PERSONAL NETWORKS AND SOCIAL INTERACTIONS IN THE GREATER TOKYO AREA: AN EXPLORATORY ANALYSIS

Giancarlos TRONCOSO PARADY¹, Kiyoshi TAKAMI² and Noboru HARATA³

¹Member of JSCE, Assistant Professor, Department of Urban Engineering, The University of Tokyo
(Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8656, Japan)
E-mail: gtroncoso@ut.t.u-tokyo.ac.jp
²Member of JSCE, Associate Professor, Department of Urban Engineering, The University of Tokyo
(Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8656, Japan)
E-mail: takami@ut.t.u-tokyo.ac.jp
³Member of JSCE, Professor, Department of Urban Engineering, The University of Tokyo
(Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8656, Japan)
E-mail: nhara@ut.t.u-tokyo.ac.jp

This article presents an exploratory analysis of personal networks and social interactions in the Greater Tokyo Area. A personal networks survey was conducted in Japan using an unrestricted name generator and interpreter. Social network characteristics in the sample were analyzed, focusing on network size, composition and spatial distribution, as well as interaction frequency patterns among network members. In addition, two models were estimated: (i) a multilevel probit model of friendship formation, used to evaluate factors associated with generation of new network ties, and (ii) a multivariate hierarchical ordinal probit model to evaluate how personal and relational characteristics were associated with social interaction contact frequency by probability mode.

Key Words: egocentric social networks, social interactions, friendship probability model, contact frequency probability model

1. INTRODUCTION

In recent years, researchers in the transportation field have started to pay attention to social networks due to the relation between social interactions and socially motivated travel. Not only do social activities account for a significant portion of travel¹, but social interaction also has been shown to be the most important purpose for leisure trips². While this has been recognized in the field, transportation planning research has largely focused on work-related travel behavior³ as well as other maintenance activities, such as shopping.

While in many cases lack of data has hampered progress, in recent years several studies have attempted to bridge this gap by explicitly measuring personal networks and using these characteristics to analyze social interaction patterns both in terms of face-to-face interaction, as well as ICT-mediated interactions⁴–⁷. Such interactions are of importance to transportation planners as they are directly associated with socially motivated travel demand.

This article presents the results of an exploratory analysis of personal networks and social interactions in the Greater Tokyo Area (Tokyo Metropolitan Area, Chiba, Saitama, and Kanagawa prefectures). A survey on personal networks was conducted using an unrestricted name generator and interpreter. Although similar studies have been conducted in the past in several cities across Japan, this is, to the best of our knowledge, the first study to use a virtually unrestricted name generator. Survey data were analyzed in terms of network size, composition, and spatial distribution and interaction frequency patterns among network members. In addition to the analysis of the data characteristics, two statistical models were estimated. First, a multilevel probit model of friendship formation was used to evaluate factors associated with generation of new network ties. Such a model is of interest for its potential use in generating population-wide networks in large-scale models, such as agent-based travel simulators incorporating social interactions⁸. Second, in order to analyze how personal and rela-
tional characteristics were associated with social interaction frequency, and potentially with socially motivated travel demand, a multivariate hierarchical ordinal probit model of social contact frequency probability was estimated. Four contact modes were modelled jointly: face-to-face contact, phone, E-mail/SMS, and social networking services (SNS).

The rest of the article is structured as follows: Section 2 discusses the existing literature from both the social sciences and the social interactions and travel perspectives. Section 3 describes the survey design and execution details. Section 4 summarizes the personal network characteristics of the sample, while Section 5 reports the results of the friendship probability and contact frequency probability models. Finally, Section 6 presents the article’s conclusion, and discusses further research directions.

2. LITERATURE REVIEW

(1) Social networks in the social sciences

Social networks research has been until recently in the domain of the social sciences. In the context of Japan, research has focused on issues such as personal network characteristics, the social isolation problem, and use patterns of new ICT technologies, and their effects on social interactions.

The so-called muen-shakai or isolated society problem, caught the Japanese public’s attention with the release of a documentary of the same title by NHK, Japan’s public broadcaster. The main argument is that the basic intermediate social groups that supported social relations in Japan, namely, family relationships, work relationships, and territorial relationships, are collapsing; but the alternative pure relationships that should emerge out of this “liberation” from traditional structures have yet to materialize.

Ishida used longitudinal data from national attitude surveys and opinion polls to show (i) the increase of unmaried persons and single households even while the desire to marry has not changed very much from the past; (ii) the fast disappearance of the traditional lifetime-employment system and the loss of a sense of belonging to one’s company; and (iii) the reduction in participation levels in local associations, and reduced contact with neighbors as evidence of the isolation phenomenon. This approach has, however, been criticized given that the data to back up these assertions were in most cases based on social attitudes rather than direct measurements of such relationships and their changes.

Research on the effect of new ICT technologies on interaction gained momentum with the popularization of internet-capable mobile phones in the late 90s amid concerns of weakening of social networks in society. A key finding was that these new modes actually became complements to face-to-face communication rather than operating at the expense of it.

(2) Measuring personal networks

Traditionally, the study of personal communities in Japan has focused on kinship. Otani was among the first researchers to focus comprehensively on personal networks in Japan, based on the work of Wellman on egocentric networks.

Otani reported from a cross-sectional random-sample survey of several Japanese regional cities average network sizes of 13.30 for Hiroshima (year 1999, population: 1.04 million), and 14.40 for Matsuyama City (year 1999, population: 0.42 million). To measure social network size, respondents were asked “How many intimate kin, co-workers, neighbors, and friends do you have (people who you get together with frequently)?”.

Some studies have also measured longitudinal changes in personal networks over time. Iwata used data from a survey of two nationwide independent cross-section random samples (2001 and 2011) asking respondents to state the “number of intimate friends you meet often with” and showed that the average number friends reduced from 11.59 to 9.42. All age cohorts showed reductions in network size, except the 10s cohort, which slightly increased from 17.29 to 18.04. Ishiguro also used data from a survey of two nationwide independent cross-section random samples (1993 and 2014) to measure social network size change in married couples living in Yamashishi city, a local city in north-eastern Japan, and Asaka, a suburb of Tokyo. He reported changes in network sizes from 16.6 to 14.5 in males, and from 14.7 to 13.3 on females. To measure social network size, respondents were asked “How many relatives who are not living with you do you feel intimate with and rely on in your daily lives (include your parents if they are not living with you)?”. Both studies seem to suggest a reduction in average network size over time.

