Network Analysis in Geographic Information Science: Review, Assessment, and Projections

Kevin M. Curtin

ABSTRACT: Network data structures were one of the earliest representations in geographic information systems (GIS), and network analysis remains one of the most significant and persistent research and application areas in geographic information science (GIScience). Network analysis has a strong theoretical basis in the mathematical disciplines of graph theory and topology, and it is the topological relationships inherent in networks that led to revolutionary advances in GIS data structures. Networks can represent an alternative datum for geo-location in the context of linear referencing and support a set of tools for graphical display known as dynamic segmentation. Many network location problems are among the most difficult to solve in terms of their combinatorial complexity and, therefore, provide both a challenge and an opportunity for GIScience researchers. Because elements of network analysis appear in a wide range of academic disciplines—from physics, to sociology, to neurobiology—there are ample opportunities for interdisciplinary investigations of emerging research topics.

KEYWORDS: Network, graph theory, topology, linear referencing, location science

Introduction

Network data structures were one of the earliest representations in geographic information systems (GIS), and network analysis remains one of the most significant and persistent research areas in geographic information science (GIScience). This paper will describe the theoretical basis for network data structures and review several major types of network data structures as they have historically been implemented in GIS. This is followed by a concise but comprehensive review of the current capabilities for network analysis in GIS, and the consequent deficiencies in GIS implementations of networks. A set of challenges is suggested for network analysis in GIS, through increased implementation of existing network theory, through expansion of existing theory and practice in the areas of network design and location, and through interactions with a wide variety of other disciplines.

The Theoretical Basis for Network Analysis in GIS

Network analysis in GIS rests firmly on the theoretical foundation of the mathematical subdisciplines of graph theory and topology. Any graph or network (the terms are synonymous in this context) consists of a set of vertices and the edges that connect them. Within graph theory there are methods for describing, measuring, and comparing graphs, and techniques for proving the properties of individual graphs or classes of graphs. Some elements of graph theory are not concerned with the cartographic characteristics (e.g., shape or length) of the features that comprise a network but, rather, with the topological attributes of those features. The topological invariants of a network are those properties that are not altered by elastic deformations. Therefore, properties such as connectivity, adjacency, and incidence are topological invariants of a network, since they will not vary if the network is deformed by a cartographic process, such as a projection. The permanence of these properties allows them to serve as a basis for describing, measuring, and analyzing networks.

Graph theoretic descriptions of networks can range from simple statements of the number of features in the network, the degree of the vertices of the graph, or the number of cycles in a graph, to more complex descriptions based on structural characteristics of networks. In some cases these network structures can be classified into idealized network types (e.g., tree networks, hub-and-spoke networks, Manhattan networks, etc.). In turn, these

Kevin M. Curtin, Program in Geographic Information Sciences, School of Economic, Political and Policy Sciences (GR31), University of Texas at Dallas, P.O. Box 830688, Richardson, TX 75083. Tel: (972) 883-6237; Fax: (972) 883-6297. E-mail: <curtin@utdallas.edu>.

ideal types may be proven to have properties that encourage their use for particular applications.

Moving beyond description, quantitative measures of the properties of graphs can be computed through network indices. Measures such as the Beta, Alpha, and Gamma indices (Kansky 1963) measure the relative connectivity of a network by comparing the number of edges to the number of vertices (in the case of the Beta Index), or by comparing proven properties of graphs to observed properties. Additional measures and analytical techniques exist within graph theory for applied instances of networks, which depend on non-topological properties such as edge length or capacity (Rodrigue et al. 2006). Although this is not the appropriate venue to review them there are many more advanced graph theoretic techniques for describing networks, for categorizing them, and for proving their properties (Harary 1982; Wilson 1996).

Implementations of Network Data Structures in GIS

While the graph theoretic definition of a network remains constant, the ways in which networks are structured in computer systems have changed dramatically through the history of GIScience. Network data structures must store the edge and vertex features that populate these network datasets, the attributes of those features, and—most importantly for network analysis—the topological relationships among the features. The choice of a network data structure can significantly influence the analyses that can be performed.

