

ランダムサンプリングによる日本人の知人数推定

「連絡の取れる人」の数のインターネット調査との比較

吉田 暁生

東京大学大学院情報学環・学際情報学府

[E-mail: yoshida.akio@iii.u-tokyo.ac.jp](mailto:yoshida.akio@iii.u-tokyo.ac.jp)

要約：口コミの影響力を考えるにあたって、ランダムサンプリングによる質問紙調査を実施し、新たな知人の定義のもとに電話帳法による知人数推定をおこなったところ、先行研究のインターネット調査との異同がみられた。日本人の「連絡の取れる人」の数の平均は129.9人だったが、インターネット利用者のほうがその人数は多かった。

キーワード：口コミ（クチコミ）、電話帳法による知人数推定、連絡の取れる人の数

1. はじめに

マスメディアが高度に発達した現代においても、個人間の口コミが消費行動に与える影響は見逃せない。携帯電話や電子メールが広く利用されるようになり、電子掲示板、ブログ、SNSによる情報交換も活発になった現在、口コミの力はむしろ増大していると考えられる。

その影響力を調べるにあたって、具体的にわれわれ個人がどれだけのネットワークサイズを持っているのか、すなわち1人につき何人くらいの知人がいるのかという知見は、たいへん重要である。ところが、正確な知人数を数え上げるのは非常に困難である。質問紙調査で「あなたには知り合いが何人いますか?」と聞くとする。回答者の立場としては、何らかの手がかりから概数をイメージして答えることになる。しかし、それでは測定誤差が大きく、分析しやすいデータが得られない。したがって、質問者は、回答者の想起のしかたを統制し、知人数の「推定値」を求めるというアプローチを取る必要がある。その方法にはいくつかあ

¹ 本稿は、第24回社会ネットワーク研究会の修士論文概要報告「日本人の知人数・他者への信頼とその規定要因」の一部を改稿したものである。報告の機会を与えてくださった社会ネットワーク研究所長の安田雪特任助教授をはじめ、たくさんのコメントをくださった研究会の方々、掲載に際しあたたかいお言葉をくださった『赤門マネジメント・レビュー』編集長の新宅純二郎助教授に謝意を表したい。

るが、大規模なアンケート調査をおこなう場合に有力なものとして電話帳法がある。²

2. 方法

電話帳法では、回答者にそのまま知人を思い出してもらうのではなく、いくつかの苗字を提示して、その苗字の知人がいるかどうかをたずねる。そして、電話帳から苗字の存在比率を割り出し、実際の知人数を推定する。この方法は Pool and Kochen (1978) によって考案され、Freeman and Thompson (1989) や Killworth, Bernard, Shelley, and McCarty (1990) によって改良された。本研究では、最も精度の高い Killworth, et al. (1990) の推定式を用いる。提示された苗字について知人が何人いるというそれぞれの回答の合計を F とすると、知人数の推定値 N_Q は以下のように算出される。

$$N_Q = F \times \frac{\text{電話帳全体の苗字数}}{\text{提示する苗字が電話帳に含まれる数}} \quad \dots \text{式 1}$$

日本における追試³には、辻・松山・針原 (2002) が東京大学の心理学研究室と社会心理学研究部の有志、和光大学経済学部の有志を中心とした大学生・大学院生を対象に 99 の苗字を用いておこなった調査、辻・針原 (2003a) が東京都板橋区と新潟県栃尾市を対象に 10 の苗字を用いておこなった調査、辻・針原 (2003b) が千葉県の千葉市中央区・松戸市・柏市・銚子市・小見川町・一宮町を対象に 10 の苗字を用いておこなった調査、そして大戸 (2005) がインターネット調査会社に登録済みの全国の回答者を対象⁴に 30 の苗字を用いておこなった調査がある。本研究では、コニカミノルタテクノロジーセンター株式会社イメージング文化研究所と東京大学大学院情報学環橋元研究室でおこなった「情報化に関する調査」(以下情報化調査)のデータを用いる。

情報化調査では、「家族・親戚をのぞいて、これまでに会って会話を交わしたことがある知り合い」⁵をたずねたのち「そのなかで、連絡をとろうと思えば、とくに苦勞せずに連絡がとれる知り合い」をたずねた。この知人の定義は、「家族、親戚、友人、その他顔と名前が一致していて相手も貴方を認識でき、必要があれば連絡の取れる人」という大戸 (2005) の定義と非常に似かよっている。大戸 (2005) は『赤門マネジメント・レビュー』第4巻6号(6月号)に掲載された価値ある先行研究であるが、その異同について、社会ネットワーク研究会では軽く触れるにとどめざるをえなかった。しかし、中には大戸 (2005) の元となっ

