Far East Journal of Applied Mathematics
Volume 39, Number 1, 2010, Pages 53-69
Published Online: March 9, 2010
This paper is available online at http://www.pphmj.com

CALCULATING EDGE LENGTH BIAS AND THE RELATIVE SIZE OF A 'NEIGHBORHOOD' IN SPATIAL SOCIAL NETWORKS

JAN RYCHTÁŘ, ERIC C. JONES, RICK L. BUNCH and ROBERT P. GOVE

Department of Mathematics and Statistics
University of North Carolina at Greensboro
Greensboro, NC 27412, U. S. A.
e-mail: j_rychta@uncg.edu

Department of Anthropology
University of North Carolina at Greensboro
Greensboro, NC 27412, U. S. A.

Department of Geography
University of North Carolina at Greensboro
Greensboro, NC 27412, U. S. A.

Department of Computer Science
University of Maryland
College Park, MD 20742, U. S. A.

Abstract

Many spatial social networks have the property that nearby nodes are more likely to be connected than are nodes that are farther apart. We develop a characteristic of spatial graphs that captures whether or not shorter distance ties are preferred over longer distance ties, and the degree to which this edge length bias occurs. This allows us to estimate what is far and what is close – what we call neighborhood radius – for any randomly generated spatial network. Results from Monte Carlo Markov

2010 Mathematics Subject Classification: 91D30.

Keywords and phrases: spatial model, community structure, GIS, geographic distance, personal networks.

Received November 28, 2009

Chain simulations presented similar distribution of edge length bias to data from personal networks from a neighborhood in New Orleans, Post-Katrina, although the latter presented greater variation.

1. Introduction

Interest in combining spatial analysis with network analysis has grown in part due to increased sophistication and affordability of both computer hardware and software. More importantly, the interest also stems from the potential applications of understanding the spatial structure of human relationships. Theoretically, such a marriage provides the opportunity to answer variants of the question "Does distance matter?" (Mok et al. [11]). Mok et al. [11] found specific drops in frequency of face-to-face contact between ego and alters in a network as distance increased between ego and each of those alters (for likelihood of interaction across space, see also Axhausen et al. [1], Carrasco et al. [3], Coburn and Russell [4], Porta et al. [14]). In an economic geography study, Pitts [13] found that trade networks over space can behave as an aggregate of actual geographic path distances, and Jones [9] found that village network structural measures vary with location in a regional economic system. Other studies about the relationship of social networks to geography have included likelihood of political activities (Johnston et al. [8], Nicholls [12]), the flow/exchange of agricultural resources across space based on variation in agro-ecological activities (Faust et al. [7], Zimmerer [18]), and effects of spatial aspects of social networks on health-related outcomes (e.g., Bates et al. [2], Cravey et al. [6], Rothenberg et al. [15], Wylie et al. [17]). However, a body of literature has yet to emerge that analyzes spatial social networks in a general sense.

In order to control for homophily in social networks due to geographic closeness, Wong et al. [16] developed an elegant model that yielded a random graph embedded in geographic space with the property that connections between nearby nodes are more likely to occur than connections between nodes that are further apart. They have also shown that such graphs share many characteristics of a social network when the 'neighborhood radius' – the cutpoint between far and near c is sufficiently large; and they also found that social networks have characteristics similar to graphs generated by their method.

In this paper, we study the inverse problem, i.e., given a spatial social network, we will provide a reliable estimate of the spatial length cutoff after which the probability of a connection between two nodes decreases significantly. One major application that would derive from solving this problem is being able to say, for any given social network, how far is far?, i.e., what length delineates between far and near vertices in the related geographic space. Our purpose is to evaluate the utility of the edge length bias algorithm and understand the distribution across networks of whether or not people in any ego's personal network in this post-disaster setting tend to interact with members of that personal network who live nearer to them or farther from them.

2. Spatial Social Network Model

In this section, we summarize the main ideas of the model given by Wong et al. [16]. A social network will be represented by a non-directed graph, i.e., a pair (V, E), where $V = \{v_1, v_2, \ldots, v_N\}$ is the set of all nodes representing the individuals in the network and $E = \{e_1, e_2, \ldots, e_L\}$ is the set of edges representing the social connections between individuals. The (random) graph is generated from the parameters $N, H, p, p_b,$ and a technical parameter Δ – which can be calculated from the previous ones – where:

- N is the number of vertices of a graph,
- H is the neighborhood radius, i.e., a number that indicates what does it mean to be nearby (if the geographical distance between two nodes is smaller than H, then they are considered nearby),
- p is the probability that any two nodes are connected,
- p_b is the proximity bias that indicates the increase in probability that two nearby nodes are connected,
- Δ indicates the decrease in probability that two far apart nodes are connected.

