Cartography and Geographic Information Science

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tcag20

SNoMaN: a visual analytic tool for spatial social network mapping and analysis

Sichen Jin, Alex Endert & Clio Andris

CARTOGRAPHY AND GEOGRAPHIC INFORMATION SCIENCE
https://doi.org/10.1080/15230406.2024.2413600

SNoMaN: a visual analytic tool for spatial social network mapping and analysis

Sichen Jin [a], Alex Endert [a] and Clio Andris [a,b]

[a]School of Interactive Computing, College of Computing, Georgia Institute of Technology, Atlanta,GA, USA; [b]School of City & Regional Planning, College of Design, Georgia Institute of Technology, Atlanta, GA, USA

ABSTRACT

Spatial social networks (SSNs) are node-link structures that evidence interpersonal or inter-organizational relationships, where nodes and edges have a defined geographic location. To model SSNs, users need both geographic and social network metrics. However, there are few GUI-based analytic tools that enable simultaneous spatial and social network exploration. In this paper, following the research framework of Exploratory Spatial Data Analysis (ESDA) and design principles of social network analysis tools, we derived three design goals of exploratory spatial social network analysis (SSNA). Guided by these design goals, we provide a visual analytic tool, SNoMaN, which links network and geographical layouts and helps users conduct SSNA by interactively computing and visualizing SSN metrics, describing spatial distributions, exploring associations, and detecting anomalies. We introduce new types of visual diagrams, including Cluster–Cluster Plots, Centralization Plots, on-the-fly mapping of geometrically bounded network modules, and Route Factor Diagrams. We illustrate these new approaches using use case studies of a 1960s network of Mafia members, a global flight network, and a food donation-sharing network in southwestern Virginia. We find that SNoMaN can be used to generate data insights that fuse a system's spatial and social dimensions that are hard to obtain otherwise.

ARTICLE HISTORY
Received 7 November 2023
Accepted 1 October 2024

KEYWORDS
Geovisualization; visual analytics; software; social networks; spatial networks; community detection

1. Introduction

A social network is a network structure where nodes represent social actors (such as individuals or organizations) and edges represent the social interactions between actors. Social network analysis (SNA) aims to investigate and characterize the social structures formed by nodes and ties (Wasserman & Faust 1994). Social networks are largely affected by geographic factors, including distance, natural barriers, geographical boundaries, access to resources, etc. For example, the number of connections between nodes typically decreases as the distance between the nodes increases (Laniado et al., 2018; Lengyel et al., 2015; Liu et al., 2017; Onnela et al., 2011; Scellato et al., 2010; Verdery et al., 2012). In addition to distance, natural barriers and geographical boundaries can also segregate network entities. For example, 75% of connections on Twitter are between accounts that are in the same country (Takhteyev et al., 2012) and county borders can reduce information sharing (Sohn et al., 2020). Other geographical influences include access to resources. For instance, obesity not only depends on social network influence (Christakis & Fowler, 2008) but also is influenced by access to fast food (Block et al., 2004).

Social networks are usually modeled outside of a GISystem, which does not provide researchers with the opportunity to measure the influence of underlying geography on social connections (Andris et al., 2018; Luo et al., 2011; Ye & Liu, 2018). Although it is beneficial to model social networks in a GISystem, there are few GUI-based analytical tools that simultaneously support social network and geographic analysis, in part because social network analysis (SNA) and GISystems matured in the separate domains of sociology and geography, respectively (Andris et al., 2018). Popular GUI-based social network analysis tools, including Gephi (Bastian et al., 2009), UCINET (Borgatti et al., 2002), and Pajek (Batagelj & Mrvar, 2004) do not have interactive mapping components and thus, research is often conducted without visualizing nodes on a map. Conversely, researchers cannot easily calculate social network metrics, such as network diameter, presence of cliques, or eccentricity on nonplanar networks in a GISystem, although such values have been integrated into GIS-based infrastructure analysis (Sevtsuk & Mekonnen, 2012). In practice, network metrics might be precalculated in SNA software (such as Gephi), and necessary spatial analysis and visualization may be performed in ArcGIS. Analysts have to manually translate

CONTACT Sichen Jin sjin86@gatech.edu
Supplemental data for this article can be accessed online at https://doi.org/10.1080/15230406.2024.2413600

data by constantly switching between loosely-coupled environments, which can hinder exploratory studies. Alternatively, a researcher may use R or Python-based IDEs to calculate spatial and social network statistics in the same workflow, but this method does not provide GUIs that support a wide set of exploratory spatial data analysis (ESDA) functions.

The goal of this research is to provide spatial social network analysts with a tool that synchronizes a social network visualization and a map visualization and provides brushing and linking capabilities to allow for simultaneous interaction between the representations. We use an ESDA research framework (Anselin, 1996; Haining et al., 1998; MacEachren & Kraak, 1997; Wei, 2022) and practices inherent to the development of standalone geospatial software (as in McKenzie et al., 2023). The exploratory and interactive process can help users quickly see spatial distributions of social networks, identify associations among social network and geographic measurements, and detect anomalies.

Accordingly, we built a tool called Social Network Mapping Analysis (SNoMaN) (Figure 1). SNoMaN is a free, open-source, web-based tool where users can upload a spatial social network as a text file and analyze it using social network metrics, spatial metrics, and combined social-spatial metrics. SNoMaN plots the network on a map, as is traditional in geo-network visualization, and links the map to a force-directed node-link visualization (called a sociogram) of the network in feature space. Because the sociogram is force-directed, it can show who is central or peripheral in the network, what chains and cliques form, whether there are disconnected groups, etc. These aspects might not be as apparent when nodes are fixed within their geographic positions. In addition, SNoMaN includes new types of visual diagrams, including a a) Centralization Plot, b) Route Factor Diagram, c) Cluster–Cluster Plot, and d) mapping of network communities, to interactively visually explore associations between geographic and network metrics and identify anomalies.

We demonstrate how SNoMaN can be used by exploring three example networks: an American Mafia network (DellaPosta, 2017), a world flight network (OurAirport, 2017), and a food sharing organization network (Edwards, 2020). Our aim is to illustrate that we can learn new things about these networks that would have been difficult to discover in separate SNA and GISystems software environments. While the tool is not intended to capture all methods of SSN exploration, it acts as a starting point for users who want to visually explore spatial social networks.

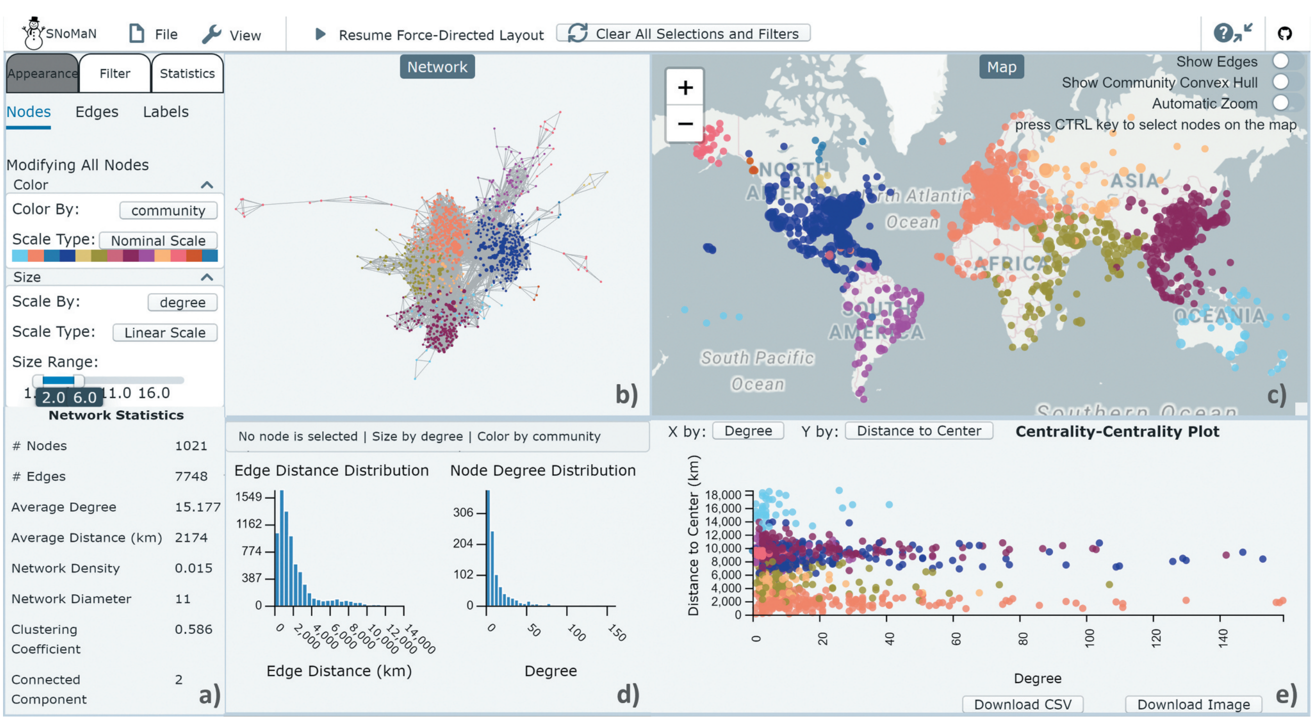


Figure 1. The SNoMaN interface, showing a world flight network of 1,022 airports (i.e. nodes) and 7,748 direct flights (i.e. edges) (OurAirport, 2017): a) Configuration and function panel includes the network appearance panel, filter panel, and analysis panel. b) Network view presents the force-directed layout of the network. c) Map view presents the geographic layout of the network. d) Detail view dynamically calculates and displays network and geographical metrics of users' selection. e) Comparison view supports comparing SSNA metrics.

The paper is structured as follows. Section 2 provides an overview of relevant research. Section 3 presents three design goals that reflect exploratory SSNA objectives. Section 4 details the user interface, and Section 5 explores three case studies. In the final section, we discuss drawbacks and future work, and conclude.

2. Related work

Our work is motivated by and situated within prior studies in the domains of visual analytic tools for networks, spatial network visualization, and spatial social network analysis.

2.1. Visual analytic tools for social networks

To make social network analysis (SNA) more accessible and efficient, a growing number of network analysis tools have been created for SNA, namely, UCINET (Borgatti et al., 2002), Gephi (Bastian et al., 2009), Pajek (Batagelj & Mrvar, 2004), NetworkX (Hagberg et al., 2008), Node XL (Smith et al., 2009), and MultiNet (Richards & Seary, 2003). Among these tools, some are of general use (e.g., Gephi, Pajek, NetworkX which cover basic functionalities of network analysis, and some are advanced in large-scale and complex analysis tasks (e.g., MultiNet). Others are more computational-oriented (e.g., NetworkX, igraph, etc.) or visually interactive (e.g., Gephi, Node XL).