² 大戸 (2005) のネーム・ジェネレーター方式に同じ。

³ いずれも式 1 を用いて推定値を求めている。

⁴ 大戸 (2005) と同一調査の辻 (2004) の記述による。表 1 の調査実施時期についても同様。

⁵ この定義による分析は、吉田 (2006) で詳述。

表1 先行研究との調査方法の比較

	大戸(2005)	情報化調査
実施時期	2003年9月	2005年4月13日～14日
地域	全国	首都圏30km圏内(50地点)
年齢	10代～50代	満13歳～69歳
抽出法	インターネット調査会社登録者	層化二段階無作為抽出
回収法	インターネット上	調査員による個別訪問留め置き法
回収率		1000票配布のうち639票回収で、63.9%
分析回答数	1002	545
苗字数	30	10
第1定義	家族、親戚、友人、その他	家族・親戚をのぞいて
第2定義	顔と名前が一致していて相手も書方を認識でき	これまでに会って会話を交わしたことがある(中で)
第3定義	必要があれば	連絡をとろうと思えば、とくに苦勞せず
第4定義	連絡の取れる人	連絡がとれる知り合い

た第19回社会ネットワーク研究会⁶にも参加し、この点に関心の高い会員も多かった。そこで、あらためて本稿において大戸(2005)の調査結果との比較に焦点を当てることにする。

表1に調査方法の異同を整理した。実施時期には約1年半の差があるが、その間に社会的ネットワークが大きく変化したとは考えにくい。対象地域は全国と首都圏という違いがあるが、辻・針原(2003a)によれば都市部と村落部で知人数に大きな差はない。年齢については少なくとも60代を含むかどうかという母集団の範囲に違いがある。知人数推定の分析に用いる回答数はもちろん多いに越したことはなく、苗字数もどのような苗字を用いるかにもよる⁷が多いほうが望ましい。しかし、先行研究の抽出法や回収法についてはさまざまなバイアスが想定される。

つぎに、知人数の質問文における定義を細分化して比較していく。第1定義では、当然ながら家族・親戚を含めないよりも含めたほうが知人数は多くなる。また、回答者の苗字が提示した苗字のなかに含まれていたかどうかで推定値は左右され、その家族・親戚が大人数が

⁶ 大戸紹子「日本における社会ネットワーク調査 2004 結果報告」(2005年4月26日)

⁷ 吉田(2006)で詳述。存在比率の多い苗字を選べば、少ない苗字で効率的な推定ができる。

どうかにも左右されるので、測定誤差が大きくなる。

第2定義の部分は、「これまでに会って会話を交わしたこと」がなくても「顔と名前が一致していて相手もあなたを認識でき」る場合もあるかもしれないが、基本的には情報化調査のほうが緩やかな制限であり、知人数は多く推定されると考えられる。「顔と名前が一致して」や「相手も貴方を認識でき」は、制限が厳しいだけでなく多くの判断を必要とする条件でもあり、回答時の認知的な負荷になる。負荷が高いと回答の精度が落ちるうえ、回答率自体も低下するので、できるかぎり回答しやすい質問文を作ることが望ましい。同様に、第3定義の「必要があれば」という条件でも、回答者が「必要といってもその程度にはいろいろある。ここでいうのはどれくらいの必要なのか」⁸と悩めば、認知的負荷が高まる。

さて、ここまでは、両者ともに回答者の知人観に左右されないような客観的な指標を追及している。それに対して、第3定義の「連絡をとろうと思えば、とくに苦労せず」は、回答者の主観を反映する要素を含んでいる。回答者によっては、小学校の卒業生名簿を探すくらいは朝飯前だという人もいれば、携帯電話のメモリーに登録されていても探すのを面倒に思う人もいるだろう。紐帯を現実には生かすかどうかという行動特性は、個々のネットワークサイズを規定する。本研究では、イノベーションの普及をもたらすような口コミのネットワークを抽出するため、基本的にはGranovetter (1973) のいうところの「弱い紐帯」を想定している。しかし、いくら弱い紐帯といっても一度会ったきりのような薄いつながりではなく、多くの人が自信を持って「知り合い」と言えるような常識的な基準を明確にしたい。「連絡をとろうと思えば、とくに苦労せず」はそうした意図のもとにおかれた条件である。「必要があれば」と違って強制的な要素はないので、知人数はより少なく見積もられるだろう。

最後の第4定義の部分はほぼ同じで、「人」か「知り合い」かの違いと考えてよい。知人の定義に敢えて「知り合い」という表現を含めたのには、例えば「サイン会で挨拶を交わしただけの有名人であっても、公開されている連絡先にコンタクト可能」といった特異な状況を排除する目的がある。

