Exploring the impact of the social network geography on the individual's activity space using structural equation models

Benjamin Gramsch Calvo*[1] and Kay W. Axhausen[1]

[1]Institute for Transport Planning and Systems, ETH Zurich, Switzerland

*Corresponding author: bgramsch@ethz.ch

Short summary

Most leisure travel has social motivations, one of them is to be in contact with people from their social network, this means that the decision does not only depends on individual preferences and restrictions but also on the other person (or persons) involved in the activity. This means that the places an individual visits for leisure are strongly correlated with the geographic location of their social network. This hypothesis is tested with a structural equation model that includes social needs and mobility demand as latent variables. The model shows a strong correlation between these two variables, showing a positive impact on the geographic distribution of social networks and the number of contacts with the area of leisure activity space, and the number of frequently visited locations. This model shows the social network's importance in individual mobility decisions and patterns.
Keywords: Social Networks, Activity Space, Leisure Travel.

1 Introduction

In travel demand, leisure travel plays an important and often overlooked role, which can negatively impact the functioning of the overall transport system. Also, leisure travel is primarily social travel as a small share of leisure is solitary (Axhausen, 2005) leading to an increasing interest in understanding leisure travel demand and the influence of social needs and activities on travel decisions. Studying social network geographies can help understand how the geographic distribution of social networks impacts daily mobility patterns, opening new perspectives to transport modeling.

Social and leisure travel is more complex than work-related travel as it has many variables that influence it while being flexible in time and space (Ruiz et al., 2016). Therefore, the individual's social network geography is an essential determinant of the social travel decision process (Axhausen, 2006). For this reason, there has been a growing amount of literature on the impact of social networks on individual mobility behavior, as between 10 and 30% of all human travel can be explained by social relations and spatial characteristics of social networks (Cho et al., 2011; Axhausen, 2008). For this reason, transport modelers have started to include social networks in their models (see Axhausen (2005)). One of the essential findings is the negative correlation between the probability of being friends and the geographical distance between two people (Liben-Nowell et al., 2005; Kowald, 2013), as the opportunity cost of the meeting is lower (Arentze et al., 2013). Thus, individuals' social networks are essential to urban mobility and access to the opportunities the city generates (Guidon et al., 2017).

One of the main differences between social travel and other types of travel is the motivation to maintain individual social connections. This motivation directly impacts the process of choosing a leisure activity and destination, as it involves not only personal preferences for the characteristics of the location but the preference and geographical location of the alters. Therefore, people with more extensive social networks tend to have higher heterogeneity in the type of locations visited and to perform more socially motivated travel (Baburajan, 2019). Moore et al. (2013) has studied the link between "longer-term" (social networks) and "shorter-term" (social activities) decisions, showing that the social ties and network density of an individual explain the activity duration and distance.

This paper looks to contribute to understanding how social networks impact individual mobility patterns by analyzing the impact of the geography and structure of social networks on the number of regular leisure locations and the size of the activity space of the ego. This question is relevant

as the alters' home locations can be considered as *anchor points* of mobility in the city, acting as pivotal places in the ego's activity space and the selection of leisure activities. To analyze this hypothesis, we use a structural equation model that includes three latent variables, *social needs*, *relationship strength*, and *mobility demand*. This model shows the relation between the unobservable variables described above by analyzing the covariances of observable variables that are associated with the unobservable variables.
The paper continues as follows: section 2 describes the methodology used in more detail. Section 3 describes the results of the models estimated. Finally, section 4 finishes with some conclusions of the paper.

2 Methodology

The data was collected through a survey conducted in Zurich, Switzerland, that included an egocentric social network and questions on regularly visited leisure venues or locations through a *place generator* and *place interpreter* Gramsch Calvo & Axhausen (n.d.). This information creates the link between the geographic distribution of social networks and the geographic distribution of leisure activities in the city. A structural regression model with three endogenous variables is used to analyze this relationship, following the work by Washington et al. (2003). To estimate the model, we define the latent variables as follows:

- Social needs (exogenous): This variable represents the individual's sociability. It explains the number of contacts in the individual's social network, the size of the social network geography, and how often the individual meets with its alters.
- Relationship strength (endogenous): This variable shows the individual's proximity to their social network. It is measured by the average trust they have in their social network and the average capacity to ask for favors to them.
- Mobility demand (endogenous): This variable shows how much the individual moves and uses the city. It explains the number of leisure locations frequently visited and the size of the leisure activity space.

