[FC-04-019] Networks

Abstract

A network is a widely used term with different definitions and methodologies depending on the applications. In GIS, a network refers to an arrangement of elements (i.e., nodes, links) and information on their connections and interactions. There are two types of networks: physical and logical. While a physical network has tangible objects (e.g., road segments), a logical network represents logical connections among nodes and links. A network can be represented with a mathematical notion called graph theory. Different network components are utilized to describe characteristics of a network including loops, walks, paths, circuits, and parallel edges. Network data are commonly organized in a vector format with network topology, specifically connectivity among nodes and links, whereas raster data can be also utilized for a least-cost problem over continuous space. Network data is utilized in a wide range of network analyses, including the classic shortest path problem.

Keywords: costs, flow, matrix, movement, network analysis, transportation

Author & citation

Chun, Y, and Kim, H. (2018). Networks. The Geographic Information Science & Technology Body of Knowledge (4th Quarter 2018 Edition), John P. Wilson (Ed.).
DOI: 10.22224/gistbok/2018.4.3

Explanation

1. Definitions
2. The Concept of Network
3. Graph Theory
4. Network Data Structure in GIS
5. Application: The Shortest Path Problems

1. Definitions

Distance is a numerical measure for the amount of space between two objects. It commonly refers to physical length on the Earth's surface. There are different types of distance including Euclidean distance, spherical distance, and street network distance. While Euclidean distance is measured as a straight line distance on a flat surface, spherical distance is measured considering the curvature of the Earth. Street network distance is measured with a travel pathway on a street network. Alternatively, street networks can also use time as a proxy for distance if hindrances on the network greatly affect the measurements. Also, non-physical distance, e.g., cognitive distance (Golledge 1987), is also considered in applications.

Network Topology is an arrangement of features that participate in a network. The connectivity of nodes and links is a fundamental building block for GIS network analysis for utilities, for example. Also direction is relevant topological information on a network.

2. The Concept of Network

The term network is widely used in various research fields and/or applications, and its meaning could be different from one's expectation. Popular examples include telecommunications networks, transportation networks, street networks, social networks, electrical circuits, and biology networks. Although the terminologies refer to different concepts or phenomena, there is a coherent and common framework that networks are composed of objects (or entities) and linkages among the objects. For example, in a simple computer network, computers can be recognized as objects that are connected through a common communication line (or wireless signals), which can be referred to as links. In GIS, these objects are often represented as points or nodes emphasizing their functional role. Similarly, the linkages are represented as lines and called edges.

A network is commonly perceived and analyzed as a system rather than an individual object. That is, an object interacts with and/or works together with other objects in a network. Such interactions can evolve into systematic activities or events that are understood through their magnitudes, causes, and impacts. In addition, a network is assumed to be a complete and closed (rather than open) system depending on the scope of analysis. This means that such activities and/or events occur only in a given network and are affected by other objects in the system. Accordingly, no impact and phenomenon beyond the system are taken into account. For example, for population migration in a country, domestic migration is considered while international migration is not.

In GIS applications, networks are either physical or logical. A physical network is composed of tangible objects, i.e., nodes and links. A street network is a typical example of a physical network since network objects, i.e., nodes and links, physically exist in geographical space. In a physical network, each node is connected to a small number of spatially proximal nodes. A map of the physical network of highways in the Dallas-Fort Worth (DFW) area shows how nodes (i.e., intersections) and links (i.e., road segments) are connected to each other (Figure 1a). Other notable examples include utility networks such as electric distribution systems or water distribution networks. Infrastructure networks such as telecommunications and the Internet backbone can be represented with this structure.

In contrast, a logical network does not have a tangible structure and focuses on transitions between nodes. Generally, each node can be linked to any other nodes in such a network, and links often cross other links. Inter-county migration flows among counties in the DFW metropolitan area are an example (Figure 1b). The top 10% inter-county migration flows are drawn with straight lines that connect centroids of corresponding origins and destinations. In this migration system, movements or transitions between areas are the main interest rather than actual or "true" migration paths. Inter-regional commodify flows are another example of a logical network, as are social networks among people or connections on social media.