Despite some specialization differences, previous reviews and comparative studies of the aforementioned SNA tools (Akhtar, 2014; Combe et al., 2010; Majeed et al., 2020; Oliveira & Gama, 2012; Van Duijn, 2005) reveal a consistency in the core functionalities supported by these tools. This consistency can largely be attributed to a widely recognized set of network statistics (e.g. node degree/betweenness/closeness centralities, and network density) and algorithms (e.g. shortest path functions and community detection algorithms), as outlined by Wasserman and Faust in their seminal SNA textbook (1994). For instance, Combe et al. (2010) conducted a comparative analysis of SNA software and identified three main functionalities expected of network analysis tools: visualization, indicator-based network statistics (e.g. degree, closeness, betweenness centrality), and functions to detect communities. Their arguments resonated with Oliveira and Gama (2012), who provided an overview of social network analysis (SNA). This work highlighted that despite SNA tools' diverse characteristics, most share common functionalities, including computing descriptive metrics at both local (actor) and global (network) levels, visualizing networks, and detecting communities. Similarly, Mark and Van Duijn (2005) evaluated 23 SNA programs' functionalities in terms of visualizing networks, calculating descriptive network statistics (e.g. centralities), running iterative algorithms (e.g. community detection), and performing statistical modeling based on probability distributions (e.g. permutation tests). Their evaluation criteria were based on the methodological framework established by Wasserman and Faust in their SNA textbook. These observations helped define the network metrics that SNoMaN incorporates.

2.2. Spatial network visualization

Geographic layouts, where nodes are pinned to longitude/latitude coordinates, are common for spatial network visualization (Andrienko & Andrienko, 2010; Guo, 2009; Rae, 2009). While creating a spatial network is a typical task in GIS software, few social network systems support geographic layouts (via pinning nodes to coordinates). However, Gephi (2010) provides a location-based network visualization and several studies have used this plugin for spatial network analysis (Babic et al., 2017; Brinkley et al., 2021; Broux, 2016; Tiwari & Aljoufie, 2021; Wang et al., 2020). Specifically, there are tools and visual techniques designed for flow network mapping such as FlowMapper (Koylu et al., 2023), and a force-directed layout of origin-destination flow maps to avoid flow cross (Jenny et al., 2017, 2018). Hale et al. (2017) discussed the use of force-directed and geographic layouts in a series of tasks. Their controlled experiment results suggest that the choice of network layout depends on the task to be performed, and that a variety of layouts should be available. Schöttler et al. (2021) further reviewed visualization techniques that aim to balance geographic space and network topology, including combined or distorted geographic and force-directed layouts that preserve both forms of information. Finally, Godwin (2022) visualized a network path with a juxtaposed map and force-directed view to support the combination of network layouts for spatial network visualization.

A problem inherent to geographic network layouts is that the visual emphasis of edges is not proportional to importance but is determined by edge distance, which can mislead the viewer. It is also difficult to preserve a global overview and provide local detail (Zou & Brooks, 2019). To avoid difficulties due to scale change and distance constraints, Guo (2007) proposed an optimized matrix layout with spatially dominated ordering to interpret population flow patterns among locations. Matrix views, where origins are rows and destinations are columns in the matrix (or vice versa), are also used to reduce link overlap and show clustering patterns in geospatial networks (Hadlak et al., 2011; Wood et al., 2017; Yang et al., 2016).

2.3. Spatial social network analysis

Spatial Social Network Analysis (SSNA) has broad applications in areas such as criminal networks (Papachristos et al., 2013; Radil et al., 2010; Walther et al., 2023), epidemiology (Emch et al., 2012), accessibility (Li, Walch et al. 2022; Block et al., 2004), humanities (Giordano et al., 2022), and telecommunication networks (Wang et al., 2015), and public health (Christakis & Fowler, 2008; Forati et al., 2023). Spatial social network analysis is part of a growing field with an annual publication growth of 20% per year for the past 20 years (Wu et al., 2022). Andris (2016) facilitated a guide for SSNA by combining social networks as a layer into standard GISystems. They listed methodologies involving node spatialization, node characterization, social and spatial grouping, and topology spatialization. To address the discordance between network analysis and GISystems, Luo & MacEachren (2014) extended the definition of nearness and relationships over the social and geographic space. Sarkar et al. (2016) further characterized existing SSNA methods into three categories: nodal, topographic, and spatial. They also discussed three core different meanings of three co-existing concepts in the two fields, including distance, communities, and scales. In follow-up studies (Sarkar et al., 2019), they formulated several network-level and node-level spatial social network metrics to measure the network expanse and node importance with social and spatial properties together. These works lay theoretical foundations and provide valuable methods and perspectives for SSNA.

Our tool builds on prior GUI-based tools and visual designs that fuse social network analysis with geographical analysis. TwitterHitter (White & Roth, 2010) retrieves Twitter users' spatiotemporal records and their social networks, then uses a map-timeline view and a network analysis view to address different criminal investigation tasks. Similarly, GeoSocialApp (Luo et al., 2011) can be used to visualize flows (such as international trade flows) with a node-link diagram and a bivariate choropleth map. In addition, Sarkar & Yadav (2021) proposed a donut visualization to preserve both topological and spatial structures of spatial social networks. Le et al. (2022) developed PhyloView for phylogenetic data visualization using a linked a propagation tree layout and a geographic layout. These tools provide helpful steps toward software for general SSNA, although they use specific, built-in datasets and offer a limited range of visual analytic components for computational network characterization and correlation analysis.

3. Design goals

In this section, we list three design goals (DGs) for assisting users with exploratory SSNA, including linking network and geographic layouts and contexts, providing established spatial social network analysis metrics, and providing joint visualizations to associate social network with geographic measurements (Table 1). These design goals are derived using an ESDA research framework (Anselin, 1996; Haining et al., 1998; MacEachren & Kraak, 1997; Wei, 2022) and considering tools in exploratory geospatial software, such as GeoDa (Anselin et al., 2009). Anselin defined

Table 1. Three design goals of SNoMaN facilitate investigating SSN by linking, calculating, displaying, and associating SSNA representations and metrics. The SNoMaN column lists components of our tool that support these goals. The SSNA metrics and representations column lists network and geographic metrics, algorithms, concepts, and representations that are employed in these views. The use cases column gives use cases (denoted as C1-C4) of SNoMaN regarding three examples of SSN datasets.

Design Goals	SNoMaN	SSNA Metrics and Representations	Use Cases
DG1	Network View, Map View, Configuration and Function Panel	Force-Directed Layout + Geographic Layout	Food sharing network C2
	Network View, Map View, Configuration and Function Panel	Network Position + Geographical Location	American Mafia network C2
DG2	Statistics Panel	Network Density, Network Diameter, Clustering Coefficient	American Mafia network C1
	Detail View	Node Degree Distribution + Edge Distance Distribution	American Mafia network C4
	Configuration and Function Panel	Degree, Closeness, Betweenness, PageRank + Distance to Center	American Mafia network C2
	Map View, Configuration and Function Panel	Network Community + Convex Hull	Flight network C2 Food sharing network C2
DG3	Centralization Plots	Degree, Closeness, Betweenness, PageRank + Distance to Center	Food sharing network C1
	Route Factor Diagram, Network View, Map View	Network Distance + Euclidean Distance	Flight network C1
	Cluster-Cluster Plot	Group Network Density + Standard Distance	American Mafia network C3

ESDA as "a collection of techniques to describe and visualize spatial distributions, identify atypical locations or spatial outliers, discover patterns of spatial association, clusters or hot spots, and suggest spatial regimes or other forms of spatial heterogeneity" (Anselin, 1996, p. 258). ESDA is especially useful when no strong prior theoretical framework exists, a common scenario in interdisciplinary social science analysis (Anselin, 1999), including SSNA. Following these ESDA paradigms, we derived three design goals, which can be employed to form hypotheses, discover patterns, and suggest associations at the early exploratory stage of SSNA.

3.1. DG1: link network and geographic layouts and contexts

By linking network and geographic contexts and layouts, users can cross-filter network entities of interest within both contexts to examine their locations and geographic surroundings on maps and their connections and positions in networks simultaneously. This is helpful for identifying geographical clusters and isolated nodes, and for examining whether the presence of certain geographical entities (e.g., rivers, mountains) influences the spatial distribution of nodes. Network structures show patterns of connections within a network, such as hierarchies, chains, triads, and rings. Geographically locating nodes and edges can help users contextualize nodes within their surroundings and within the landscape, and connect nodes with their geographical surroundings.

3.2. DG2: provide established spatial social network analysis metrics (e.g., degree, betweenness, closeness centralities)

To facilitate exploratory analyses, the tool should dynamically calculate and report key network and geographic metrics (Section 2.1), such as network density, edge distance distribution, and node degree distribution, on a user-defined sub-sample of nodes or the global network. The tool should calculate the network density of a particular region of interest (as selected by the user), show the average distance between a focal node and connections in an ego-centric network, and highlight a set of connections within a defined distance range. In addition, the tool should support symbolizing these metrics with different visual properties (e.g., color, size, position, and location) and highlight special nodes, paths, communities, and geographic or network regions. For example, showing nodes sized by their degree centrality on the map indicates the locations of powerful, or well-connected nodes. By presenting network paths on the map, they can be interpreted as geographic transfer routes. Coloring nodes by their community (a network-derived metric) and outlining the community's spatial expanse using a convex hull on the map shows the geographical regions of communities.

3.3. DG3: provide joint visualizations to associate social network and geographic metrics

This design goal is derived to facilitate exploring the relationship between network and geographic metrics within SSNs, which usually requires plotting metrics that were computed from separate network and spatial analysis platforms on the same plot statistically; and even then, they are rarely performed. For instance, by comparing nodes' network and geographical centralities, users can explore whether being geographically central yields more connections (as in Andris et al., 2021; Onnela et al., 2011) and search for atypical cases. Associating network and geographical distances can point out whether the number of hops between two nodes will often rise with physical separation (Leskovec & Horvitz, 2014) and discern any anomalies where a new connection could potentially increase travel efficiency. By comparing the spatial dispersion of community members and their level of connectivity, users can discover if groups with dispersed members tend to be loosely connected compared with spatially concentrated groups. To assist in investigating these relationships, the tool should support associating network and geographic metrics on the same plot while allowing users to interactively select interesting nodes or groups of nodes to explore.

This list is not fully representative of all social and spatial network analysis combinations, as it does not include inferential statistics, temporality, simulation, and other intermediate and advanced SSNA goals. However, these goals are established to include primary descriptive statistics for both spatial and social analysis. In the next section, we address these design goals with new diagrams and functionalities implemented in SNoMaN.

4. SNoMaN

SNoMaN has one Configuration and Function Panel (Figure 1a) and four primary views: 1) Network View (Figure 1b), 2) Map View (Figure 1c), 3) Detail View (Figure 1d), and 4) Comparison View (Figure 1e).

4.1. Import dataset

The dataset import panel allows users to load example datasets or import their own spatial network datasets to SNoMaN. The supported network dataset format is a comma separated value (CSV) file for node and edge lists. For the node list file, each row is one node entity and each column represents one attribute of the node. Three attributes are required for a node list CSV file: ID, longitude, and latitude, although more node attributes are welcome. The edge list file needs two columns called "source" and "target" (based on Gephi's standards), where the first column contains the ID of a node of interest, and the second column contains the ID of a node that it is connected to. The tool currently only supports unweighted, undirected networks.