Non-Topological Data Structures

The earliest computer-based systems for automated cartography stored network edges as independent records in a database. Each record contained a starting and ending point, and the edge was defined as the connection between those points. Attribute fields were associated with each record, and some implementations included a list of "shape points" that approximated curvature. This structure did not contain any information regarding the topological properties of the edges and was therefore termed the non-topological structure (colloquially, the "spaghetti" data model).

Non-topological data models had the advantages of being easy to understand, allowing straightforward capture of spatial data through digitizing, and being efficient in terms of digital cartographic display. The latter advantage led to the wide acceptance of this data structure for computer-aided drafting software packages. The disadvantages of non-topological data structures included the tendency for duplicate edges to be captured, particularly coincident boundaries of polygonal features, which in turn led to sliver errors. Most importantly for this discussion, the lack of topological information made these structures essentially useless for network analysis.

Due to these disadvantages, non-topological data models were essentially abandoned in mainstream GIS, until a variant data structure became extremely popular in the mid-1990s. The shapefile is a non-topological data structure developed by Environmental Systems Research Institute (ESRI). It was designed primarily to allow for rapid cartographic rendering, and the structure performs admirably in that respect. Although specialized tools have been developed to compute topological relationships "on-the-fly," and therefore it is possible to perform some network analysis, the shapefile is generally considered to be inefficient for network analysis.

Topological Data Models

There is broad recognition that maintenance of topological properties is an important element for many GIS functions, including network analysis. As has been well documented elsewhere, the Census Bureau is primarily responsible for the inclusion of topological constructs in GIS data structures due to the development of the Dual Incidence Matrix Encoding (DIME) data structure (Cooke 1998). "Dual Incidence" refers to the capture of topological information between nodes and along lines. The DIME data structure evolved into the Topologically Integrated Geographic Encoding and Referencing (TIGER) structure that is still in use today.

There were two elements of this advance that profoundly influenced the ability to conduct network analysis in GIS. First, the DIME structure captures incidence which is one of the primary topological properties defining the structure of networks. The graph theoretic methods developed over several centuries could now be employed inside GIS. Second, many of the features captured by the Census Bureau were streets or other transportation features. Since the Census Bureau has a mandate that covers the entire United States, a national transportation database was made available for use in GIS, and this database was captured in a structure that could support high-level network analysis. Thus this structure became the *de facto* standard for vector representations in GIS.

However, the topological data model also imposes difficult constraints on network analysts. In order to explicitly define polygons with which populations could be associated, the data model had to enforce planarity. Thus, at every location where network features cross, a vertex must exist in the database. This is true whether or not a true intersection exists between the network features, and it is most problematic when modeling bridges or tunnels. Moreover, planar enforcement demands that network features be divided at every intersection, causing a single feature to be represented as a series of records in the database. This repetition can increase the database size many times over and encourage errors in the database.

Pure Network Data Models

The limitations on the ability to perform network analysis when using early GIS data structures have necessitated the development of pure network data models. These include non-planar data structures that relax planarity requirements in order to more realistically model real world networks (Fohl et al. 1996); data structures that support turns and directed edges in order to model the impedances encountered when moving between and along network features; and perhaps most importantly, data structures (such as the star data structure) that allow more efficient operation of network analysis procedures (Evans and Minieka 1992). Pure network data structures have, in the last five years, become widely available in commercial-off-the-shelf (COTS) GIS software packages such as the Geometric Network and Network Data Set in ESRI's ArcGIS product, and the Geographic Data Object Networks built within Intergraph's Transportation Manager product.

Current Cartographic and Analytical Capabilities in Network GIS

The most common and familiar implementations of network models are those used to represent the networks with which much of the population interacts every day: transportation and communications networks. Cartographic conventions (such as using blue lines to represent rivers) serve to reinforce the interpretation of the functions of these networks (Figure 1). Many other types of networks can be modeled in GIS, including utility networks (electricity, telephone, cable), other transportation networks (airlines, shipping lanes, transit routes), and even networks based on social connections—if there is a geographic component.

