Relevant Literature for Examining Social Networks in Geographic Space

Clio Andris

SFI WORKING PAPER: 2014-04-010

Relevant Literature for Examining Social Networks in Geographic Space

Clio Andris
Santa Fe Institute
1399 Hyde Park Rd.
Santa Fe, NM 87501
clio@santafe.edu

Abstract

Although interpersonal relationships (modeled as social networks) form, persist and dissolve within the provisions and confines of geographic space, they are rarely modeled within Geographic Information System (GIS) environment. This disconnect may be a result of a lack of research questions that ask about how social life is intertwined with the built environment, or lack of integrated modeling environments.

Instead, researchers focus on either social discoveries (such as, the number of contacts an average agent has) or spatial discoveries (such as, the venues where agents are located on a Saturday night), though it beneficial to ask questions that combine each variable (does one's choice of locale on a Saturday night depend on how many contacts one has?) in order to learn more about socialization, human behavior and quality of life in geography.

Many fields of research play a role in the combination of Social Network Analysis and GIScience. These include theory and practice of modeling flows in GIS, how social connectivity data and social agents have been embedded in GISystems in the past (topics: physical networks, flow visualization, social regionalization, agent-based models). We review how complex network methods have been applied to spatial features (topics: social regionalization, transportation networks and studies of distance decay). Finally, we list topics that describe human interpersonal behavior as social systems (social network analysis), within proximal environments (community sociology) and with focus on communication and psychology (studies of interpersonal relationships). These topics are chosen based on their ability to illuminate parts of the study of social systems within geographic space that can show how, where and why we form relationships over geographic space.

Keywords: Geography, GIS, Flows, Connectivity, Personal Relationships, Location-Based Social Networks (LBSN)

1 Introduction

In our daily lives, we seamlessly combine spatial variables (such as traffic) and social network variables (such as a niece's birthday party) to produce decisions ("attend party despite the cost of traffic"). Social and spatial variables can be considered econometric in nature. We naturally forfeit the benefits of geospatial infrastructure and environment for social betterment, and vice versa. In this mindset, one may ask: Do marriages suffer in certain cities because one spouse may not find a fulfilling social network? Is it better to live in a place with a strong Internet backbone for video chatting or to live in a place with many proximal friends who meet face-to-face?

Unlike our minds, our current geographic models do not integrate interpersonal relationships (as modeled by the social network) into geographic space (as modeled by GISystems, henceforth GIS) to show simultaneous social and spatial variables of place and space.

In this vein, spatial modelers do not tend to ask such questions about why humans connect and how these relationships affect place and space. For example: are certain communities at risk of being under or over-connected with other communities? Are there places where people are more prone to connect (e.g. are more social), and why? How much are we influenced by others to use certain geographies? What happens to geographies (such as cities) when external connections become more or less viable? Yet, these types of questions are necessary for the understanding human behavior within the built environment, understanding how to build cities, and fostering communities, social capital and quality of life.

1.1 One dataset, many views

Researchers from diverse fields often study the same events, such as friends going to meet in a city, family calling one another from two separate countries, co-workers emailing one another from their homes. Yet, what is deemed "important" about of each event differs with the researcher's background and research questions. Using the same geolocated telephone call data set, one may be interested in how spouses communicate during the workday, while another asks how the economic region will grow in the future based on commuter flows. In reality, the strength (or weakness) of marriages may a crucial factor in economic growth.

Furthermore, methodologies to findings patterns in such data are diverse. The sociologist or social network analyst may find how "central" each family member is to his or her wider circle of family, using a mathematical formula to calculate one's network "centrality" among a network of a large extended family (Jackson 2008). The political geographer may be interested in the magnitude of telephone flows that connect two countries, as this may be an indicator of political ties, potential remittance flows, and foreign policy considerations. A GIS analyst may be interested in visualizing the magnitude of emails to find trends on who is connecting where (Luo and MacEachren 2014). An expert in agent-based modeling (ABM) may use this information to program agents to exhibit the monthly contact frequencies revealed in the data, so that agents can exhibit "typical" human behavior. A computer scientist may be interested in developing the fastest ways to catalog, retrieve and synthesize a large datasets that holds many cases of connectivity over geographic space (Doytsher et al. 2010).

1.2 Combining social and spatial

Despite efforts to integrate social variables into GIS (Goodchild et al. 2000; Yu and Peuquet 2009; Goodchild and Janelle 2010; Torrens 2010), research findings on interpersonal connections tend to either consider geographic topology or social relationships. As a result of this segregation, we find, for example, that obesity correlates with social network configuration (Christakis and Fowler 2007) and correlates with city form and accessibility (Ketsens et al. 2012, Ingame et al. 2009). Yet, we do not interact with others without geography; these variables should be used concurrently in a united model, to understand obesity. Also, two separate studies reveal the communication and network behavior of college students (Eagle et al. 2009) and how the campus is designed to host social interaction (Sevtsuk et al. 2009). As a further testament of the lack of cross-cutting integration, Sevtsuk et al. and Eagle et al.'s studies were performed on the same college campus during the same timeframe.

