual Meeting of the North American Chapter of the Association for Computational Linguistics*.
- Culotta, A., R. Bekkerman and A. McCallum, 2004, "Extracting Social Networks and Contact Information from Email and the Web," *In Proceedings of the 1st Conference on Email and Anti-Spam*.
- Furukawa, T., M. Ishizuka, I. Ohmukai and K. Uchiyama, 2007, "Analyzing Reading Behavior by Blog Mining," *In Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence*, 2: 1353-58.
- Java, A., X. Song, T. Finin and B. Tseng, 2007, "Why We Twitter: Understanding Microblogging Usage and Communities," *In Proceedings of the 9th WebKDD and 1st SNA-KDD2007*, 56-75.
- 金英子・松尾豊・石塚満, 2007, 「Web 上の情報を用いた企業間関係の抽出」『人工知能学会論文誌』22 (1): 48-57.
- Lewis, K., J. Kaufman, M. Gonzalez, A. Wimmerb and N. Christakis, 2008, "Tastes, Ties, and Time: A New Social Network Dataset Using Facebook.com," *Social Networks*, 30: 330-42.
- 松尾豊・友部博教・橋田浩一・中島秀之・石塚満, 2005, 「Web 上の情報からの人間関係ネットワークの抽出」『人工知能学会論文誌』20: 46-56.
- 松尾豊・安田雪, 2007, 「SNS における関係形成原理: mixi のデータ分析」『人工知能学会論文誌』22: 531-41.
- 小川雄大・柄沢聡太郎・橋口誠, 2010, 『パーフェクト PHP (PERFECT SERIES 3)』技術評論社。
- Sakaki, T., M. Okazaki and Y. Matsuo, 2010, "Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors," *In Proceedings of the 19th International Conference on World Wide Web*, 851-60.
- Tong, S. T. B., Van Der Heide, L. Langwell, J. B. Walther, 2008, "Too Much of a Good Thing? The Relationship between Number of Friends and Interpersonal Impressions on Facebook," *Journal of Computer-Mediated Communication*, 13: 531-49.
- Tyler, J., D. Wilkinson and B. Huberman, 2003, "Email as Spectroscopy: Automated Discovery of Community Structure within Organizations," *Kluwer, B. V.*, 81-96.

参照 URL

- Facebook : <http://www.facebook.com>
- Gephi : <http://gephi.org>
- Gmail : <https://mail.google.com>
- Institute for Web Science and Technologies : <http://west.uni-koblenz.de>
- KONECT : <http://konect.uni-koblenz.de>
- mixi : <http://mixi.jp>
- Twitter REST API Documentation : <https://dev.twitter.com/docs/api>
- Twitter : <http://twitter.com>
- University of Koblenz-Landau : <http://www.uni-koblenz-landau.de>
- Yahoo! Japan : <http://www.yahoo.co.jp>
- Yahoo! Mail : <https://mail.yahoo.com>