To construct the graph from the above parameters, the nodes $V = \{v_1, v_2, \ldots, v_N\}$ of the graph are embedded in the plane with a distance function d. We assume that the points are distributed in space according to a homogeneous Poisson point process (see Cox and Isham [5]) with a parameter ρ (as shown in Wong et al. [16], parameters ρ and H are conjugate to each other and it is therefore enough to consider only one of them).

Once the nodes are embedded into the space, an edge between two nodes v_i, v_j is generated with probability $f(d(v_i, v_j))$, where $d(v_i, v_j)$ denotes the geographical distance between embedded nodes v_i, v_j and f is a function defined by

$$f(x) = \begin{cases} p + p_b, & (x \le H), \\ p - \Delta, & (x > H). \end{cases} \quad (1)$$

Wong at al. [16] have shown that graphs generated in this way share many common characteristics with social networks (when H is sufficiently large; in their case 3/2) and vice versa, that spatial social networks can often be seen as a result of the above procedure. Figure 1(a) shows an example of a spatial random graph generated by the above procedure.

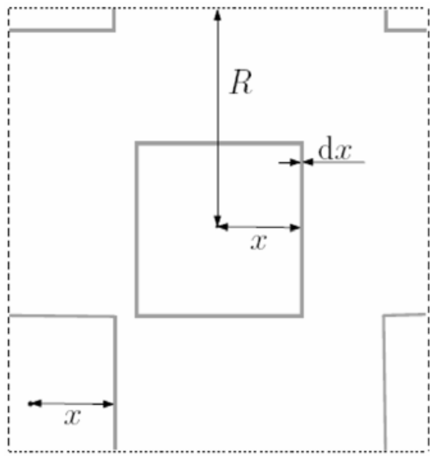

Figure 1. Periodic square with size $2R \times 2R$. The dotted and the dashed lines are to be identified. The gray square in the middle represents all points with the distance $d \in [x, x + dx]$. The gray lines in the corners represent the same square shifted towards boundaries. When dx is small, the gray area of the circumference of one of the squares is $8xdx$.

3. Inverse Problem-determining Network Parameters from the Network

In the previous section, we described how to construct a graph from the four parameters N, H, p, p_b. The main aim of this paper is to provide a procedure to reverse the process, i.e., for an empirical spatial social network, we want to determine parameters N, H, p, p_b.

The estimate of N is simple, it is exactly the count of the nodes of the network. The best estimate of p is the Maximum Likelihood Estimate given by the edge density, i.e., the total number of edges L divided by the number of potential connections $N(N-1)/2$. Similarly, if H is known, then the best estimate for p_b would be based on the density of the short edges (edges shorter than H). Consequently, what remains, is the nontrivial estimation of H, which is the neighborhood radius based on all connections in the network.

Wong et al. [16] estimated H by searching for the step function f from (1) as the least square error fit to the edge length distribution function. In the following sections, we present another, presumably simpler way to estimate H.

4. Edge Length Bias

In this section, we introduce a characteristic β (we refer to as edge length bias) of a spatial graph and also calculate β in full generality for graphs that were generated by the model in Section 2. The characteristic β should capture the bias (towards a particular length of an edge) and is defined as a ratio of the average edge length to the average (geographical) distance between any two nodes. Mathematically,

$$\beta = \frac{\dfrac{\sum_{i=1}^{N}\sum_{j=i+1}^{N} x_{ij} d(v_i, v_j)}{\sum_{i=1}^{N}\sum_{j=i+1}^{N} x_{ij}}}{\dfrac{\sum_{i=1}^{N}\sum_{j=i+1}^{N} d(v_i, v_j)}{\sum_{i=1}^{N}\sum_{j=i+1}^{N} 1}}, \tag{2}$$

where $x_{ij} = 1$ if nodes v_i and v_j are connected and $x_{ij} = 0$, otherwise.

The formula (2) is simple and straightforward to program in a computer and thus β can be calculated for any empirical spatial graph.