Nevertheless, we find a number of successful studies of interpersonal relationships in geographic space. Emch et al. (2012) simultaneously model a social network of individuals and each's position in geographic space to determine that the spatial closeness of two agents, over their levels of interpersonal interaction, is a stronger cause of disease spread. Other successful models of disaggregate (e.g. agent-based) social systems in a geographic setting include those of urban gangs (Hipp et al. 2010; Radil et al. 2010; Papachristos et al. 2013) transportation choice (Frei and Axhausen 2011; Arentze et al. 2012).

1.3 Characteristics of a holistic field

This field should be able to integrate social networks into a GIS environment in order to show human connections over space. It should be able to utilize human behavioral data, such as the choice to ride a subway, in order to reflect the decisions of the individual. The holistic field should also be equipped for studies that can respond to social theory questions.

A holistic view of social and spatial variables must be able to capture how humans use contacts and geography simultaneously (as in Ryley and Zanni 2013). It should leverage information about the magnitudes of connectivity on edges of the networks of places. Heavy flows of people (such as migrants) or information (such as telephone calls) interchange between certain locations, indicate that locations well-connected. Such connections reveal social desire to connect with humans or the built environment at the destination and are often formalized as formulas predicting the decay of connectivity with increased distance between points of interest.

Flow models and distance decay laws are integral to understanding how social choices in geography, but fall short of revealing the role of social variables in spatial decision-making, and vice versa, because they employ aggregate data of flows between cities that remove the individual's path. Thus, agents or unique individual IDs should be used to connect to individual's geographic behavior to his or her social behavior in terms of relationships.

A holistic view of social and spatial variables must also have quantitative information about interpersonal relationships. The interpersonal relationship is often qualitative, private and ill-defined, presenting a challenge for its inclusion as a variable in spatial models. Yet, studies in Social Network Analysis (SNA) and computational social science have been successful at quantifying relationships and the relationship structures of groups, and such numeric representation facilitates the use of the relationship as a quantitative variable.

However useful, SNA simplifies complex relationships to a variable defining the presence or absence of a relationship between two agents. This representation needs to be bolstered by a deeper theoretical understanding about dyadic interpersonal relationships that can describe relationship types, inception and maintenance techniques. Regarding groups, a similar problem of minimization in representation arises, but can be remedied with rich accounts of urban and community dynamics—how humans behave in groups and within institutions. This field is also a natural fit for integration in a GIS because it often focuses on relationships in neighborhoods and cities.

1.4 Topics and Agenda

In this article, we outline bodies of literature that are useful for understanding social systems and social networks within the context of geographic space. We aim to help researchers in both social and spatial fields to avoid assumptions about the alternative variable. For instance, as a computational social scientist, to forgo modeling a social system within a spatial environment would be to assume that agents congregate and form ties without the help of the built environment. As a geographer, to disregard valuable data on (geolocated) social relationships would be to assume that places are not affected by their connections to other places or by their quality of social life. Neither of these assumptions will lead to salient findings about social behavior in geography.

In the following section, we outline bodies of research that can be helpful in this endeavor. A goal for the field of integrated social and spatial research is to use geo-located human behavioral data with GIScience and Social Network Analysis (SNA) techniques in order to respond to questions about social theory. Table 1 summarizes the strengths of each topic listed in the following section. The presence of a number indicates that the topic provides a technological, methodological, and/or theoretical contribution. The number of the value is only to designate a list. Topics are listed, roughly, from focus on the spatial environment (1) to focus on the human (9).

Table 1. Topics of interest and respective characteristics for the spatial analysis of social systems.

Topic	Uses GIS	Uses SNA	Uses Human Behavioral Data	Responds to Social Theory Questions
Physical Networks	1			
Transportation Networks		2	2	
Flow Visualization	3		3	
Social Regionalization	4	4	4	
Agent-Based Models	5		5	
Social Distance Decay		6	6	6
Social Network Analysis		7	7	7
Community Sociology			8	8
Interpersonal Relationships			9	9

Numbers represent the presentation order of the topic.

Each topic informs an integral part of the larger picture of social relationships in geographic space. Notably, some topics are more advanced than others for this goal, and when *combined* with another topic, a more holistic picture of actions in the built environment arises.

For example, research on distance decay (section 2.6) that uses Euclidean distance to signify distance between two people may be improved by accounting for travel time in the built environment with transportation network analysis (section 2.5). Efforts that show movement flows (section 2.2) could be enriched by considering why connections are made with those at the other end of a trajectory (section 2.9). The development of such models would result in enhanced scholarship on social theories, understanding of the human aspect of place and space, urban planning initiatives, and policy in health, education and transportation sectors.