To measure social network characteristics, these studies have largely relied on asking respondents directly to count the number of social network members. Some surveys, such as the Urban Life and Family Survey, and the 2013 Japan General Social Survey, also used name generators to gather more detailed information on the characteristics of some network alters. A name generator is a question that aims at eliciting members composing a person's social network. The name generators in these studies were limited to two and four network members, respectively. This approach allows researchers to observe lower bounds of the network, such as the presence of at least
one important social contact. For example, using data from the 2013 Japan General Social Survey, Ishida\textsuperscript{(1)} estimated the percentage of isolated persons, defined as persons who have no person to discuss important problems with, and have no emotional-support-providing relationships at 8.9%. He also found certain socio-demographics to be associated with isolation, specifically males, elders, divorcees, widowers, and residents of small villages. Consistent with these findings, Ishiguro\textsuperscript{(2)} reported a doubling of the isolation rate from 1993 to 2014, from 2.5% to 5.7% for males, and from 0.8% to 1.7% for females. For elders (age 65 or older) the isolation rate increased from 4.8% to 10.5% for males, and 1.3% to 2.6% for females.

(3) The network geography dimension

From a transportation research perspective, a key criticism of the social sciences literature is that it has largely focused on network topology and paid little attention to fine-grained geographical characteristics of networks\textsuperscript{(3),(4)}. This is particularly relevant to travel behavior, given potentially strong associations between network geography and mobility patterns. In addition, worth highlighting in the Japanese literature is the use of upper-bounded name generators. These bounds limit the possibility of mapping more accurately network geographies or observing fine-grained relational characteristics. A number of studies in Europe and the Americas have overcome similar limitations by using considerable higher upper bounds for name generators and detailed name interpreters\textsuperscript{(5),(6),(23),(24)}. In line with these international studies, this survey attempts to overcome such limitations in the context of the Greater Tokyo Area.

(4) Interaction frequency, ICT, and distance

Another object of interest in many social networks studies has been interaction frequency between network members. A particularly consistent finding is the complementarity effect between face-to-face interaction and other contact modes\textsuperscript{(5),(6),(23),(24)}. As mentioned earlier, such complementarity effects have also been reported in the context of the rapid expansion of internet-capable mobile phones in Japan in the late 90s\textsuperscript{(14),(20)}.

This complementarity effect has also been found to be mediated by face-to-face interactions\textsuperscript{(25)}, suggesting that occasional face-to-face meetings are still necessary to maintain ties and contacts\textsuperscript{(26)}. As a comprehensive review was recently conducted on social networks and travel behavior, we refer the interested reader to the work of Kim et al.\textsuperscript{(27)}. We will, however, discuss the results of this study in the context of the literature findings.

(5) Summary

This section presented a summary of the literature on social networks and social interactions with particular emphasis on the Japanese context. Key limitations identified were the use of direct counts to measure social networks, and the use of limited name generators to gather information on social network characteristics. This limits the knowledge that can be elicited about ego-centric social network, in particular related to the geographical features of the network, and fine-grained relational characteristics, of interest to transportation planners due to their relationship with human mobility, specifically with socially motivated travel.

3. SURVEY DESIGN

To overcome some of the limitations identified in the Japanese literature, an original web-based survey was conducted. This is, to the best of our knowledge, the first study to use a virtually unrestricted name generator in the context of Japan. Capturing a larger number of network members results in a higher resolution of network characteristics. Furthermore, by gathering detailed geographical information on ego and alter locations, the geographical aspects of the network, which are very relevant to travel behavior, can be better captured.

The survey was conducted between February and March of 2018, targeting adult residents of the Greater Tokyo Area (defined as the Tokyo Metropolitan Area, and the prefectures of Chiba, Saitama, and Kanagawa). It consisted of three sections. The first section focused on the respondents’ (hereinafter, ego) individual and household characteristics:

1. Household size
2. Relationship type, age, gender, and employment status of all household members
3. Education level of ego
4. Driver’s license ownership status
5. Car and bicycle ownership status
6. Residential and employment location
7. Household income

The second section focused on eliciting ego’s personal network via a set of questions called name generators, which aim at helping egos elicit the persons in his/her network that meet specific criteria. These elicited persons are called hereinafter, alters. The name generators used in this study are adapted from Kowald and Axhausen\textsuperscript{(4)}:

1. Name the persons whom you spend your free time with (Examples of free time activities include sports, cultural events, club activities, eating and drinking out, and outings during holidays, etc.)
2. In addition to the persons named in the previous questions, please name the persons with whom you discuss important private problems unrelated to work.

The system allowed naming 25 alters per name generator, totaling a maximum of 50 alters. The name interpreter then elicited the following information from all alters:

1. Relationship type
2. Relationship length
3. Gender
4. Age
5. Employment status
6. Marital status
7. Contact frequency by mode (face-to-face, phone, e-mail and short message services (SMS), social networking services (SNS))
8. Residential and employment location

Based on the elicited names, respondents were then asked to create groups of alters that know each other and usually spend their free time together and name each group (hereinafter cliques). A maximum of eight cliques could be formed. Alters could belong to more than one clique. For each clique, the following information was elicited:

1. Place when members meet often (maximum of three places per clique), if any
2. Type of activities usually conducted at that place

The third section focused on questions regarding community trust, but are out of the scope of this article, hence are not discussed. The survey system was developed in-house, and respondents were recruited from an opt-in consumer panel maintained by a survey research firm in Japan accredited by the Japan Marketing Research Association (JMRA). Two pre-tests were conducted. The first was conducted using a sample of students and staff of the Urban Engineering Department at the University of Tokyo. The main objective was to check for errors in the survey system operation. The second test (target n=100) was done using as a sampling frame the same consumer panel as the main survey and aimed at evaluating quality of responses and completion rates, particularly in terms of quality of location data (i.e., residential location and work location) and to assess the effect of the burden of the name generator and interpreter on response quality. Since the survey system was developed in-house, we could also observe data of respondents that did not complete the survey. A valid sample size of n=97/100 was gathered out of 132 persons who started the survey yielding an overall completion rate of 75.7%.