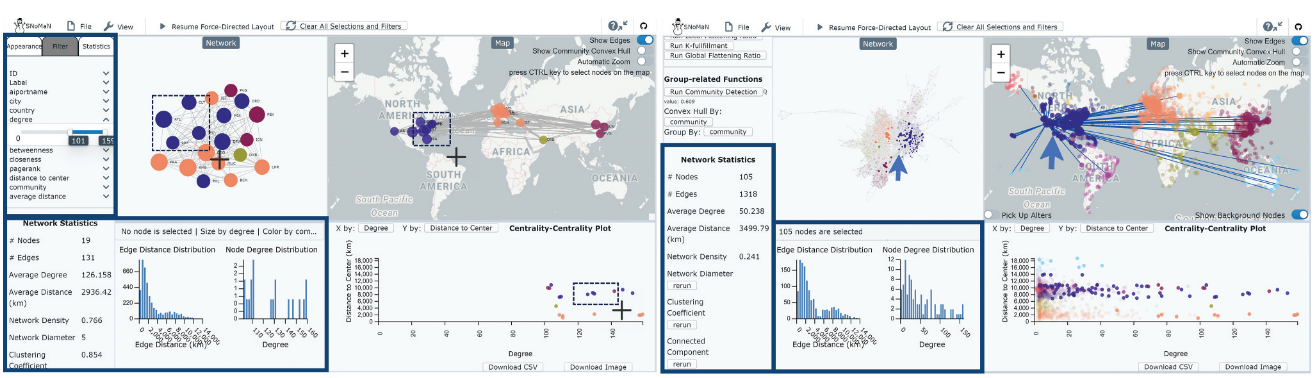
4.2. Network view

The Network View (Figure 1b) presents a force-directed layout (Dwyer, 2009) of the network (DG1). A force-directed layout, also known as a sociogram, is a common method for displaying network topological structures in social network analysis. It is calculated by assigning spring-like forces on edges. The force simulation pulls connected nodes together, pushes disconnected nodes apart, and tends to place highly connected nodes toward the center of the drawing (Bannister et al., 2013). The relative network position of nodes in a force-directed layout algorithm reflects topological closeness. In addition, the Network View supports dragging and pinning nodes to modify the layout as needed. Users can drag to select a set of topologically close (i.e., network position) nodes, click to select a node and its connections (i.e., an ego-centric network (Perry et al., 2018)), and hover to highlight nodes and connections across views. SSNA metrics (e.g., edge distance distribution, node degree distribution, network density) of users' above selections will be computed and presented under the Detail View (Figure 2) (DG2). In this way, users can calculate and reflect on statistics about, for instance, a node's set of friends, including their interconnectedness (known as a clustering coefficient), the average distance between friends, and edge distance distribution of friends.

4.3. Map view

The Map View (Figure 1c) presents the Geographical Layout of the network on a basemap showing the underlying geography. A geographical layout plots nodes on a basemap, allowing users to see the spatial distributions of network entities and examine the network's dynamics in a place of interest (Faust et al., 2000; Loginova et al., 2020; Radil et al., 2010) (DG1). Different basemap options are provided from the OpenStreetMap database, such as terrain and transportation. The Map View provides options to hide or show labels, edges, and convex hulls when needed and to brush or draw polygons on the map to select nodes.

In contrast to the relative positions of nodes in the Network View, the absolute locations of nodes on maps are determined by their coordinates (i.e., longitude and latitude). The fixed locations do not allow interactions like dragging and pinning to adjust the layout. The Map View supports similar interactions as the Network View. For



a) b)

Figure 2. Users can use filter widgets, brushing and clicking interactions to cross-filter network entities of interest within both contexts to examine their locations and geographic surroundings on maps and their connections and positions in networks simultaneously (DG1). In addition, SSNA metrics (e.g., edge distance distribution, node degree distribution, network density, network diameter) of users' selections will be computed and presented under the detail view (DG2). a) The filter panel allows users to select and group nodes based on different criteria, such as labels, centrality ranges, and community associations. Users can also brush in the network view or map view to select a set of topologically or geographically close nodes. b) Users can click to select a node and its connections (an ego-centric network).

instance, users can brush on the map to select nodes that are geographically proximate or click to select a node and its connections. Network and geographic metrics of users' selections are subsequently automatically computed and presented under the Detail View (DG2) (Figure 2). When the "Automatic Zoom" feature is on, the map will automatically zoom to the bounding box area of the nodes within the current selection.

4.4. Detail view

The Detail View (Figure 1d) presents network statistics as text and histograms including the number of selected nodes and edges, the average distance, network density, network diameter, edge distance distribution, and node degree distribution of a user-defined subnetwork by selection and other filter interactions in other views (DG2).

Network density is the proportion of connections in a network relative to the total possible connections; the value ranges from 0 (no connections) to 1 (all possible connections are present, i.e., a clique). Network density provides an overview of how well connected the network is. Network diameter is the maximum (or mean) number of "hops" it takes to reach any two nodes in a network and is a measure of reachability in SSNs. Diameter is also a key factor in the detection of Small World networks (Watts & Strogatz, 1998), which indicate groups of relationships with interconnections between the groups.

An edge distance distribution is often represented by a histogram of edges binned by their lengths. This histogram can suggest whether nodes are dispersed on the map and how distance affects the formation of social ties (as in Scellato et al., 2010). For example, a uniform distribution suggests a weak relationship between distance and connection frequency. Gaps between bars suggest possible natural barriers and boundaries between nodes. A node degree distribution summarizes the frequency of node degrees of a network. A long tail distribution provides evidence of preferential attachment (Barabási & Albert, 1999), which suggests that the network has a few nodes with many connections and many nodes with few connections. A left-skewed distribution suggests the opposite. The histogram supports a brush interaction to highlight edges within a certain distance range or degree range.

4.5. Configuration and function panel

4.5.1 Compute SSNA metrics

After an SSN dataset is imported, SSNA metrics of the whole network and individual nodes are automatically computed and added to the dataset and users can subsequently compute metrics for communities with user-specified parameters using the group-related functions under statistics panel (DG2). Network statistical summaries (e.g. network density, network diameter) are computed and displayed under the statistics panel.

4.5.1.1 A) Calculate and encode nodes' network and geographical centralities (e.g. degree, closeness, betweenness centralities, and distance to center) (Figure 3a).

Standard node centralities in social network analysis include degree, closeness, betweenness, and PageRank centralities (Section 2.1). Different types of node network centrality metrics can be used to identify the network role of a node, such as nodes that serve as hubs (i.e., high degree centrality), or those that serve as a bridge or connector (i.e., with high betweenness centrality). Closeness centrality measures the number of hops, on average, a node is from all other nodes in a network. It is calculated as the reciprocal of the average of the length of the shortest paths between the node

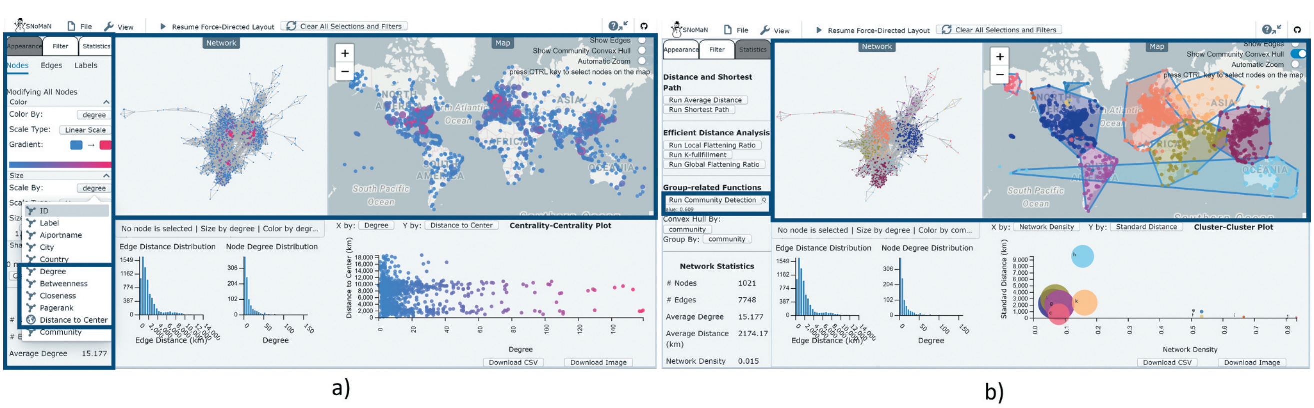


Figure 3. Users can compute SSNA metrics under the configuration and function panel (DG2). a) After an SSN dataset is imported, SSNA metrics of the whole network and individual nodes will be automatically computed and added to the dataset. Users can access these metrics through the configuration and function panel for visual encoding. b) Users can run the community detection algorithm to identify network communities and visualize their spatial expanses using the statistics panel.

and all other nodes in the graph. Betweenness centrality is the number of shortest paths between any two nodes that pass through a focal node. It measures the extent to which a node is important for connecting other nodes. PageRank is calculated by the number and value of incoming links to a node to estimate its relative importance. The more high-value nodes that link to the focal node, the higher its PageRank score will be.

A node's *geographic* centrality is a measure of a node's relative geographic location to the network. The distance to center metric measures a node's distance to the network's mean center (i.e. a conceptual center calculated as the average longitude and latitude of all nodes). A low value of distance to the center can indicate that the node has better access to other nodes and can be more easily reached, although exceptions are common.

4.5.1.2 B) Identify network communities and visualize their spatial expanses using the statistics panel (Figure 3b). Community SSNA metrics are not automatically computed after the dataset is imported because they are based on user-specified parameters. A network community can be identified a-priori by categorical features, or communities can be defined using community detection algorithms (see Fortunato & Hric, 2016). These algorithms (which are also referred to as modularity or graph partitioning algorithms) are used to divide networks into subgraphs wherein the subgraphs contain nodes that are more likely to connect within the subgraph than with nodes in another subgraph. For SSNs without predefined communities, users can detect communities using the Louvain algorithm (Blondel et al., 2008) by clicking on the Community Detection button. The output of the algorithm is as follows: a Q-value that represents the strength of the partition, ranging from −1 to +1; labels for nodes signifying which community (i.e. subgraph) they are assigned to; how many subgraphs exist, and how many nodes each subgraph has. After running the community detection algorithm, the community association result will be assigned to each node and presented by color coding in all views. The node color will then change in the Network View and the Map. In addition to using the community detection algorithm to partition networks, users can also use a predefined community category, i.e., a categorical node attribute, for community analysis.

The spatial expanse of a community is defined by the activity spaces or locations of its members (Hu et al., 2020), representing the geographical reach of the community. Showing community expanses can help users identify where communities' activity areas overlap and complement each other (Wang et al., 2015). The convex hull button implements an algorithm (Barber et al., 2011) to outline the spatial expanse of each community (DG1) given a user-selected attribute for grouping nodes or based on the network partition results. The algorithm returns a set of polygons that encapsulate the nodes of each group (i.e., the minimal convex set containing all group members on the map (Preparata & Shamos, 2012)). To avoid overfitting a shape by including outlier nodes that stretch the convex hull, we use a standard cutoff value of Z-score >3 to eliminate outliers for each group before drawing the convex hull.

4.5.2 Network Appearance Panel

This panel allows users to specify the visual encoding for network entities (DG2), including node color and range, node size and scale, node shape, edge color, and label size. After metrics are computed, the results (e.g. node degree, node community association, edge distance) are assigned to each node or edge and can be accessed through this panel for visual encoding. Changes to the visual encoding will update all three views.