3. 結果

先行研究で報告されている結果と順に比較していくことにする。表2は知人数の平均値をまとめたものである。先行研究と同様のインターネット利用者の知人数の平均値は、142.3人とやや多くなった。これは、知人数を多く推定するような表1の第2定義の差異が、知人数を少なく見積もるような第1定義および第3定義の差異にもかかわらず、非常に強く影響した結果だと考えられる。調査対象者全体では129.3人、調査実施月の「全国・年齢(5歳

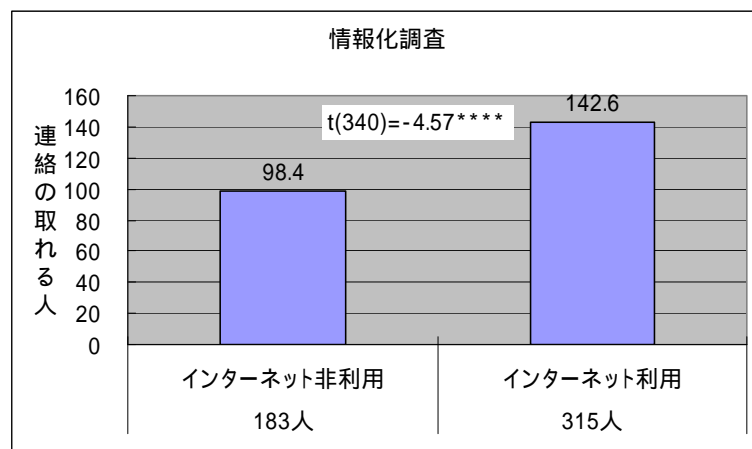
⁸ 必要とは必ず要ると書くが、現実には「必要不可欠」までの厳密な意味で用いられない場合も多い。

表2 先行研究との知人数の平均値の比較

種別	平均値	標準誤差	標準偏差	最大値
大戸(2005) インターネット利用者	129.8	4.7	148.1	2759.4
インターネット利用者	142.6	8.4	149.1	937.0
情報化調査 調査対象者全体	129.3	6.3	147.8	1183.6
当月人口比補正	129.9			

注) 補正值は、男女別5歳ごとの知人数に対象年齢中の人口比を乗じて合計したもの。男女別13~14歳の人口比は10~14歳の5分の2とした。

図1 インターネット利用による知人数の平均値差



注) 平均値差の検定結果は、知人数を対数変換しておこなったもの。
**** : $p < .0001$

階級), 男女別推計人口」(総務省統計局, 2005) によって平均値の補正をおこなうと 129.9 人であり、インターネット利用者と非利用者の平均値差は、図1にみるように高度に有意である。⁹ 一方、最大値は先行研究の3分の1程度にとどまった。これは主にサンプル数の違いによるものと考えられる。また、分析に用いた全国順位40~100位前後(存在比率0.2%前後)の苗字ひとつに対する回答知人数は、通常0~3におさまるが、先頭に提示した苗字についてだけ極端に多い回答がみられた。これと同様の回答傾向は、辻・針原(2006)でも報告されている。本研究では、提示した10の苗字のうち先頭に提示した苗字についてのみ20人以上の知人がいると回答した4サンプルを異常値とみなして、分析対象から外す手続きを

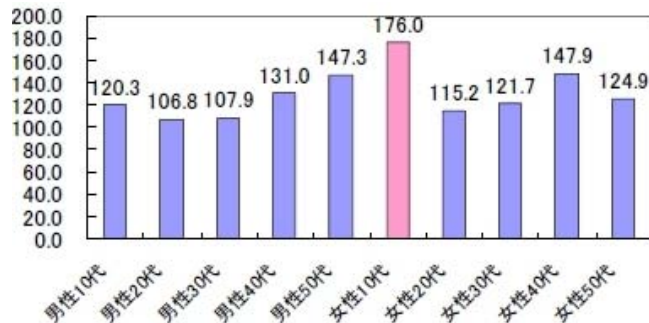
⁹ 知人数の度数分布の全体は、べき分布ではなく対数正規分布に近似できる(吉田, 2006)。統計的検定については知人数を対数変換した上でおこなっており、以下の分析でも同様である。

