

# Analyzing changes in social networks over time using panel data:

## A case study of international students in Japan

### パネルデータを用いたソーシャルネットワーク変化の分析

#### —訪日留学生を対象として—

Department of Urban Engineering, The University of Tokyo 37-176954 Swarnali Dihingia

Social networks have attracted attention in different fields of research in recent years and the role of social networks in shaping decision in a variety of context has been studied. This incorporation of social dimension in modelling has created the need to understand how social network evolve over time and in particular which social contacts are retained. However, existing studies have failed to capture this change over time in social network in real time. This study presents an exploratory analysis in the formation of ties and its retention of social networks by using the data conducted by a booklet diary survey, carried for 21 days in the case of new international students at The University of Tokyo. Some research question regarding the relational attributes and activities are put forward to study the retention of ties over time and a panel data probit has been used for the friendship retention estimation.

## 1.Introduction

### 1.1 Rationale

Over the last decades, transportation researchers have started to consider individual needs in conducting activities as an explanation of their travel choice behavior. These studies have been shifting from trip based to activity based and from individuals to ones' social network<sup>1)</sup>. There have been many studies done to measure how these social networks of an individual influence mobility decisions as it is clear that social interactions enable travel and vice versa. For travel behavior research, the social network approach is where an individual is taken as a unit of observation and the relationships connecting these actors are studied<sup>2)</sup>.

Nonetheless, a key characteristic of social networks is that they are not static but evolve with time and also with life events occurring in the ego's life. The process of network formation and change over time is complex and depends on various characteristics of the different individuals involved, and its study requires adequate data<sup>3)</sup>. This move from static to dynamic, longitudinal approaches of network analysis allows a corresponding move from purely descriptive to explanatory terms<sup>4)</sup>. First attempts to model the dynamics of social network and their interaction with life-cycle events have been made by Sharmeen<sup>5)</sup> who have extended their method towards dynamic analysis and prediction of friendship formation, its maintenance and its retention, although

through retrospective data collection and not real time.

Keeping the above limitation in mind, this study focuses on the dynamics of ego-centric social network over time, particularly (i) how ties are generated and (ii) their retention along with their social interactions and activity space in case of new students, who are triggered by a life event: joining a new university (in a foreign country).

As mentioned earlier, the previous studies conducted were cross-sectional and retrospective in nature and the key contribution of this study is the usage of real panel data to observe the evolution and the change in real time, which to the best of my knowledge has not been done until now in this field. This study used a Friendship Diary to collect data for the study of the evolution of the social network and its subsequent changes in real time.

Based on the above rationale, this study aims to answer the following research questions:

1. How do social networks and activity patterns change given a life event change (i.e., joining a university)?
2. What is the effect of relational attributes on tie retention?
3. What is the effect of joint activity types on tie retention?

In this study, since the students surveyed are new international students, it is assumed that their social network starts from scratch in Japan. Based on this assumption, two hypotheses on social network and

activity has been postulated.

1. The size of social network is unstable at the start of a life cycle event and stabilizes after an initial period of adaptation.
2. The size of the local activity space of an individual stabilizes after an initial exploration.

One point to note is that Hypothesis 2 was postulated by Axhausen<sup>6)</sup>. However, the hypothesis is theoretical in nature and has not yet been empirically validated.

## 1.2 Importance in Planning

In behavioral science, behavior and outcomes are shaped by social interactions between agents. The relevance of this study and model should be appealing not only to researchers but also to practical analysts and policy evaluators as when considering travel demand forecasting, the need to simulate and predict change of individuals with time is very well recognized by Bhat<sup>7)</sup>. It is a well-known fact that social network generates and influences activity and travel patterns. Therefore, the evolution of social network over time can be used as a function to understand the change in travel behavior. At the same time there is a potential to incorporate these changes in agent-based model of social interactions and activity generation.

## 2. Survey and Data collection

### 2.1 Introduction

To understand the afore-mentioned changes, a survey was conducted at the start of the Spring semester in April 2019 by distributing a handy, A6 size booklet named, "Friendship Diary." The participants were studied for a month and as such, the participants were in possession of the diary for the whole period of April 2019. The booklet survey carried out for this study was approved by the School of Engineering's research ethics committee at the University of Tokyo (Certificate No. KE18-59).

The bonus points of the diaries are that they refer to any daily event-recording procedures that tracks "information in relation to the passage of time" or any method of data collection that "entails respondents to record information about their subjective experiences, behaviors, cognitions, and social interactions linked to a temporal framework"<sup>8)</sup>.