To evaluate whether or not respondents with larger networks were less likely to complete the survey, we conducted group comparisons via group mean t-test and Wilcoxon non-parametric test. Both tests showed no statistically significant differences in social network sizes between the respondents who completed the survey and those who dropped out after completing the name generator (t-value 0.99, df=27.78, p-value 0.33; Wilcoxon test 1033, p-value 0.58). This suggests that respondents who reported large networks were as likely to complete the survey as those who reported smaller networks.

The main survey was conducted in March 2018. The sampling method was quota sampling to match the gender and age distribution of the Greater Tokyo Area. Participation rate, that is, the rate of respondents who started the survey, was 9.7%. Completion rate among participants was 65.2%. Effective sample size was 347 egos.

4. SAMPLE PERSONAL NETWORK CHARACTERISTICS

(1) Ego personal characteristics
Ego-level characteristics are summarized in Table 1 while Fig. 1 shows the network size histogram. Average network size was 4.60 (median: 3) alters (including zero-sized networks) with a maximum of 28 alters. Women’s networks were on average larger than men’s, while egos over 60 years old exhibited considerably larger networks than other age groups.

The number of isolated egos, that is, respondents who explicitly reported not having any members in their networks was 7.2%. Similar to findings in the literature, men were overrepresented in the isolates sample making up 68% of the isolates against 47% in the non-isolate subsample. Worth highlighting is the fact that the average network size in the sample differs considerably from the values reported in the literature, although given the network size measurement methodological differences, comparisons are difficult. When compared to the average network sizes reported the networks of the Japanese sample are considerably smaller.

![Network size histogram](image-url)
Table 1 Descriptive statistics of ego-level characteristics.

| Variable                      | Share  |
|-------------------------------|--------|
| Ego is male                   | 49.0%  |
| Age:                          |        |
| 20–29                         | 11.8%  |
| 30–39                         | 17.6%  |
| 40–49                         | 19.6%  |
| 50–59                         | 14.7%  |
| Over 60                       | 36.3%  |
| Employment:                   |        |
| Full-time employee            | 40.6%  |
| Freelancer                    | 2.0%   |
| Part-timer                    | 13.3%  |
| Student                       | 1.7%   |
| Home maker                    | 23.9%  |
| Unemployed/retired            | 16.1%  |
| Other                         | 2.3%   |
| Education:                    |        |
| High school and under         | 26.8%  |
| Technical college             | 20.7%  |
| Bachelor’s degree or higher   | 52.4%  |
| Household income (JPY):       |        |
| Not disclosed                 | 16.7%  |
| Under 5 million               | 29.4%  |
| 5–8 million                   | 21.9%  |
| 8–11 million                  | 13.3%  |
| Over 11- million              | 8.1%   |

| Variable                      | Mean   | S.D.  | Med. | Min. | Max. | Skew. | Kurt. |
|-------------------------------|--------|-------|------|------|------|-------|-------|
| Household size                | 2.46   | 1.10  | 2    | 1    | 7    | 0.69  | 0.42  |
| Network size                  | 4.60   | 4.54  | 3    | 0    | 28   | 2.16  | 6.06  |
| Female (n=177)                | 5.31   | 4.45  | 4    | 0    | 27   | 1.5   | 2.77  |
| Male (n=170)                  | 3.86   | 4.52  | 3    | 0    | 28   | 2.98  | 10.83 |
| 20–29 (n=41)                  | 3.85   | 3.15  | 3    | 0    | 14   | 1.2   | 0.84  |
| 30–39 (n=61)                  | 3.75   | 3.96  | 2    | 0    | 19   | 1.81  | 3.01  |
| 40–49 (n=68)                  | 3.57   | 3.5   | 2    | 0    | 17   | 2.05  | 4.05  |
| 50–59 (n=51)                  | 3.75   | 2.92  | 3    | 0    | 14   | 1.26  | 1.75  |
| Over 60 (n=126)               | 6.16   | 5.69  | 5    | 0    | 28   | 1.79  | 3.52  |
| Under 8 million (n=178)       | 4.44   | 4.4   | 3    | 0    | 27   | 2.18  | 6.16  |
| 8–11 million (n=46)           | 4.28   | 3.73  | 4    | 0    | 18   | 2.08  | 4.62  |
| Over 11- million (n=29)       | 5.28   | 3.84  | 4.5  | 0    | 14   | 0.8   | -0.32 |

Kowald et al.\textsuperscript{28} reported average network sizes of 23.8 for Toronto, 11.9 for Zurich, 23.9 for Eindhoven, 20.9 for Concepcion, and 21.6 for Switzerland.

(2) Ego-alter relational characteristics

Ego-alter relational characteristics are summarized in Table 2. In terms of the spatial distribution of ego and alters, the average geodesic distance was 51 km, while the median value was 6 km. Fig. 2 plots the ego-alter distance distribution.

Kowald et al.\textsuperscript{28} reported that the tie distance distribution of social contacts followed a power law distribution. Although there are several ways to fit a power law distribution, in their study they estimated the exponent coefficient $\gamma$ via a linear regression model defined as $\text{probability}_{\text{tie}} \sim \text{distance}^{-\gamma}$ and distances were capped at 100 km. Using the exact same method on the Tokyo sample (Fig. 2), we got a significant exponent coefficient of 1.27 ($p < 0.000$), and in range with their reported values (1.58 – 1.08). This method, however, has not only been shown to generate significant and systematic large errors, but also suffers from three main issues: (i) assumptions to estimate standard errors from linear regression are not valid; (ii) linear regression fits can account for a large share of the variance even when the data do not follow a power law distribution; and (iii) fits extracted from a linear regression are not valid probability distributions\textsuperscript{29}.