とった。¹⁰

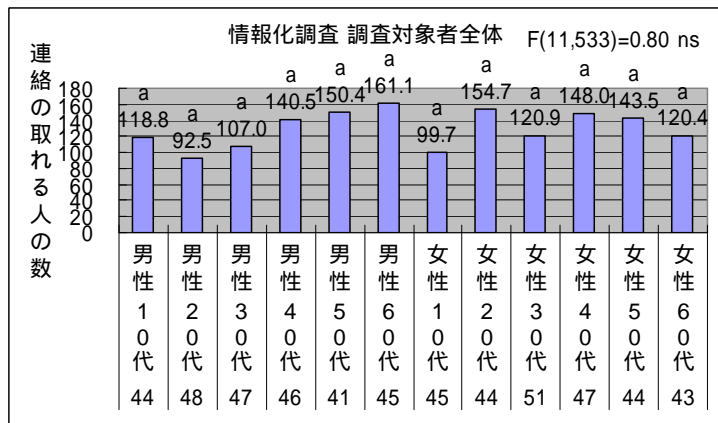
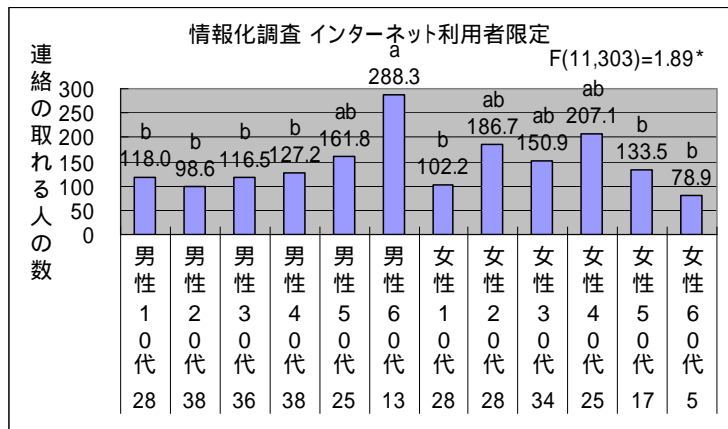
先行研究の分析方法は定かではないが、偶然に自分の苗字や親戚の苗字が提示された回答者から、極端に大きい推定値が得られた可能性がある。図2では、性・年代別の知人数の平均値を比較した。男性は、いずれも20代で減ったあと年代が高くなるにつれて徐々に増加する傾向にある。一方、女性の傾向は先行研究と異なり、10代と20代の平均値が大きく逆転している。図2のほかに二元配置分散分析もおこなったが、インターネット利用/非利用の主効果は強かった ($p < .0001$) もものの、性・年代の主効果はみられず、交互作用も傾向程度だった ($p < .10$)。女性だけに限った場合も同様である。交互作用があったとしても、女性20代が多く、女性10代が少ないことに変わりはない。

図3では、最終学歴別¹¹の知人数の平均値を比較した。先行研究では高校が最も高くなったが、高校はむしろ低い部類になった。この傾向はイ

図2 性・年代別の知人数の平均値の比較



出所) 大戸 (2005)



注) 平均値差の検定は、知人数を対数変換しておこなったもの。
* : $p < .05$ ns : no significance 数値上のアルファベットは同符号間で Tukey の多重範囲検定による有意差がないことを示す。

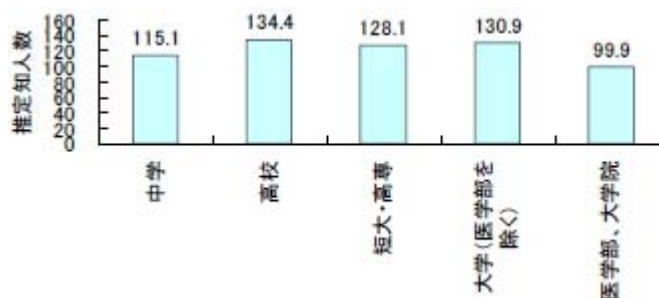
¹⁰ 吉田 (2006) で詳述。最高裁判所判事の国民審査の投票結果との比較もおこなった。