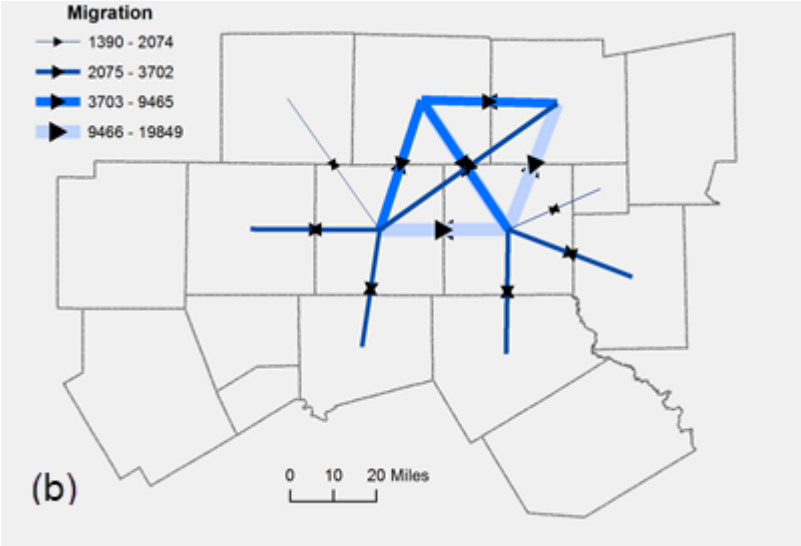


Figure 1a and 1b. Highway network (a) and migration flows (b) in the Dallas-Fort Worth metropolitan area. Source: author.

In a network, cost refers to the value of travel through fixed network objects such as links, and can be represented in various forms. The most basic and common form of cost is geographic distances of links such as street segment length in a street network or interstate distances in a migration system. Travel time, traffic congestion, speed limit on a street network, or walking distance on a trail are alternatives. Cost can be static and unchanging over time, or dynamic with time-varying values, such as real time traffic flows. Data on dynamic costs have become increasingly available and are essential for many transportation applications such as an intelligent transportation system.

3. Graph Theory

Any network can be represented via a standard mathematical notion called graph theory. In general, a graph G = (V, E) is comprised of two elements: a set V of vertices v and a set E of edges e. Vertices v and edges e are preferably and interchangeably called nodes and

links in a spatial network. In general, the terms vertices and edges refers to geometric elements rather than abstract representation of real phenomena. In graph theory, a distance between vertices usually refers to a topological distance while it is represented in the network as a metric such as cost among nodes (based on time or physical distance). Only the existence of connections between nodes is of concern, rather than accounting for attributes on them. A network can be classified into two distinct types: planar (Figure 2a) and non-planar (Figure 2b). Planar networks are drawn in two-dimensional plane space. In Figure 2a, a node (v7) is made where two links (e13 and e25) intersect. In contrast, non-planar networks are flexible in creating links among nodes, allowing edges to cross and structuring the form in a three-dimensional space. In general, planar graphs may represent spatial network systems such as street networks and highway networks (e.g., Figure 1a). Non-planar graphs are used to represent logical networks such as telecommunications, the Internet backbone, and air transportation routes where logical network structure is of concern (e.g., Figure 1b).

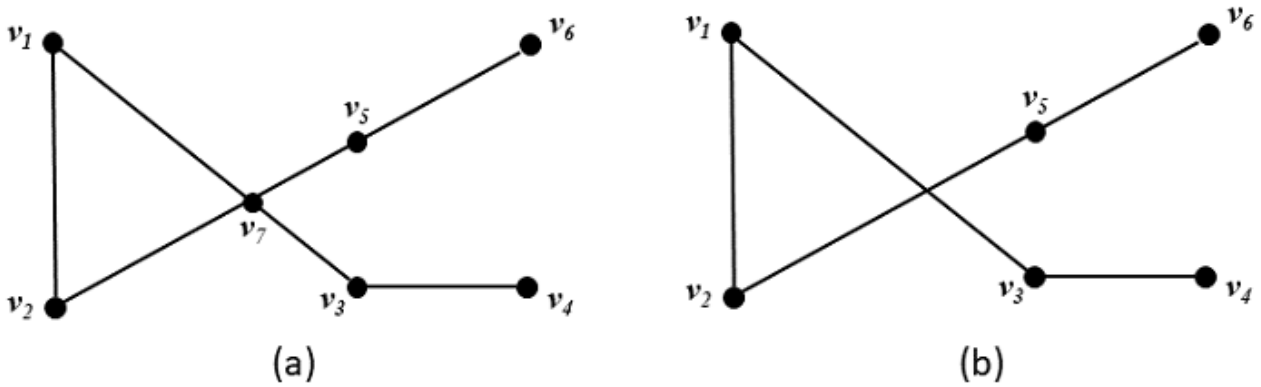


Figure 2. Network types: (a) planar and (b) non-planar. Source: author.

