# Visualizing Spatial Dependencies in Network Topology

**Jamie Olson** **Kathleen M Carley**

July 12, 2010
CMU-ISR-09-127

Institute for Software Research
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Center for the Computational Analysis of Social and Organizational Systems

CASOS technical report

This work was supported in part by the National Science Foundation under the IGERT program (NSF 045-2598) for training and research in CASOS, the Office of Naval Research (ONR N00014-06-1-0104, ONR N00014-08-1-1186) and the Army Research Office (W911NF0710317). Additional support was provided by CASOS - the Center for Computational Analysis of Social and Organizational Systems at Carnegie Mellon University.

**Keywords:** network analysis, geographic information, spatial networks, visualization

**Abstract**

Increasingly, the data available to network analysts includes not only relationships between entities but the observation of entity attributes and relations in geographic space. Integrating this information with existing dynamic network analysis techniques demands new models and new tools. This paper introduces extensions to the ORA dynamic network analysis platform intended to meet this need. In particular, we present new visualization techniques for displaying the network topology of large, noisy datasets embedded in geographic space. We present these extensions and demonstrate them on some sample datasets.

# Contents

# 1 Introduction

Traditional social network analysis (SNA) focuses on the measurement and analysis of relationships between a single set of agents [26]. In contrast, dynamic network analysis (DNA) is characterized by the opportunistic use of the currently available data to model all visible aspects of a system [2]. Typically, this includes multiple node types, node attributes, and snapshots of relationship structure over time [1].

We consider such datasets augmented with geographic location information and find a need for new network measures, visualization techniques and analytical methods to enable the analysis of these spatially embedded networks. Relational information is increasingly labelled with information regarding the spatial locations of the entities involved. As researchers and analysts explore this spatially embedded network information, they will increasingly require the ability to integrate the spatial analysis and network analysis into a single framework. By explicitly modeling both network topological information and spatial dependencies, we attempt to produce a visualization of the network that is simple, intuitive and informative. The rest of the paper is organized as follows. In section 2, we survey the prior work in areas related to spatially embedded networks. Section 3 gives an overview of the Ora-GI tool and its features. Then, in section 4, we describe a new method for discovering spatial dependencies in spatially embedded networks. Section 5 shows how these techniques can be used in analyzing a real-world dataset, concluding with a summary and a discussion of limitations and future work.

# 2 Background

## 2.1 Geospatial Information

The recent proliferation of sensor systems has produced many datasets which feature spatial information in addition to the attributes and relationships typically measured. Examples include data from GPS sensors embedded in vehicles or cellular phones, logs of online activities, and information collected from intelligence networks. Although such data is often collected as the flow of entities through space, additional information can often be extracted. For example, the Reality Mining project[7] begins with such flow data and infers interactions between collocated individuals. In this way, events can be inferred from the initial flow data. In addition to sensor systems there are also situations in which more static networks are augmented with location information. For example, in text analysis, location information can be inferred by leveraging information from gazeteers and spatial databases by matching placenames with latitude/longitude pairs. This produces a multi-mode netork in which some subset of the nodes have been labeled with spatial information. Regardless of how this spatially embedded relational data is collected, it is increasing in availability, in size and in scope.

The incorporation of spatial locations into network data requires dealing with a fundamental disconnect between spatial information and relational information. Spatial information is funda-

[1] Subsets of these augmentations have been discussed within SNA literature. DNA can be thought of as the explicit study of new challenges arising from modeling all simulatenously.

mentally continuous, whereas, relations, in contrast, are defined as existing between pairs(sets) of discrete entities. This presents a barrier to the unification of spatial analysis and network analysis. Although we can break geographic space down into discrete locations [21], it is ultimately a continuous dimension, and any partitioning necessarily results in lost information. This loss of information increases the risk of the ecological fallacy [23] and the related modifiable areal unit problem [19].

The ecological fallacy is based on the observation that any calculation on aggregated data carries the risk that subsequent results may be an artifact of the aggregation. This happens because once data are aggregated, any subsequent analysis assumes that the data within the aggregate unit are homogeneous [23, 19, 18] and any individual differences are unimportant. This is an open problem, but one way to minimize the risk associated the ecological fallacy is to explore a wide variety of different levels and methods of aggregation. Results that are consistent regardless of the specific aggregation are more likely to be due to the properties of the actual phenomena. Although we will not explore the implications of the ecological fallacy for spatially embedded networks in this work, it is an important problem to acknowledge and consider in performing analysis.

Spatial dependence is a key concept in the analysis of spatial information. A set of spatially embedded random variables is spatially dependent if each random variable is dependent on nearby random variables. For example, spatial dependencies exist when nearby observations are more likely to be similar than would be expected from independent observations.

## 2.2 Relational Information

Prior work involving the interaction between spatial and network analysis has not sufficiently integrated the two types of information. Previous work on spatial networks by geographers has primarily focused either on storage and retrieval [24, 20] or on calculating shortest paths in transportation or distribution-type data [14]. Research in the geography of social networks has focused on case studies [E.g. 9] and on the influence of propinquity [1]. The tendency of individuals to associate with other nearby individuals, or propinquity, has been widely observed in a variety of different contexts [1]. This foundational research exploring the influence of geography on human interactions is of the utmost importance, we also need tools and techniques for analyzing the data produced by such studies.

We believe that more attention is needed on the general problem of analyzing spatially embedded networks. In particular, many of the basic exploratory techniques commonly used in network analysis seem to be somewhat less useful in analyzing networks in space. In particular, visualizations, which "have provided investigators with new insights about network structures and have helped them to communicate those insights to others" [10] do not appear to be as informative for spatially embedded networks. For example, one way in which visualizations yield useful information is through the use of layout mechanisms designed to highlight topological properties of a network, such as centrality, cohesive subgroups, etc. The most straightforward way of visualizing spatially embedded networks is a simple projection of the observed nodes onto a map based on their observed locations. By a priori choosing to locate nodes according to their observed spatial attributes, we lose any opportunity to arrange the network according to its topology. This makes visualization of spatially embedded networks less effective in illuminating and communicating the network structure. Furthermore, experience with real-world networks suggests that a small amount of noise or background activity can further decrease the utility of these visualizations. In large, noisy networks this simple projection may not yield an effective visualization.