We develop the alternative formula for β under the following assumptions on the graph and its spatial embedding:

1. All nodes lie in the square with dimensions $2R \times 2R$.

2. The square has periodic boundaries (i.e., the geographical space looks like a doughnut) and distance between two points with coordinates (x_1, y_1) and (x_2, y_2) is given by

$$d = d_{\infty, per} = \max\{P_x, P_y\},$$

where

$$P_x = \min\{|x_1 - x_2|, |x_1 - x_2 + 2R|, |x_1 - x_2 - 2R|\}$$

and

$$P_y = \min\{|y_1 - y_2|, |y_1 - y_2 + 2R|, |y_1 - y_2 - 2R|\}.$$

Consequently, any point can be considered as a geographical center of the square.

3. The nodes are distributed by a Poisson point process with parameter ρ; in particular, any region of an area A contains on average $A\rho$ points.

4. Any two nodes v_i, v_j are connected with probability $f(d(v_i, v_j))$, where f is a function with values between 0 and 1.

The situation is depicted in Figure 2. The parameter R plays a role of a “radius of visibility” − for any particular individual, “nothing” really exists farther than the distance R from an individual. This assumption does not unreasonably bound or limit the space for each graph because all finite networks lie in a finite geographical area. The use of periodic boundaries and non-Euclidean metric may seem highly unrealistic, but it is only a technical assumption that allows much easier analytical calculation. The computer simulations presented in the Results section show that the characteristic β is robust and practically independent of whether the region has periodic boundaries or not and even what type of the metric is used to measure the distances between the nodes.

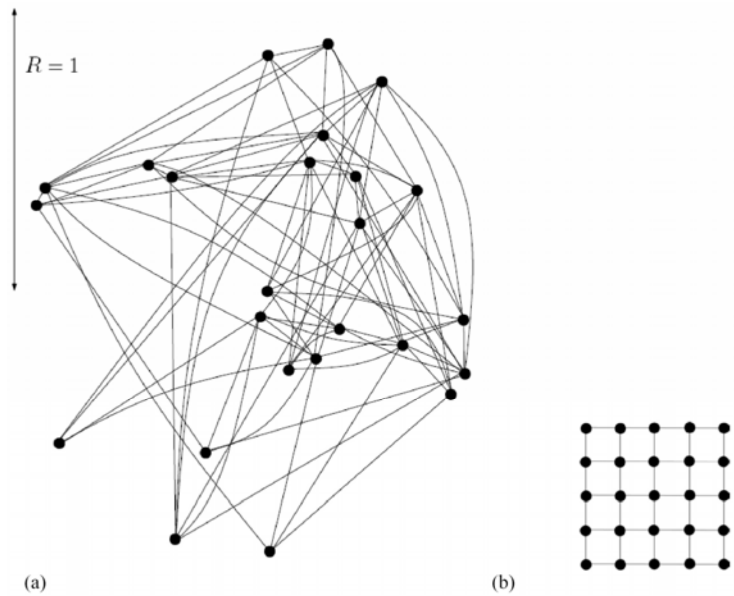

Figure 2. (a) Spatial random graph generated with $N = 25$, $R = 1$, $H = 0.25$, $p + p_b = 0.5$, $p - \Delta = 0.25$. The estimated neighborhood radius was $\sigma = 0.247$, and bias $\beta = 0.777$; (b) Square lattice with 25 vertices.

Due to the homogeneity of the Poisson point process, the average distance between any two nodes of the graph can be ascertained simply by calculating the average distance from a given fixed node to all other nodes. Because of the periodic boundaries, we may assume, without loss of generality, that the fixed node is positioned in the center of the square. Also, the Poisson point process yields that, for small dx, there are (on average) $\rho 8xdx$ points whose distance to the center is between x and $x + dx$. This means that the average distance to the center is

$$\frac{\int_0^R x \cdot \rho 8xdx}{\int_0^R \rho 8xdx} = \frac{2}{3} R, \tag{3}$$

where the bottom part represents the total number of points $(N = 4R^2\rho)$ and the top part represents the sum of the distances of all points to the center. Similarly, the average length of an edge can be calculated by

$$\frac{\int_0^R f(x)\cdot x\cdot \rho 8x dx}{\int_0^R f(x)\cdot \rho 8x dx} = \frac{\int_0^R f(x)x^2 dx}{\int_0^R f(x)x dx}. \tag{4}$$