2 Relevant Literature

2.1 Geographic Networks

Geographers traditionally model physical, not social, networks and flows on these networks within a GISystem. These traditional networks include those of roads and water streams (Haggett and Chorley 1969). The routing of goods and commodities has been of interest to GIScientists, and has resulted, for instance, in the widespread use of Esri ArcMap's Network Analyst software (Esri 2013). However, these flows tend to follow road networks, whereas social networks do not. Moreover, the breakthroughs in early computer mapping of migration flow currents (Tobler 1959; 1978), have not found their way into mainstream GIS software. An exception is the work of Sevtsuk and Mekonnen (2012), in building tools to apply network methods to city form. Their research can be used, for instance, to find the building that is most likely to be passed by an automobile in New York City. Mainstream software does provide powerful tools for the analysis of networks applications of non-social flows such as fright, supply and traffic management. These not comprised of individual decisions and so, do not reflect social networks and social decisions. Yet, the basic infrastructure, such as the GIS' ability to read Origin/Destination (OD) matrices and calculate network variables (such as shortest path), should be leveraged for networks that indicate social interaction.

2.1 Transportation Networks

There are few geographic features that are approached with complex network metrics that are also used in social networks. Transportation networks, geographic in nature, are often analyzed with such metrics, but analysis tends to reflect upon the network's structural qualities (Barrat et al. 2004), without regard to geographic location, though in one case, the structure is considered as a factor of local or regional coverage (Derrible and Kennedy 2010). Also the dynamics of flows (such as number of airline passengers) on the network are often disregarded in favor for the analysis of the structure. Barthélemy (2011) provides an excellent review of spatial network types, structures and metrics. We note that although the networks presented in this piece are geographic in nature, there is no mention of integration with GISystems to correlating such networks with other features in the landscape (such as demographic data or natural features) for inquiry and exploration.

On one hand, systems are sometimes described as Small World (Watts and Strogatz 1998) Scale-Free (Barabasi and Albert 1999), or as hub and spoke configuration (O'Kelley 1998). These distinctions are drawn from evidenced by the network's properties, such as its distribution of nodes with certain numbers of neighbors (degrees). In applications of commuters (DeMontis et al. 2010), railways (Sen et al. 2003) and subway systems (Latora and Marchiori 2001) small-world indicators are tested. Also, the nodes of transportation systems are characterized (Derrible 2012) with network analysis measures.

Such studies reveal that travelers can have faster (though longer) traversal distances in small-world road networks, though traffic jams can rapidly form on these structures (Xu and Sui 2007). Also, airline networks show many triangles but maximize movement efficiency (Guimera et al. 2005; Xu and Harriss 2008). Although these transportation analyses are currently not integrated with GISystems, they represent a step towards this possibility since they overtly apply complex network theory to geospatial features (e.g. an airport or train station) or by assessing the national and local places that have transportation networks with different (or similar) features (Xu and Sui 2007**.** The beneficial layering and operational techniques found in GIS would help examine such spatial entities.

2.3 Flow Data Visualization

The visualization of flow data has become a topic of interest for geographers and cartographers interested in spatial interaction. Software projects are dedicated to help users find patterns in flow data: these advancements include Flowmapper software (Tobler 2004), which allow for symbology (such as color and size) to visualize flow data by weight or type; flow mapping (Guo 2010) which allows for unsupervised classification of flows in the vein of (Takatsuka and Gahegan 2002). Recent analyses of migrants in the UK, commutes and holidays in Germany, France and Denmark (Rae 2009; Limtanakool et al. 2007; 2009) leverage visualizations of flows to explain spatial patterns, rather than rely on covariance or eigenvalue statistics of an OD matrix. Others search for new visualization methods in OD matrices, such as representing flows as cells that connect between two spatial grids (Wood et al. 2010) or as 'star' like features where each central node is classified geometrically as in Weighted Radial Variation (WRV) (Andris 2011). While prescriptions for flow software and data structures that can satisfy social network representation have been made (Glennon 2010; Kim et al. 2012) there little is available to analyze, visualize, and program a social network of agents in a GIS setting. One promising example is SNA software studio Gephi (Bastian et al. 20009) which has a number of useful plugins that allow visualized social networks to be explored as georeferenced GIS files. Yet, large scale flow visualization is important, as it provides fast and reliable renderings that give the investigator clues as to which places he or she may want to investigate, and to give the investigator an overview of well-connected and poorly-connected pairs of places.

2.4 Social Regionalization

Social interaction data have also been used to partition geographic regions, such as the urban hierarchy in Montana (Davies 1979), and the persuasions of Connecticut residents towards New York or Boston (Green 1955). Tangentially, the partitioning of a graph of similar voting patterns between U.S. county centroids (Guo 2008) follows work on partitioning space by clustering minimum spanning tress (Assunção et al. 2006). Regions segmented based on similarities (used to form a region) and differences (used to split regions) in surnames (Cheshire et al. 2010), commuters (Nielsen and Hovgensen 2008), currency circulation (Brockmann et al. 2006), and telephone calls (Ratti et al. 2011; Calabrese et al. 2011) have produced new 'social' partitions that map culture hearths and peripheries. This integration teaches us about the nature of socialization and its boundaries in geographic space, and brings us closer to the creation of “functional” regions instead of those that rely on arbitrary political divisions or natural features. These studies illuminate the physical, demographic, infrastructural and political boundaries that serve as obstacles for socialization, and may prevent the spreading of many things, from political preferences to diseases. These methods use aggregate data flows, and can be improved by using individual social networks, which show connections in more detail and sidestep subjective decisions such as how to define a network node (as a city, county, nation, etc.), and issues that arise from representing a distribution of human behavior as a summary statistic.