Although there has been excellent work on effective visualizations of spatial networks [22, 21, 3], there has little effort to portray the underlying network topology. Specifically, these approaches all assume that the spatial network consists of connections between places(e.g. transportation networks) rather than connections between distinct entites which happen to be embedded in space(e.g. communication between people located in space). Because of this, these technique tend not to take into account the contribution of local connections to the network topology. For example, in a collaboration network, this is effectively ignoring collaborations between collocated individuals, potentially misrepresenting the overall structure of the network. A visualization technique that only takes into account edges between locations will necessarily fail to capture certain characteristics of the network.

Social network analysis has produced a wide range of network statistics that attempt to characterize a node's position within a network. Popular measures include *betweenness centrality*, which counts the number of shortest paths through a node, *degree centrality*, which count the number of edges a node has, and *eigenvector centrality*, which recursively defines important nodes as nodes connected to other important nodes[26]. Note that we do not restrict ourselves dyadically independent statistics as in exponential random graphs.

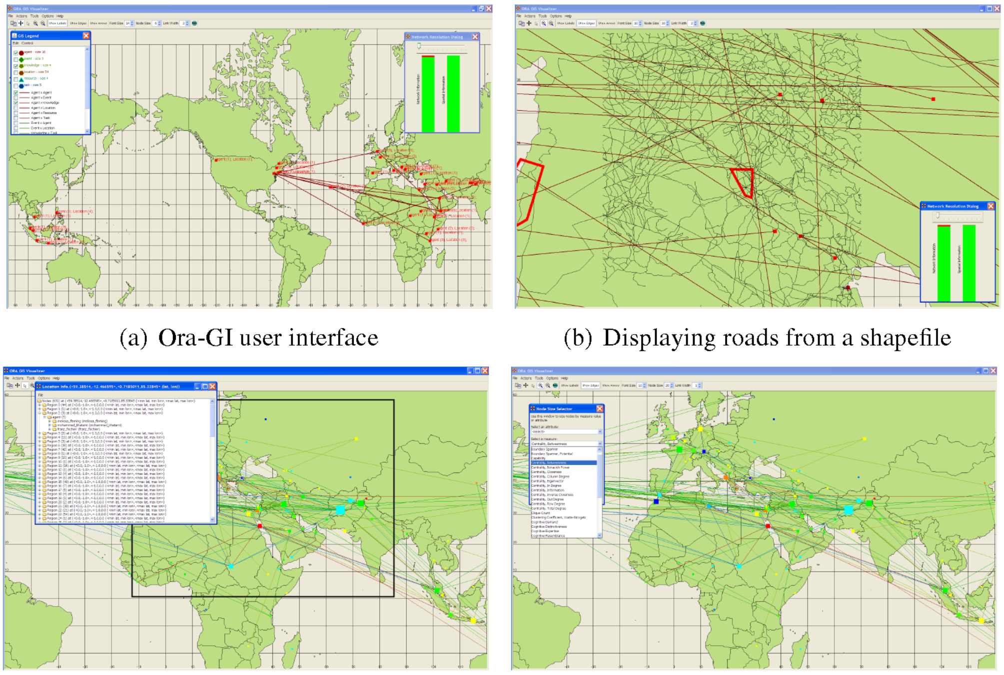
We develop a technique for visualizing spatially embedded networks based on explicitly portraying the interaction between network topology and geographic location. We assume a priori that there exist spatial dependencies in the topology of spatially embedded networks. Given that dependency, we model the relationship between the geographic location of a node and its topological properties.

# 3 ORA-GI

The Organizational Risk Analysis tool (Ora) is an integrated collection of tools for the analysis of relational information. It combines data transformation, visualization, network analysis, graphics generation tools, temporal analysis, and a simulation engine for short-term prediction in social networks. ORA with Geospatial Information (Ora-GI) allows the integration of geospatial information into the analysis of relational data, leveraging the existing tools in Ora. Ora-GI supports many navigational features that are common in geospatial analysis such as zoom, pan and select. Figure 1(a) shows the Ora-GI user interface. Most of these translate intuitively from the geospatial domain into the network domain (e.g. zoom, selection(Figure 1(c))).

## 3.1 Visualization

In addition to basic navigational features, Ora-GI supports a a variety of common network analysis methods. Ora-GI allows analysts to adjust the visual components (e.g. color, size) of places as related to network properties (see Figure 1(d)). In the example shown, places are colored according to the Newman-Girvan grouping algorithm, which attempts to discover densely subgroups within a larger network [16]. Places are also sized in relation to betweenness centrality, the extent that they lie between other nodes in the network [11]. Users also have the ability to perform further analysis on selected areas by saving a selection as a new new metamatrix (see Figure 1(c)). In doing this, the full capabilities of the Ora suite can be leveraged.



(a) Ora-GI user interface

(b) Displaying roads from a shapefile

(c) Selecting a region for analysis

(d) Analyzing groups and betweenness centrality

## 3.2 Resolution and Scale

Just as many standard techniques for geospatial analysis can be leveraged in geospatial DNA, several potential pitfalls of geospatial analysis also appear in geospatial network analysis. In particular, the modifiable area unit problem is of particular relevance [18]. The modifiable area unit problem is the danger that the results of a particular analysis are not representative of the actual data but simply an artifact of the aggregation and segregation of geospace into comparable units.

In discussing the aggregation and segregation of geospace, it becomes necessary to differentiate between *locations* and *places*. We refer to *locations* as precise positions in geospace, most commonly a <latitude, longitude> pair. In contrast, we use *places* to refer to the meaningful regions in which a research question is posed. For example, a social network may contain individuals labeled with home addresses *locations*. We may use this dataset to determine the US cities that are most connected in this social network. In this context, we would use US cities as *places* for the analysis.