When using the more adequate maximum likelihood estimation method\textsuperscript{29}, the estimate of the exponent coefficient is 4.05, a considerably different estimate. To calculate goodness of fit, we followed the approach proposed by Clauset et al.\textsuperscript{29} (See Appendix A for the details of the estimation method) with 2,500 synthetic power law distributed data sets.
Table 2. Descriptive statistics of ego-alter relational characteristics.

| Variable                      | Share(n)          |
|-------------------------------|-------------------|
|                               | All samples       | Male ego | Female ego |
| Homophily:                    |                   |          |            |
| Gender                        | 67.5%             | 60.6%    | 72%        |
| Employment status             | 44.0%             | 47.7%    | 42%        |
| Age cohort                    | 47.1%             | 50.1%    | 45%        |
| Relationship length:          |                   |          |            |
| 1 year or less                | 1.7%              | 2.0%     | 1%         |
| 1–5 years                     | 12.9%             | 10.6%    | 14%        |
| 5–10 years                    | 13.1%             | 12.4%    | 14%        |
| Over 10 years                 | 72.3%             | 75.0%    | 70%        |
| Relationship type:            |                   |          |            |
| Family                        | 40.4%             | 41.3%    | 40%        |
| Work                          | 10.2%             | 14.1%    | 8%         |
| School                        | 10.8%             | 10.1%    | 11%        |
| Neighbor                      | 5.2%              | 5.4%     | 5%         |
| Club or circle friend         | 10.4%             | 6.9%     | 13%        |
| Other                         | 23.0%             | 22.3%    | 23%        |
| Strong tie alters:            |                   |          |            |
| Tier 1                        | 21.3%             | 18.6%    | 23%        |
| Tier 2                        | 18.3%             | 18.6%    | 18%        |
| Tiers 1 and 2                 | 12.7%             | 12.6%    | 13%        |
| Ego-alter distance            |                   |          |            |
| All samples                   | 51.12             | 415.10   | 6.06       |
| Ego is male                   | 31.69             | 106.49   | 5.96       |
| Ego is female                 | 64.40             | 531.06   | 6.14       |
| Contact frequency (All)       |                   |          |            |
| Face-to-face                  | 102.13            | 145.64   | 18.00      |
| Phone                         | 31.77             | 80.90    | 2.00       |
| E-mail or SMS                 | 37.71             | 88.42    | 2.00       |
| SNS                           | 22.66             | 73.54    | 0.00       |
| Contact freq.(male ego)       |                   |          |            |
| Face-to-face                  | 110.18            | 151.98   | 18.00      |
| Phone                         | 31.74             | 79.21    | 2.00       |
| E-mail or SMS                 | 33.14             | 84.68    | 2.00       |
| SNS                           | 17.92             | 66.68    | 0.00       |
| Contact freq.(female ego)     |                   |          |            |
| Face-to-face                  | 96.63             | 140.97   | 18.00      |
| Phone                         | 31.78             | 82.07    | 2.00       |
| E-mail or SMS                 | 40.83             | 90.81    | 2.00       |
| SNS                           | 25.90             | 77.75    | 0.00       |

We also tested the power law hypothesis without a distance cutoff, that is, using all data in the sample. As shown in Fig. 3 and Fig. 4, goodness of fit tests suggest that the power law distribution is not a plausible fit to the data. Note that with this method, significant p-values are indicative of poor fit.

Contact frequency by mode was measured with an ordinal variable with eight levels: never, 1–3 times per year, 4–11 times per year, 1–2 times per month, 3–4 times per month, 1–3 times per week, 4–7 times per week, several times per day. To convert to a continuous measure of annual contact frequency, the midpoint of each frequency range was set as the frequency value. For the never case, frequency was set at 0, and for the several times per day case the frequency was set at 365.

Of all modes, face-to-face was the most frequent contact mode (mean: 102, median: 18). As Fig. 5 shows, all modes tend to decrease with distance. Face-to-face contact frequency shows the sharpest decrease, as a result of higher communication costs given distance, while ICT-based mode exhibits a less steep decline. It can also be seen that most interactions occur within a 50 km distance range, even though for ICT-based modes the marginal communication costs are considerably lower than those for face-to-face contact, showing that distance still imposes limits on interactions, even in the age of low-cost communications.

When segmenting contact frequency by relationship type, as shown in Fig. 6, family and work friends account for the largest share of contact, particularly
for face-to-face contact.

Regarding social isolation measures, in addition to the size issue discussed above, among egos who reported at least one alter in their network, we estimated the share of the sample who reported no contact or low contact for each mode. Egos with no contact were defined as those for which the sum of all annual contact frequencies with alters was zero. Egos with low-contact frequencies were defined as those for which the sum of the annual contact frequencies with all alters was less or equal to 12. That is, roughly one interaction per month or less in total. As shown in Fig. 7, 6% of egos had low face-to-face contact frequencies with their network, while 2% had no contact, or less than once per year.

When looking at ICT contact frequencies, 13% of the sample had low ICT contact frequencies and 6% had no contact. When considering both face-to-face and ICT, only 3% had low contact levels, and 1% no contact at all.

![Fig.2 Log-log plot of relationship probability and distance. Linear regression fit, 100km cutoff, $\gamma=1.27$ (p<0.00).](image1)

![Fig.3 Log-log plot of relationship probability and distance. Maximum likelihood fit, 100km cutoff, $\gamma=4.05$, $k_{min}=38$km (p<0.00).](image2)

![Fig.4 Log-log plot of relationship probability and distance. Maximum likelihood fit, no distance cutoff, $\gamma=1.88$, $k_{min}=15$km (p<0.00).](image3)

![Fig.5 Ego-alter yearly contact frequency plots.](image4)
Fig. 6 Ego-alter yearly contact frequency by relationship type.

Fig. 7 Percentage of egos with low- and no-contact with network members.

Table 3 Descriptive statistics of clique composition by type.