4.5.3 Filter Panel

The Filter Panel allows users to select nodes based on different criteria, such as labels, centrality ranges, and community associations (Figure 2a). Example tasks include identifying the geographic location of high-degree nodes or a geographically peripheral network community and calculating the average distance between high-degree nodes (DG2). Users can use numeric sliders or category drop-down menus to filter nodes by attribute. When multiple filters are applied, the intersection of the results will be shown.

4.6. Comparison view

The Comparison View (Figure 1e) is a scatterplot, where color, size, and position along the x-axis and y-axis are used to encode different variables. This element is helpful for comparing geographic metrics with network metrics simultaneously. Users can select metrics of interest from a drop-down menu. The plot supports brush selection to filter nodes in both Network View and Map View (DG3). Below, we list a set of three types of special plots that fuse SNA and GIS themes under the concepts of being central, separated, and clustered, respectively.

4.6.1 Centralization plot (Figure 4a)

The Centralization Plot compares nodes' network (e.g., degree, betweenness, closeness centralities) and geographic (e.g., distance to center) metrics. Each dot

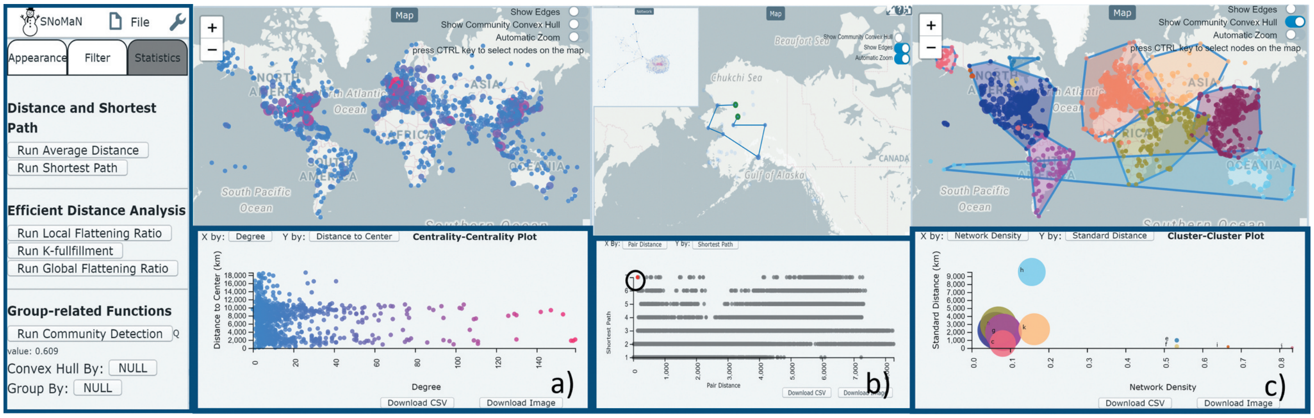


Figure 4. The comparison view implements a scatter plot to assist users in exploring relationships between different network and geographic metrics (DG3). The plot supports brush and click interactions to highlight cases of interests in both network view and map view. a) The Centralization plot compares nodes' network (e.g. degree) and geographic (e.g. distance to center) metrics. b) The Route Factor Diagram compares network distance (i.e. shortest path) and Euclidean distance between pairs of nodes. c) The Cluster–Cluster plot shows relationships between group size, group spatial dispersion, and group network density.

represents one node and shares the same color encoding as other views. In Figure 4a), the x-axis encodes the node's distance to the network center, and the y-axis encodes the node's degree. A downward trend implies that well-connected nodes are in the "thick of things," in the middle of the map, while an upward trend implies that well-connected nodes are mostly located in peripheral areas. This plot helps users explore how geographic convenience might be linked to an entity's power.

4.6.2 Route factor diagram (Figure 4b)

This plot associates the network distance (i.e. shortest path) and Euclidean distance between each pair of nodes (Sarkar et al., 2019). A path between (any) two nodes can be measured by Euclidean distance (or cost distance) or by "hops" in a network, which is known as network distance, i.e. a discrete integer measurement signifying the length of the shortest path (in hops) between two nodes (Bouttier et al., 2003) (Sometimes the latter metric is called a "geodesic" (Bouttier et al., 2003), which may likely confuse readers who try to reserve geo-terms for geographic features). The ratio of geographic distance to cost network distance is also known as a route factor (Hay, 1973; Sarkar et al., 2019).

Euclidean distance and network distance are automatically computed for every pair of nodes. Users can create a Route Factor Diagram, where each dot represents a route between all pairs of nodes. The x-axis encodes the length of the shortest path and the y-axis encodes the geographic distance between two nodes. An upward trend implies that social closeness is positively correlated with geographical nearness, while interesting anomalies might be found at the top-left corner (i.e., geographically close but topologically far) and at the bottom right corner (i.e., geographically far but topologically close). Nearby neighbors who are separated by many hops (i.e., cases that fall into the top-left corner) could be good candidates for a new connection, as this would reduce the network diameter with a small change. Users can click on individual cases to interactively highlight the route in both Network View and Map View. This allows them to see routes between nodes that are disconnected by many or few hops and, by examining the rendered geographic path, the geographic features near the path.

4.6.3 Cluster–cluster plot (Figure 4c)

This plot helps users explore the relationship between group size, group spatial dispersion, and group network density. We use standard distance (Flury & Riedwyl, 1986) to measure a group's spatial dispersion. This scalar value is the standard deviation of the distance of each group member from the group mean center. It tells us whether members within a community tend to be concentrated near each other or scattered. After running the Cluster–Cluster algorithm based on user-specified group categorization, the results will be presented in Comparison View. Each dot in this plot represents one group. The x-axis encodes the network density of the group and the y-axis encodes the standard distance of the group. The dot size encodes the group size and the color encodes the community name (the same color scheme as other views). Spatially clustered and densely connected communities can be found in the chart's lower right corner. On the contrary, spatially dispersed and loosely connected communities can be found in the chart's upper left corner. A downward trend line indicates that geographically dispersed community members are not inter-connected.

4.7 Implementation

SNoMaN is implemented using D3.js, React.js, and MobX for front-end development and Flask for the back-end framework. It is built on Argo Lite (Li et al., 2020), a network visualization tool. NetworkX (Hagberg et al., 2008) is used for network metrics computation. SNoMaN is deployed as an open-source tool and is freely accessible here: https://snoman.herokuapp.com/.

5. Use cases

We use three datasets: an American Mafia network, a world flight network, and a food sharing network to demonstrate how users can use SNoMaN to conduct SSNA. We describe the datasets, analysis tasks, and insights derived under each dataset briefly in Table 2.

5.1. The American mafia network

Our data set on the American Mafia Network shows connections within a criminal network of mafiosi in the United States in the 1960s (DellaPosta, 2017). It includes 680 Mafia members and 2,699 edges signifying criminal associations. More than half of mafioso (around 350 members) live in New York City. This network was created from the coding of paper dossiers that were compiled by the United States Federal Bureau of Narcotics. Mafia member location information is derived from their residential addresses, and their connections are their "known associates."

5.1.1. C1 network statistics

After the dataset is uploaded, SNoMaN automatically computes metrics for the whole network and individual nodes. The Mafia network has an average degree of 7.9, the network density is 0.012, the network diameter is 8, the average clustering coefficient is 0.375, and it contains two connected components. From the Map View (Figure 5a), users can see that Mafia families are spatially clustered, suggesting that distance is an obstacle for growing groups.

5.1.2. C2 the network hub, New York City

New York City is home to many members of the Mafia. Users can label these nodes and apply the degree filter to gray out low-degree nodes. It becomes clear that all high-degree nodes (degree > 40) are in New York City, including the family bosses (e.g. John Ormento, Vito Genovese, Anthony Strollo, Salvatore Santoro) (Figure 5b). Shown on the Cluster-Cluster View, the top three largest families are the Genovese, Lucchese, and Gambino families (Figure 5d). Upon clicking on

Table 2. Three SSN datasets, use cases, and insights derived from the tool.

Spatial Social Network Dataset	Questions within use cases, and insights derived from the tool
The American Mafia network (680 nodes, 2699 edges) (DellaPosta, 2017) **Nodes**: mafioso **Edges**: criminal associations **Node Attributes**: ID, Name, Family Name, Location (Longitude, Latitude)	*C1: Are mafiosi densely connected? Are family members geographically clustered?* Mafia families are densely connected. Most intra-family connections happen within the same city. *C2: Where is the organizational and geographical center of the network?* The organizational and geographical center is in New York City. *C3: Are geographically clustered families more dense than geographically dispersed families? How does family size affect group density?* Spatially clustered small families are more densely connected than large families, with a few exceptions. *C4: Which mafiosi live far from their criminal associates?* Some families have satellite members, e.g. Joe Bonnano.
The global flight network (1,022 nodes, 7,748 edges) The U.S. flight network (184 nodes, 1,320 edges) (OurAirport, 2017) **Nodes**: airports **Edges**: direct flights **Node Attributes**: ID, Name, Location (Longitude, Latitude)	*C1: Are geographically close airports also efficiently connected in the flight network?* Seven hops are required to reach a nearby airport in Alaska but one direct flight connects Hawaii with most major airports. *C2: Do the airline connections form regional communities?* Yes, country and continent boundaries reduce connection frequency across the administrative borders.
Food sharing network (105 nodes, 153 edges) (Edwards, 2020) **Nodes**: food sharing organizations **Edges**: collaborative connections **Node Attributes**: ID, Name, Location (Longitude, Latitude)	*C1: How should organizations coordinate locations for better logistical efficiency?* Moving "Feeding America of Southwest Virginia" organization from the network periphery to the center can reduce the travel costs of the network. *C2: How do organizations collaborate to form service areas?* Nearby organizations work closely to serve different areas. One community covers the southwest triangle area.