The implementation of GIS-based network representations for transportation applications has increased dramatically over the past decade, with nearly ubiquitous availability of location services and address-based driving directions through such internet services as Mapquest and Google Maps. Although the importance of the widespread acceptance of this most common application of GIS network analysis cannot be overstated, the focus in this limited space will be on advances in basic cartographic and geographic research.

Linear Referencing

Network GIS is the only sub-discipline within GIScience (outside of pure geodesy) that has developed a method for redefining the spatial reference system on which locations are specified (Scarponcini 2002). For many applications, it is the network itself that acts as the underlying datum rather than a coordinate system designed to locate objects on the Earth's surface. The process of using a network for reference is termed linear referencing. The most frequently recognized application of linear referencing is the mile marker system employed along U.S. highways (Federal Transit Administration 2003).

Linear referencing can be applied to any network-based phenomenon. It has been applied for mapping accidents, traffic stops, or other incident locations, displaying traffic counts along streets, maintaining the location of fleet vehicles, and performing asset management functions such as the recording of pavement conditions or the location of street signs, bridges, exits, and many other traffic-related objects (Federal Highway Administration 2001).

There are a myriad of linear referencing systems (Fletcher et al. 1998; Scarponcini 2001), methods (Noronha and Church 2002; Nyerges 1990), and data models (Curtin et al. 2001; Dueker and Butler 2000; Koncz and Adams 2002; Sutton and Wyman 2000; Vonderohe et al. 1997) available for implementation. Because linear referencing developed independently in practice over a period of several decades, recent efforts have focused on the development of generic tools for its implementation, and on the specifica-