2.5 Agent-Based Models

When the human is integrated into a spatial model as an individual, not as part of a census, he or she is usually done so as part of an agent-based model (ABM). Agents are programmed to reflect a typical decision-making person through advancements in Computational Social Science (Torrens 2010). Agent-based models can be programmed to enact the observed social behavior “rules” found in ethnographic and data-driven studies. An elderly agent might be programed to e-mail his or her friends four or five times more often than his or her family (Sayago and Blat 2010). Moreover, agents should have specific goals, cultures, customs and plans (Yu and Peuquet 2009) that reflect cultural or social norms. While ABMs represent the individual, and thus, can represent a number of different perspectives in the landscapes (Couclelis 1997), current use of ABM do not often use individuals who are represented as part of a larger social network, which prohibits analysis about social influence, such as the spreading of information. Yet, this field could integrate this characteristic, and should be studied and leveraged because it can show the individual creating, maintaining and dissolving relationships in a social network model as linked to his or her behavior in a GIS model. The disaggregate nature of these types of models give an extra degree of freedom that can be used to show, perhaps whether a disintegrated friendship (as shown in the social network) might correlate with reduced visits to a certain part of a city (through the agent’s behavior in a GIS). Promisingly, the functionality of some software such as NetLogo (Tisue and Wilensky 2004) can successfully integrate a social network with GIS layers and can be explored.

2.6 Social Distance Decay

Researchers show how communication flows and likelihood of friendship between a pair of people decays with the Euclidean distance between their origin and destination, but without account for underlying geography of the pair’s situation, including political, physical and social boundaries.

For instance, the probability of acquaintanceship for two random people drops by approximately 90% for each additional 100 minutes of car travel time between the (random) pair (Lee et al. 2010) On the Twitter network, 80% of friends are separated by over 1,000 kilometers (km), while over 50% of friends on the FourSquare check-in network are separated by less than 1 km (Scellato et al. 2010). Online chat conversations are longer, but less frequent when the pair is distant (Leskovec and Horvitz 2008). Communication probability also decreases with distance in a cell phone network, and friendship probability follows suit in a blog community (Liben-Nowell and Kleinberg 2008). The average duration of phone calls increases with distance, reaching a plateau when partners are 40 km apart (Lambiotte et al. 2008). Online photo sharing data has also been used to show that probability of friendship decreases with distance (Crandall et al. 2010). Others show that travel distances decay as a power law, movement and diffusion exhibits mathematical scaling properties such as Lévy flight behavior (Gonzalez et al. 2008).

These types of analyses are performed by those with high computing power, physicists, computer scientists, which can shape the research questions (boyd and Crawford 2011), and thus, have left little room for integration of the large agent-based networks and GIS. In fact *geography* has been equated with distance, as seen in many titles, including (Barthélemy 2011) who describes "geography in social networks" with a list of distance decay functions. Others name their work "geography" when rendering the Euclidean distance values that exist between every individual pair of geo-located nodes, shown as probability distributions (probability density function) and lauded for their 'fat tails' or log normal behavior, as a physical law (e.g. Scellato et al. 2010). They are not sufficient for geographic analysis because they lack the underlying spatial landscape and topology, city form, accessibility and real-life issues of spatial configuration and geometry (as echoed by Loglisci et al. 2010). Understanding these laws in the context of a GIS can provide more information about how relationships fare given, obstacles to face-to-face meetings, different places from which a member of a relationship hails, or interaction spaces.

2.7 Social Network Analysis

A social network can be thought of as a systematic representation of connections between agents (people), manifested in a graph of nodes their connections. The complex calculations and treatment of these data structures is generally referred to as social network analysis (SNA), a field now considered to be within the domain of Computational Social Science (CSS) (Lazer et al. 2009). SNA is a robust, burgeoning field with dedicated journals and textbooks, but has interacted little with GISystems (Andris 2011).

Most social network analysis produces valuable knowledge about human behavior without using geographic space as a variable in their findings. Social network analysis can human behavior on topics of teamwork and incentives (Kerns et al. 2006), trust propagation (Watts and Dodds 2007), innovation (Bettencourt et al. 2007), Pluralistic Ignorance (Centola and Wilier 2009), the benefit of weak ties (Granovetter 1973) and strong ties (Centola and Macy 2007), and social influence (Salganik and Watts 2008). Some social network analysis claims to predict obesity, depression, anxiety, loneliness, drinking, eating, exercise within contacts (such as Christakis and Fowler 2007), or future friendships (Clauset et al. 2008). Additionally, the public can learn about a new fashion or innovation, but it is only when friends or a critical mass of people display the novelty or fashion that the adoption is realized "complex contagion" (Centola and Macy 2007). These human behaviors and 'rules' for social systems are integral for understanding how decisions such as attending a concert or traveling to a new university, are made in the built environment as based on the influence of one's social network. These are especially important in policy: perhaps members of a social network are susceptible to a certain communicable disease because a friend has been diagnosed. Due to privacy issues, a public health agency could not contact the friends of this agent (as retrieved from phone call records). However, using a GIS, the neighborhoods of the friends can be mapped, and policy can be applied to certain geographic areas—such as townships or school districts.