This distinction is important in discussing geospatial network analysis. The domain of geospace is continuous and as such, geospatial data is likely to be continuous. The relational data underlying network analysis, is defined by connections between discrete entities and/or attributes. In order to include geospatial places in network analysis it is necessary to aggregate continuous locations into meaningful places.

The importance of aggregation for geospatial network analysis means that any system for geospatial DNA must have some method of assessing the modifiable areal unit. Ora-GI uses user-adjustable geospatial clustering [8] to dynamically aggregate locations into places. This allows analysts to not only select their perceived appropriate level of analysis, but also to easily perform sensitivity analysis by increasing and decreasing the resolution level.

In addition to sensitivity analysis, Ora-GI also provides a quantitative measure of information loss due to geospatial aggregation of the network[17]. This is presented as the proportion of network's information content that is preserved in the current aggregation. By combining sensitivity analysis with the information loss metric, Ora-GI helps analysts to make more informed decisions about the most appropriate level of analysis.

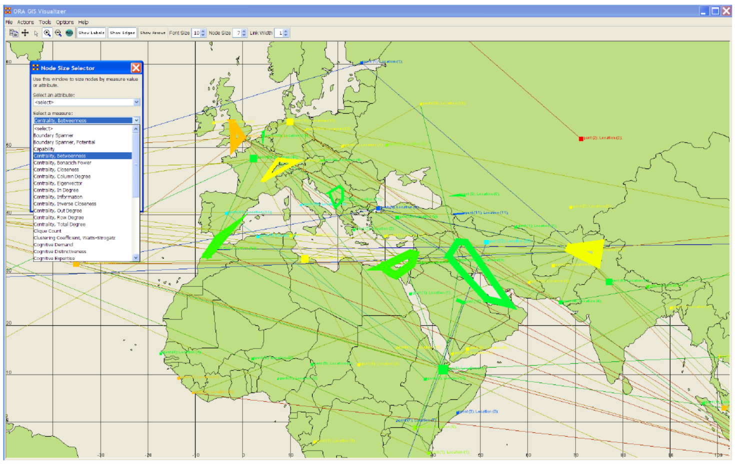


**Figure 1:** Clustered places colored according to a Newman-Girvan grouping and sized according to betweenness centrality

# 4 Exploring Spatial Dependencies with Kernel Smoothing

## 4.1 Notation and Assumptions

Consider a set of locations, $S = \{s_1, s_2, ..., s_n\}$ paired with a network, $\mathbb{G}$, defined by an adjacency matrix, $X$. We interpret this pairing as meaning that node $i$ in $\mathbb{G}$ having outgoing connections $X_{i*}$ is embedded at the location defined by $s_i$. We primarily consider $S$ as containing point location, but the technique is not liminted to such data. Any spatial information on which a distance measure can be reasonably defined can be used.

We assume spatial dependencies in the topology of the network. We mean not that simply that there are spatial dependencies in the distribution of connections but that there are spatial dependenies in the network positions of the nodes. Using betweenness centrality as an example, may mean that certain locations are more likely to contain nodes with high betweenness centrality in the network.

This spatial dependence assumption is an extremely strong assumption that is doubtless violated in many real-world networks. Unfortunately, there are currently no statistical tests for spatial dependencies in the network topology to inform us in making this assumption. Although there are tests of spatial dependence in general[15, 12], the importance of propinquity in spatially embedded social networks is a potentially confounding variable. For example, consider spatial grid, where a node is located at each intersection and it is connected to its four immediate neighbors. Nodes towards the spatial center of the grid will necessarily have certain network topological properties(e.g. high betweenness centrality) simply due to the interaction between the spatial distribution of nodes and propinquity. Because propinquity is a simple, well-documented phenomena with sociological theory supporting it, we should always begin by considering it as an explanation rather than some more complicated spatial dependency. A desirable test for spatial dependencies in the network topology would control for effects due to propinquity. Nevertheless, we can use general tests for spatial dependence [15, 12], but we must remember that we are not accounting for propinquity.

## 4.2 Kernel Density Estimation

One technique for visualizing large spatial data sets is the use of kernel density estimation to interpolate the point intensity across the spatial region of interest. Kernel smoothing uses a kernel function, $k$, to interpolate the intensity, $\lambda$, of a phenomena based on the observed set of discrete observations. For a target location, $s$:

$$\hat{\lambda}(s) \quad = \sum_{i:D(s,s_i)<\tau} \frac{1}{2\tau^2} k\left(\frac{D(s,s_i)}{\tau}\right) \tag{1}$$

where $\tau$ is a bandwidth parameter and $D(s, s_i)$ is a distance function. The bandwidth parameter, $\tau$, represents the fundamental tradeoff between a smooth interpolated function and a loss of information and oversmoothing. Although the choice of a kernel function may appear to be a difficult and important decision, it is considered to have relatively little impact on the interpolated function [25]. This intensity estimation procedure is frequently used to create heatmap-style images but a variation called kernel smoothing [13].

## 4.3 Kernel Smoothing

Kernel smoothing expands this method to smooth values rather than intensities [13]. Intuitively, each observation is weighted in proportion to its proximity to the target location. If $s_i$ are discrete observations and $y(s_i)$ are some function or attribute of each observation, then the interpolated value for a target location $s$ is:

$$\hat{y}(s) = \frac{\sum_{i:D(s,s_i)<\gamma} y(s_i) \frac{1}{2\gamma^2} k\left(\frac{D(s,s_i)}{\gamma}\right)}{\sum_{i:D(s,s_i)<\gamma} \frac{1}{2\gamma^2} k\left(\frac{D(s,s_i)}{\gamma}\right)} \tag{2}$$

Rather than use smooth some spatial property of the discrete observations, $s_i$, we smooth some network statistic, $z(X, i)$ across the spatial area. For a network statistic $z(X, i)$ and a target location, $s$:

$$\hat{z}(s) = \frac{\sum_{i:D(s,s_i)<\gamma} z(X,i) \frac{1}{2\gamma^2} k\left(\frac{D(s,s_i)}{\gamma}\right)}{\sum_{i:D(s,s_i)<\gamma} \frac{1}{2\gamma^2} k\left(\frac{D(s,s_i)}{\gamma}\right)} \tag{3}$$

This methodology has been incorporated into the larger tools described in [5] and we use that tool to demonstrate the new approach.