All these latent variables are constructed with observed variables from the social network structure and distribution of leisure activities. The model solves four equations simultaneously:

$$X_i = \Lambda_x \xi + \delta_i \tag{1}$$

$$Y_{ji} = \Lambda_{y_j} \eta_j + \varepsilon_{ij} \tag{2}$$

$$\eta_1 = \Gamma_1 \xi + \zeta_1 \tag{3}$$

$$\eta_2 = B\eta_1 + \Gamma_2 \xi + \zeta_2 \tag{4}$$

Equation 1 corresponds to the estimation of the dependent variables by the latent dependent variable, and equation 2 is the estimation of the observed dependent variables by the latent dependent variables. Λ_i is the coefficient of X and Y_i or matrix of loadings corresponding to the latent dependent variables, η is the vector of latent dependent variables, and ξ is the vector of latent dependent variables. Finally, equations 3 and 4 are the structural equations, where B and Γ_i are the weights predicted between the independent and dependent variables.
The model used in this paper uses the social network geography as the independent latent variable, which explains the observable variables *number of contacts* and *area of geographical distribution*. The two dependent latent variables are relationship quality and mobility patterns. The first variable explains the observable variables *social capital*, *meeting frequency*, and *trust level*. The second explains *number of places regularly visited* and *area of activity space*. Table 1 explains the type of variables used and their description.
To better explain the relationship between variables in the model, figure 1 shows a path diagram with the variables of interest and their relationship, adapted from the diagrams presented in Lin (2021). The circles represent latent variables, and squares are observable variables; vectors represent the direction of the correlation, single-pointed arrows represent the direction of the effect, and double-pointed arrows represent residuals of the covariance.

Table 1: Description of variables considered in the model

Name of Variable	Mean	SD	Type	Description
# of contacts	11.94	6.13	Ordinal	Number of contacts specified in the *name generator*.
Area of social network	6.30	3.37	Continuous	Logarithm of the area in km^2 of the convex hull generated by the geographical distribution of the social networks' house locations.
Trust level	82.62	19.21	Continuous	Percentage of the ego's alters with whom they discuss important problems or can ask for help.
Social capital	16.61	16.61	Continuous	Average social capital levels on the ego's social network, measured by the number of contacts that the ego would ask for help in different situations.
Meeting frequency	18.03	10.31	Continuous	Total number of time the ego meets with their alters in an average week.
Area of activity space	3.07	2.29	Continuous	Logarithm of the area in km^2 of the convex hull generated by the leisure locations visited by the ego.
# places visited	7.21	4.03	Ordinal	Number of places mentioned by the ego in the *place generator*