¹¹ 吉田 (2006) で詳述。

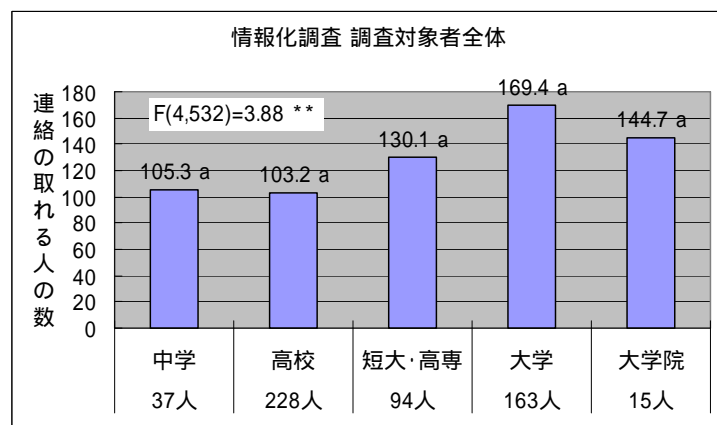
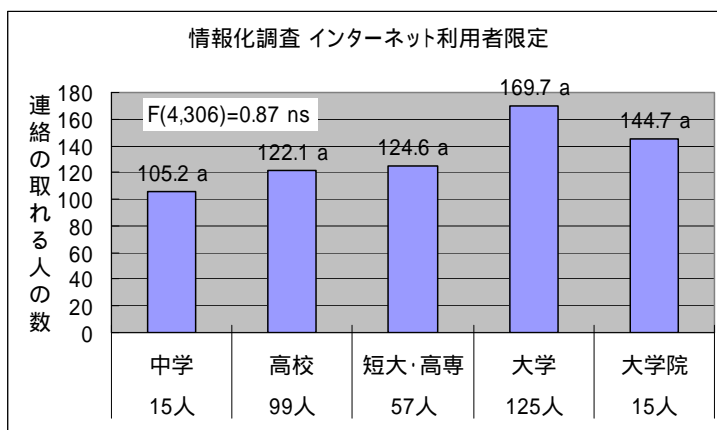
インターネット利用者に限定しても変わらなかった。一方、図4では配偶者の有無による知人数の平均値、図5では職業別の知人数の平均値を比較したが、ともに分散分析による有意差は見られなかったものの、先行研究とほぼ同様の傾向がみられた。ただし、先行研究の「かつて仕事をしていた」は、無職か専業主婦のいずれかにほぼ振り分けられている。大戸(2005)は、配偶者の有無、職業および該当者の年齢から専業主婦のネットワークサイズの小ささを予見したが、まさにそのとおりの結果が本研究によって確認されたことになる。また、図5に加えておこなった二元配置分散分析で、インターネット利用/非利用の主効果($p < .0001$)、職業の主効果(有意差なし)はともかく、傾向程度ながら交互作用が認められた($p < .10$)ことに注目したい。

さらに、大戸(2005)は最終学歴、配偶者の有無、職業の有無による結果を総合して「最終学歴が高校の仕事をしている既婚者の推定平均知人

図3 最終学歴別の知人数の平均値の比較



出所)大戸(2005)

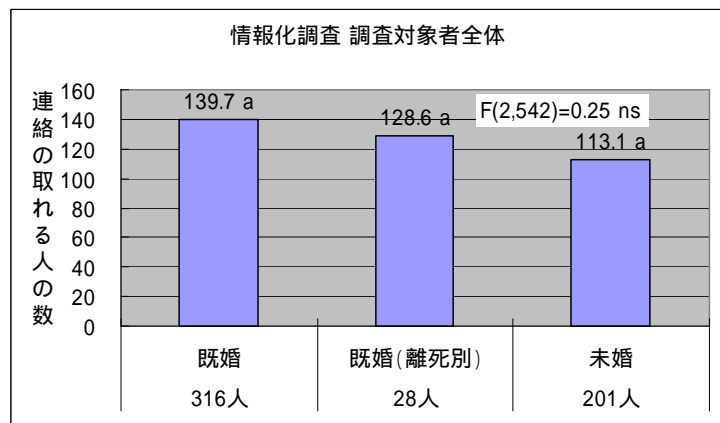
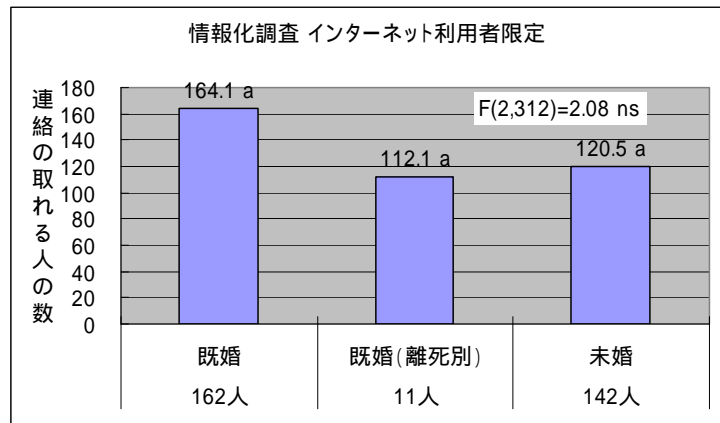
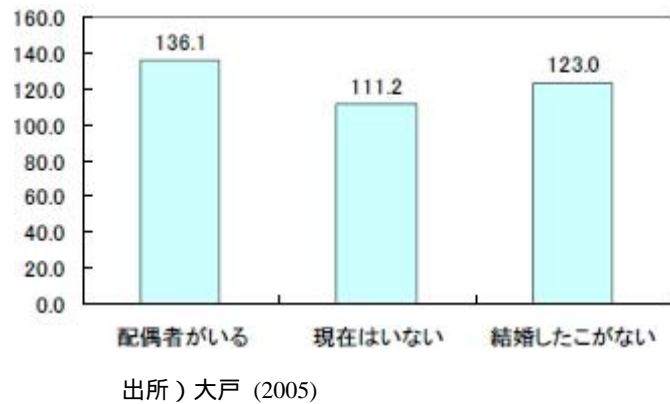