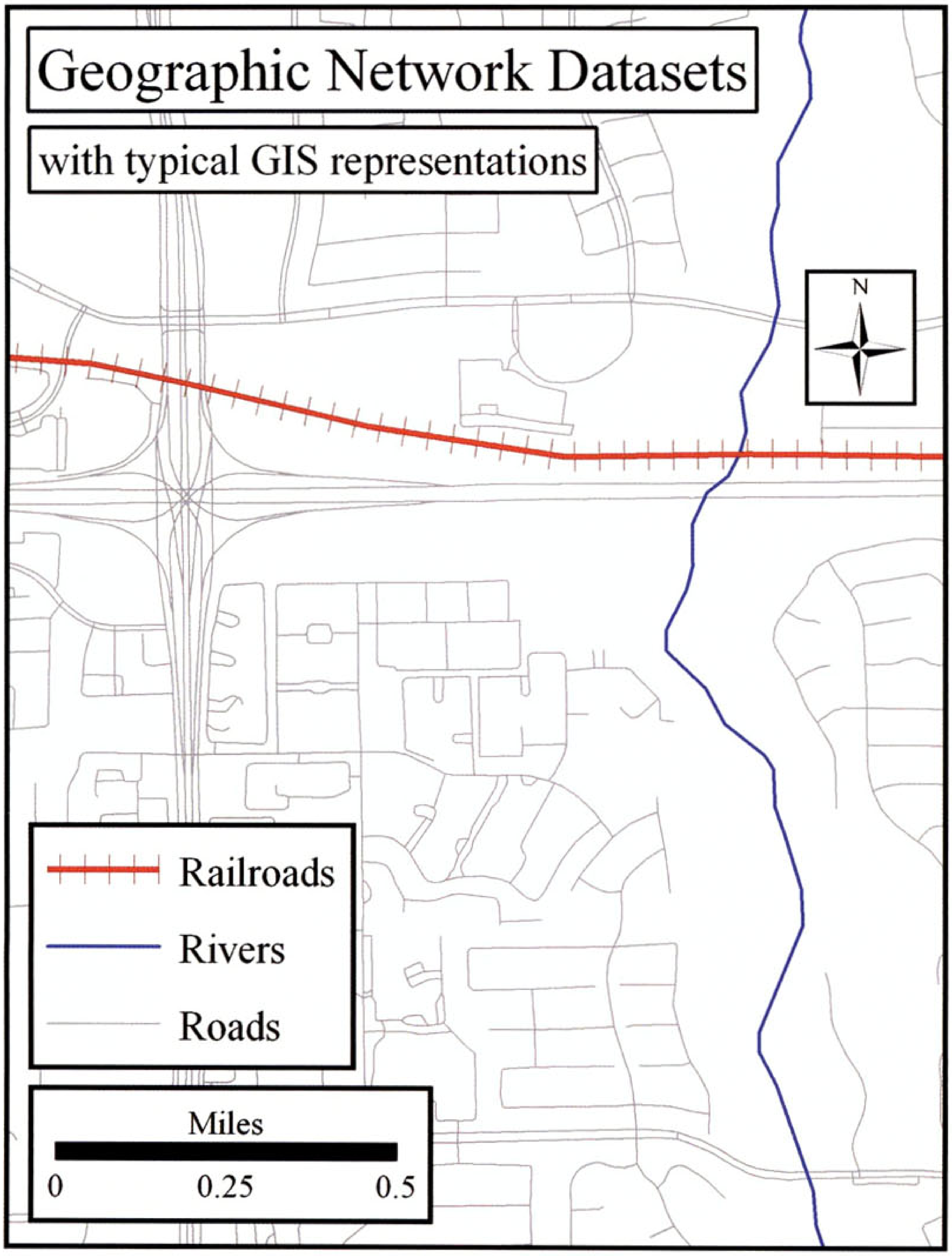

Figure 1. Typical network GIS representations.

tion of well defined processes for successful implementation (Curtin and Nicoara 2005).

Routing Across Networks

Currently implemented network GIS tools are dominated by routing functions. Routing is the act of selecting a course of travel, and it is arguably the most fundamental logistical operation in network analysis. Without question, the most common objective in routing across networks is to minimize the cost of the route. Cost can be defined and measured in many ways, but it is frequently assumed to be a function of distance, time, or impedance in crossing the network. There are several extremely efficient algorithms for determining the optimal route, the most widely cited of which was developed by Edsgar Dijkstra (1959).

Shortest path routing has been fully integrated into commercial-off-the-shelf GIS software packages. In the most recent release of the industry standard network analysis package there are four fundamental operations that can be performed, all of which are derivatives of route finding algorithms. These functions are:

1. Finding a route between point locations;
2. Determining the service area for a facility;
3. Finding the closest facility across the network (see Figure 2); and
4. Creating an origin–destination matrix.

The determination of a service area for a facility is simply a matter of finding the shortest path from the facility to demand points on the network, and allocating demand points (or the associated network locations) to their nearest facility. Finding the closest facility for any location is similarly a matter of finding the shortest path from the demand to each possible facility and choosing the shortest of these solutions. Lastly, an origin–destination matrix is defined as a table of shortest paths between all origins and destinations, and such a matrix is generated with an "all shortest paths" algorithm variant of the standard shortest path algorithm (Dantzig 1966). Therefore, the entire suite of state-of-the-art network analysis functions implemented in GIS is composed of derivatives of the shortest path algorithm. There are, of course, many parameters that can be set in order to define more complex versions of shortest path problems, and there are innovative ways in which users can combine these operations to conduct more complex analyses, but the underlying dominance of shortest path routing remains clear.

Although network analysis in GIS has been largely limited to the simplest routing functions, the recent past has seen the development of object-oriented data structures, the introduction of dynamic networks (Sutton and Wyman 2000), the ability to generate multi-modal networks, and the use of simulation methods to generate solutions to network problems. Some network flow modeling functions have also been implemented, although there are substantial opportunities for additional theoretical advances and diversified application.

Challenges and Opportunities for Network Analysis in GIS

There are several advancements in network GIS that could improve—or even revolutionize—our ability to research network based problems in the near term. The greatest potential over the next five years lies in the implementation of known network analytical and logistic algorithms that have until now escaped the attention of GIS software developers. As one example, all of the graph theoretic indices and measures based on connectivity discussed above could be implemented today with very little difficulty. More importantly, there are substantial benefits to be had through research in the areas of network location and network design. Beyond these areas, the greatest challenge for network GIS is to integrate geographic research with that being undertaken in heretofore unrelated disciplines.