2.8 Urban Sociology

In urban and community sociology, social networks are considered important parts of a city or neighborhood. Examples of community case studies (Chrisman 1970; Wellman 1979; Fischer 1982) provide expertise fraternization of race and gender and socioeconomic groups in communities. These types of in-depth case studies illuminate valuable detail of understanding nuances and qualitative metrics such as trust and support that one cannot ascertain from big data studies. The factors that lead to growth, development, change, and deterioration of communities is a specialty in this field that is not found in other topics, and should be applied to spatial models. GIS can benefit this field, as many personal social networks are illuminated in one place (such as rural Thailand (Verdery et al. 2012)) and may not be generalizable across communities, unless other communities are found to have the same variables (such as a fishing industry). Studies that generalize across communities show how migration and telecommunication is an economic driver of globalization (Sassen 2000; Massey 2005; Castells 2009; Rojas 2010) but do not incorporate empirical models of physical space in their theoretical contributions to worldwide flow systems. GIS analysts can better understand the flows of remittances from these studies on globalization, and add these variables as informal economic indicators of places.

2.9 Interpersonal Relationships

The study of interpersonal relationships has much to offer GIS analysts, and gain from modeling spatial topology. Though relationships cannot escape the provisions and obstacles posed by the built environment, the field tends to be focused on the relationship itself. Research topics in this robust, interdisciplinary field classify salient types of relationships, and illuminate processes of relationship progression, features of successful marriages (Gottman et al. 2005), and verbal and non-verbal communication mechanisms (Duck et al. 1991). In addition, studies on relationship maintenance and commitment (such as Rusbult et al. 2004) can inform researchers studying massive telephone call or instant message datasets of the role of the telephone in maintain social ties, how frequently these actions should take place, and the benefits and drawbacks. That long distance relationships (LDRs) and derivatives such as long distance romantic relationships (LDRRs) (Stafford 2005) have not embraced as an examinable feature of the landscape is unfortunate. Authors of these studies find, for example, that of LDDR partners that become geographically reunited, one-third end their relationships within 3 months of the reunion (Stafford et al. 2006). Studies of geolocated LDRs can be extremely useful to geographers, especially those interested in connectivity, flows, ICTs, and transportation. Also, the study of Geography is GIS can enrich studies of LDRs by describing the obstacles and travel costs to face-to-face meeting, and enumerating the "competing" options available to each party in their respective social environments (Stouffer 1940). In order to integrate relationships into GISystems, we should look to mathematically-modelled relationships found in Gottman et al. (2005).

Although we find many helpful pieces in the story of how socialization occurs within the context of geographic space a clearer path is needed for comparing one's connectivity behavior to one's social relationships. We have discussed the prospects and drawbacks of certain literatures that pertain to social systems within geography. How these literatures will fit together, complement one another, and be applied to research endeavors is yet to be seen.

3 Discussion and Conclusions

By pulling from the aforementioned fields, researchers may be able to approach a more unified understanding of socialization in geographic space. The above fields each contribute important paradigms, theories, models and terminology necessary for understanding social systems within geography, or as the technical convergence of SNA and GIS. Future social/spatial models will show how people behave in geographic space *due to one another*, and can be both informed by and used to benefit the scholarship of many topics in geography. A small sample includes *chain migration*: how we move to a new place due to our family and friend relationships; *discrete choice models* at the household level: how we decide on where to live and which modes of transportation to use based on our family members; *spatial cognition*: how we are drawn to places where our contacts have live(d), work(ed), or visit(ed).

This review leaves us with a number of issues, namely that the GIS or SNA researcher is still left without a technical “how-to” for social networks in GIS. We believe researchers might benefit from (a) more concrete examples of integration, (b) a taxonomy of different types of spatially-embedded social networks and models (e.g. networks where each node is one city); (c) an outline of best practices, new metrics and methods; (d) the creation of a better terminology for describing the behavior of convergent social and spatial systems (e.g. the term “cluster” is used in both fields to mean different things); (e) an agenda of research questions to test with models of human socialization in geographic space; (f) tips for managing big data in this domain. We hope to address these issues in future work.

Acknowledgements

We thank the Santa Fe Institute, The John Templeton Foundation (Grant 15705), Luis M.A. Bettencourt and Geoffrey West.