| Variable                                 | Share (n) |
|------------------------------------------|-----------|
|                                          | All (405) | Family (110) | Work (45) | School (51) | Other (199) |
| Strong tie alters:                       |           |               |           |             |             |
| Tier 1 (Would discuss important issues with) | 22.4%    | 29.6%         | 16.9%     | 20.9%       | 20.0%       |
| Tier 2 (Would ask for help in an emergency) | 18.6%    | 26.6%         | 14.6%     | 19.9%       | 14.8%       |
| Homophily:                               |           |               |           |             |             |
| Age cohort                               | 48.0%     | 20.2%         | 40.4%     | 90.0%       | 54.3%       |
| Gender                                   | 72.6%     | 48.2%         | 80.0%     | 92.5%       | 79.3%       |
| Employment status                        | 45.5%     | 22.8%         | 67.5%     | 49.8%       | 52.0%       |
| Civil status                             | 52.4%     | 43.1%         | 57.4%     | 55.6%       | 55.6%       |
| Clique activity purpose (1)              |           |               |           |             |             |
| Eating/drinking out                      | 43.0%     | 23.6%         | 66.7%     | 72.5%       | 40.7%       |
| Physical activity/sports                 | 7.2%      | 0.0%          | 2.2%      | 0.0%        | 14.1%       |
| Other hobbies, leisure                   | 10.9%     | 2.7%          | 0.0%      | 3.9%        | 19.6%       |
| Shopping                                 | 2.5%      | 6.4%          | 0.0%      | 2.0%        | 1.0%        |
| Volunteering                             | 2.7%      | 0.0%          | 4.4%      | 0.0%        | 4.5%        |
| Socializing                              | 6.4%      | 7.3%          | 4.4%      | 7.8%        | 6.0%        |
| Other                                    | 27.4%     | 60.0%         | 22.2%     | 13.7%       | 14.1%       |
| Clique activity purpose (2)              |           |               |           |             |             |
| Eating/drinking out                      | 9.9%      | 5.5%          | 15.6%     | 15.7%       | 9.5%        |
| Physical activity/sports                 | 2.5%      | 1.8%          | 0.0%      | 0.0%        | 4.0%        |
| Other hobbies, leisure                   | 2.2%      | 0.0%          | 4.4%      | 2.0%        | 3.0%        |
| Shopping                                 | 2.5%      | 5.5%          | 0.0%      | 2.0%        | 1.5%        |
| Volunteering                             | 0.7%      | 0.0%          | 0.0%      | 0.0%        | 1.5%        |
| Socializing                              | 1.2%      | 1.8%          | 0.0%      | 0.0%        | 1.5%        |
| Other                                    | 8.1%      | 16.4%         | 6.7%      | 9.8%        | 3.5%        |

| Clique size including ego                | Mean   | S.D. | Med | Min. | Max.  | Skew. | Kurt. |
| All cliques (405)                        | 4.14   | 1.80 | 3    | 3    | 14    | 2.60  | 8.50  |
| Family cliques (110)                     | 4.24   | 1.87 | 4    | 3    | 13    | 2.63  | 8.53  |
| Work cliques (45)                        | 4.18   | 2.33 | 3    | 3    | 14    | 2.68  | 7.28  |
| School cliques (51)                      | 3.71   | 1.06 | 3    | 3    | 7     | 1.37  | 0.83  |
| Other cliques (199)                      | 4.20   | 1.77 | 4    | 3    | 14    | 2.25  | 6.26  |
(3) Clique characteristics

Regarding the number of cliques per ego, 12% of the sample with network size 1 or larger did not report any cliques. From the remaining sample, 28% reported cliques composed of only one alter, that is, a dyad. In this section we consider only the cliques composed of cliques composed of two alters or more. In the subsample of egos with cliques of two alters or more, the average number of cliques was 2.4 (median 2) with a maximum of six cliques, while the average size of clique when including ego was 4.14 (median 3). In total there were 405 cliques reported. **Table 3** summarizes the clique composition characteristics by clique type.

Based on the names respondents gave to the cliques they reported in the survey, we classified cliques into four groups: family cliques (keywords: family, relatives, spouse, siblings, etc.); work cliques (workmates, former workmates, company, firm, etc.); school cliques (junior high school, high school, college, grad school, etc.); and other cliques.

5. MODELING FRIENDSHIP AND CONTACT FREQUENCY PROBABILITY

To further analyze the relationships between personal networks and social interactions we built two statistical models. First, we modelled friendship probability as a function of ego-alter characteristics, and then we modelled ego-alter contact frequency probability by modes.

The friendship probability model is of interest for its potential use in generating population-wide networks in large-scale models, such as agent-based travel simulators incorporating social interactions.

To model friendship probability and contact frequency by mode, a multilevel binary probit was estimated. The dependent variable was a binary variable \( y \), indicating the existence of a tie between ego and alter.

Since in this survey we did not observe non-ties, in addition to the observed ties (\( y=1 \)), we randomly sampled 25 egos from the data and added them to the data with dependent variable valued \( y=0 \). A similar resampling method has been used by Arentze, Kowald, and Axhausen. Given the existing limitations to observe non-ties, this approach provides a practical, if somewhat ad hoc, way to simulate random encounters for ego with persons with different sociodemographic characteristics. **Table 4** summarizes the results of bootstrap estimates of the friendship formation model with 1000 replications, each with a different set of randomly sampled non-ties. For each ego, relatives were excluded from the model.

| Variable                      | Coef. | t-stat |
|-------------------------------|-------|-------|
| Constant                      | -2.96 | -37.12|
| Ego and alter are both male   | 0.72  | 9.56  |
| Ego and alter are both female | 0.88  | 12.74 |
| Same civil status             | 0.38  | 7.04  |
| Same employment status        | 0.48  | 9.09  |
| Same age cohort               | 1.27  | 23.85 |
| Inverse of geodesic distance  | 1.03  | 15.05 |
| Sigma                         | 0.81  | 13.31 |

All parameters are statistically significant at the 0.01 level in all realizations. All measures of homophily are positively associated with friendship probability, while distance is negatively associated with it.

To model contact frequency probability, a multivariate hierarchical ordinal probit model was estimated. This model structure allows not only to measure correlation among contact modes but also accounts for the hierarchical nature of the data, which include both ego-level personal and network characteristics and ego-alter relational characteristics. Fixed effect variables included in this model are variables identified in the literature to be associated with contact frequencies among network members. Random effects are specified to capture the variance and covariance of contact modes among egos, as well as their correlations. **Table 5** summarizes the contact frequency probability model results.

Regarding fixed effects, compared to the 60 years and older cohort in the sample, younger cohorts have lower phone contact frequency probabilities, with the largest effect observed for the 20–29 years cohort. Similarly, this cohort has lower E-mail/SMS contact frequency probability than all other cohorts, as it is likely that this group relies largely on SNS for ICT-based communications. Other than that, although there are some statistically significant coefficients that, with few exceptions, are related to face-to-face contact frequency probability, for ego personal characteristics, no clearly discernible pattern can be observed.

Social network size is negatively associated with contact frequencies for all modes except SNS. This suggests that egos with larger networks tend to adjust contact frequencies with each alter to account for a larger number of alters in their network, as suggested by Dist. Note that although this suggests fewer contact for each ego-alter pair, given larger networks, total contact frequencies at the ego level might be higher.