Note that for (4) we used an expression that is basically identical to the left hand side of (3) with the addition of the term $f(x)$ that denotes a probability that there is an edge between the two nodes having geographical distance x. Combining (3) and (4) yields an alternative formula for β, which is

$$\beta = \frac{3}{2R}\frac{\int_0^R f(x)x^2 dx}{\int_0^R f(x)x dx}. \tag{5}$$

5. Estimating Relative Neighborhood Size: Neighborhood Radius

Now, we can proceed to the second and main task: given a specific graph, determine what is close, or how far is far. Substituting (1) for $f(x)$ in (5), we get

$$\beta = \frac{p - \Delta + (p_b + \Delta)\sigma^3}{p - \Delta + (p_b + \Delta)\sigma^2}, \tag{6}$$

where $\sigma = \frac{H}{R}$ represents the relative size of the neighborhood, or neighborhood radius. Assuming σ to be small, the homogeneity of the Poisson process yields that majority of the nodes are outside of one's neighborhood and thus $\Delta \approx 0$. Hence, we can approximate

$$\beta \approx \frac{p + p_b\sigma^3}{p + p_b\sigma^2} \approx \frac{p}{p + p_b\sigma^2} \approx 1 - \frac{p_b}{p}\sigma^2. \tag{7}$$

This immediately yields an estimate for σ based on β, namely,

$$\sigma \approx \sqrt{\frac{p}{p_b}(1 - \beta)}. \tag{8}$$

The parameter p can be simply estimated by $p = 2L/N^2$, where L is the number of edges and N is the number of nodes. Estimate for p_b is impossible to get without the knowledge of σ, but simply assuming high density of connections within the neighborhood (assuming that distance matters, or that the closest nodes are most likely to be connected), we may estimate σ by

$$\sigma \approx \frac{\sqrt{2L(1-\beta)}}{N}. \tag{9}$$

Formula (9) thus provides an estimate for the relative neighborhood size for an empirical spatial social network and, as such, requires knowledge of the number of nodes N, the number of edges L and the edge length bias β that can be easily calculated using formula (2).

6. Theoretical Results

We ran Monte Carlo Markov Chain simulations to generate spatial social networks for parameters N, H, p, p_b (we embedded the graph into a square with a standard Euclidean metric). We tested more than 100,000 graphs with $R = 1$, $N \in \{25, ..., 100\}$, $H \in [0.075, 0.5]$, $p + p_b \in [0.5, 1]$, $p - \Delta \in [0.1, 0.4]$. For each graph, we calculated edge length bias β using formula (2) and using either periodic or non-periodic boundaries; and either standard Euclidean metric or ell infinity metric, i.e., a metric such that

$$d_\infty((x_1, y_1), (x_2, y_2)) = \max\{| x_1 - x_2 |, | y_1 - y_2 |\}. \tag{10}$$

We concluded that there have been no significant differences between our various approaches to calculation of β. In fact, in all instances, for the same graph, the edge length biases calculated under different metric or boundary conditions did not differ by more than 5%.

It should be noted that the robustness of β relies heavily on the randomness and resulting irregularity of the graphs we investigated. Once we are interested in graphs that are regular (for example, the square lattice in Figure 2(b)), then β will become sensitive to boundary conditions and the used metric. In the particular case of a square lattice, the bias calculated for the Euclidean metric is significantly smaller than for d_∞ metric and having periodic boundaries increases the bias even further. This phenomenon is caused largely by the regular nature of the graphs – all of the edges have the same length (regardless of what metric we use to measure that), but the distance between unconnected edges decreases if measured by d_∞ metric or with periodic boundaries. Real world networks look more like the random graphs we considered than like the square lattice, and thus we do not expect the outcomes to be skewed in real world networks.

We also used β (calculated from (2) using Euclidean metric) and then estimated H using formula (9) from the results of the simulations. We concluded that (9) is a reliable estimate of H. Results are demonstrated in Figure 3. We plotted the frequency of the ratio between estimated neighborhood radius based on (9) and the correct one based on (1). It can be seen that our estimate was within the range 50%-150% of the correct neighborhood radius. The histogram shows that our method slightly overestimates the neighborhood radius, i.e., β was slightly smaller than it should have been. Also, in about 4% of the cases, the calculated neighborhood radius had value larger than 1 (and hence it could not be used in (9)); those cases are recorded as N/A.