# 5 Evaluation

## 5.1 Data

From the 25th to 30th of June 2005, a sensor network queried Automated Identification System (AIS) transponders on merchant marine vessels conducting exercises in the English Channel, recording navigational details such as current latitude and longitude, heading, speed, reported destination, and several forms of identifying information. In total, movements of over 1700 vessels were recorded, with activities ranging from simple shipping lane traversals to apparently complex itineraries with stops at multiple ports of call. The dataset we analyzed includes 42869 AIS reports from approximately 1729 distinct vessels, over a large geographic range that suggests multiple polling stations, shown in Figure 2.

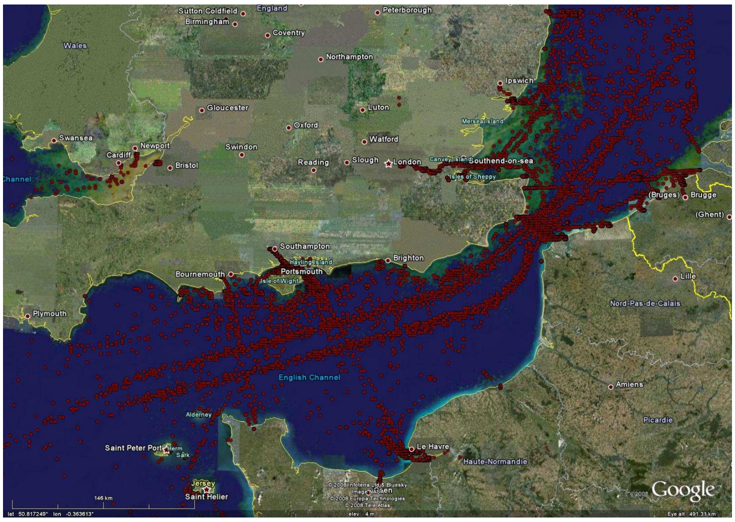



**Figure 2:** Maritime data collected by the Automated Identification System

Although the specific format of the message is standardized, several factors limit the consistency and precision of the interpretation of AIS reports. The numerical precision of the geographic location fields is fixed but distances over degrees of latitude and longitude vary around the globe, resulting in actual physical precision of the readings are also inconsistent. In the English Channel area, the effective sensor resolution was approximately 1100 meters, or .6 nautical miles, meaning that smaller differences in location could not be accurately distinguished. Because of this, it is not possible to examine movement patterns at a higher geographic scale with AIS data. Another limitation of AIS data is the polling frequency and duration. Although it varies somewhat across the sampled region, queries appeared to be conducted at approximately 40 minute intervals, meaning that activities on a similar timescale might be unrecorded or almost impossible to identify. For these reasons we focus on patterns at a low geographic scale, across the entire sampled region. More information as well as an in depth analysis of this dataset can be found in [4].

## 5.2 Analysis

We examine this AIS dataset using the kernel smoothing technique as implemented in the ORA-GI Figure 3 shows a simple projection of the network in a geographic context.

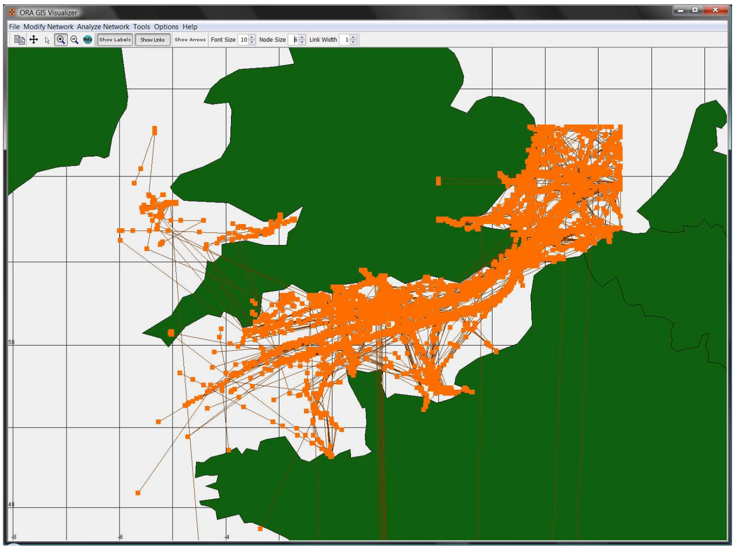
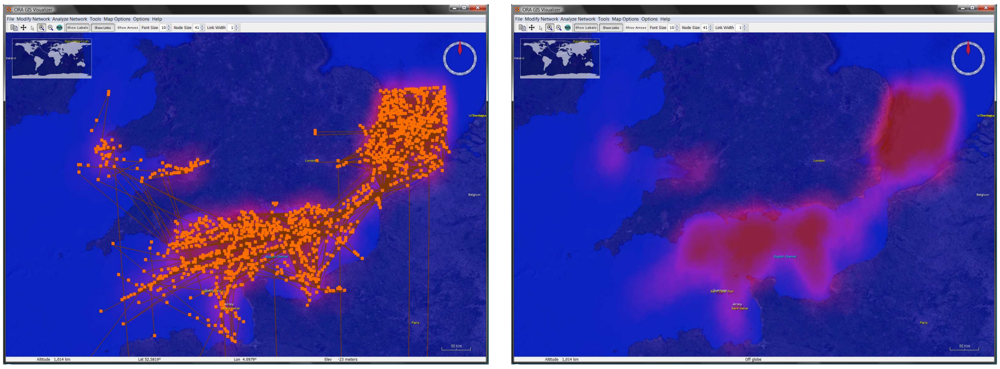


**Figure 3:** Initial ORA-GIS Visualization

This, however, obscures any information concerning the number of observations at each locations. Figure 5.2 shows a heatmap of the ship observations.