Network Design Problems

The routing problem described in the previous section is one of the simplest of a class of problems known as network design problems (Ahuja et al. 1993). Moreover, the shortest path algorithm is among a small group of efficient algorithms that exist for this class of problems. However, it is not the only one. There are, for example, extremely efficient heuristics for the minimum spanning tree problem (Kruskal 1956; Prim 1957) that have surprisingly eluded network analysts involved in the development of GIS software. Perhaps even more important would be the implementation of a method to solve minimum cost flow problems such as the classic transportation problem. Variants of the simplex algorithm to solve this problem optimally have appeared in the literature for decades (Ford and Fulkerson 1956), and this problem can be applied in many disciplines, yet it has not been a topic of development for GIS. Similarly, maximum flow algorithms would seem to be of interest for many transportation- and utility-based GIS applications, but the optimal solution procedure for this problem has not found its way from the literature into GIS practice.

There are, of course, many important network design problems that are very difficult to solve optimally due to their combinatorial complexity (Magnanti and Wong 1984). The Traveling Salesman Problem (TSP) is the most notable of these; it is widely regarded as the most important problem in combinatorial optimization (Applegate et al. 1998). Implementations of the TSP do appear in GIS software packages, but they are problematic. The solution procedures are necessarily heuristic and not guaranteed to determine the optimal solution which, in and of itself, is not surprising. However, a recent review of these implementations (including implementations in both the ArcGIS