注)平均値差の検定結果は、知人数を対数変換しておこなったもの。
 **: $p < .01$ ns: no significance 数値右のアルファベットは同符号間で Tukey の多重範囲検定による有意差がないことを示す。

数が最も多い」としているが、本研究では、高校・職業あり¹²・配偶者ありのグループ(N=78)の知人数の平均値は122.0となり、調査対象者全体の平均値を下回った。同様の分類で多いグループは、¹³ 大学・職業あり・配偶者ありのグループ(N=73)の190.5人、大学・職業なし・配偶者ありのグループ(N=16)の203.4人などだった。これは図3に示した最終学歴の結果が先行研究と異なることと対応している。ここまでの主な結果は、以下の4点にまとめられる。

1. インターネット利用者のほうが非利用者よりも知人数が多い。
2. 「これまでに出て会話をしたことがある」知人数のほうが「顔と名前が一致して相手も貴方を認識でき(る)」知人数よりも多い。
3. 性・年代による分析において、先行研究では女性10代の知人数が最も多く、女性20代が少なかった

図4 配偶者の有無による知人数の平均値の比較



注) 平均値差の検定結果は、知人数を対数変換しておこなったもの。
 ns: no significance 数値右のアルファベットは同符号間で Tukey の多重範囲検定による有意差がないことを示す。

¹² フルタイムとパート・アルバイトが職業あり、学生・生徒、専業主婦、無職が職業なし。

¹³ セル数が10以上のものに限る。

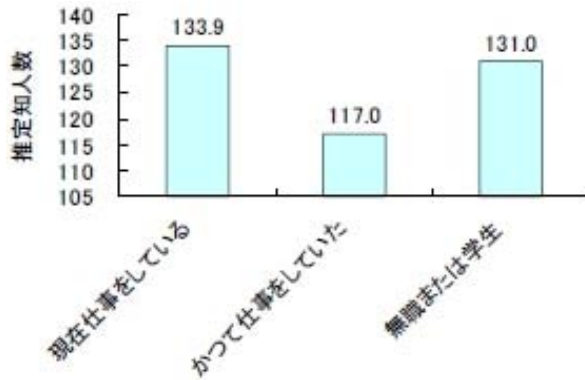
が、本研究では比較的に女性10代は少なく、女性20代は多かった。

4. 最終学歴による分析において、先行研究で最も知人数が多いのは高校だったが、本研究では大学が最も多かった。

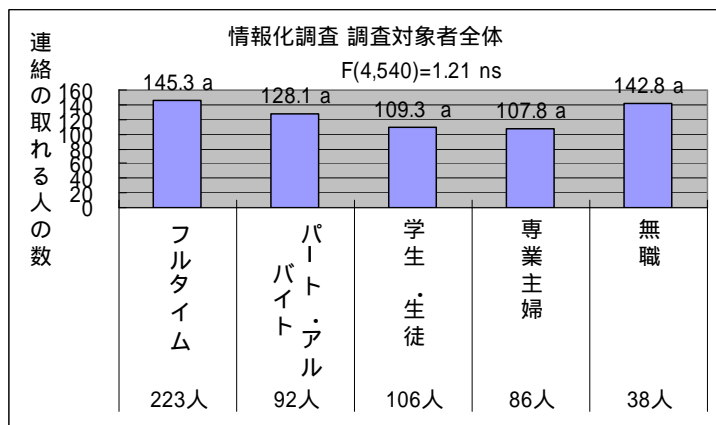
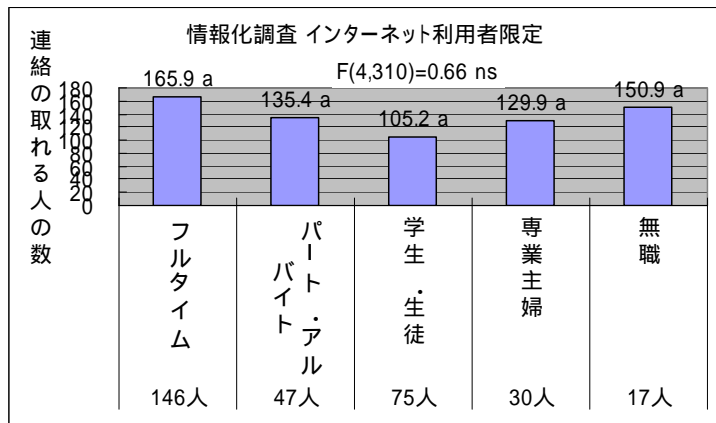
4. 考察