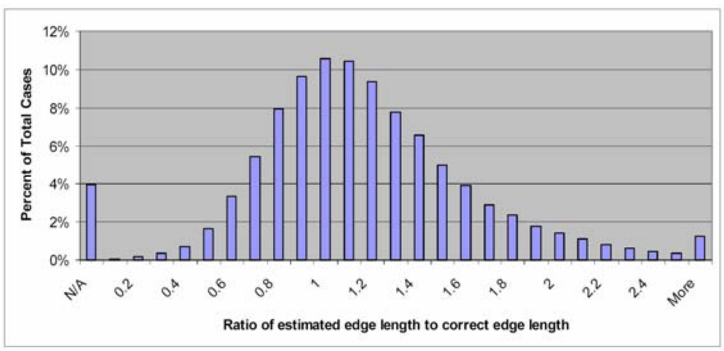

Figure 3. Correctness of the edge length bias estimate. We plotted the frequency of the ratio between estimated edge length bias (by formula (9)) and the correct one. N/A indicates that in some cases, we calculated $\beta > 1$ and thus could not use (9). The histogram shows aggregate results for all parameter values tested.

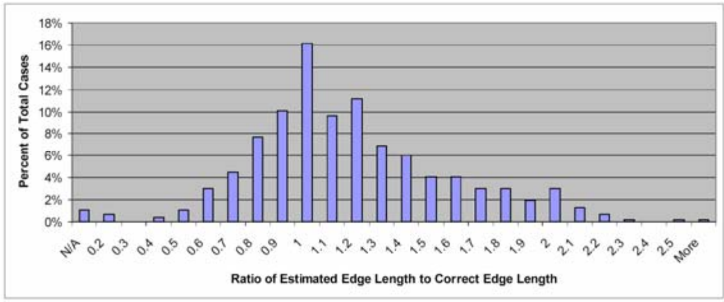

Figure 4. Correctness of the edge length bias estimate for large graphs $(N = 100)$.

As further seen in Figure 4, once we restrict ourselves to relatively large graphs $(N = 100)$, the estimation by formula (9) becomes more precise. Basically, Figures 3 and 4 say how often we were correct (i.e., how often we over/underestimated and by how much) about the neighborhood radius that was practically known to us.

Both the overestimation of the neighborhood radius and cases where $\beta > 1$ were more common for parameters $p - \Delta \in [0, 0.1] \cup [0.4, 1]$. Those parameters represent graphs with either too few or too many edges outside of the local neighborhood. If $p - \Delta$ is small, then the better approximation of formula (6) would be

$$\beta = \frac{p - \Delta + (p_b + \Delta)\sigma^3}{p - \Delta + (p_b + \Delta)\sigma^2} \approx \frac{(p_b + \Delta)\sigma^3}{(p_b + \Delta)\sigma^2} = \sigma \tag{11}$$

rather than using formula (7). Formula (11) was the original motivation for introducing the characteristic β; yet its use is limited to cases when there are almost no edges outside the neighborhood (and thus the neighborhood radius could be estimated heuristically). On the other hand, if $p - \Delta$ is large, then there are too many connections outside of the local neighborhood and it is therefore hard, if not impossible to say where exactly the neighborhood ends.

A similar phenomenon happens when H is too small or too big. It has been observed already in Wong et al. [16] that for too small H the graph is more like the standard Erdős-Renyi random graph; and it is clear that too big H has a similar effect. We therefore had to restrict ourselves on the intermediate values of H.

7. Empirical Results

Our real world data comes from interviews of a random sample of people living in a single neighborhood in New Orleans that was heavily impacted by Hurricane Katrina. We were interested in how spatial aspects of relationships affect disaster recovery. The following description does not address that question, but, rather, provides a dataset for evaluating the edge length bias algorithm and a comparison for our theoretical results from above. In December 2008, we asked respondents $(n = 60)$ to list 45 people they "know by sight or by name with whom they have had contact in the past two years or could have had contact if they wanted to", which is a sufficiently general question so as not to introduce interviewer bias in the listing of names. This type of network is called a *personal network*, in contrast to whole networks, which are theoretically bounded groups whose members are more likely to interact with one another than with random people outside of the group.