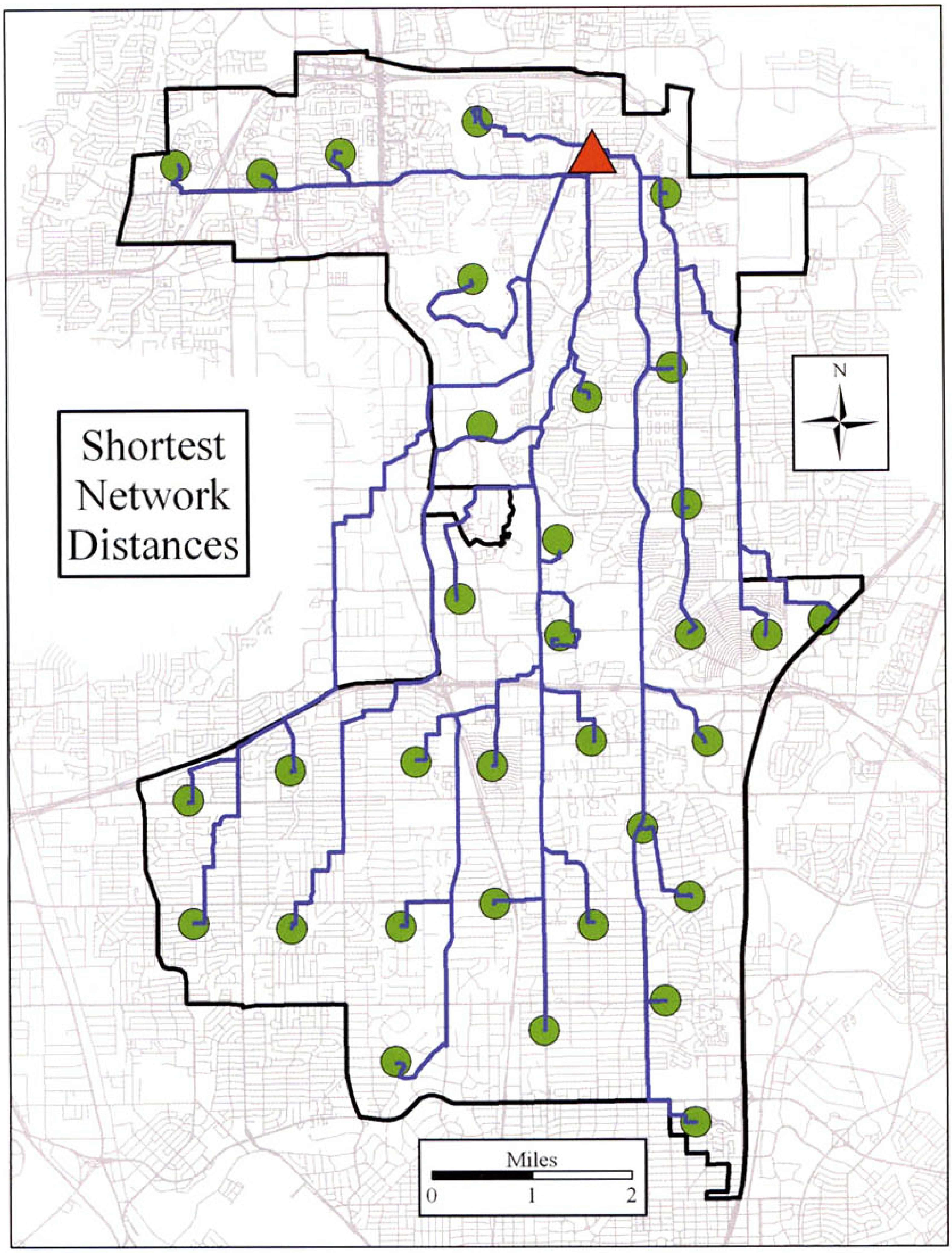

Figure 2. Shortest network paths to a facility.

and Intergraph software packages) (Curtin and Voicu 2007) found that the users are not well informed about the facts that heuristic methods are used (if they are informed at all), that there is no way to know how far from optimal the solution may be, and that optimal solutions for many mid-sized problems could be obtained through the integration of GIS and integer programming solution software.

Although finding optimal solutions to the TSP is a lofty and long-term goal, the ability to implement alternate objective functions—such as improving accessibility (Murray and Wu 2003), maximizing coverage (Current and Schilling 1994), or defining multiple objectives (Ceder 2001)—should be well within reach. More advanced methods of determining convex costs on network edges (in order to allow more realistic congestion conditions on networks) should also be seen as a reasonable short-term goal for improving network design analysis in GIS.

Network Location Problems

Location on networks involves selecting network locations on an existing network such that an objective is optimized. Just a few examples of such problems are the P-Median problem which seeks to locate facilities in such a way that the demand-weighted distance is minimized, the P-Center problem which seeks to locate facilities so that the maximum distance between a demand point and a facility is minimized, the maximal covering problem (and the related flow covering problem) that seeks to serve as much demand as possible within a given service distance, or maximal dispersion problems which seek to separate facilities as much as possible.

Due to the extreme difficulty in finding optimal solutions to these highly combinatorially complex problems there have been very few implementations in GIS, and those that have appeared have been necessarily heuristic in nature. Some recent attempts have been made to integrate GIS and optimal solution software in order to solve these problems optimally (Curtin et al. 2005). Research advances in this area could transform the sub-discipline of facility location science, but in order to do so the GIScience community must look beyond the current capabilities of existing software and recognize that related disciplines such as operations research and management science have a set of complementary techniques that can be readily applied to spatial problems.

Interdisciplinary opportunities for Network Analysis in GIS

The network is a compelling research paradigm because its form can so intuitively represent complex systems. The ability to comprehend the complex systems around us—whether they are transportation systems, intricate communication systems such as the internet, or interactions at the cellular level—is of increasing importance in an increasingly complex world. Since networks are fundamentally spatial there is a clear opportunity for research in network analysis that could prove valuable across a wide range of disciplines.

It has been suggested that in the long term, network thinking will be integral to all branches of science as complex systems pervade research agendas (Strogatz 2001). The delineation of contact networks in order to analyze the spread of disease within the field of microbiology is one research area that could benefit not only from network analysis in GIS, but from geography's substantial set of methods for modeling diffusion processes (Wallinga et al. 1999). Micro-geographic analyses of the network connections in the cerebral cortex have already employed graph theoretic measures of connectivity (Sporns et al. 2000) and could perhaps benefit from applications of the flow functions developed for network GIS. Although GIS has already influenced research across the social sciences, opportunities for additional cross-fertilization are abundant. For example, criminologists recognize that the journey to crime occurs across a network (Groff and McEwen 2005). Social network analysis is being widely used in disciplines as diverse as medicine (Finnvold 2006), psychology (Walker et al. 2006), urban planning (Toccolini et al. 2006), and computer science (Bera and Claramunt 2005) to name only a few, and the spatial components of these analyses are frequently paramount. Even networks of scientists themselves can be analyzed to determine the nature of research collaboration (Newman 2004).