### 2.2 Diary log- Friendship Diary

The Friendship Diary consists of five Sections.

Section I consists of the individual and the household

characteristics of the participants, hereafter ego. This section also ask the egos about the frequency of activities they did before coming to Japan, which is important as the basic idea of activities done by the ego can be understood and accordingly how conducting these activities in a new environment plays a role in the generation, maintenance and the retention of the ties can be observed.

Section II and III provide essential information for social ties and personal networks, where the second section reflects on the circumstance of knowing the person with whom leisure activity was conducted through the name generator, "*Please name the people in Japan with whom you spend your leisure time until now.*" The third section consists of the daily meeting diary where the egos note down the name of the people they meet every day at the end of each day, and this was done for a month. In view of the originality of the study being real time panel data, this daily meeting diary appease the need of it. This section also captures the nature and variation of encounters or interpersonal contacts and the place of the activity done.

To understand the spatial distribution of the activities done, the egos were requested to fill in the name of the nearest station the activity was done. Moreover, to bring in an insight about the tie maintained and retained, the egos were also requested to rate the overall experience about each activity with alters done, keeping in mind the alter and the place, on a scale of 0-5, 0 being awful and 5 being amazing.

The fact that such a task is quite tedious, demanding, and time consuming, leads to the question of how long one should keep the diary. The optimal duration of diary keeping is a trade-off between costs and benefits<sup>9)</sup>. A diary that lasts less than two weeks might apparently be too short to assess the entire change in the social network and so the survey was carried out for a month. The first week, however, was conducted on a retrospective basis as mentioned earlier through name generator and the other three weeks were conducted on daily basis for 21 days.

Section IV consists of the names of the clubs or groups to which their alters belong along with the circumstance in which the planned activity took place. The main purpose of this is to understand the cliques of friends and to capture the names of the contacts of the egos who knows each other.

Section V, the last and the most important section captures the tie retention. Here the egos mention the name of the alters that they still have contact with at the end of the survey period. The answer to this section is the input to the model estimated in Chapter 4 as the dependent variable, whether the tie is retained or not.

### 2.3 The Fieldwork of Diary Keeping

Participants were recruited from new international students who entered the University of Tokyo as graduate students or exchange students in April 2019, based on convenience sampling through the Japanese classes at various campuses and the dormitories which are located at Komaba, Kashiwa and Shirokanedai. Japanese classes seem to be an ideal platform for the students to conduct in social activities and thus bring any changes to their recent, developing and unstable social network.

Given that the task of diary keeping for a month is highly demanding and thus require a strong commitment from the participants and to ensure that the participants will complete their diaries, a monetary incentives of a prize money of ¥3,000 Amazon gift card was promised to the participants after successful completion of the Friendship Diary as also a mean of encouraging them to keep their dairy-keeping routine.

### 2.4 Response Rate

A total of 32 booklets were distributed at Komaba International Dormitory and 11 at Kashiwa International Lodge and 21 booklets have been received duly filled by the participants from Komaba and five from Kashiwa respectively, thus resulting in a response rate of 65% and 45%. Four out of seven booklets were collected from Japanese classes giving a response rate of 57%. So, as a whole, 30 completed booklets were collected thus obtaining a total response rate of 60%.

## 3. Descriptive analysis

### 3.1 Ego personal characteristics

Out of 30 booklets, two were discarded/decided not to be used for further analysis because of its negative/foul response results, upon inspection.

As can be seen from Figure 1, the percentage of male is slightly higher than the female and 20 egos are master-course students for a period of two years in Japan, with five students being exchange students and three doctor-course students. Most egos except three live in the dormitories and their age cohort falls in the age group of

21-30 years.

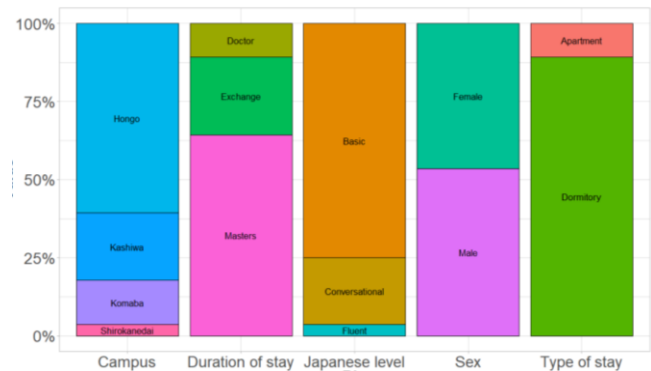


Figure 1: Ego personal characteristics

### 3.2 Social network size of ego and its basic pattern

The basic pattern in the social network change in the 28 egos on a daily basis could be obtained through the survey for the whole 21 days period to test Hypothesis 1. As seen from Figure 2, nine egos' social network size kept on increasing until the end of the survey period, and 12 egos stabilized in the last few days. The rest seven egos' social networks stabilized at a quite early period of time. Figure 3 shows the mean number of new alters met and the mean number of the alters retained, the trend towards stabilization could be observed. Concluding from Figure 2 and 3, the hypothesis postulated is true in the case of this study.