References

1. Andris, C. (2011). Methods and metrics for social distance. Unpublished Dissertation, Department of Urban Studies and Planning, Massachusetts Institute of Technology.
2. Arentze T., van den Berg P., & Timmermans H. (2012). Modeling social networks in geographic space: approach and empirical application. *Environment and Planning A*, 44(5), 1101–1120.
3. Assunção, R., Neves, M., Câmara, G., & Freitas, C. (2006). Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. *International Journal of Geographical Information Science*, 20(7), 797-811.
4. Barabasi, A., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509-512.
5. Barrat, A., Barthelemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, 101(11), 3747-3752.
6. Barthélemy, M. (2011). Spatial networks. *Physics Reports*, 499(1), 1-101.
7. Bastian, M.; Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. Presented at the international AAAI conference on weblogs and social media (pp. 361-362). San Jose, CA, 17-20 May, 2009.
8. Bettencourt, L., Lobo, J., & Strumsky, D. (2007). Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size. *Research Policy*, 36(1), 107-120.
9. boyd, d., & Crawford, L. (2011). Six Provocations for Big Data. A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society. Oxford U.K., Oxford Internet Institute. 21 Sep, 2011.
10. Brockmann, D., Hufnagel, L., & Geisel, T. (2006). The scaling laws of human travel. *Nature*, 439(7075), 462-465.
11. Calabrese, F., Dahlem, D., Gerber, A., Paul, D., Chen, X., Rowland, J., Rath, C., & Ratti, C. (2011). The Connected States of America: Quantifying Social Radii of Influence. The Third IEEE International Conference on Social Computing (SocialCom2011). Boston, MA, USA, 9-

11 Oct., 2011.

12. Castells, M. (2009). *Communication power*. Oxford: Oxford University Press.
13. Centola, D., & Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3), 702-734.
14. Centola, D., Willer, R., & Macy, M. (2005). The Emperor's Dilemma: a computational model of self-enforcing norms. *American Journal of Sociology*, 110(4), 1009-1040.
15. Cheshire, J. A., Longley, P. A., & Singleton, A. D. (2010). The surname regions of Great Britain. *Journal of Maps*, 6(1), 401-409.
16. Chrisman, N. (1970). Situation and social network in cities. *Journal of Canadian Review of Sociology and Anthropology*, 7(4), 245-257.
17. Christakis, N., & Fowler, J. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, 357(4), 370-379.
18. Clauset, A., Moore, C., & Newman, M. (2008). Hierarchical structure and the prediction of missing links in networks. *Nature*, 453(7191), 98-101.
19. Couclelis, H. (1997). From cellular automata to urban models: new principles for model development and implementation. *Environment and Planning B,* 24(2), 165-174.
20. Crandall, D., Backstrom, L., Cosley, D., Suri, S., Huttenlocher, D., & Kleinberg, J. (2010). Inferring social ties from geographic coincidences. *Proceedings of the National Academy of Sciences*, 107(52), 22436-22441.
21. Davies, W. (1979). Urban connectivity in Montana. *The Annals of Regional Science*, 13(2), 29-46.
22. De Montis, A., Chessa, A., Campagna, M., Caschili, S., & Deplano, G. (2010). Modeling commuting systems through a complex network analysis: A study of the Italian islands of Sardinia and Sicily. *Journal of Transport and Land Use*, 2(3), 39-55.
23. Derrible, S., & Kennedy, C. (2010). Characterizing metro networks: state, form, and structure. *Transportation*, 37(2), 275-297.
24. Derrible, S. (2012). Network centrality of metro systems. *PloS one*, 7(7), e40575.
25. Doytsher, Y., Galon, B., & Kanza, Y. (2010). Querying geo-social data by bridging spatial networks and social networks. In: Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 39-46). San Jose, CA, 3 -5 Nov, 2010.
26. Duck, S., Rutt, D. J., Hoy, M., & Strejc, H. H. (1991). Some evident truths about conversations in everyday relationships all communications are not created equal. *Human Communication Research*, 18(2), 228-267.
27. Eagle, N., Pentland, A., & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences,* 106(36), 15274
28. Emch, M., Root, E., Giebultowicz, S., Ali, M., Perez-Heydrich, C., & Yunus, M. (2012). Integration of spatial and social network analysis in disease transmission studies. *Annals of the Association of American Geographers*, 102(5), 1004-1015.
29. Environmental Systems Research Institute (Esri), (2013). ArcGIS 10.2 Extension: Network Analyst. Redlands, CA.
30. Fischer, C. (1982). *To dwell among friends: personal networks in town and city*. Chicago: University of Chicago Press.
31. Frei, A., & Axhausen, K.W. (2011). Modeling spatial embedded social networks, Working Paper Transport and Spatial Planning, 685, IVT, ETH Zurich, Zurich.
32. Glennon, A. (2010). Creating and validating object-oriented geographic data models: Modeling flow within GIS. *Transactions in GIS*, 14(1), 23-42.
33. González, M., Hidalgo, C., & Barabási, A. (2008). Understanding individual human mobility patterns. *Nature*, 453(7196), 779-782
34. Goodchild, M., & Janelle, D. (2010). Toward critical spatial thinking in the social sciences and