橋元 (2005) は、Kraut, Patterson, Lundmark, Kiesler, Mukhopadhyay, and Scherlis (1998) および Kraut, Kiesler, Boneva, Cummings, Helgeson, and Crawford (2002) の一連の研究を受けて、インターネットを利用することによって、対人関係が阻害されるという「インターネット・パラドックス」について日本での検証をおこなっている。その結果、インターネットを利用するからといって社会的ネットワークの規模が小さくなるわけではないとしている。本研究でインターネット利用者の知人数が多いことは、それと一貫した知見である。¹⁴ ところが、The World Internet Project Japan (2004) によれば、現状、

図5 職業別の知人数の平均値の比較



出所) 大戸 (2005)

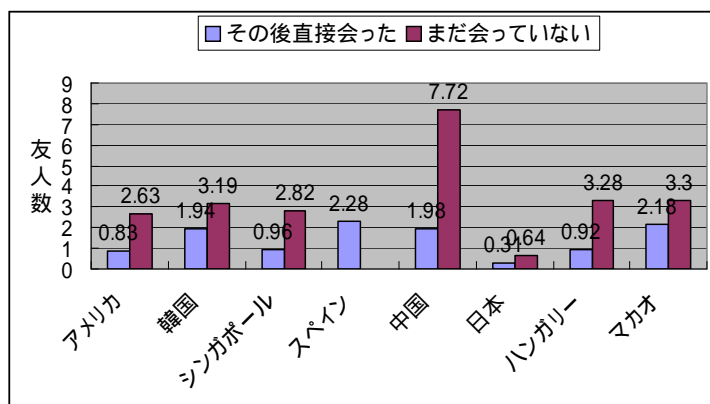


注) 平均値差の検定結果は、知人数を対数変換しておこなったもの。
 ns: no significance 数値右のアルファベットは同符号間で Tukey の多重範囲検定による有意差がないことを示す。

¹⁴ ただし、1日1時間以上のインターネット利用者の知人数は多くない (吉田, 2006)。

インターネット上で知り合いができることは非常に少ない(図6)。したがって、インターネットの利用は知人数を増大させる他の変数の擬似相関であるか、もしくは知人数が多い人ほどインターネットを利用する状況におかれやすいのだと考えられる。しかし、外出の機会が少ないと思われる男性60代や専業主婦においてインターネット利用者の知人数の多さが顕著であることも合わせて考えると、少なくともインターネットはネットワークサイズの維持に貢献していると言えそうである。反対に、セル数が少なく先行研究の対象に含まれてもいないが、女性60代のインターネット利用者で知人数が少ないこと、あるいはセル数が少ないものの先行研究と一貫した結果である、既婚者(離死別)のインターネット利用者で知人数が少ないことも注目値する。表3に示すように女性や60代はインターネットの利用率が低い、60代女性で「一人暮らしになってパソコンを娘から買い与えられた」という事例も報告されており

図6 インターネットを通じて知り合った友人の数



出所) The World Internet Project Japan (2004) から新たにグラフを作成

注) 回答は0も含めて平均値を算出。日本はパソコンによるインターネットに限定。

表3 性・年代別インターネット利用率

	インターネット		
	利用	非利用	合計
男性	205 73.2%	75 26.8%	280 100.0%
女性	155 56.6%	119 43.4%	274 100.0%

$\chi^2(1, N=554) = 16.9$ ****

	インターネット		
	利用	非利用	合計
10代	67 69.8%	29 30.2%	96 100.0%
20代	75 76.5%	23 23.5%	98 100.0%
30代	77 77.8%	22 22.2%	99 100.0%
40代	70 72.2%	27 27.8%	97 100.0%
50代	50 59.5%	34 40.5%	84 100.0%
60代	21 26.3%	59 73.8%	80 100.0%