Although this data is represented as a network, it is not initially particularly useful. This is due to the relatively high precision achieved by the AIS tracking system. Since every coordinate is interpreted as a distinct location, the trails are degenerate in the sense that no two ships visit the same location or revisit their own path. Figure 5 shows in general how the high resolution of the data can result in a network that is not spatially meaningful. To yield a meaningful network across space, we use density-based clustering [8] to merge nearby points into a smaller set of aggregated meta-locations. Figure 6 shows the results of a clustering of points based on geospatial density. The



(a) Heatmap with network  (b) Heatmap without network

**Figure 4:** Heatmap of ship intensity both with(a) and without(b) the network visible

meta-locations discovered by the clustering algorithm correspond to the major ports and shipping lanes in the dataset.

Now that the locations have been clustered, we can examine the topology of this network as it is embedded in space. Figure 5.2 shows various the network of meta-locations with each meta-network sized according to various network measures.

As prior experience has shown us in [5], although this can be a valuable measure for discovering certain pieces of information, it does not clearly elucidate the general spatial dependencies in the network topology. This is because two meta-locations near each other may either tend to have similar measures, as in hub centrality, or they may not, as it would appear in both eigenvector centrality and betweenness centrality. To visualize these spatial dependencies, it can be valuable to use the kernel smoothing methodology proposed earlier. Figures 5.2 and 5.2 show the same network measures using the kernel smoothing technique.

Table 5.2 shows Moran's I spatial autocorrelation statistics, suggesting that all of these measures appear to have spatial dependencies. This supports our assumption of spatial dependence in the topology of the network, but it also confirms the need to develop new statistical tests for spatial dependence in network structure. It is not clear to what extent the observed spatial dependence is caused by propinquity as opposed to specific spatial dependencies in the network topology.

This brief analysis was meant to simply demonstrate how the tools can be chained in practice. For a more detailed examination of this dataset please see [6].

# 6 Conclusion and Future Work

The intersection of spatial analysis and network analysis is an exciting field demanding new tools, techniques and visualizations. In particular, the projection of a network onto a spatial layout dramatically increases the complexity of the visual information. We proposed a new strategy for approaching the analysis of spatially embedded networks based on discovering the spatial dependencies of the network topology. We also present the beginnings of a methodology for discovering these spatial dependencies through statistical tests and visualization based on kernel smoothing.

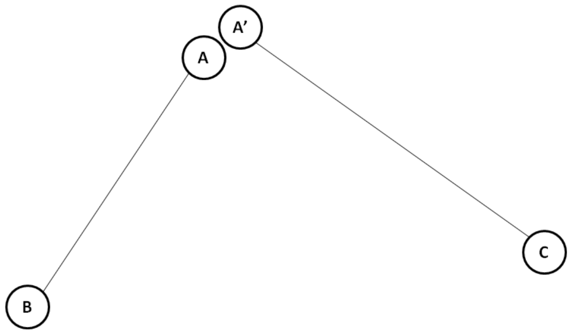



**Figure 5:** Discrete entities fail to effective capture spatial proximity

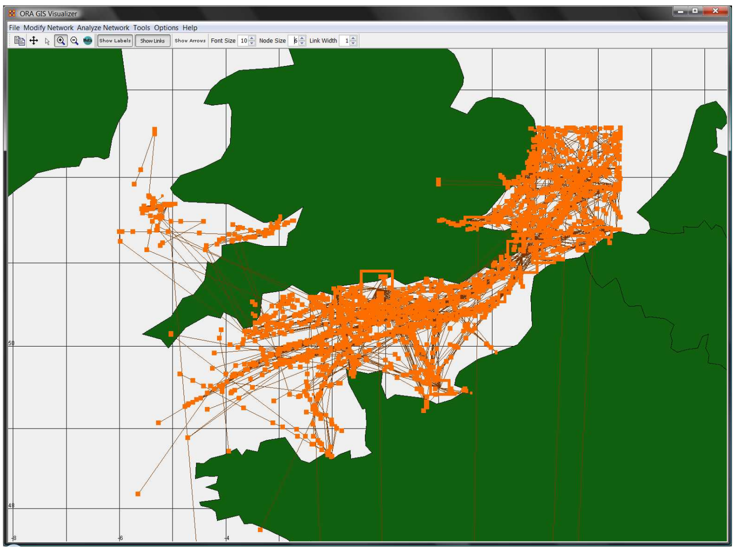



**Figure 6:** Meta-locations defined by density-based clustering

| Measure | Moran's I Statistic |
|---|---|
| Boundary Spanner | 8.78 |
| Capability | 512.66 |
| Authority Centrality | 16.19 |
| Betweenness Centrality | 494.03 |
| Bonacich Power Centrality | 481.58 |
| Closeness Centrality | 13.30 |
| Eigenvector Centrality | 113.10 |
| Hub Centrality | 20.10 |
| Information Centrality | 24.19 |
| Degree Centrality | 479.91 |
| Clustering Coefficient | 19.86 |
| Constraint | 21.42 |
| Distinctiveness Correlation | 309.54 |
| Expertise Correlation | 21702.68 |
| Resemblance Correlation | 154049.32 |
| Similarity Correlation | 26.77 |
| Effective Network Size | 1013.76 |
| Exclusivity | 284.76 |
| Simmelian Ties | 221.00 |

**Table 1:** Spatial autocorrelation statistics for various node-level network topology statistics

We explore this general framework and the new tools developed in the context of one sample dataset drawn from observations of ships over time. We find that in general the visualization techniques appear to be informative, but additional interpretation of the results is necessary in order to understand the advantages of the kernel smoothing technique over simple size-by visualizations. Furthermore, it is not clear that the statistical test we applied is appropriate in this context. Although the statistical tests showed every topological statistic to be significantly spatially dependent, the visualization did not imply strong dependencies for some of the statistics. Propinquity can likely explain a large portion of the variance in network topology across space without assuming spatial dependencies.