Thus there is a wide variety of disciplines with interests related to network analysis, and a well established ability within GIS to model networks and perform such analyses. This strongly suggests that network GIS will be a dynamic sub-discipline within geographic information science for the foreseeable future. The growing interest in Open Source GIS may lead to an increased pace of innovation in this field and more rapid integration of scientific advancements into software packages. The extent to which geographers will be able to influence this process will depend on the robustness of the methods developed and the strength of the basic research agenda.

REFERENCES

Ahuja, R.K., T.L. Magnanti, and J.B. Orlin. 1993. *Network flows: Theory, algorithms, and applications*. Upper Saddle River, New Jersey: Prentice-Hall, Inc.

Applegate, D., R. Bixby, V. Chvtal, and W. Cook. 1998. On the solution of traveling salesman problems. *Documenta Mathematica*: Extra Volume—Proceedings of the International Congress of Mathematicians: 645-56.

Bera, R., and C. Claramunt. 2005. Connectivity inferences over the Web for the analysis of semantic networks. In: Proceedings, 5th International Workshop on *Web and Wireless Geographical Information Systems*.

Ceder, A. 2001. Operational objective functions in designing public transport routes. *Journal of Advanced Transportation* 35 (2):125-44.

Cooke, D.F. 1998. Topology and TIGER: The Census Bureau's contribution. In: T. W. Foresman (ed.), *The History of Geographic Information Systems*. Upper Saddle River, New Jersey: Prentice Hall.

Current, J.R., and D.A. Schilling. 1994. The median tour and maximal covering tour problems: Formulations and heuristics. *European Journal of Operations Research* 73:114-26.

Curtin, K. M., and G. Nicoara. 2005. Linear referencing as a process: Moving beyond data models and objects. Paper read at GIS Planet 2005, at Estoril, Portugal.

Curtin, K. M., V. Noronha, M. F. Goodchild, and S. Grise. 2001. *ArcGIS Transportation Data Model*. Redlands, California: Environmental Systems Research Institute.

Curtin, K. M., F. Qiu, K. Hayslett-McCall, and T. Bray. 2005. Integrating GIS and maximal covering models to determine optimal police patrol areas. In: F. Wang (ed.), *Geographic Information Systems and Crime Analysis*. Hershey, Pennsylvania: Idea Group Publishing.

Curtin, K. M., and G. Voicu. 2007. An analysis of traveling salesman implementations in geographic information systems. Paper read at Annual Meeting of the Association of American Geographers, San Francisco, California.

Dantzig, G.B. 1966. All shortest routes in a graph. Paper read at International Symposium, Rome, Italy.

Dijkstra, E.W. 1959. A note on two problems in connexion with graphs. *Numerische Mathematik* 1: 269-71.

Dueker, K. J., and J. A. Butler. 2000. A geographic information system framework for transportation data sharing. *Transportation Research Part C-Emerging Technologies* 8(1-6): 13-36.

Evans, J.R., and E. Minieka. 1992. *Optimization Algorithms for Networks and Graphs*, 2nd ed. New York, New York: Marcel Dekker.

Federal Highway Administration. 2001. *Implementation of GIS based highway safety analysis: Bridging the gap*. U.S. Department of Transportation, McLean, Virginia.

Federal Transit Administration. 2003. Best practices for using geographic data in transit: A location referencing guidebook. U.S. Department of Transportation, Washington, D.C.

Finnvold, J. E. 2006. Access to specialized health care for asthmatic children in Norway: The significance of parents' educational background and social network. *Social Science & Medicine* 63(5): 1316-27.