$\chi^2(5, N=554) = 69.9$ ****

注) **** : $p < .0001$

(橋元, 金, 吉田, ファン, 2005)、パーソナルネットワークが小さくなるときの補完としてインターネットを導入する状況は十分に考えられる。

知人数の定義の違いについては、先述したように予想通りの影響がみられた。ネットワーク分析においては、目的に応じて知人を適切に定義し、可能な限り正確な知人数を求めることが必要である。

性・年代による分析や最終学歴による分析において先行研究との食い違いが見られた原因については、先行研究の調査対象がインターネット利用者に限定されていることに求めることができない。女性10代にとっては「顔と名前が一致している」や「相手も貴方を認識でき」といった条件が他の属性に比べて障壁にならなかった可能性が考えられるものの、同一調査内で定義を変えた比較ではないため、そうした交互作用は定かではない。当面、抽出法、回収法、サンプル数、推定に用いた苗字数（あるいは苗字の種類）、質問文の認知的負荷などの差異が影響したものと考えざるをえないが、最終的な結論を出すためにも、今後、知人数推定の調査を重ねていくことが望まれる。

参考文献

- Freeman, L. C., & Thompson, C. R. (1989). Estimating acquaintanceship volume. In M. Kochen (Ed.), *The small world* (pp. 147-158). Norwood, NJ: Ablex.
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360-1380.
- 橋元良明 (2005)「インターネット・パラドックスの真偽」橋元良明 編『講座社会言語科学2 メディア』(pp. 156-181). ひつじ書房.
- 橋元良明, 金相美, 吉田暁生, ファン ヨンソク (2005)「東京とソウルの家庭に見る情報生活の風景 メディアライフの今と将来」共同プロジェクト 日韓家庭訪問調査報告書, コニカミノルタテクノロジーセンター株式会社イメージング研究所 & 東京大学大学院情報学環橋元研究室.
- Killworth, P. D., Bernard, H. R., Shelley, G. A., & McCarty, C. (1990). Estimating the size of personal networks. *Social Networks*, 12, 289-312.
- Kraut R., Kiesler, S., Boneva, B., Cummings, J., Helgeson, V., & Crawford, A. (2002). Internet paradox revisited. *Journal of Social Issues*, 58(1), 49-74.
- Kraut, R., Patterson, E. M., Lundmark, V., Kiesler, S., Mukhopadhyay, T., & Scherlis, W. (1998). Internet paradox: A social technology that reduces social involvement and psychological well-being? *American Psychologist*, 53(9), 1017-1032.
- 大戸紹子 (2005)「日本における社会ネットワーク調査 2004 結果報告」『赤門マネジメント・レビュー』4(6), 313-322. <http://www.gbrc.jp/GBRC.files/journal/AMR/AMR4-6.html>

Pool, I. de S., & Kochen, M. (1978). Contacts and influence. *Social Networks*, 1, 5-51.

総務省統計局 (2005) 「人口推計 月報 平成 17 年 8 月(平成 17 年 4 月確定値,平成 17 年 8 月概算値)」
<http://www.stat.go.jp/data/jinsui/200508/zuhyou/05k2-1.xls>

辻 竜平 (2004) 「知人数に関わるネットワークの諸特性 インターネット調査より」日本社会心理学会 第 45 回大会.

辻 竜平, 針原素子 (2002) 「都市と村落の社会的ネットワークと一般的信頼」日本社会心理学会 第 43 回大会.

辻 竜平, 針原素子 (2003a) 「『小さな世界』における信頼関係と社会秩序」『理論と方法』18, 15-31.

辻 竜平, 針原素子 (2003b) 「日本人の人間関係と知人数の規定因 『小さな世界』モデルの検証」日本社会心理学会 第 44 回大会.

辻竜平, 針原素子 (2006) 「知人数の推定と補正」数理社会学会 第 41 回大会.

辻竜平, 松山久美, 針原素子 (2002) 「日本における知人・友人数の推定」数理社会学会第 33 回大会.

The World Internet Project Japan. (2004). *Internet usage trends in Japan: Survey report 2003*.

吉田暁生 (2006) 「パーソナルネットワークの社会心理学的分析」東京大学大学院学際情報学府社会情報学コース修士論文.

赤門マネジメント・レビュー編集委員会

編集長 新宅 純二郎

編集委員 阿部 誠 粕谷 誠 片平 秀貴 高橋 伸夫 藤本 隆宏

編集担当 西田 麻希

赤門マネジメント・レビュー 5巻5号 2006年5月25日発行

編集 東京大学大学院経済学研究科 ABAS/AMR 編集委員会

発行 特定非営利活動法人グローバルビジネスリサーチセンター

理事長 高橋 伸夫

東京都文京区本郷

<http://www.gbrc.jp>