Regarding relational characteristics, gender and
Table 5  Multivariate hierarchical ordinal probit of contact frequency by mode.

| Fixed effects                  | Face-to-Face | Phone | E-mail/SMS | SNS |
|-------------------------------|--------------|-------|------------|-----|
|                               | Coef. t-stat | Coef. t-stat | Coef. t-stat | Coef. t-stat |
| Gender                        |              |        |            |     |
| Ego is male                   | 0.24 2.19    | 0.01 0.05 | -0.21 -1.11 | -1.76 -3.94 |
| Ego is female (reference)     | -            | -      | -          | -   |
| Age of ego                    |              |        |            |     |
| 20–29                         | 0.18 0.96    | -0.91 -3.34 | -1.00 -2.99 | 0.95 1.41 |
| 30–39                         | 0.40 2.87    | -0.55 -2.72 | -0.27 -1.10 | 0.88 1.62 |
| 40–49                         | 0.09 0.71    | -0.47 -2.49 | -0.38 -1.62 | 0.66 1.22 |
| 50–59                         | 0.24 1.75    | -0.59 -2.91 | -0.10 -0.40 | -1.56 -2.74 |
| 60 and over (reference)       | -            | -      | -          | -   |
| Employment status of ego      |              |        |            |     |
| Full-time employee            | 0.54 4.15    | -0.03 -0.13 | -0.05 -0.22 | -0.83 -1.50 |
| Freelancer                    | 0.32 1.15    | 0.36 0.82 | 0.60 1.12  | -0.96 -0.80 |
| Part-timer                    | 0.43 2.65    | <0.00 0.01 | -0.04 -0.13 | -1.71 -2.58 |
| Homemaker                     | 0.49 3.21    | -0.02 -0.11 | 0.17 0.62  | -1.93 -2.95 |
| Student                       | 1.02 2.93    | 0.18 0.33 | -1.06 -1.39 | 0.09 0.06 |
| Unemployed/retired/other       | -            | -      | -          | -   |
| Civil status of ego           |              |        |            |     |
| In a relationship             | 0.03 0.12    | 0.56 1.37 | 0.92 1.80  | -0.17 -0.17 |
| Married with children         | 0.44 3.34    | <0.00 -0.02 | -0.19 -0.85 | -0.78 -1.55 |
| Married with no children      | 0.31 1.99    | -0.17 -0.79 | 0.32 1.21  | -0.89 -1.45 |
| Divorced                      | 0.31 1.12    | -0.17 -0.40 | -0.28 -0.57 | -0.51 -0.46 |
| Widowed                       | 0.38 1.67    | 0.42 1.22 | 0.10 0.24  | -0.52 -0.50 |
| Single (reference)            | -            | -      | -          | -   |
| Social network size           | -0.02 -3.10  | -0.06 -4.48 | -0.06 -3.75 | -0.04 -1.21 |

Relational characteristics

| Relative                      | 0.34 1.50    | -1.46 -5.98 | -2.74 -8.81 | -3.11 -5.13 |
| Immediate family              | 0.90 5.41    | -0.08 -0.47 | 0.22 1.12   | -0.26 -0.88 |

Homophily

| Gender (non-relative)         | -0.94 -11.29 | -0.66 -7.50 | -0.28 -2.97 | -0.50 -3.91 |
| Gender (relative)             | 0.03 0.27    | -0.01 -0.04 | 0.07 0.53   | -0.23 -1.13 |
| Age (non-relative)            | -0.11 -1.45  | -0.19 -2.27 | -0.18 -2.09 | -0.29 -2.40 |
| Age (relative)                | -0.35 -1.90  | -0.35 -1.77 | -0.69 -3.19 | -0.66 -2.06 |
| Occupation (non-relative)      | -0.03 -0.40  | 0.12 1.40   | -0.06 -0.73 | 0.05 0.37  |
| Occupation (relative)         | <0.00 -0.03  | 0.04 0.23   | 0.12 0.72   | 0.38 1.62  |
| Civil status (non-relative)    | 0.39 5.07    | 0.11 1.27   | 0.08 0.92   | 0.05 0.39  |
| Civil status (relative)       | -0.44 -3.13  | -0.11 -0.78 | 0.08 0.51   | 0.19 0.86  |
| Strong tie (non-relative)      | 0.28 3.59    | 0.31 3.72   | 0.18 2.07   | 0.10 0.87  |
| Strong tie (relative)         | 0.01 0.10    | 0.36 2.79   | 0.48 3.47   | 0.43 2.19  |
| Relation duration Over 10 yrs (non-relative) | -0.04 -0.47 | 0.09 1.06 | -0.02 -0.22 | -0.17 -1.37 |
| Relation duration Over 10 yrs (relative) | -0.58 -2.62 | 1.62 7.18 | 2.18 7.44 | 3.06 5.70 |
| Log of geodesic distance (non-relative) | -0.19 -11.52 | -0.02 -1.10 | <0.00 -0.07 | <0.00 0.05 |
| Log of geodesic distance (relative) | -0.39 -14.01 | 0.04 1.55   | 0.03 0.88   | -0.12 -2.84 |

Thresholds

| Coef. t-stat | Random effects |
|--------------|---------------|
| 1/2          | -1.75 -8.93   | Var. S.D. Correlations |
| 2/3          | -1.04 -5.34   | F2F 0.17 0.41 F2F Phone E-mail |
| 3/4          | -0.58 -3.02   | Phone 0.74 0.86 51 |
| 4/5          | 0.05 0.27     | E-mail 1.28 1.13 0.40 0.47 |
| 5/6          | 0.37 1.90     | SNS 6.02 2.45 0.07 0.24 -0.06 |
| 6/7          | 0.99 5.08     |               |
| 7/8          | 1.44 7.34     |               |

Model Comparison (against final model)

| Par. | AIC | L.R. Test | LR Test | d.f. | p-value |
|------|-----|-----------|---------|------|---------|
| No mode segmentation, no interactions | 42 18813 | -9365 | 1209.9 | 103 | <0.001 |
| Segmentation, no interactions | 117 18113 | -8940 | 359.32 | 28 | <0.001 |
| Final model (fixed effects only) | 135 19329 | -9530 | 1540.1 | 10 | <0.001 |
| Final model | 145 17809 | -8760 | - | - | - |

While in the case of age homophily among relatives, age homophily, coefficients are negatively associated with contact frequency probability, with no significant effect for gender homophily with relatives.
given that families are by definition multigenerational, the negative coefficients are not unexpected, for the case of non-relatives, the effect direction is rather counterintuitive. It is worth mentioning that negative gender\(^6\) and age\(^{25,31}\) homophily for face-to-face contact has been reported in the literature, but no distinctions were made by relationship type, which makes comparison difficult. It is also important to note that, as shown in Table 4, age and gender homophily were both positively associated with friendship probability. As such, the dynamics of homophily in terms of contact frequency deserve further investigation.