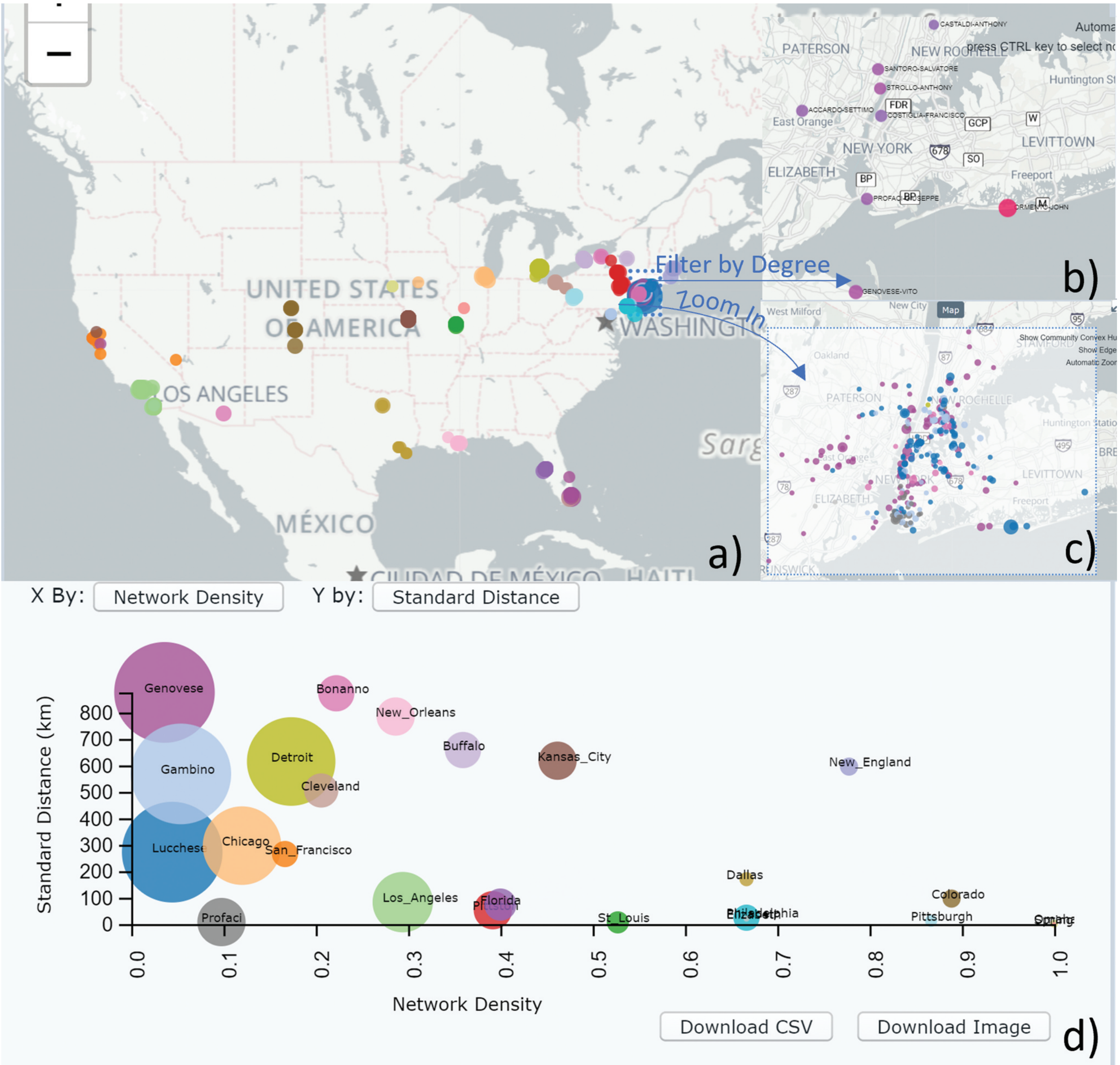


Figure 5. a) The map view shows locations of mafiosi (node color by family name, size by degree value). b) All high-degree nodes (degree > 40) are in NYC. c) The top three largest families, the Genovese, Lucchese, and Gambino families, are all dominant in NYC. d) The Cluster–Cluster view plots the standard distance and network density of each mafia family (node color by family name, size by family size).

a family dot, the map automatically zooms to fit the active area of the family, and users can see these families are all dominant in NYC (Figure 5c).

5.1.3. C3 family tightness and spatial dispersion

The Cluster–Cluster Plot (Figure 5d) shows that almost all densely connected groups are spatially clustered small groups (shown at the right-bottom corner of the chart). One exception is the New England family, whose members are scattered across New England, as well as Miami, Florida. Family size also affects the network density; most large families tend to maintain footholds in a variety of locales, resulting in loose connections across cities. SNoMaN facilitates this analysis by computing the density of the subgroups on the fly and populating this scalar value as an option on the scatterplot as an independent variable.

5.1.4. C4 Joe Bonnano, an exiled Mafia Capo

Most families have their own city bases, and family members are clustered geographically within the city

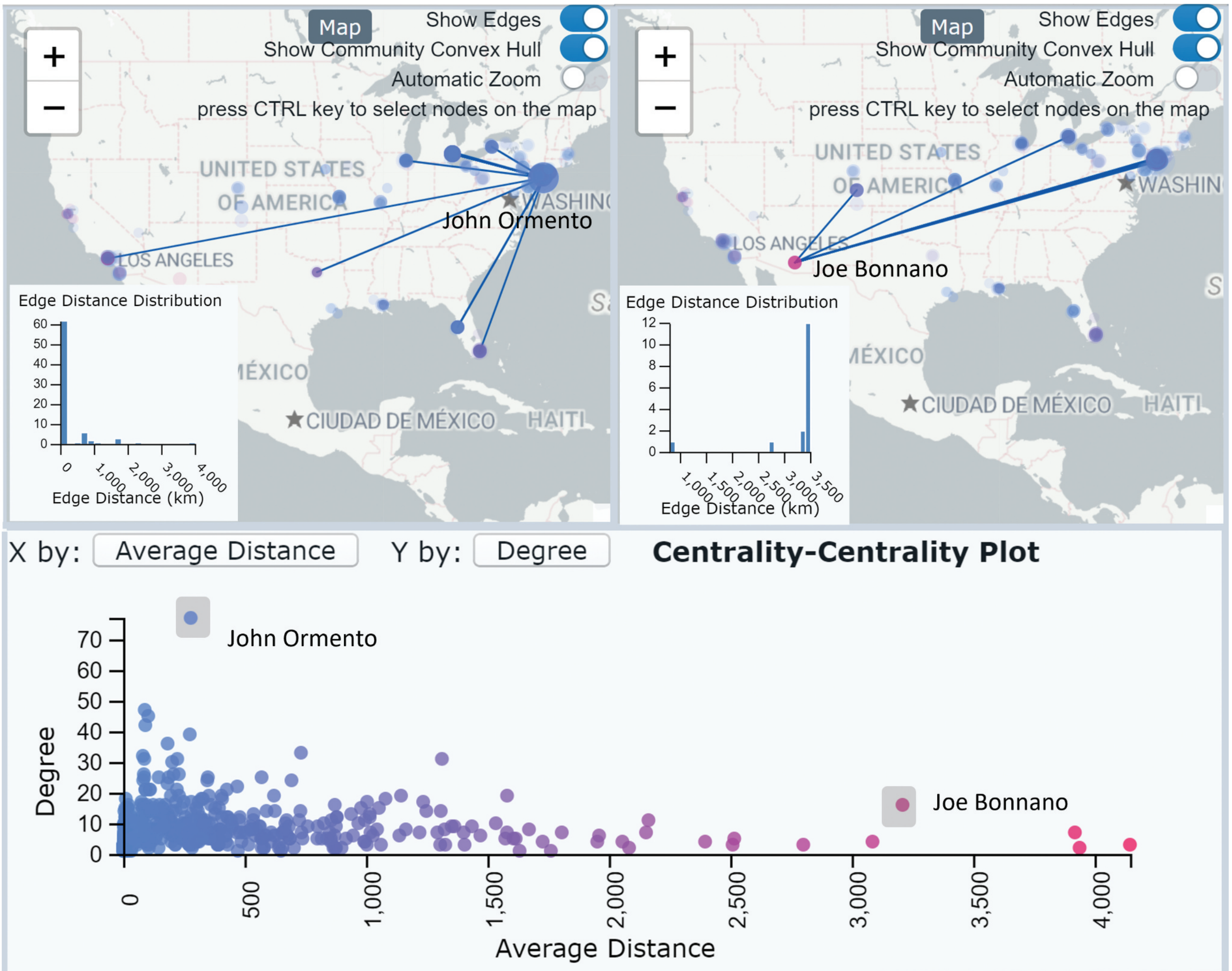


Figure 6. The x-axis of the scatterplot displays the average Euclidean distance of nodes' connections. John Ormento exemplifies a scenario where the connected mafiosi are mostly close to the focal node, illustrated by its edge distance distribution histogram as well. In contrast, Joe Bonnano exemplifies a scenario where most of its connections are far from it, also illustrated by its edge distance distribution.

(Figure 6). As a result, individuals' connections tend to be located close to them. However, there are some "satellite" members who live farther away from the center of the network than their connections. Users can compute and plot the average Euclidean distance from each node to its connections in the software. In Figure 6, the x-axis of the scatterplot displays the average Euclidean distance of nodes' connections. A user can see that John Ormento's connections are mostly close to him, illustrated by his edge distance distribution as well. In contrast, Joe Bonnano is an exception, as most of his connections are distant; Outside evidence suggests Joe Boannano moved to Arizona after conflicts with the Bonnano family. By selecting the node, the edge distance distribution can be used to investigate the distances of Joe Bonnano's connections.

5.2. The flight network

The world flight network (OurAirport, 2017) includes 1,022 airports (geolocated by longitude/latitude coordinates) and 7,748 edges indicating direct flights between airports. The data was collected by assembling government datasets and individual contributions. The flight network is not explicitly a social network, but we include it to help teach users to use our tool. Most users have some intuition about flights (which allows them to explore preconceptions and hypotheses) but may lack intuition about smaller case studies.

5.2.1. C1 exploring the route factor diagram, Alaska chains vs Hawaii direct flights

Network distance and geographical distance are expected to be positively correlated, but there are often

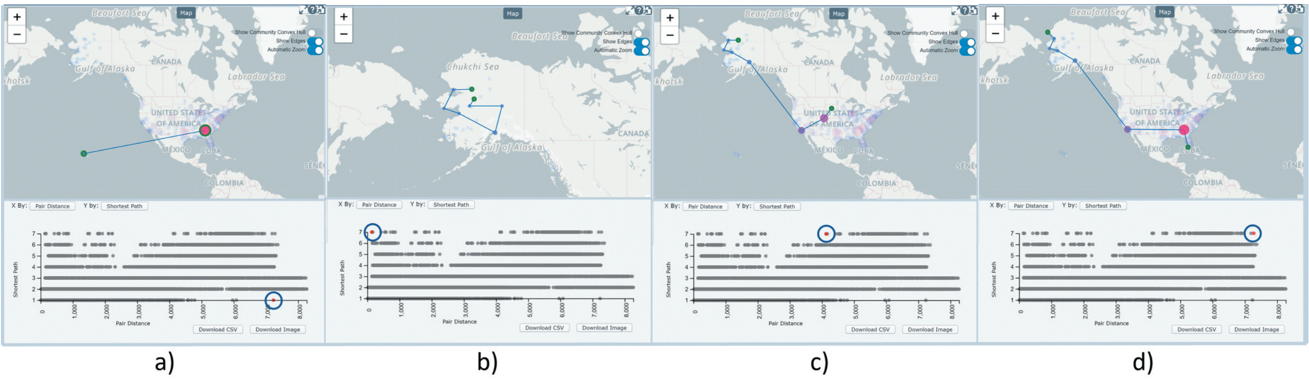


Figure 7. The route factor diagram compares network distance with geographical distance. Example paths are highlighted on map view. a) “far friends:” spatially distant airports, Hawaii and Atlanta, are connected by direct flights; b) “near strangers:” nearby airports in Alaska are separated by several network hops; c),d) the chain pattern in Alaska increases the network diameter to seven. (nodes on the map represent airports, color and size signify degree; nodes on the route factor diagram represent paths, interactively highlighted as red.)

exceptions. The Route Factor Diagram (Figure 7) is a scatterplot comparing pair distance (in kilometers) with shortest path (in hops) to help users evaluate the network’s spatial efficiency and find potential missing connections. We zoom into the U.S. flight network (184 nodes, 1,320 edges), to detect chain patterns using the Route Factor Diagram. In this diagram, each dot is an edge. Upon hovering over a dot, the diagram shows that the most distant direct flight connects Atlanta and Hawaii (Figure 7a). The chart also shows airports that are spatially close but topologically far (i.e. “Near Strangers”: low geographic distance but high network distance) (Figure 7b). The route is highlighted in Alaska, and users can see the chain patterns: though two airports are close to each other, there is no direct flight between them. These airports are connected subsequently in a long independent chain, without direct flights or central hubs for transfer. The chain pattern increases the network diameter because almost all the paths that are more than three hops are to or from Alaska (Figure 7c) and d). The findings suggest potential solutions to increasing spatial efficiency, such as developing a local transfer hub in Alaska.