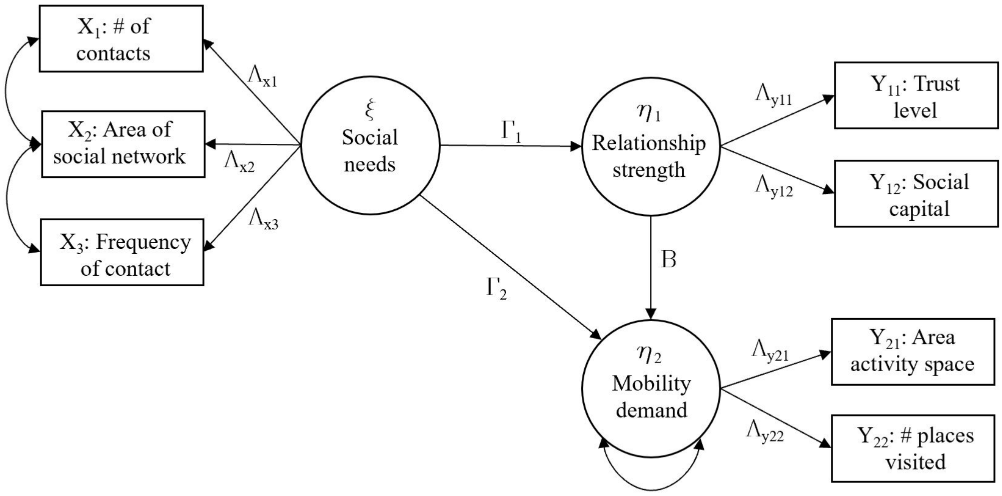

Figure 1: Specification of variables of the model and their effects

3 Results and discussion

To measure the impact of the social needs of an individual on their mobility demand, we have estimated two models following the framework explained in section 2. The first one includes all the variables except the latent variable *relationship strength* and its endogenous variables *trust level* and *social capital*. The second model includes all the variables of the model. Table 2 presents the estimate of the loadings. The latent variable *social needs* influences the three measured variables compared to the number of contacts. The social network area is explained approximately two times more by social needs, while the frequency of contact is explained by a factor of approximately three and a half. Analyzing the *relationship strength*, we can see that both *social capital* and *average trust* is explained similarly by *mobility demand*. *Mobility demand* also has a positive impact on *leisure activity space*, and *number of places visited*, with the former variable explained twice as much as the latter. Both models show similar loading factors with a difference of 0.315 in the impact of *social needs* on *social network area* and 0.228 on *frequency of contact*. The covariances of the model have the expected direction. There is a negative correlation between the *social network area* and the *frequency of contact*, as the farther away the ego's social connections live, the costlier it is to visit them. Therefore there is a lower number of face-to-face meetings. Conversely, the more contacts the ego has, the more frequently they meet with people.
Regarding the model's goodness-of-fit, models 1 and 2 have similar indices. The Comparative Fit Index and the Tucker-Lewis Index are one or higher, and the Root Mean Square Error is close to zero. The only noticeable difference is in the Chi-square, in which model two has a robust value of 8.939 against 1.079.

Table 2: Loadings of the measurement model

A. Estimates of loadings							
		Model 1			Model 2		
Latent variables	Measurement variables	Loadings	Standardized estimates	*p-value*	Loadings	Standardized estimates	*p-value*
Social needs	Number of contacts	1.000	0.826		1.000	0.960	
	Social network area	2.734	2.259	*<0.001*	2.025	1.944	*<0.001*
	Frequency of contact	4.286	3.541	*0.004*	3.451	3.313	*<0.001*
Relationship strength	Social capital	-	-	-	1.000	9.502	
	Average trust	-	-	-	0.978	9.293	*<0.001*
Mobility demand	Leisure activity space	1.000	1.609		1.000	1.614	
	Number of places visited	0.498	0.801	*<0.001*	0.495	0.798	*<0.001*
B. Covariances of measurement error							
Measurement variables			Covariances	*p-value*		Covariances	*p-value*
Social network area	Frequency of contact		-7.415	*0.009*		-5.856	*0.002*
Number of contacts	Frecuency of contact		2.015	*0.100*		1.761	*0.085*
C. Goodness-of-fit							
Name of test			Standard	Robust		Standard	Robust
Chi-square	Test statistic		0.269	1.079		5.155	8.939
	Degrees of freedom		2	2		9	9
	p-value (chi-square)		*0.974*	*0.583*		*0.821*	*0.443*
Goodness of fit index	Comparative Fit Index (CFI)		1.000	1.000		1.000	1.000
	Tucker-Lewis Index (TLI)		1.009	1.006		1.008	1.000
Root Mean Square Error Approximation	RMSEA		0.000	0.000		0.000	0.000
	p-value RMSE		*0.980*	*0.909*		*0.999*	*0.991*

Table 3: Estimates of the structural model

A. Regression weights								
		Model 1			Model 2		
Endogenous variables	Explanatory variables	Estimate	Standardized estimates	*p-value*	Estimate	Standarized estimates	*p-value*
Mobility demand	Social needs	0.589	0.302	*<0.001*	0.526	0.313	*<0.001*
	Relationship strength	-	-	-	0.012	0.068	*0.362*
Relationship strength	Social needs	-	-	-	-4.131	9.502	*<0.001*