All of the techniques used in the analysis have been incorporated incorporated into the ORA tool for analyzing network data. This allows other researchers to easily explore these new techniques and to incorporate them into their analysis.

In general, the idea of exploring the spatial dependencies of the network structure is only reasonable if there are in fact spatial dependencies in the network structure. Without improved statistical tests, it is impossible to know how reasonable this approach is in general.

We did not explore the effect that the level of aggregation has on any of these results. Because the network statistics computed on the meta-location network may be fairly complex, it is possible that the level of aggregation could have a strong impact on the results. An improved understanding of the influence that the aggregation step has is crucial for the results drawn from this type of analysis to be generalizable. Other future work includes improvements to the statistical tests for spatial dependence, the exploration of alternative spatial visualization and the incorporation of temporal information into the analysis.

Although there are serious limitations to the specific methodologies presented here, we believe that analysis of spatially embedded networks through spatial dependencies in the network topology is a promising new approach to analyzing such networks.

# References


[1] C.T. Butts. *Spatial models of large-scale interpersonal networks*. PhD thesis, Carnegie Mellon University, Pittsburgh, May 2002.

[2] Kathleen Carley. Dynamic network analysis. In *Committee on Human Factors*, pages 133–145. National Research Council, 2004.

[3] Kenneth C. Cox, Stephen G. Eick, and Taosong He. 3d geographic network displays. *SIGMOD Rec.*, 25(4):50–54, 1996.

[4] George B. Davis and Kathleen M. Carley. Computational analysis of merchant marine global positioning data. Technical Report CMU-ISRI-07-109, Institute for Software Research, Carnegie Mellon University, 2007.

[5] George B Davis, Jamie F Olson, and Kathleen M Carley. Oragis and loom: Spatial and temporal extensions to the ora analysis platform. Technical Report CMU-ISR-08-121, Carnegie Mellon University, 2008.

[6] James A Davis, Tom W. Smith, and Peter V. Marsden. General social surveys, 1972-2006. Technical Report ICPSR04697-v2, National Opinion Research Center, Chicago, IL, 2007.

[7] N. Eagle and A. Pentland. Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):255–268, 2006.

[8] Martin Ester, Hans-Peter Kriegel, Jorg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In Evangelos Simoudis, Jiawei Han, and Usama Fayyad, editors, *Second International Conference on Knowledge Discovery and Data Mining*, pages 226–231, Portland, Oregon, 1996. AAAI Press.

[9] Katherine Faust, Barbara Entwisle, Ronald R. Rindfuss, Stephen J. Walsh, and Yothin Sawangdee. Spatial arrangement of social and economic networks among villages in nang rong district, thailand. *Social Networks*, 21(4):311 – 337, 2000.

[10] L. C. Freeman. Visualizing social networks. *Journal of Social Structure*, 1(1), 2000.

[11] Linton C. Freeman. Centrality in social networks: Conceptual clarification. *Social Networks*, 1(3):215–239, 1979.

[12] RC Geary. The contiguity ratio and statistical mapping. *The Incorporated Statistician*, pages 115–145, 1954.

[13] T. Hastie, R. Tibshirani, J. Friedman, T. Hastie, J. Friedman, and R. Tibshirani. *The elements of statistical learning*, chapter 6. Springer New York, 2001.

[14] M. Kolahdouzan and C. Shahabi. Voronoi-based k nearest neighbor search for spatial network databases. In *Proceedings of the 30th international conference on Very Large Databases*, volume 30, pages 840–851. VLDB Endowment, 2004.

[15] PAP Moran. Notes on continuous stochastic phenomena. *Biometrika*, pages 17–23, 1950.

[16] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69(2):026113, Feb 2004.

[17] Jamie F Olson and Kathleen M Carley. Summarization and information loss in network analysis. In *Workshop on Link Analysis, Counter-terrorism, and Security held in conjunction with the SIAM International Conference on Data Mining (SDM)*, April 2008.

[18] S Openshaw and S Alvandies. *Geographical information systems : principles, techniques, management, and applications*, volume 1, chapter Applying geocomputation to the analysis of spatial distributions. John Wiley & Sons, New York, 2 edition, 1999.

[19] S. Openshaw and PJ Taylor. The modifiable areal unit problem. *Quantitative Geography: A British View*, pages 60–69, 1981.

[20] D. Papadias, J. Zhang, N. Mamoulis, and Y. Tao. Query processing in spatial network databases. In *Proceedings of the 29th international conference on Very Large Databases*, volume 29, pages 802–813. VLDB Endowment, 2003.

[21] Doantam Phan, Ling Xiao, R. Yeh, and P. Hanrahan. Flow map layout. pages 219–224, Oct. 2005.