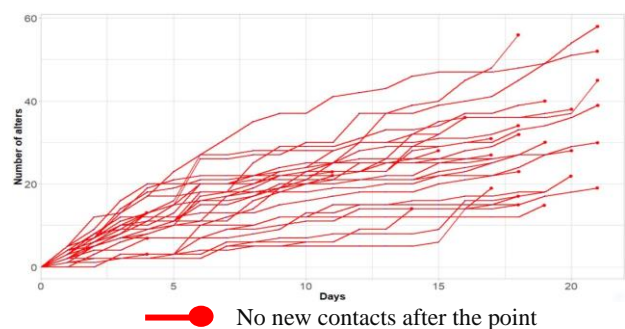


Figure 2: Cumulative plot of new alters met

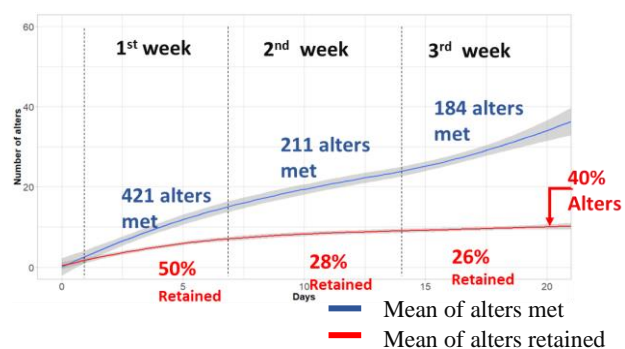


Figure 3: Mean of alter met vs retained

Also as seen from Figure 3, the intensity of meeting people is higher in the first week than in the second and the third week, and the same goes for the tie retention. As a whole, 40% of the ties has been retained for the whole three weeks.

### 3.3 Ego-alter activities characteristics

For Hypothesis 2, the confidence ellipse was calculated to observe the growth of the activity area. However, as seen from Figure 4 and 5, the activity area still keeps on growing and as such Hypothesis 2 could not be validated.

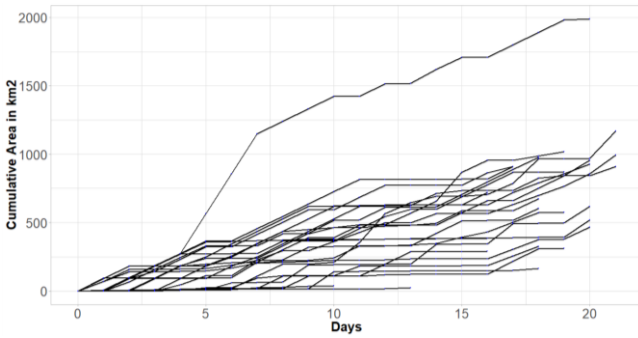


Figure 4: Cumulative growth in activity space

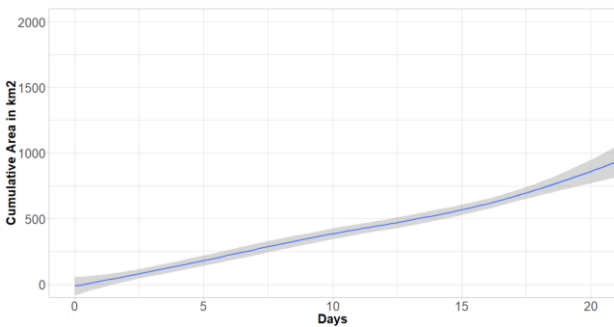


Figure 5: Mean of activity area growth

## 4. Model Formulation and Estimation Procedure

### 4.1 Introduction

The objective of this chapter is to model the retention of friends and to measure what factors play a role in the formation of tie, their retention or the loss of ties for the specific egos. The panel data probit model was built on the 28 egos and their 713 alters, as it accounts for the panel group structure which is multiple observations of the same ego with different alters.