Table 3 shows the results of the regression of the latent variables. The difference between the estimations of model 1 and model 2 is non-significant. The latter model standardized estimate shows an increase of 0.09. There is a significant impact of *social needs* in *mobility demand*, but *relationship strength* shows a small non-significant impact. Individuals with more alters tend to visit more places for leisure and to have a higher leisure activity space. However, this mobility demand does not depend on the strength of the relationship with those alters. Finally, the more social needs a person has, the lower the ego's average *relationship strength*; this is explained because individuals with more contacts tend to have proportionally more peripheral alters, reducing the average social capital and trust level of the network.

4 Conclusions

Leisure travel is highly influenced by the individual's friends, family, and acquaintances. As most leisure travel has a social motivation associated, the places a person regularly visits are directly affected by where the individual's social network lives. This paper tries to understand that correlation by comparing the geographic size of the social network with the leisure activity space. The results show a correlation between these two variables. The latent variable *social needs* explains the number of contacts in the ego's social network and their distribution in space. At the same time, it also significantly impacts *mobility demand*. Also, we can see that *relationship strength* does not impact mobility demand.

The results of this paper contribute to the literature on social networks and mobility, as it links the number of friends and their geographic distribution with the number of leisure locations a person visits and their leisure activity space. These two variables of mobility demand can help explain many other travel decisions, such as mode choice and mobility tool ownership. These variables will be included in future work related to social networks and leisure activity spaces.

References

Arentze, T. A., Kowald, M., & Axhausen, K. W. (2013). An agent-based random-utility-maximization model to generate social networks with transitivity in geographic space. *Social Networks*, *35*(3), 451–459. doi: 10.1016/j.socnet.2013.05.002

Axhausen, K. W. (2005). Social networks and travel: Some hypotheses.. doi: 10.3929/ETHZ-A-004663201

Axhausen, K. W. (2006). Social factors in future travel: A qualitative assessment. *Research Collection, ETH Zurich*, 29.

Axhausen, K. W. (2008). Social networks, mobility biographies, and travel: Survey challenges. *Environment and Planning B: Planning and Design*, *35*(6), 981–996. doi: 10.1068/b3316t

Baburajan, V. (2019). An investigation into the influence of land-use, social networks and information and communication technologies on destination choice for social activities. *Transactions of the Association of European Schools of Planning*, 14.

Cho, E., Myers, S. A., & Leskovec, J. (2011). Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '11* (p. 1082). San Diego, California, USA: ACM Press. doi: 10.1145/2020408.2020579

Gramsch Calvo, B., & Axhausen, K. W. (n.d.). Place generator & place interpreter: A new survey method to understand regular destination choice. *Research Collection, ETH Zurich*. doi: 10.3929/ETHZ-B-000561743

Guidon, S., Wicki, M., Axhausen, K. W., & Bernauer, T. (2017). Investigating the relationship between individuals' social networks, mobility behaviour, and social capital: Survey method and first results. *Research Collection, ETH Zurich*.

Kowald, M. (2013). Focussing on leisure travel: The link between spatial mobility, leisure acquaintances and social interactions. *PhD Thesis, ETH Zurich*.

Liben-Nowell, D., Novak, J., Kumar, R., & Raghavan, P. (2005, June). Geographic routing in social networks. , *102*(33), 11623 – 11628. doi: https://doi.org/10.1073/pnas.0503018102

Lin, J. (2021). *Introduction to structural equation modeling (SEM) in R with lavaan.* Los Angeles, California,.

Moore, J., Carrasco, J.-A., & Tudela, A. (2013). Exploring the links between personal networks, time use, and the spatial distribution of social contacts. *Transportation*, *40*(4), 773–788. doi: 10.1007/s11116-013-9467-4

Ruiz, T., Mars, L., Arroyo, R., & Serna, A. (2016). Social networks, big data and transport planning. *Transportation Research Procedia*, 7.

Washington, S., Karlaftis, M. G., & Mannering, F. L. (2003). *Statistical and econometric methods for transportation data analysis.* Boca Raton: Chapman & Hall/CRC.