5.2.2. C2 geographical divisions in network communities

The community detection algorithm on the world flight network returns 12 communities with a Q value of 0.602 and each airport is colored by its community. The results of community detection were performed relying solely on connections between airports and using no spatial information. The community detection results represent groups of nodes that are more tightly interconnected with each other than with nodes outside the community. Despite the lack of spatial information about the nodes during the community detection process, mapping the results shows that community boundaries largely align with country and continent borders. This observation suggests that most flights stay within local regions and geographical boundaries tend to separate airport communities into different groups (Figure 1c). Moreover, airports near another country’s borders may be close to each other but often belong to different communities; this provides evidence of how continental administrative boundaries can affect network connections. In addition, we find community heterogeneity within Africa. North African airports often connect with European airports, while southeast African airports connect mostly with Southwest Asian airports, and West African airports form a separate community.

5.3. The food sharing network

The Thrive network (Edwards, 2020) is a collaborative network of food sharing organizations in rural, southwestern Virginia. It includes 105 organization nodes and 153 edges representing collaboration relationships. Organizations are mapped by longitude and latitude. SNoMaN identifies service areas of communities and the relationship between node location and node role in this dataset.

5.3.1. C1 coordinate location with importance

The Centralization Plot (Figure 8a) illustrates that nodes with higher degree are closer to the network center, which suggests that geographic closeness may facilitate collaboration. For instance, the Community Foundation of the New River Valley (CFNRV), which has the highest network degree, is located at near the network center (Figure 8b). However, one exception, Feeding America

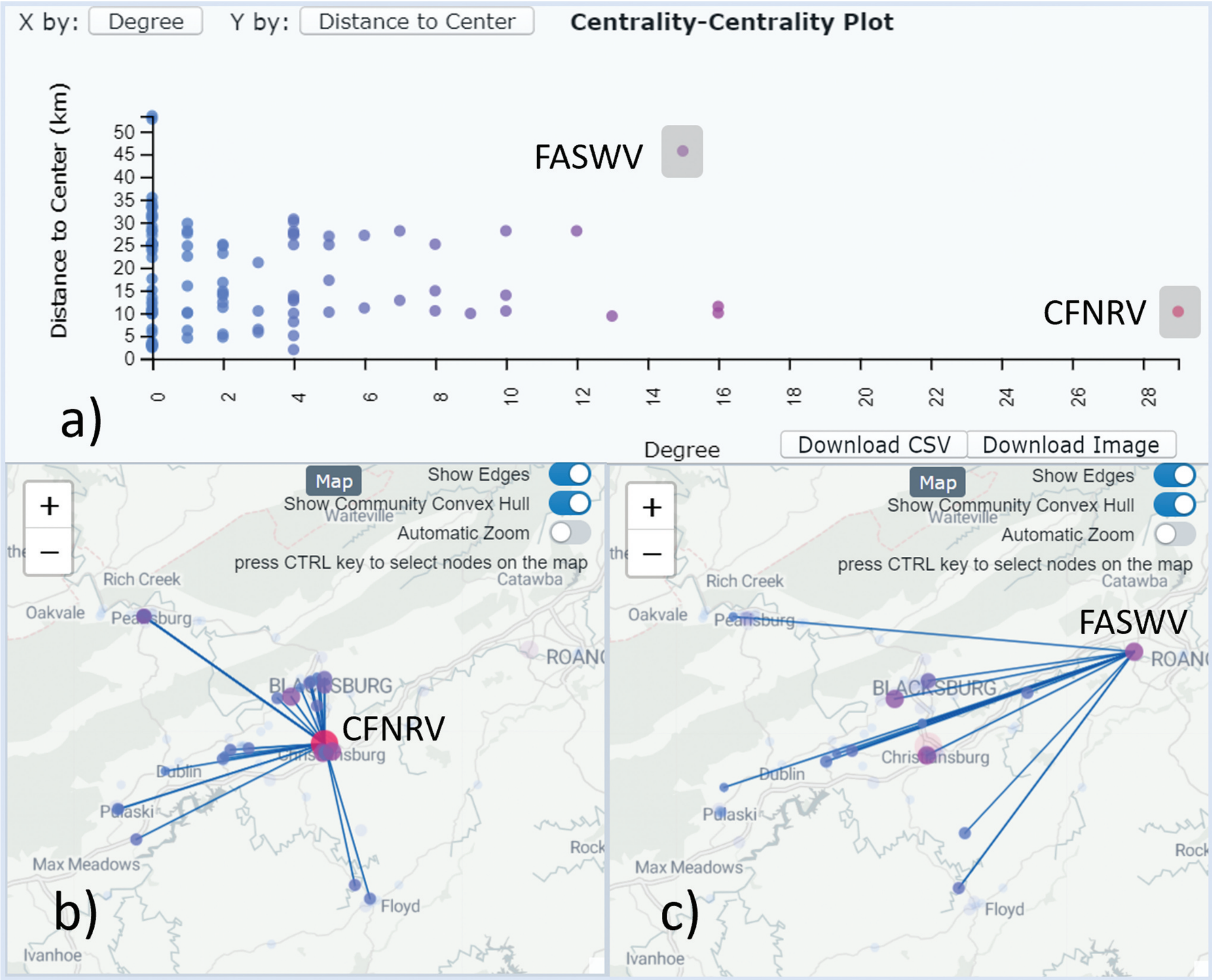


Figure 8. a) The centralization plot suggests node degree is negatively correlated with distance to the network center. b) The node with the highest degree, Community Foundation of the New River Valley (CFNRV), is at the network center. c) One exception is feeding America of Southwest Virginia (FASWV) in the northeast.

of Southwest Virginia (FASWV), despite having many connections, is located at the periphery of the geographic map area (Figure 8c). In this case, we may suggest that this organization moves closer to the city center to reduce the cost of transporting food between organizations, or plan for direct transport to an intermediary.

5.3.2. C2 service area of the community

Using the community detection function and convex hull visualization, users can explicitly show the service areas of certain groups. The community algorithm partitions the network into seven communities with a Q value of 0.452. Users can choose to display the convex hull of each community. Users can see that the orange community forms a "service area" in the south part of the study area that overlaps with the green community (Figure 9). The overlaps between their service areas indicate that these areas may not be coordinated for optimal logistics, but are predicated on personal connections or other factors.

6. Discussion

After creating and exploring SNoMaN, we note some reflections. First, regarding the design needed to fuse two mature analysis methods (SNA and GIS), we found that network (sociogram) node attributes were more natural to display on the map than the display of geographic attributes in the network view. For example, it was more facile to color map nodes by their network

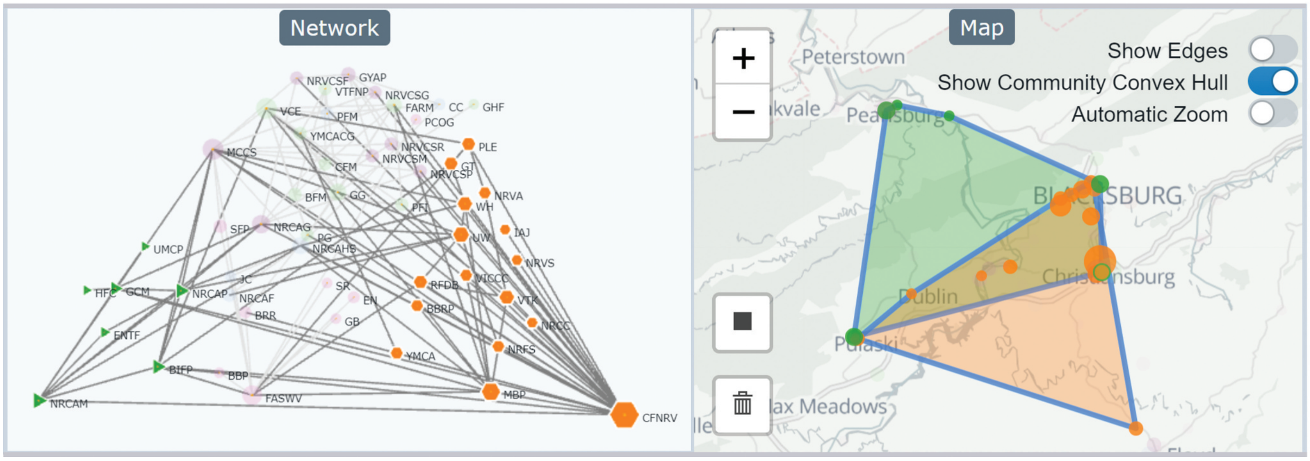


Figure 9. The convex hull in the Map View outlines the "service area" of a set of organizations that share food. The nodes in the Map View and network View are colored by their community association and sized by degree.

betweenness centrality than to encode distance onto the sociogram graph. However, this observation can be balanced as we refine and add more features to the tool. For example, we can experiment with encoding distance as edge thickness on the sociogram.

Secondly, we have observed that a lack of a lingua franca and the interchanging use of similar concepts with different meanings in GIS and SNA could hinder the effectiveness and users' interpretation of our new types of visual diagrams. For example, in the Centralization Plot, the "distance to center" metric can confuse users, as the term "center" may suggest the geographical center of a study area. However, it refers to the network mean center, calculated as the average longitude and latitude of all nodes. It is a conceptual center that shifts toward where most nodes are located. Similarly, users unfamiliar with SNA techniques may need clarification on whether spatial information is used in detecting and mapping network communities. For instance, how does the algorithm partition networks into groups, and is their alignment with geographical boundaries is the cause or the effect? However, providing visual depictions can aid in explaining the concepts in these diagrams. For the Cluster–Cluster Plot, visualizing the convex hull of each community can help users with little GIS background understand the standard distance, where a long stretched convex hull corresponds to a high standard distance, and a concentrated small convex hull means a low standard distance. The Route Factor Diagram tends to be the most straightforward and easy-to-understand diagram. It generates new insights that surprise users, such as how two geographically close nodes can be distant in the network. Visualizing the paths between two nodes on a map makes what the two distance metrics and the plot represent easier to grasp.

Third, we found that it was beneficial to include sample datasets at very different geographic scales, where the connections have different meanings (e.g. food sharing vs. a flight). Not only did this allow us to test the tool with different map scales, but it revealed that not all tasks were useful for each network. For instance, community detection showed key structures in the Mafia network, but was less helpful for uncovering patterns in the food sharing network – as the network is not built around distinctive groups; the Route Factor Diagram was helpful in the food sharing network and flight network, but not for the Mafia network – as the goal of the Mafia network is not to improve global efficiency, per se. This was in interesting finding that will help us develop better rules about which methods are appropriate for which types of networks, given the network's assumptions and structures.

Fourthly, in addition to the three case studies, we have collaborated with SNA and GIS researchers who analyze SSNs to conduct workshop user studies for evaluating the tool. The results of the user study are omitted in this paper as there is more information to be reported from that dataset than what can be fit into one single paper. Potential study questions include how users' expertise in SNA and GIS affects their usage experience of the tool in this dual disciplinary area, the advantages and disadvantages of SNoMaN compared to existing tools for SSNA (e.g. Gephi, ArcGIS), and what new insights users can generate when they interact with the network using both geographic and connectivity measurements.