We created a random seed list of 25 numbers between 1 and 45 and used those seeds to select 25 alters from each personal network, since work by McCarty et al. has shown that around 20-25 alters randomly chosen out of a personal network of 40+ typically preserves structural integrity. From the respondents providing a personal network, we asked several things about the members of their network-including the distance, time and direction that the respondent perceived each of them to live (or work) from the respondent's home. Direction was given as one of eight cardinal directions, but typically was N, E, W or S. Distance was given in miles or blocks, the latter of which we converted to miles. We collected perceived time of ego's estimate of how long it takes to get to each alter's house. Time was given in hours or minutes, the latter of which we converted to hours. The units are not important in this study, as our purpose is to evaluate the utility of the edge length bias algorithm and understand the distribution across networks of whether or not people in any ego's personal network in this post-disaster setting tend to interact with other members of that personal network who live nearer to them or farther from them. Nonetheless, future research will involve analysis of the role of actual distance in predicting the distributing of edge length bias, thus giving insight into debates on scale in networks.

We then asked the respondent if each of the 25 people in the network knew one another well, a little bit or not at all, to the best of their knowledge-research on personal network tie prompts shows that a general prompt with few categories like ours produces reliable data while more specific prompts with many categories produces less reliable data (Killworth et al. [10]). The solicitation of the personal network ties produces, in this case, a theoretical maximum of 300 ties or $((25 \times 25) - 25)/2$.

Using direction and distance of alters from ego, we created X, Y coordinates for each network member in a personal network using geometry. We then calculated the extent to which the people who lived near each other were likely to have a network tie between them or, put another way, we calculated to what extent a personal network was likely to have people with a 'preference' for ties shorter than average possible geographic distance between members of that personal network. To do this, we used a command line software program that we developed which employs an algorithm that we call the edge length bias algorithm. The software calculated distances between each alter based on X, Y coordinates from ego $(0, 0)$. We completed this same procedure of creating X, Y coordinates separately for time as a distance and for geographic distance, and used the software to calculate the spatial bias based on each personal network based on time.

The results of applying the edge length bias methodology to real world data are presented in Figures 5 and 6. Figure 5 is a plot of the 60 personal networks against the edge length bias. In this analysis of real world data, we did not calculate neighborhood radius, but only edge length bias.

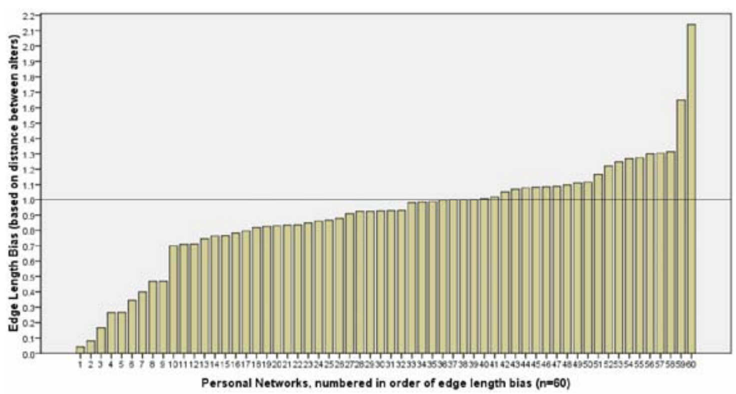

Figure 5. Plot of edge length bias values. The value of 1 means that people in that network tended to have ties within that network that averaged out to equal the average edge length of *possible* ties.

Figure 6 is a summary of values at each 0.1 interval of edge length bias. For Figures 5 and 6, the ties of low interaction and frequent interaction were collapsed so that we had binary graphs instead of valued graphs. The software can calculate edge length bias using valued graphs, but we do not yet know what scale is appropriate for stipulating the mathematical difference between no tie (0), low interaction (1) and much interaction (2). The graphs and X, Y coordinates did not include ego (the interviewee); only the 25 alters in each personal network were the basis of these calculations. We chose to not include ego in calculations because we are interested in the flow of resources and information in a personal network toward ego when ego has limited capacity for reaching out, such as following a disaster.