The dependent variable being whether the tie is retained (1) or not (0), the independent variables being ego-alter characteristics or homophily, frequency of activities conducted by the ego with the alters (location based), and the mean of the ratings (0-5) they have rated

to the activities done with the alters at that particular place.

The estimation is done by two model pooled probit and random effect probit to check and understand which model accounts for the panel group structure.

The model specification used is as follows:

$$y_{ij}^* = \beta X_{ij} + \alpha_i + u_i$$

For individuals  $i = 1, 2, \dots, n$  and for alters  $j = 1, 2, \dots, m$ . The random effect  $\alpha_i$  is assumed to be uncorrelated with  $X_{ij}$ . The  $u_i$  are the i.i.d normal errors. The sign of the latent variable determines the observed binary outcome variable, i.e.  $y_{ij} = 1$  if  $y_{ij}^* > 0$ ;  $y_{ij} = 0$  otherwise.

### 4.2 Estimation results

In this section, the estimation result of the friendship retention model, the discussions and the implications are presented to seek for the answers to the last two research questions framed in Section 1.1.

Out of all the independent variables, the ones which produced rational and satisfactorily significant results, with the backing of intuition, were retained in the final model. Positive coefficients of variables indicate a higher propensity to retain a tie with respect to that variable; negative coefficients show a lower propensity to retain a tie. Table 1 shows the result.

The rho-squared for the pooled model is 0.20 whereas for the random effect model is 0.21, thus suggesting an acceptable goodness-of-fit for both models. Both models also show almost similar significance of estimates, with gender (both female), nationality and activity frequencies showing statistically significant positive effects. However, the sigma in random effect probit model, which accounts for the standard deviation of the individual effects (intercepts), is significant which justifies the use of the panel group structure of similar ego with different alters. Ego-alter being of the same nationality has a higher chance of retaining friendship between them. In case of the gender homophily, both ego and alter being female shows a higher propensity of retaining ties than otherwise. The frequency of activities done with the alters shows positive effect, however, the effect is different for different types of activities.

**Table 1: Estimation of the Friendship Retention model**

		POOLED PROBIT		RANDOM EFFECT PROBIT	
		Estimate	t-value	Estimate	t-value
<b>Intercepts</b>		<b>-1.408</b>	<b>-3.923</b>	<b>-1.296</b>	<b>-3.190</b>
<b>Both female</b>		<b>0.291</b>	<b>2.166**</b>	<b>0.262</b>	<b>1.813 *</b>
<b>Both male</b>		0.094	0.773	0.141	1.009
<b>0-3 years difference</b>		-0.128	-1.049	0.119	-0.898
<b>Same nationality</b>		<b>0.506</b>	<b>3.546***</b>	<b>0.501</b>	<b>3.320***</b>
<b>Home based activity frequency</b>		<b>5.978</b>	<b>4.167***</b>	<b>5.413</b>	<b>3.799***</b>
<b>School based activity frequency</b>		<b>5.191</b>	<b>4.793***</b>	<b>5.133</b>	<b>4.505***</b>
<b>Public-space based activity frequency</b>		<b>3.516</b>	<b>5.893***</b>	<b>3.548</b>	<b>5.840***</b>
<b>Experience rating</b>		<b>0.142</b>	<b>1.762*</b>	0.122	1.358
<b>Sigma</b>		-		0.489	<b>3.851***</b>
<b>Groups</b>		LL(initial)	<b>-475.493</b>	<b>-475.493</b>	
<b>Egos</b>	28	LL(maximum)	<b>-381.696</b>	<b>-375.831</b>	
<b>Alters</b>	713	Rho sq.	<b>0.20</b>	<b>0.21</b>	

Significant codes: 0.10\*, 0.05\*\*, 0.01\*\*\*

## 5. Conclusion

### 5.1 Research Questions and their Inferences

This section summarizes the inferences drawn for each of the pre-identified research questions from the analysis of the research data collected.

*1. How does social network and activity pattern change given a life event change (i.e., joining a university)?*

As identified in Section 3.3 and 3.4, social network increases at the start of joining university and later kind of stabilizes. The activity space kept on increasing until the end of the survey and so the final stabilization could not be captured.

*2. What is the effect of relational attributes on tie retention?*

As identified in Section 4.2, nationality, ego-alter being both being female shows a higher propensity to retain ties between them. The positive effect of the same nationality homophily could be explained by the cultural similarity and a feeling of home that might/could be achieved by the egos and the alters with each other. Another point might be the activities conducted by the country specific clubs which play a role in helping the egos in retaining ties with the alters. Also, at the same time, students tend to go for suggestions and help when in a new country from their own country.