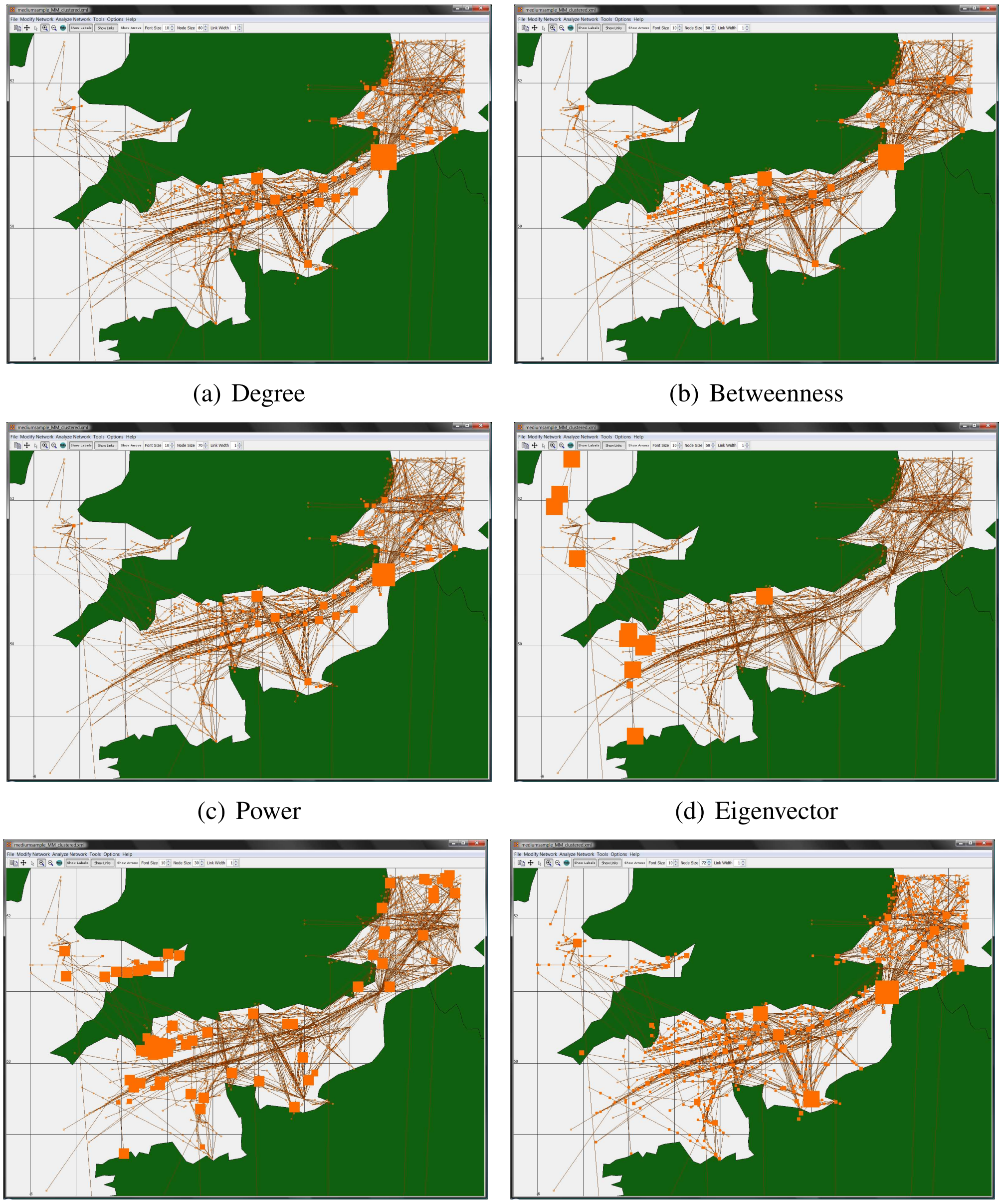
[22] F.H. Post, B. Vrolijk, H. Hauser, R.S. Laramee, and H. Doleisch. The state of the art in flow visualisation: Feature extraction and tracking. 22(4):775–792, 2003.

[23] W.S. Robinson. Ecological correlations and the behavior of individuals. *American Sociological Review*, 15(3):351–357, 1950.

[24] J. Sankaranarayanan, H. Alborzi, and H. Samet. Efficient query processing on spatial networks. In *Proceedings of the 13th annual ACM international workshop on Geographic information systems*, pages 200–209. ACM New York, NY, USA, 2005.

[25] L.A. Waller and C.A. Gotway. *Applied spatial statistics for public health data*. Wiley-Interscience, 2004.

[26] Stanley Wasserman and Kathleen Faust. *Social Network Analysis*. Cambridge University Press, Cambridge, MA, 1994.



(a) Degree

(b) Betweenness

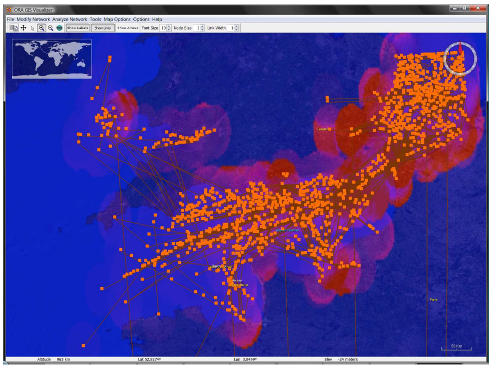
(c) Power

(d) Eigenvector

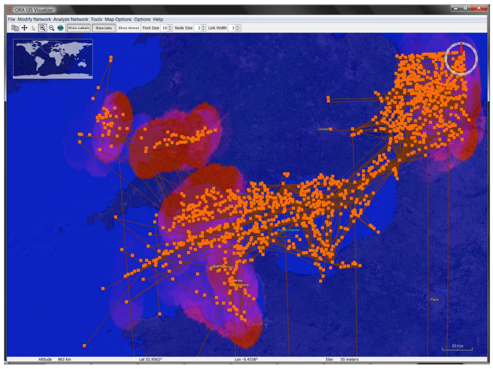
(e) Hub

(f) Exclusivity

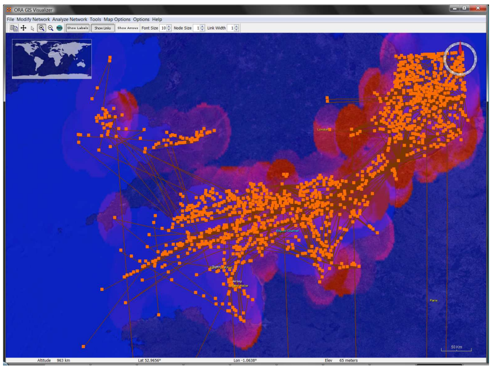
**Figure 7:** Assorted maps showing meta-location networks with nodes sized by various topological measures.