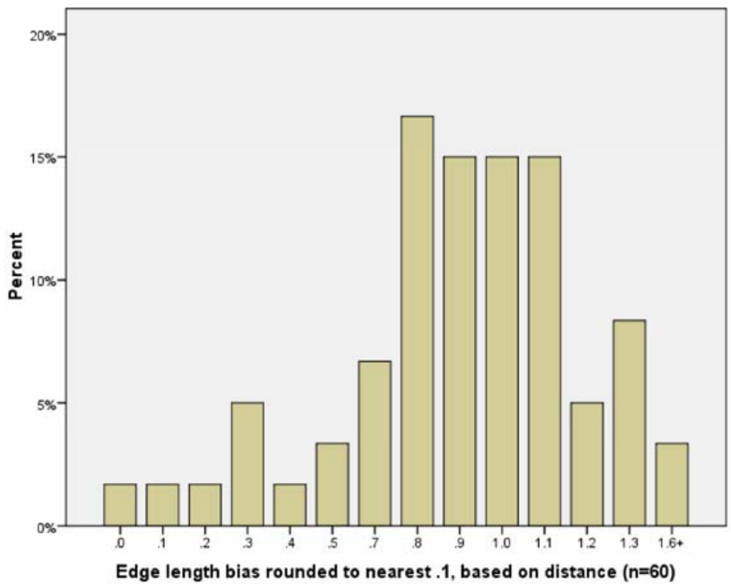

Figure 6. Summary distribution of edge length bias values, by percent of total, for 60 personal networks.

Figure 5 shows that approximately two-thirds of the personal networks in this sample were characterized by *actual* edge lengths that averaged less than the average edge length of *possible ties* within each of those networks. Around 17% of networks had average edge lengths less than half of the average edge length of possible ties, and yet only 3% of networks had average edge lengths 150% of the average edge length of possible ties.

Figure 6 presents the same data as in Figure 5, but in a summary of the distribution of incidences at intervals of 0.1. Figure 7 is also a summary of edge length bias values and is based on data from approximately 35,000 runs of randomly generated networks, thus providing a comparison to Figure 6. It bears remarkable resemblance to the real world data from Figure 6.

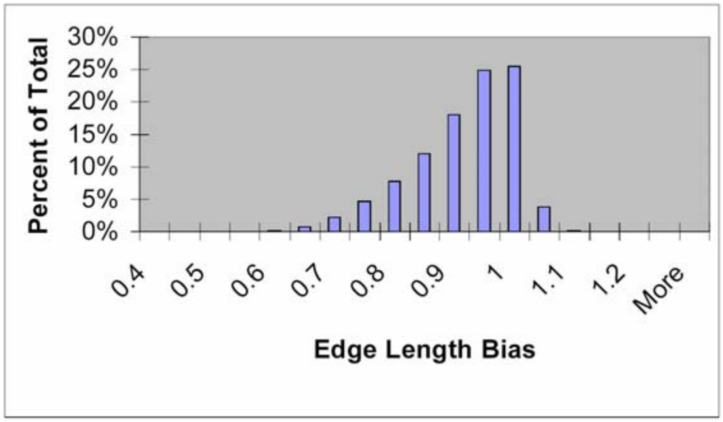

Figure 7. Distribution of edge length bias values for spatial networks produced by Monte Carlo Markov Chain simulations.

8. Conclusions and Discussion

In this note, we devised a characteristic of a spatial graph that allows us to quantify the likelihood of shorter connections between nodes over longer ones. The characteristic is robust in the sense that it is scalable (does not matter what units are used for measurements) and, more importantly, it allows one to compare spatial graphs with very different geographical sizes if theoretically permitted by the hypothesis being evaluated (e.g., local village networks might not be compared to author citation networks among scientists). Also, it is practically metric independent as it does not matter what metric one uses to measure the distances. This robustness is of substantial theoretical importance, since periodic boundaries and d_∞ metric are very easy to deal with analytically on paper, without any need of computer simulation.

In addition, and perhaps more importantly, we developed a formula that uses edge length bias to estimate the effective geographic size of one's personal neighborhood. This procedure should provide the ability to estimate what is far and what is close for any randomly generated spatial network. The prediction by the formula is surprisingly good for spatial graphs that are generated by the Poisson point process, and the formula works extremely well for large graphs ($N = 100$ or more).

Our procedure should also work for real world spatial networks, which are not uniformly distributed, but which, in the New Orleans case, showed similar distributions to the theoretical results although slightly more varied. It will be important to conduct many studies to start to develop typical ranges of variation in spatial networks – in our case, personal networks in post-disaster settings will be the priority along with comparisons to non-affected people with similar sociodemographic characteristics.