Fletcher, D., J. Expinoza, R. D. Mackoy, S. Gordon, B. Spear, and A. Vonderohe. 1998. The case for a unified linear reference system. *URISA Journal* 10(1): 7-11.

Fohl, P., K.M. Curtin, M.F. Goodchild, and R.L. Church. 1996. A non-planar, lane-based navigable data model for ITS. Paper read at International Symposium on Spatial Data Handling, Delft, The Netherlands.

Ford, L. R., and D. R. Fulkerson. 1956. Solving the transportation problem. *Management Science* 3: 24-32.

Groff, E., and J. T. McEwen. 2005. Disaggregating the journey to homicide. In: F. Wang (ed.), *Geographic Information Systems and Crime Analysis*. Hershey, Pennsylvania: Idea Group Publishing.

Harary, F. 1982. *Graph Theory*. Reading, U.K.: Addison Wesley.

Kansky, K. 1963. Structure of transportation networks: Relationships between network geography and regional characteristics. In: *Research Papers*. Chicago, Illinois: University of Chicago.

Koncz, N. A., and T. M. Adams. 2002. A data model for multi-dimensional transportation applications. *International Journal of Geographical Information Science* 16(6): 551-69.

Kruskal, J. B. 1956. On the shortest spanning subtree and the traveling salesman problem. In: *Proceedings of the American Mathematical Society* 7: 48-50.

Magnanti, T.L., and R.T. Wong. 1984. Network design and transportation planning: Models and algorithms. *Transportation Science* 18(1): 1-55.

Murray, A.T., and X. Wu. 2003. Accessibility trade-offs in public transit planning. *Journal of Geographical Systems* 5: 93-107.

Newman, M.E.J. 2004. Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences of the United States of America* 101: 5200-205.

Noronha, V., and R. L. Church. 2002. Linear referencing and other forms of location expression for transportation. Vehicle Intelligence & Transportation Analysis Laboratory, University of California, Santa Barbara, California.

Nyerges, T. L. 1990. Locational referencing and highway segmentation in a geographic information system. *ITE Journal* (March), pp. 27-31.

Prim, R. C. 1957. Shortest connection networks and some generalisations. *Bell System Technical Journal* 36: 1389-1401.

Rodrigue, J., C. Comtois, and B. Slack. 2006. *The Geography of Transport Systems*. London, U.K.: Routledge.

Scarponcini, P. 2001. Linear reference system for life-cycle integration. *Journal of Computing in Civil Engineering* 15(1): 81-8.

Scarponcini, P. 2002. Generalized model for linear referencing in transportation. *Geoinformatica* 6(1): 35-55.

Sporns, O., G. Tononi, and G.M. Edelman. 2000. Theoretical neuroanatomy: Relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cerebral Cortex* 10(2): 127-41.

Strogatz, S. H. 2001. Exploring complex networks. *Nature* 410(8): 268-76.

Sutton, J.C., and M.M. Wyman. 2000. Dynamic location: An iconic model to synchronize temporal and spatial transportation data. *Transportation Research Part C-Emerging Technologies* 8(1-6): 37-52.

Toccolini, A., N. Fumagalli, and G. Senes. 2006. Greenways planning in Italy: The Lambro River Valley Greenways System. *Landscape and Urban Planning* 76(1-4): 98-111.

Vonderohe, A., C. Chou, F. Sun, and T. Adams. 1997. *A generic data model for linear referencing systems*. National Cooperative Highway Research Program, Transportation Research Board, Washington D.C.

Walker, L.R., M. Mason, and I. Cheung. 2006. Adolescent substance use and abuse prevention and treatment: Primary care strategies involving social networks and the geography of risk and protection. *Journal of Clinical Psychology in Medical Settings* 13(2): 131-9.

Wallinga, J., W. John Edmunds, and M. Kretzschmar. 1999. Perspective: Human contact patterns and the spread of airborne infectious diseases. *Trends in Microbiology* 7(9): 372-7.

Wilson, R.J. 1996. *Introduction to Graph Theory*. Essex, U.K.: Longman.

■