Regarding occupational and civil status homophily, civil status and face-to-face contact were the only significant coefficients, with relative civil status homophily showing a positive association, and non-relative homophily showing a negative association. Effect directions are reasonable, since, similar to the argument for age homophily, familial ties are by definition diverse, while for non-relative ties, egos are likely to interact more with alters that are in similar life stages.

Consistent with findings from the literature, Tie strength (tier 1 or 2) is, with some exceptions, significantly and positively associated with higher contact frequency probability for both face-to-face\(^6,32\) and ICT-based contacts\(^6\). Relationship duration is not significant for non-relatives, but negatively associated with face-to-face and positively associated with ICT contact frequency probability with relative. Relationship duration for relatives is somewhat hard to interpret, but family members known for less than 10 years include young children (under 10 years old), which might explain the higher face-to-face contact frequency probability and negative effects for ICT modes.

Finally, ego-alter distance is negatively associated with face-to-face contact frequency probability, with larger effects observed for relatives, than for non-relatives. Regarding ICT modes, SNS contact with relatives is negatively associated with distance, although the coefficient size suggests a smaller effect than face-to-face. In this sample, however, no substitution effect was observed between face-to-face and ICT modes given increases in distance.

Regarding random effects, all modes exhibit positive correlations with each other except for the cases of SNS with face-to-face and SNS with E-mail/SMS, which exhibited negative correlations. These correlations are, however, very small. This suggests that in the sample, phone and E-mail/SMS are complementary modes to face-to-face as reported in the literature, while SNS is, to a very small extent, a substitute. The likelihood ratio test p-value of the final model against the fixed-effects-only model was \(<0.001\).

6. DISCUSSION AND CONCLUSION

In this article we presented the results of the first ego-centric personal network survey in Japan using an almost unrestricted name generator, which allowed for better capturing network geographies and fine-grained relation attributes. Although due to the sampling method, population representativeness cannot be claimed, we summarize the results of this study regarding personal networks in Japan, and social networks in general when possible, and discuss some potential avenues for further research. As an exploratory study, the results presented here should be validated in future research with a probability sample.

1. Are Japanese personal networks smaller than European and Pan-American networks?

Using a methodology comparable to recent studies in Europe and the Americas, the average network size in this sample was 4.6, a considerably smaller network than the reported values in international literature. Even reported values in the Japanese literature using direct counting of network members suggest smaller networks. Although there are sampling and methodological differences preventing correct comparisons, we hypothesize that Japanese networks are on average smaller than European and Pan-American networks. Several factors can explain this phenomenon, such as resource constraints (i.e., time constraints for workers) and personality traits, but the very nature of Japanese social relationships might be a key factor, in particular, group consciousness and their vertical nature. High levels of group consciousness in Japan, as Nakane\(^33\) points out, strengthens sense of unity and group solidarity, but also exacerbates the awareness between “our people” and “outsiders.” Furthermore, we argue that the vertical nature, specifically the strong awareness of ranking order in Japanese relationships might impose additional costs to creating and maintaining ties, which might help explain smaller networks. Even if traditional social groups were to be collapsing as it has been argued, there is no reason to believe that these features of Japanese social interactions will not exist in some form or another in voluntary groups and pure relationships outside of traditional family, work, or local relationships.

That being said, it has been pointed out by Miyata et al. that the Japanese might display greater reticence to disclose personal information, which might make network size comparisons with Western countries more difficult\(^14\). However, it remains to be seen to what extent this holds true for anonymous survey panels, such as the one used in this study. At any rate, a probability survey is necessary to obtain comparable and generalizable results.
2. Do ego-alter distances follow a power law distribution?

Although it has been reported frequently in the literature, goodness of fit tests rejected the hypothesis that ego-alter distances follow a power law distribution in our sample, even though the most frequently used, yet inaccurate method to fit a power law distribution yielded a significant exponent coefficient. Thus, we would be cautious about the power-law distribution argument and recommend additional testing with the independent samples. Finding a parametric form of social network distribution is of interest as such a parametric form can be useful for simulating population-level social networks, and potentially implemented in agent-based simulation models that incorporate social network interactions.

3. What about social isolation?

While smaller networks do not necessarily imply isolation, in our sample, 7.2% did not state having a person to spend free time with, or to discuss important problems with. Data from the 2013 Japan General Social Survey also reported a value 8.9%. Furthermore, among those respondents who reported at least one alter, the share of persons with no face-to-face contact was 2% and with no contact whatsoever by any mode was 1%. These values might suggest some level of social isolation. However, two important issues must be pointed out. First, if this measure of social isolation was valid, whether this was a direct result of the collapse of intermediate social groups, as suggested by the proponents of the isolated society theory remains an open question. Second, it remains to be validated whether or not this is a true measure of social isolation or just a statistical artifact. Similar proportions of respondents not reporting any alters have been observed in surveys outside Japan. For example, in the snowball sample study conducted in Switzerland by Kowald, 6.2% of the sample reported no alters at all. And as Kowald points out, with the current instruments it is difficult to differentiate between people who have no social contacts and people who did not report them. As such, a better instrument design might be necessary to adequately capture social isolation.

4. What about friendship formation and contact frequency probability?

The results of the friendship formation model suggest that homophily and distance are important factors in friendship probability. However, a key limitation of this model is that in the survey, non-ties are not observed, so there is room for improvement in terms of survey methods. For example, a meeting diary might be a good strategy to document potential new ties. Such a method could be combined with a natural experiment (i.e., persons moving to a new city, starting a new job, etc.) to observe transformations in social networks that allow for the estimation of better parameters with true panel data.

In terms of contact frequency probability, no clear patterns were observed for personal-level characteristics. On the other hand, in terms of effect directions, relatively consistent patterns were observed for network size and relational characteristics, such as age and gender homophily, tie strength, and relationship duration. Some effects, such as the negative effects of age and gender homophily with non-relatives were rather unexpected, suggesting the need for further investigation on the dynamics of homophily in terms of contact frequency.

Distance was negatively associated with face-to-face contact frequency and no substitution effect was observed given increases in ego-alter distances in the sample.