Finally, SNoMaN has a number of limitations that we hope to address in our future work. First, we chose a web-based approach to reduce users' setup and installation efforts and increase tool accessibility. However, processing and rendering data on websites does not protect data privacy, which may be a significant limitation for some social network analysts. The software is also poorly suited for large networks, which GIS analysts often encounter (such as origin-destination data), which can hinder certain types of analysis. Second, the map view can be easily cluttered. We hope to create a more effective map visualization by using multi-scale views, inset maps, edge bundling, and the ability to group and collapse nodes in a force-directed layout based on geographical hierarchy (i.e. county, city, state, country). We can also adopt a matrix network representation alongside the map and network. Lastly, while SNoMaN is an initially promising step toward supporting low-level exploratory SSNA tasks, there are many more sophisticated metrics and modeling techniques for high-level SSNA tasks. These include simulation of networks over time, the ability for nodes to absorb spatial data, the execution of computational models, and the calculation of distance metrics using travel time instead of Euclidean distance. Currently, the program supports visualization goals, but does not provide inferential statistics. It also is not built for saving routines that can be re-run, which is a benefit of scripting approaches.

7. Conclusion

Spatial social network analysis encompasses a burgeoning set of tasks, methods, workflows, and metrics that can help researchers learn about how power is distributed over geographic space and how connections may be impacted by distance and spatial impedance factors. The small but growing field of analysis combines social network analysis and GIS to uncover new principles about how people and organizations assemble within the boundaries and provisions of geographic space. However, SSNA research lacks formalized procedures and tools that can support simultaneous network and geographical visual analytics.

Here, our goal was to facilitate SSN exploration with a tool that supports metrics from both social and spatial disciplines. The result is a more cohesive view of two worlds that explore networks. Our use cases demonstrate how SNoMaN can facilitate ESDA and generate insights for three SSNs. We hope this work will increase awareness of how geographical variables are intertwined with small social network connections, and support users as they explore their own SSN datasets in SNoMaN.

Acknowledgments

The authors wish to thank Shao-Peng Wu for his assistance in enhancing the functionality of our tool, which is important for the successful development of this software. In addition, we thank Jaimie Kelly, Daniel DellaPosta, and Patrick Park for providing helpful sample data. Thank you for your support and collaboration.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Directorate for Social, Behavioral and Economic Sciences [NSF BCS-SBE-2045271].

ORCID

Sichen Jin http://orcid.org/0009-0006-3471-4716
Clio Andris http://orcid.org/0000-0002-8559-5079

Data availability statement

The information of the tool, code, dataset that support the findings of this study is available at 10.6084/m9.figshare.26307751. Specifically, the three datasets used to demonstrate the use cases of our tool in this paper are from the following resources:

- The Mafia network dataset supporting the findings of this study is available within the article (DellaPosta, 2017) and its supplementary materials.
- The flight network dataset is openly available in the repository "OurAirports" at https://ourairports.com/data/.
- The food sharing data supporting the findings of this study is available within the article (Edwards, 2020) and its supplementary materials.

References

Akhtar, N. (2014). Social network analysis tools. In *2014 Fourth International Conference On Communication Systems And Network Technologies* (pp. 388–392). https://doi.org/10.1109/CSNT.2014.83

Andrienko, N., & Andrienko, G. (2010). Spatial generalization and aggregation of massive movement data. *IEEE Transactions on Visualization and Computer Graphics*, *17* (2), 205–219. https://doi.org/10.1109/TVCG.2010.44

Andris, C. (2016). Integrating social network data into gisystems. *International Journal of Geographical Information Science*, *30*(10), 2009–2031. https://doi.org/10.1080/13658816.2016.1153103

Andris, C., DellaPosta, D., Freelin, B. N., Zhu, X., Hinger, B., & Chen, H. (2021). To racketeer among neighbors: Spatial features of criminal collaboration in the American mafia. *International Journal of Geographical Information Science*, *35*(12), 2463–2488. https://doi.org/10.1080/13658816.2021.1884869

Andris, C., Liu, X., & Ferreira, J., Jr. (2018). Challenges for social flows. *Computers, Environment and Urban Systems*, *70*, 197–207. https://doi.org/10.1016/j.compenvurbsys.2018.03.008

Anselin, L. (1996). *Interactive techniques and exploratory spatial data analysis*. *Regional Research Institute Working Papers*.

Anselin, L. (1999). The future of spatial analysis in the social sciences. *Geographic Information Sciences*, *5*(2), 67–76. https://doi.org/10.1080/10824000909480516

Anselin, L., Syabri, I., & Kho, Y. (2009). Geoda: An introduction to spatial data analysis. In M. M. Fischer & A. Getis (Eds.), *Handbook of applied spatial analysis: Software tools, methods and applications* (pp. 73–89). Springer. https://doi.org/10.1111/j.0016-7363.2005.00671.x

Babic, M., Fichtner, J., & Heemskerk, E. M. (2017). States versus corporations: Rethinking the power of business in international politics. *The International Spectator*, *52*(4), 20–43. https://doi.org/10.1080/03932729.2017.1389151

Bannister, M. J., Eppstein, D., Goodrich, M. T., & Trott, L. (2013). Force-directed graph drawing using social gravity and scaling. In *Graph Drawing: 20th International Symposium, GD 2012. Revised Selected Papers 20* (pp. 414–425), *Redmond, WA, USA*. September 19–21, 2012. https://doi.org/10.1007/978-3-642-36763-2 37.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509–512. https://doi.org/10.1126/science.286.5439.509

Barber, C. B., Dobkin, D. P., & Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software (TOMS)*, *22*(4), 469–483.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *Third International AAAI Conference On Weblogs And Social Media*. https://doi.org/10.13140/2.1.1341.1520

Batagelj, V., & Mrvar, A. (2004). Pajek—analysis and visualization of large networks. In *Graph drawing software* (pp. 77–103). Springer. 10.1007/978-3-642-18638-74

Block, J. P., Scribner, R. A., & DeSalvo, K. B. (2004). Fast food, race/ethnicity, and income: A geographic analysis. *American Journal of Preventive Medicine*, *27*(3), 211–217. https://doi.org/10.1016/j.amepre.2004.06.007

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory & Experiment*, *2008* (10), 10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Borgatti, S., Everett, M., & Freeman, L. (2002). *UCINET for Windows: Software for social network analysis*.

Bouttier, J., DiFrancesco, P., & Guitter, E. (2003). Geodesic distance in planar graphs. *Nuclear Physics B*, *663*(3), 535–567. https://doi.org/10.1016/S0550-3213(03)00355-9

Brinkley, C., Manser, G. M., & Pesci, S. (2021). Growing pains in local food systems: A longitudinal social network analysis on local food marketing in Baltimore County, Maryland and Chester County, Pennsylvania. *Agriculture and Human Values*, *38*(4), 911–927. https://doi.org/10.1007/s10460-021-10199-w

Broux, Y. (2016). Locating ancient sites with the help of network analysis: A lost cause? *Ancient Society*, 27–38. https://www.jstor.org/stable/44080249

Christakis, N. A., & Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *New England Journal of Medicine*, *358*(21), 2249–2258. https://doi.org/10.1056/nejmsa0706154

Combe, D., Largeron, C., Egyed-Zsigmond, E., & Géry, M. (2010, October). A comparative study of social network analysis tools. In *Web intelligence & virtual enterprises* (Vol. 2). https://hal.science/hal-00531447

DellaPosta, D. (2017). Network closure and integration in the mid-20th century American mafia. *Social Networks*, *51*, 148–157. https://doi.org/10.1016/j.socnet.2016.11.005

Dwyer, T. (2009). Scalable, versatile and simple constrained graph layout. *Computer Graphics Forum*, *28*(3), 991–998. https://doi.org/10.1111/j.1467-8659.2009.01449.x

Edwards, J. E. (2020). *Over the River and through the Woods: Examining the relationship between network structure, collaboration and geography* [Unpublished doctoral dissertation]. Virginia Tech.

Emch, M., Root, E. D., 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. https://doi.org/10.1080/00045608.2012.671129

Faust, K., Entwisle, B., Rindfuss, R. R., Walsh, S. J., & Sawangdee, Y. (2000). Spatial arrangement of social and economic networks among villages in nang rong district, thailand. *Social Networks*, *21*(4), 311–337. https://doi.org/10.1016/S0378-8733(99)00014-3

Flury, B. K., & Riedwyl, H. (1986). Standard distance in univariate and multivariate analysis. *American Statistician*, *40* (3), 249–251. https://doi.org/10.1080/00031305.1986.10475403

Forati, A., Ghose, R., Mohebbi, F., & Mantsch, J. R. (2023). The journey to overdose: Using spatial social network analysis as a novel framework to study geographic discordance in overdose deaths. *Drug & Alcohol Dependence*, *245*, 109827. https://doi.org/10.1016/j.drugalcdep.2023.109827

Fortunato, S., & Hric, D. (2016). Community detection in networks: A user guide. *Physics Reports*, *659*, 1–44. https://doi.org/10.1016/j.physrep.2016.09.002

Gephi. (2010). *The geographic network layout of gephi*. https://gephi.wordpress.com/2010/05/17/map-geocoded-data-with-gephi

Giordano, A., Cole, T., & Le Noc, M. (2022). Spatial social networks for the humanities: A visualization and analytical model. *Transactions in GIS*, *26*(4), 1683–1702. https://doi.org/10.1111/tgis.12938

Godwin, A. (2022). Paths through spatial networks. In *2022 IEEE Visualization And Visual Analytics (VIS)* (pp. 31–34).

Guo, D. (2007). Visual analytics of spatial interaction patterns for pandemic decision support. *International Journal of Geographical Information Science*, *21*(8), 859–877. https://doi.org/10.1080/13658810701349037

Guo, D. (2009). Flow mapping and multivariate visualization of large spatial interaction data. *IEEE Transactions on Visualization and Computer Graphics*, *15*(6), 1041–1048. https://doi.org/10.1109/TVCG.2009.143

Hadlak, S., Schulz, H.-J., & Schumann, H. (2011). In situ exploration of large dynamic networks. *IEEE Transactions* *on Visualization and Computer Graphics*, *17*(12), 2334–2343. https://doi.org/10.1109/TVCG.2011.213

Hagberg, A., Swart, P., & S Chult, D. (2008). *Exploring network structure, dynamics, and function using networkx. Tech. Rep*. Los Alamos National Lab.(LANL).

Haining, R., Wise, S., & Ma, J. (1998). Exploratory spatial data analysis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *47*(3), 457–469. https://doi.org/10.1111/1467-9884.00147

Hale, S. A., McNeill, G., & Bright, J. (2017). Where'd it go? how geographic and force-directed layouts affect network task performance. In *EuroRV^3@ EuroVis* (pp. 13–17). https://dl.acm.org/doi/abs/10.2312/eurorv3.20171108

Hay, A. (1973). *Transport for the space economy*.