humanities. *GeoJournal*, 75(2), 3-13.
35. Goodchild, M., Anselin, L., Appelbaum, R., & Harthorn, B. (2000). Toward spatially-integrated social science. *International Regional Science Review*, 23(2), 139-159.
36. Gottman, J., Murrey, J., Swanson, C., Tyson, R., & Swanson, K. (2005). *The mathematics of marriage: dynamic nonlinear models*. Cambridge, MA: The MIT Press.
37. Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6),1360-1380.
38. Green, H. L. (1955). Hinterland boundaries of New York City and Boston in Southern New England. *Economic Geography*, 31(4), 283-300.
39. Guimera, R., Mossa, S., Turtschi, A., & Amaral, L. (2005). The world-wide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proceedings of the National Academy of Sciences*, 102(22), 7794-7799.
40. Guo, D. (2008). Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP). *International Journal of Geographical Information Science*, 22(7), 801-823.
41. Guo, D. (2010). Flow mapping and multivariate visualization of large spatial interaction data. *IEEE Transactions on Visualization and Computer Graphics*, 15(6), 1041-1048
42. Haggett, P., & Chorley, R. (1969). *Network analysis in geography*. New York: Edward Arnold.
43. Hipp, J. R., Tita, G., & Greenbaum, R. (2009). Drive-bys and Trade-ups: The impact of crime on residential mobility patterns in Los Angeles. *Social Forces*, 87(4), 1777-1812.
44. Inagami, S., Cohen, D. A., Brown, A. F., & Asch, S. M. (2009). Body mass index, neighborhood fast food and restaurant concentration, and car ownership. *Journal of Urban Health,* 86(5), 683-695.
45. Jackson, M. O. (2008). *Social and economic networks*. Princeton, NJ: Princeton University Press.
46. Kearns, M., Suri, S., & Montfort, N. (2006). An experimental study of the coloring problem on human subject networks. *Science*, 313(5788), 824-826.
47. Kestens, Y., Lebel, A., Chaix, B., Clary, C., Daniel, M., Pampalon, R., & Subramanian, S. V. (2012). Association between activity space exposure to food establishments and individual risk of overweight. *PloS one*, 7(8), e41418.
48. Kim, K., Lee, S. I., Shin, J., & Choi, E. (2012). Developing a flow mapping module in a GIS environment. *The Cartographic Journal*, 49(2), 164-175.
49. Lambiotte, R., Blondel, V., de Kerchove, C., Huens, E., Prieur, C., Smoreda, Z., & Van Dooren, P. (2008). Geographical dispersal of mobile communication networks. *Physica A,* 387(21), 5317-5325.
50. Latora, V., & Marchiori, M. (2002). Is the Boston subway a small-world network? *Physica A*, 314(1), 109-113.
51. Lazer, D., et al. (2009). Life in the network: the coming age of Computational Social Science. *Science*, 323(5915), 721–723.
52. Lee, C.; Scherngell, T., & Barber, M. (2010). Investigating an online social network using spatial interaction models. ArXiv preprint: arXiv:0911.1229.
53. Leskovec, J., & Horvitz, E. (2008). Planetary-scale views on a large instant-messaging network. In: Proceedings of the 17th International Conference on the World Wide Web (pp. 915-924). Beijing, 21-25 Apr, 2008.
54. Liben-Nowell, & J. Kleinberg (2008). Tracing information flow on a global scale using internet chain letter data. *Proceedings of the National Academy of Sciences*, 105(12), 4633-4638.
55. Limtanakool, N., Schwanen, T., & Dijst, M. (2009). Developments in the Dutch urban system on the basis of flows. *Regional Studies*, 43(2), 179-196.
56. Limtanakool, N.; Dijst, M., & Schwanen, T. (2007). A theoretical framework and methodology for characterising national urban systems on the basis of flows of people: empirical evidence

for France and Germany. *Urban Studies*, 44(11), 2123-2145.