5. Implications to the field and future work

Through this exploratory analysis we aimed at highlighting the potential of incorporating social networks to travel behavior analysis, in particular, to study socially motivated travel which, due to higher spatio-temporal variability, is much harder to predict than dimensions more traditionally analyzed, such as commuting and shopping. To deepen our understanding of such behavior, there is need to move away from the independent decision maker assumption and work to incorporate the effects of the networks a person is embedded in.

Although the social contact measures used here are rather rough, they can be used to provide a measure of average levels of demand for social interactions. However, on a longer time horizon, future research should move towards more explicitly measuring and analyzing socially motivated travel incorporating social network interactions in dimensions such as activity choice, companionship choice, destination choice, etc.

We expect the findings from this exploratory study will motivate a deeper discussion regarding the role of social networks and social interactions on socially motivated travel in the Japanese context.

ACKNOWLEDGEMENTS: This study was supported by JSPS KAKENHI Grant No.:17K14737.

The survey presented in this article was approved by the School of Engineering’s Research Ethics Committee at the University of Tokyo. Certificate No. KE17-65.
CONTRIBUTION STATEMENTS

GTP conceived the study. GTP, KT, and NH designed the survey instrument. GTP built the survey system, analyzed the data, and wrote the first draft of the manuscript. GTP, KT, and NH all provided comments and revisions to the first manuscript and approved the final version of the manuscript.

APPENDIX A: ESTIMATING THE EXPONENT COEFFICIENT OF A POWER-LAW DISTRIBUTED VARIABLE AND EVALUATING ITS GOODNESS OF FIT

This section describes the estimation procedure for the exponent coefficient of a power-law distributed variable and its goodness of fit evaluation as defined in Clauset et al.\(^{29}\) and Barabasi\(^{39}\).

1. For a given ego-alter distance \(k\) choose a value \(k_{\min} < K_{\min} < k_{\max}\) and estimate the degree exponent \(\gamma\) as

\[
\gamma = 1 + N \left[ \sum_{i=1}^{N} \ln \frac{k_i}{K_{\min} - \frac{1}{2}} \right]^{-1}
\]

2. For the calculated pair of values \((\gamma, K_{\min})\), assume a probability distribution function PDF for \(k\) of the form

\[
p_k = \frac{1}{\zeta(\gamma, K_{\min})} k^{-\gamma}
\]

3. Estimate the respective cumulative distribution function CDF of \(k\) as

\[
p_k = 1 - \frac{\zeta(\gamma, k)}{\zeta(\gamma, K_{\min})}
\]

where \(\zeta(\cdot)\) is the Hurwitz-zeta function.

4. For the pair of values \((\gamma, K_{\min})\), use the Kolmogrov-Smirnov test to get the maximum distance \(D\) between the observed data CDF \(S(k)\) and the fitted model described in the previous step, where \(D\) is defined as

\[
D = \max_{k \geq K_{\min}} |S(k) - P_k|
\]

5. Find the value that minimizes \(D\) by scanning the whole range of \(k\). The optimal fit for \(\gamma\) is that where \(D\) is minimized, \(D_{\min}\).

6. To estimate the goodness of fit, use the estimated optimal values of \(\gamma, K_{\min}\) to generate a sequence of values \(k\) following the probability distribution function defined in step 2, and calculate \(D_{\text{synth}}\) for the synthetic data.

7. Repeat step 6 \(M\) times and get the distribution of \(D_{\text{synth}}\), \(p(D_{\text{synth}})\).

8. Plot \(p(D_{\text{synth}})\) and plot \(D_{\min}\) as a vertical line. If this line falls within the distribution \(p(D_{\text{synth}})\), then the power law is a reasonable model fit.

9. The \(p\)-value is effectively the share of cases where the resulting statistic is larger than the value for the empirical data. Barabasi\(^{39}\) suggests the threshold \(p>0.01\) to establish a plausible fit for the data, while Clauset et al.\(^{29}\) suggests a stricter threshold of \(p>0.10\).

REFERENCES

1) Axhausen, K. W.: Social networks and travel: Some hypotheses, Social Dimensions of Sustainable Transport: Transatlantic Perspectives, pp. 90–108, 2005.
2) Schlich, R., Schönfelder, S., Hanson, S. and Axhausen, K. W.: Structures of leisure travel: Temporal and spatial variability, Transport Reviews, Vol. 24, pp. 219–237, 2004.
3) Kowald, M. and Axhausen, K. W.: Social Networks and Travel Behavior, pp. 1–4, 2015.
4) Kowald, M. and Axhausen, K. W.: Focusing on connected personal leisure networks: Selected results from a snowball sample, Environment and Planning - Part A, Vol. 44, pp. 1085–1100, 2012.
5) Carrasco, J. A. and Miller, E. J.: Exploring the propensity to perform social activities: A social network approach, Transportation, Vol. 33, pp. 463–480, 2006.
6) Van den Berg, P., Arentze, T. and Timmermans, H.: Size and composition of ego-centered social networks and their effect on geographic distance and contact frequency, Transportation Research Record: Journal of the Transportation Research Board, pp. 1–9, 2009.
7) Axhausen, K. W. and Frei, A.: Size and structure of social network geographies, Arbeitsbericht Verkehrs- und Raumplanung, Vol. 444, 2007.
8) Arentze, T. A., Kowald, M. and Axhausen, K. W.: An agent-based random-utility-maximization model to generate social networks with transitivity in geographic space, Social Networks, Vol. 35, pp. 451–459, 2013.
9) Harada, K. and Sugasawa, H.: Urbanism and personal networks: Multilevel analyses of kin, neighbors, and friends, Japanese Sociological Review, Vol. 65, pp. 80–96, 2014.
10) Harada, K.: Social stratification and personal networks: Educational, occupational and income differences among men and women, Iryo to Shakai, Vol. 22, pp. 57–68, 2012.
11) Ishida, M.: Koritsu no shakaigaku: Muen shakai no shohōsen, Keisō Shobō, 2011.
12) Fujimoto, K.: From Social Isolation to Social Connection, Minerva Shobō, 2012.(In Japanese)
13) Ito, M. and Okabe, D.: Intimate connections: Contextualizing Japanese youth and mobile messaging, The Inside Text: Social, Cultural, and Design Perspectives on SMS (Harper, R., Palen, L., Taylor, A. ed.), pp. 127–145, Springer, 2005.
(Received February 8, 2019)
(Accepted September 18, 2019)