*3. What is the effect of joint activity types on tie retention?*

The magnitude of the effects of the activity frequency differed by the location where these activities were conducted, with home-based and school-based activities exhibiting the largest coefficients. This is a rather intuitive finding given that the sample is composed of university students and usually campuses becomes a place for frequency interactions that might reinforce tie strength. In addition, 25 out of the 28 egos surveyed live in university dormitories and are thus being provided with a platform to generate and maintain ties because of the presence of common kitchen and other common facilities such as lounge and laundry.

### 5.2 Key conclusion

This study presents a model that investigates social network dynamics over time and at the same time particularly focuses on what relational attributes and importance of activity incorporates in maintaining and retaining ties in case of new students, triggered by a life event, joining a new university studied for a period of three weeks. The study unveils interesting insights, showing the variables which have a significant impact on tie retention.

The model estimation is based on a sample of individuals who are new to Japan and has just joined The University of Tokyo in the month of April 2019. The first step towards testing and understanding this idea is made through an empirical model that studied what factors/variables (homophily, activity types and location) play a role in maintain and later retaining ties with the alters.

Finally, the activity frequency was positively associated with tie retention, with the magnitude of the effects differing with the location where these activities were conducted.

### 5.3 Weaknesses and strengths of the research

The findings of the study have clearly uncovered a few factors and throws light upon a few limitations. First being not able to capture the home location of the alters compared to the ego's home location as to understand the importance of activities that were carried out. This could have explained a lot the importance of the activities conducted and explained significantly the three activities based on location and their importance. As in the case of home-based activities having a higher importance in the tie retention might be an indication that the ego and the alters probably were the residents in the same place, such as residents in the same dormitories which could further explain the frequency of various activities done based on place.

Second, the data is collected from the point of view of the egos, and the whole study is relied on information provided by one of the dancers, if we use the metaphor "it takes two to tango". This creates significant scope for heterogeneity at the ego-alter level as a relationship might have ended not because the ego decided to interrupt it, but because the alter might no longer be interested in it.

Also, the sample size was small and the observation period being short, the stabilization could have been better observed.

In spite of the above-stated weaknesses, the method of data collection in this study have clearly bridged the gap of conducting study in the evolution of social network starting from scratch and conducting a panel study in real time. As such this survey can be further refined to collect a more detailed out survey keeping in mind the wants and needs of the researchers. Also, this studied allowed us to observe the whole change in social network from formation to maintenance to retention on a daily basis and also the role of activities and its location.

### 5.4 Way forward

If time and budget permit, the survey could be carried out for a longer time frame thus promising a better-quality data for panel study. Also, an ego level heterogeneity study could be carried out by adding some relevant questions in the Friendship diary and also, the survey burden could be reduced.

### Acknowledgements

I would like to thank the respondents who took part

in the survey. It is for them this study was possible.

### References

- 1) Borgers, A., Hofman, F., & Timmermans, H. (2002). Conditional choice modelling of time allocation among spouses in transport settings. *European Journal of Transport and Infrastructure Research*, 5-17
- 2) Wasserman, S., & Faust, K. (1994). *Social Network Analysis*. New York: Cambridge University Press.
- 3) Calastri, C., Hess, S., Daly, A., & Carrasco, J. (2018). Modelling the loss and retention of contacts in social networks: Therole of dyad-level heterogeneity and tie strength. *Journal of Choice Modelling*, 29, 63-77.
- 4) Scott, J. (2013). *Social Network Analysis*. Washington DC: Sage.
- 5) Sharmeen, F. (2014). An analysis of the dynamics of activity and travel needs in response to social network evolution and life-cycle events:a structural equation model. . *Transport. Res. Pol. Pract.*, 159-171.
- 6) Axhausen, K. (2008). Social networks, mobility biographies, and travel: Survey challenges. *Environment and Planning B*, 35, 981-997.
- 7) Bhat, C. (1999). Activity-Based Modeling of Travel Demand. In R. W. Hall, *Transportation Science* (pp. 39-65). Springer.
- 8) Breakwell, G. M., & Wood, P. (1995). Diary techniques. In S. H.-S. G. M. Breakwell, *Research methods in psychology* (pp. 294–302). London: Sage.
- 9) Yang, F. (2004). Contact Diaries: Building Archives of Actual and Comprehensive Personal Networks. *Field Methods*, 19; 194.