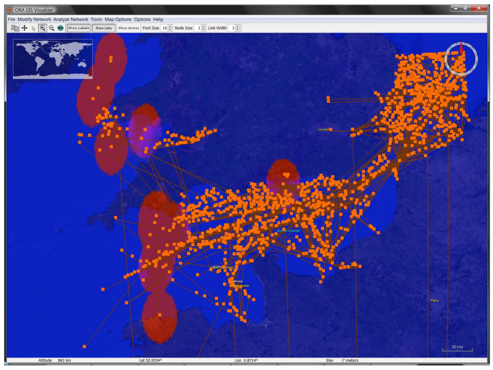
(a) Degree



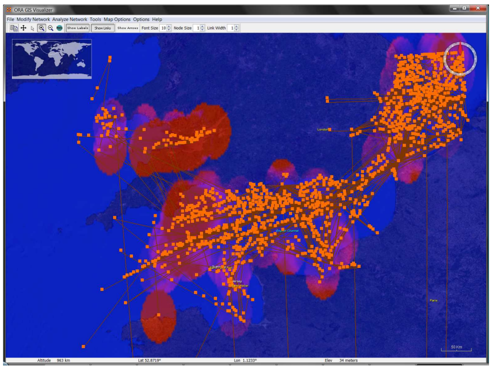
(b) Betweenness



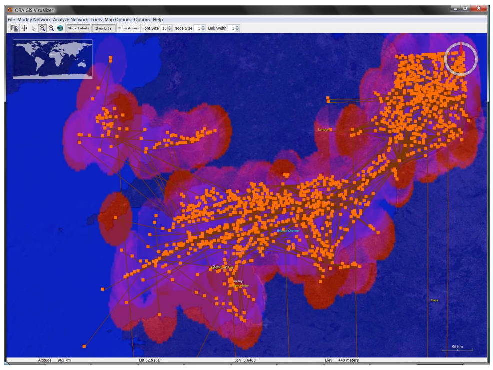
(c) Power



(d) Eigenvector

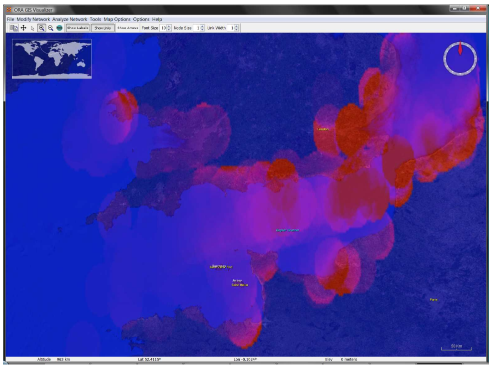


(e) Hub

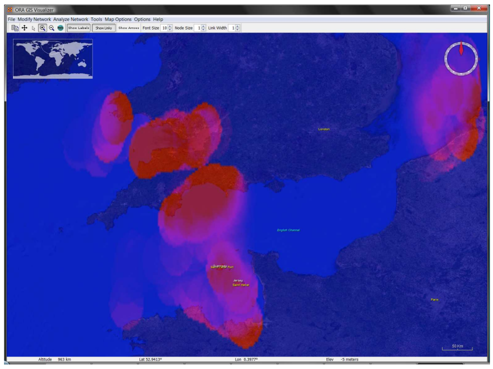


(f) Exclusivity

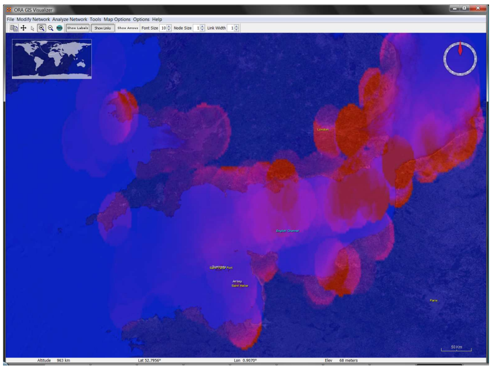
**Figure 8:** Assorted maps showing meta-location networks with the interpolated network topological measures.



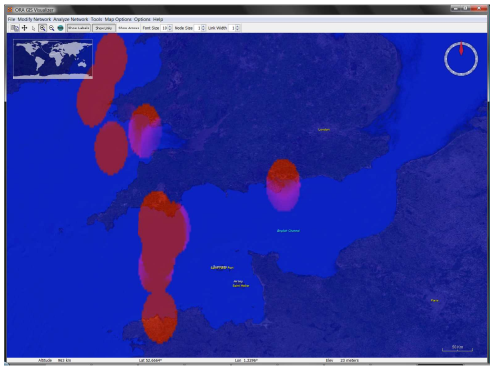
(a) Degree



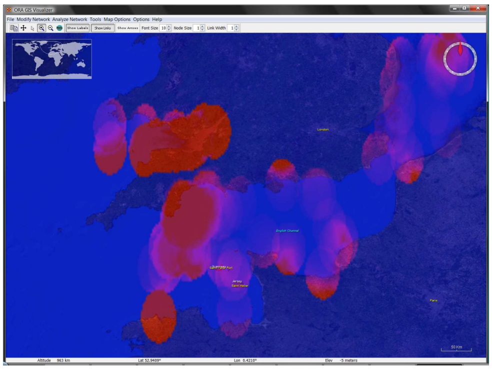
(b) Betweenness



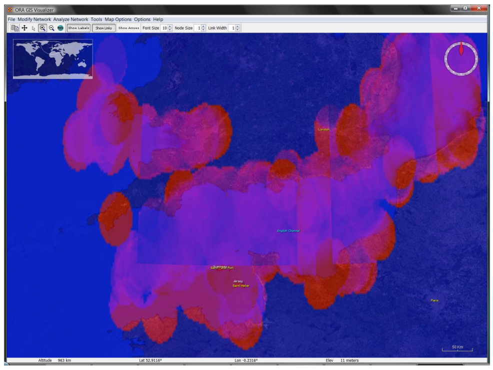
(c) Power



(d) Eigenvector



(e) Hub



(f) Exclusivity

**Figure 9:** Assorted maps showing the interpolated network topological measures.