Hu, L., Li, Z., & Ye, X. (2020). Delineating and modeling activity space using geo-tagged social media data. *Cartography and Geographic Information Science*, *47*(3), 277–288. https://doi.org/10.1080/15230406.2019.1705187

Jenny, B., Stephen, D. M., Muehlenhaus, I., Marston, B. E., Sharma, R., Zhang, E., & Jenny, H. (2017). Force-directed layout of origin-destination flow maps. *International Journal of Geographical Information Science*, *31*(8), 1521–1540. https://doi.org/10.1080/13658816.2017.1307378

Jenny, B., Stephen, D. M., Muehlenhaus, I., Marston, B. E., Sharma, R., Zhang, E., & Jenny, H. (2018). Design principles for origin-destination flow maps. *Cartography and Geographic Information Science*, *45*(1), 62–75. https://doi.org/10.1080/15230406.2016.1262280

Koylu, C., Tian, G., & Windsor, M. (2023). Flowmapper.Org: A web-based framework for designing origin–destination flow maps. *Journal of Maps*, *19*(1), 1996479. https://doi.org/10.1080/17445647.2021.1996479

Laniado, D., Volkovich, Y., Scellato, S., Mascolo, C., & Kaltenbrunner, A. (2018). The impact of geographic distance on online social interactions. *Information Systems Frontiers*, *20*(6), 1203–1218. https://doi.org/10.1007/s10796-017-9784-9

Le, M. T., Attaway, D., Anderson, T., Kavak, H., Roess, A., & Züfle, A. (2022). Phyloview: A system to visualize the ecology of infectious diseases using phylogenetic data. In *2022 23rd IEEE International Conference On Mobile Data Management (MDM)* (pp. 222–229). https://doi.org/10.1109/MDM55031.2022.00051

Lengyel, B., Varga, A., Ságvári, B., Jakobi, Á., Kertész, J., & Zhou, W.-X. (2015). Geographies of an online social network. *PLOS ONE*, *10*(9), e0137248. https://doi.org/10.1371/journal.pone.0137248

Leskovec, J., & Horvitz, E. (2014). Geospatial structure of a planetary-scale social network. *IEEE Transactions on Computational Social Systems*, *1*(3), 156–163. https://doi.org/10.1109/TCSS.2014.2377789

Li, S., Zhou, Z., Upadhayay, A., Shaikh, O., Freitas, S., Park, H., Wang, Z. J., Routray, S., Hull, M., Chau, D.H., (2020). Argo lite: Open-source interactive graph exploration and visualization in browsers. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (pp. 3071–3076). https://doi.org/10.1145/3340531.3412877

Liu, Y., Pham, T.-A. N., Cong, G., & Yuan, Q. (2017). An experimental evaluation of point-of-interest recommendation in location-based social networks. *Proceedings Of The VLDB Endowment*, *10* (10), 1010–1021). https://dl.acm.org/doi/10.14778/3115404.3115407

Li Walch, A, Patel, S, Yilmaz, M, Chambers, J., & Chambers, J. (2022). Optimal spatial resource allocation in networks: Application to district heating and cooling. *Computers & Industrial Engineering*, *171*, 108448. https://doi.org/10.1016/j.cie.2022.108448

Loginova, J., Sigler, T., Martinus, K., & Tonts, M. (2020). Spatial differentiation of variegated capitalisms: A comparative analysis of Russian and Australian oil and gas corporate city networks. *Economic Geography*, *96*(5), 422–448. https://doi.org/10.1080/00130095.2020.1833713

Luo, W., & MacEachren, A. M. (2014). Geo-social visual analytics. *Journal of Spatial Information Science*, *8*, 27–66. https://doi.org/10.5311/JOSIS.2014.8.139

Luo, W., MacEachren, A. M., Yin, P., & Hardisty, F. (2011). Spatial-social network visualization for exploratory data analysis. In *Proceedings Of The 3rd ACM SIGSPATIAL International Workshop On Location-Based Social Networks* (pp. 65–68). https://doi.org/10.1145/2063212.2063216

MacEachren, A. M., & Kraak, M.-J. (1997). Exploratory cartographic visualization: Advancing the agenda. *Computers & Geosciences*, *23*(4). Elsevier. 335–343. https://doi.org/10.1016/S0098-3004(97)00018-6

Majeed, S., Uzair, M., Qamar, U., & Farooq, A. (2020). Social network analysis visualization tools: A comparative review. In *2020 IEEE 23rd International Multitopic Conference (INMIC)* (pp. 1–6). https://doi.org/10.1109/INMIC50486.2020.9318162

Mark, H., & Van Duijn, M. A. (2005). Software for social network analysis. *Models and methods in social network analysis*, *270*, e316.

McKenzie, G., Battersby, S., & Setlur, V. (2023). Mixmap: A user-driven approach to place-based semantic similarity. *Cartography and Geographic Information Science*, *51*(4), 583–598. https://doi.org/10.1080/15230406.2023.2176930

Oliveira, M., & Gama, J. (2012). An overview of social network analysis. *Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery*, *2*(2), 99–115. https://doi.org/10.1002/widm.1048

Onnela, J.-P., Arbesman, S., González, M. C., Barabási, A.-L., Christakis, N. A., & Perc, M. (2011). Geographic constraints on social network groups. *PLOS ONE*, *6*(4), e16939. https://doi.org/10.1371/journal.pone.0016939

OurAirport. (2017). *Airport, airline and route data*. https://ourairports.com/data/

Papachristos, A. V., Hureau, D. M., & Braga, A. A. (2013). The corner and the crew: The influence of geography and social networks on gang violence. *American Sociological Review*, *78*(3), 417–447. https://doi.org/10.1177/0003122413486800

Perry, B. L., Pescosolido, B. A., & Borgatti, S. P. (2018). *Egocentric network analysis: Foundations, methods, and models* (Vol. 44). Cambridge university press. https://doi.org/10.1017/9781316443255

Preparata, F. P., & Shamos, M. I. (2012). *Computational geometry: An introduction*. Springer Science & Business Media. https://dl.acm.org/doi/10.5555/4333

Radil, S. M., Flint, C., & Tita, G. E. (2010). Spatializing social networks: Using social network analysis to investigate geographies of gang rivalry, territoriality, and violence in los angeles. *Annals of the Association of American Geographers*, *100*(2), 307–326. https://doi.org/10.1080/00045600903550428

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. https://doi.org/10.1016/j.compenvurbsys.2009.01.007

Richards, W., & Seary, A. (2003). *Multinet. version 4.38 for windows*. Simon Fraser University.

Sarkar, D., Andris, C., Chapman, C. A., & Sengupta, R. (2019). Metrics for characterizing network structure and node importance in spatial social networks. *International Journal of Geographical Information Science*, *33*(5), 1017–1039. https://doi.org/10.1080/13658816.2019.1567736

Sarkar, D., Sieber, R., & Sengupta, R. (2016). Giscience considerations in spatial social networks. In *Geographic Information Science: 9th International Conference, GIScience 2016* (Vol. *9*. pp. 85–98), *Montreal, QC, Canada*. September 27–30, 2016, *Proceedings* https://doi.org/10.1007/978-3-319-45738-36

Sarkar, D., & Yadav, P. (2021). Donut visualizations for network-level and regional-level overview of spatial social networks. *arXiv Preprint arXiv*, *2101*(929). https://doi.org/10.48550/arXiv.2101.00929

Scellato, S., Mascolo, C., Musolesi, M., & Latora, V. (2010). *Distance matters: Geo-social metrics for online social networks. WOSN*. https://dl.acm.org/doi/10.5555/1863190.1863198

Schöttler, S., Yang, Y., Pfister, H., & Bach, B. (2021). Visualizing and interacting with geospatial networks: A survey and design space. *Computer Graphics Forum*, *40*(6), 5–33. https://doi.org/10.1111/cgf.14198

Sevtsuk, A., & Mekonnen, M. (2012). Urban network analysis —a new toolbox for ArcGIS. *Revue internationale de géomatique*, *22*(2), 287–305. https://doi.org/10.3166/rig.22.287-305

Smith, M. A., Shneiderman, B., Milic-Frayling, N., Mendes Rodrigues, E., Barash, V., Dunne, C. & Gleave, E. (2009). Analyzing (social media) networks with nodexl. In *Proceedings of the Fourth International Conference On Communities And Technologies* (pp. 255–264). https://doi.org/10.1080/10447318.2011.544971

Sohn, C., Christopoulos, D., & Koskinen, J. (2020). Borders moderating distance: A social network analysis of spatial effects on policy interaction. *Geographical Analysis*, *52*(3), 428–451. https://doi.org/10.1111/gean.12218

Takhteyev, Y., Gruzd, A., & Wellman, B. (2012). Geography of twitter networks. *Social Networks*, *34*(1), 73–81. https://doi.org/10.1016/j.socnet.2011.05.006

Tiwari, A., & Aljoufie, M. (2021). A qualitative geographical information system interpretation of mobility and COVID-19 pandemic intersection in uttar pradesh, india. *Geospatial Health*, *16*(1), 1. https://doi.org/10.4081/gh.2021.911

Van Duijn, M. (2005). Software for social network analysis. *Model Methods Social Network Analysis*, *28*. https://doi.org/10.1017/CBO9780511811395.013

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. https://doi.org/10.1016/j.socnet.2011.04.003

Walther, O. J., Prieto-Curiel, R., Padron, J., & Scheuer, J. (2023). Mapping the travel geography of the 9/11 network. *The Professional Geographer*, 1–15. https://doi.org/10.1080/00330124.2023.2199324

Wang, K., Bettencourt, C., M, L., Liu, Y., & Andris, C. (2015). Linked activity spaces: Embedding social networks in urban space. *Computational Approaches for Urban Environments*, 313–336. https://doi.org/10.1007/978-3-319-11469-913

Wang, Z., L, L., X, L., Frans, Y., F, V., & Yan, J. (2020). A network perspective for mapping freshwater service flows at the watershed scale. *Ecosystem Services*, *45*, 101129. https://doi.org/10.1016/j.ecoser.2020.101129

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press. https://doi.org/10.1017/CBO9780511815478

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Wei, R. (2022). Exploratory spatial data analysis. In *Handbook of spatial analysis in the social sciences* (pp. 305–321). Edward Elgar Publishing. https://doi.org/10.4337/9781789903942.00026

White, J. J., & Roth, R. E. (2010). *Proceedings Of GIScience*. TwitterHitter: Geovisual analytics for harvesting insight from volunteered geographic information. *Proceedings of GIScience* (Vol. 2010).

Wood, J., Dykes, J., & Slingsby, A. (2017). Visualization of origins, destinations and flows with od maps. In *Landmarks in mapping* (pp. 343–362). Routledge. https://doi.org/10.4324/9781351191234-30

Wu, L., Peng, Q., Lemke, M., Hu, T., & Gong, X. (2022). Spatial social network research: A bibliometric analysis. *Computational Urban Science*, *2*(1), 21. https://doi.org/10.1007/s43762-022-00045-y

Yang, Y., Dwyer, T., Goodwin, S., & Marriott, K. (2016). Many-to-many geographically-embedded flow visualisation: An evaluation. *IEEE Transactions on Visualization and Computer Graphics*, *23*(1), 411–420. https://doi.org/10.1109/TVCG.2016.2598885

Ye, X., & Liu, X. (2018). *Integrating social networks and spatial analyses of the built environment* (Vol. 45) (No. 3). SAGE Publications Sage UK. https://doi.org/10.1177/2399808318772381

Zou, L., & Brooks, S. (2019). A dynamic approach for presenting local and global information in geospatial network visualizations. *GeoInformatica*, *23*(4), 733–757. https://doi.org/10.1007/s10707-019-00350-5