Networks comprised of human interactions, particularly, those comprised mainly of face-to-face interactions, as summarized by Wong et al. [16], are characterized by low number of ties relative to possible number of ties, relatively short number of steps between any two people, existence of subgroups, and existence of key individuals. These real world characteristics do not necessarily negatively affect the results of the neighborhood radius formula, as long as graphs are large enough and as long as the percentage of connections either outside or inside the neighborhood is not close to the extremes.

Acknowledgements

NSF grant BCS-0620213 provided support for development of the command line software developed for this study. The authors would like to thank the UNCG Office of undergraduate Research, which provided support for Robert Gove, and the UNCG College of Arts and Sciences and the UNCG office of Research for providing support for data collection in New Orleans.

References

[1] K. W. Axhausen, A. Frei and T. Ohnmacht, Networks, biographies and travel: First empirical and methodological results, Paper Presented at the 11th International Conference on Travel Behaviour Research, Kyoto, Japan, August 2006.

(http://www.ivt.ethz.ch/vpl/publications/reports/ab392.pdf)

[2] S. J. Bates, J. Trostle, W. T. Cevallos, A. Hubbard and J. N. S. Eisenberg, Relating diarrheal disease to social networks and the geographic configuration of communities in rural Ecuador, Amer. J. Epidemiology 166(9) (2007), 1088-1095.

[3] J. A. Carrasco, E. J. Miller and B. Wellman, How far and with whom do people socialize? Empirical evidence about distance between social network members, Transportation Research Record 2076 (2008), 114-122.

[4] C. E. Coburn and J. L. Russell, District policy and teachers' social networks, Educational Evaluation and Policy Analysis 30(3) (2008), 203-235.

[5] D. R. Cox and V. Isham, Point Processes, Monographs on Applied Probability and Statistics, Chapman & Hall, London, 1980.

[6] A. J. Cravey, S. A. Washburn, W. M. Gesler, T. A. Arcury and A. H. Skelly, Developing socio-spatial knowledge networks: a qualitative methodology for chronic disease prevention, Soc. Sci. Med. 52(12) (2001), 1763-1775.

[7] K. Faust, B. Entwisle, R. R. Rindfuss, S. J. Walsh and Y. Sawangdee, Spatial arrangement of social and economic networks among villages in Nang Rong district, Thailand, Social Networks 21(4) (1999), 311-337.

[8] R. Johnston, C. Propper, R. Sarker, K. Jones, A. Bolster and S. Burgess, Neighbourhood social capital and neighbourhood effects, Environment and Planning A 37(8) (2005), 1443-1459.

[9] E. C. Jones, Wealth-based trust and the development of collective action, World Development 32(4) (2004), 691-711.

[10] P. D. Killworth, C. McCarty, E. C. Johnsen, H. Russell Bernard and G. A. Shelley, Investigating the variation of personal network size under unknown error conditions, Sociol. Methods Res. 35(1) (2006), 84-112.

[11] D. Mok, B. Wellman and R. Basu, Did distance matter before the Internet? Interpersonal contact and support in the 1970s, Social Networks 29(3) (2007), 430-461.

[12] Walter Nicholls, Place, networks, space: Theorising the geographies of social movements, Trans. Inst. British Geographers 34(1) (2008), 78-93.

[13] F. R. Pitts, The medieval river trade network of Russia revisited, Social Networks 1 (1978/79), 285-292.

[14] S. Porta, P. Crucitti and V. Latora, The network analysis of urban streets: a primal approach, Environment and Planning B: Planning and Design 33(5) (2006), 705-725.

[15] R. Rothenberg, S. Q. Muth, S. Malone, J. J. Potterat and D. E. Woodhouse, Social and geographic distance in HIV risk, Sex. Transm. Dis. 32(8) (2005), 506-512.

[16] L. H. Wong, P. Pattison and G. Robins, A spatial model for social networks, Physica A: Statistical Mechanics and its Applications 360(1) (2006), 99-120.

[17] J. L. Wylie, L. Shah and A. Jolly, Incorporating geographic settings into a social network analysis of injection drug use and bloodborne pathogen prevalence, Health and Place 13(3) (2007), 617-628.

[18] K. S. Zimmerer, Geographies of seed networks for food plants (Potato, Ulluco) and approaches to agrobiodiversity conservation in the Andean countries, Society and Natural Resources 16(7) (2003), 583-601.