57. Loglisci, C., Ceci, M., & Malerba, D. (2010). Relational learning of disjunctive patterns in spatial networks. In: Proceedings on the first workshop on dynamic networks and knowledge discovery (pp. 17-28). Barcelona, 24 Sep, 2010.
58. Longley, P. A., Cheshire, J. A., & Mateos, P. (2011). Creating a regional geography of Britain through the spatial analysis of surnames. *Geoforum*, 42(4), 506-516.
59. Luo, W. and MacEachren, A. M. (2014). Geo-social visual analytics. *Journal of Spatial Information Science*, Accepted-typesetting.
60. Massey, D. (2005). *For space.* London: Sage Publications Ltd.
61. Nielsen, T. A. S., & Hovgesen, H. H. (2008). Exploratory mapping of commuter flows in England and Wales. *Journal of Transport Geography*, 16(2), 90-99.
62. O'Kelly, M. (1998). A geographer's analysis of hub-and-spoke networks. *Journal of Transport Geography*, 6(3), 171-186.
63. Papachristos, A., Hureau, D., & Braga, A. (2013). The corner and the crew: The influence of geography and social networks on gang violence. *American Sociological Review*, 78(3), 417–447.
64. Radil, S., Flint, C., & Tita, G. (2010). Spatializing social networks: Using social network analysis to investigate geographies of gang rivalry, territoriality, and violence in Los Angeles. *Annals of the American Association of Geographers*, 100(2), 307-326.
65. Rae, A. (2009). From spatial interaction data to spatial interaction information? Geovisualisation and spatial structures of migration from the 2001 UK census. *Computers, Environment and Urban Systems*, 33(3), 161–178.
66. Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., Claxton, R., & Strogatz, S. (2011). Redrawing the map of Great Britain from a network of human interactions. PLoS one, 5(12), e14248.
67. Rojas, F. (2010). New York Talk Exchange: transnational telecommunications and migration in a global city. Unpublished Dissertation, Department of Urban Studies and Planning, Massachusetts Institute of Technology.
68. Rusbult, C. E., Olsen, N., Davis, J. L., & Hannon, P. A. (2004). Commitment and relationship maintenance mechanisms. In: H. Reis, & C. Rusbult (Eds.), *Close Relationships* (pp. 287-303). New York: Psychology Press.
69. Ryley, T. J., & Zanni, A. M. (2013). An examination of the relationship between social interactions and travel uncertainty. *Journal of Transport Geography*, 31, 249-257.
70. Salganik, M., & Watts, D. (2008). Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market. *Social Psychology Quarterly*, 71(4), 338-355.
71. Sassen, S. (2000). *Cities in a world economy*. New York: Sage Publications.
72. Sayago, S., & Blat, J. (2010). Telling the story of older people e-mailing: An ethnographical study. *International Journal of Human-Computer Studies*, 68(1), 105-120.
73. Scellato, S., Mascolo, C., Musolesi, M., & Latora, V. (2010). Distance matters: Geo-social metrics for online social networks. Presented at the 3rd conference on online social networks, Boston, MA, 20 Jun, 2010.
74. Sen, P., Dasgupta, S., Chatterjee, A., Sreeram, P., Mukherjee, G., & Manna, S. (2003). Small-world properties of the Indian railway network. *Physical Review E*, 67(3), 036106.
75. Sevtsuk A., & Mekonnen, M. (2012). Urban network analysis: a new toolbox for ArcGIS. *Revue Internationale de Géomatique*, 22(2), 287-305.
76. Sevtsuk, A., Huang, S., Calabrese, F., & Ratti, C. (2009). Mapping the MIT campus in real time using WiFi. In: M. Foth (Ed.), *Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City* (pp. 326–338). Hershey, PA: IGI Global.
77. Stafford, L. (2005). *Maintaining long-distance and cross-residential relationships.* Mahwah,

NJ: Lawrence Erlbaum.

78. Stafford, L., Merolla, A. J. &, Castle, J. D. (2006). When long-distance dating partners become geographically close. *Journal of Social and Personal Relationships*, 23(6), 901-919.
79. Stouffer, S. (1940). Intervening opportunities: a theory relating mobility and distance. *American Sociological Review*, 5(6), 845-867.
80. Takatsuka, M., & Gahegan, M. (2002). GeoVISTA Studio: A codeless visual programming environment for geoscientific data analysis and visualization. *Computers and Geosciences*, 28(10), 1131-1144.
81. Tisue, S., and Wilensky, U. (2004). Netlogo: A simple environment for modeling complexity. Presented at the international conference on complex systems, Boston, MA, 16-21 May, 2004.
82. Tobler, W. (1959). Automation and cartography. Geographical Review, 49(4). 526-534.
83. Tobler, W. (1978). Migration fields. In: W. A. W. Clarke, & G. Moore (Eds.), *Studies in Geography no 25: Population mobility and residential change* (pp. 214-232). Evanston, IL: Northwestern University.
84. Tobler, W. (2004). Movement mapping. *Center for Spatially Integrated Social Science*, Regents of University of California at Santa Barbara.
85. Torrens P. (2010). Geography and Computational Social Science. *GeoJournal*, 75(2), 133-148.
86. Verdery, A. M., Entwisle, B., Faust, K., & Rindfuss, R. R. (2012). Social and spatial networks: kinship distance and dwelling unit proximity in rural Thailand. *Social Networks*, 34(1), 112-127.
87. Watts, D., & Dodds, P. (2007). Influentials, networks, and public opinion formation. *Journal of Consumer Research*, 34(4), 441-458.
88. Watts, D. J, & Strogatz, S.H. (1998). Collective dynamics of "small-world" networks. *Nature*, 393(6684), 440-442.
89. Wellman, B. (1979). The community question: the intimate networks of East Yorkers. *The American Journal of Sociology,* 84(5), 1201-1231.
90. Wood, J., Dykes, J., & Slingsby, A. (2010). Visualisation of origins, destinations and flows with OD maps. *The Cartographic Journal*, 47(2), 117-129.
91. Xu Z., & Harriss, R. (2008). Exploring the structure of the U.S. intercity passenger air transportation network: A weighted complex network approach. *GeoJournal*, 73(2), 87-102.
92. Xu, Z., & Sui, D. (2007). Small-world characteristics on transportation networks: a perspective from network autocorrelation. *Journal of Geographical Systems,* 9(2), 189-205.
93. Yu, C., & Peuquet, D. (2009). A geoagent-based framework for knowledge-oriented representation: embracing social rules in GIS. *International Journal of Geographical Information Science*, 23(7), 923-960.