Figure 3 introduces several fundamental notions from graph theory that are employed to describe the characteristics of a network. This network is called a digraph because the E(e45) and E(e54) are ordered pair of vertices, indicating that each link defines a specific direction between v4 to v5 using an arrow symbol. If no direction is specified on the edges, then the network is a simple-graph indicating both directions are allowed on the links. In a network, if two vertices are connected by an edge, their relationship is defined as adjacent. For example, v1 is adjacent to v2 and v3, but not with other vertices. If an edge with one vertex is adjacent to itself it is defined as a loop, such as e66 in Figure 3. Two distinct edges connected with an identical set of vertices are called parallel edges. For example, two edges e45 and e54 are parallel edges to v4 and v5 because they connect v4 and v5 simultaneously. Identifying the sequence of edges in a graph is a critical component with regard to network analysis and network modeling because the sequence of edges from a starting node to an ending node, also known as route, can be structured given the information of connectivity among vertices via their edges. The sequence of edges is classified into walks, paths, or circuits depending on the characteristics of the sequence. A walk refers to a finite alternate sequence of edges from a starting node to an ending node which allows the repetition of edges or vertices in creating the sequence. In contrast, a path is a sequence of edges that does not allow multiple membership of edges in the sequence. For example, the sequence {e12, e23, e36, e66, e36, e34} is recognized as one of walk

from node v1 to v4. However, it is not a path. A sequence of {e13, e34} is a path between the nodes v1 to v4. A circuit is a special case of path where the starting node and the ending node are identical. A sequence {e12, e23, e13} is a case of circuit because the path is closed.

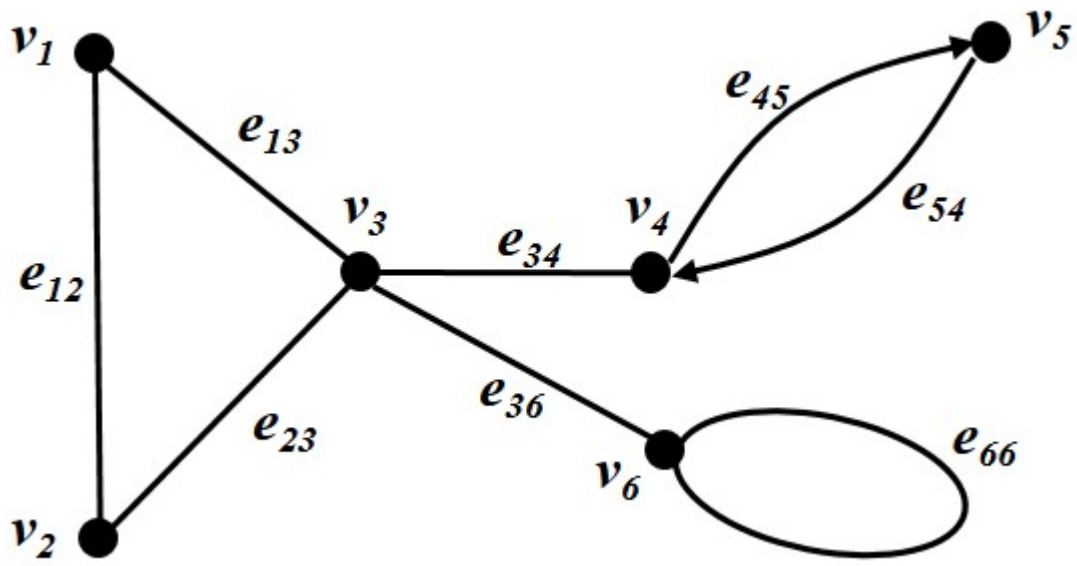


Figure 3. Network representation. Source: author.

To characterize a network, a connectivity matrix C(cij) and a node-edge adjacency table are used. Figure 4 illustrates how to build a connectivity matrix with the network of Figure 3. In the matrix, if two vertices vi and vj are directly connected by an edge eij then cij = 1, otherwise cij = 0. The matrix is symmetric if no direction is specified but can be asymmetric if the network is digraph. Based on the connectivity matrix, the degree of a node d(v) is defined as the number of edges that are incident with v, which is generally the row-sum of each v in the connectivity matrix. The node-edge adjacency table specifies the membership of edges incident to each node. The table is used for spatial networks where a C(cij) is sparse so that the retrieval of paths based on the connectivity matrix does not perform efficiently.

$C(c_{ij})$	v_1	v_2	v_3	v_4	v_5	v_6	$d(v)$
v_1	0	1	1	0	0	0	2
v_2	1	0	1	0	0	0	2
v_3	1	1	0	1	0	1	4
v_4	0	0	1	0	1	0	2
v_5	0	0	0	1	0	0	1
v_6	0	0	1	0	0	1	2

	$E(e_{ij})$
v_1	e_{12}, e_{13}
v_2	e_{12}, e_{23}
v_3	e_{13}, e_{23}, e_{34}, e_{36}
v_4	e_{34}, e_{45}, e_{54}
v_5	e_{45}, e_{54}
v_6	e_{36}, e_{66}

Figure 4. Matrix representation for a network: connectivity matrix (left) and node-edge adjacency table (right). Source: author.

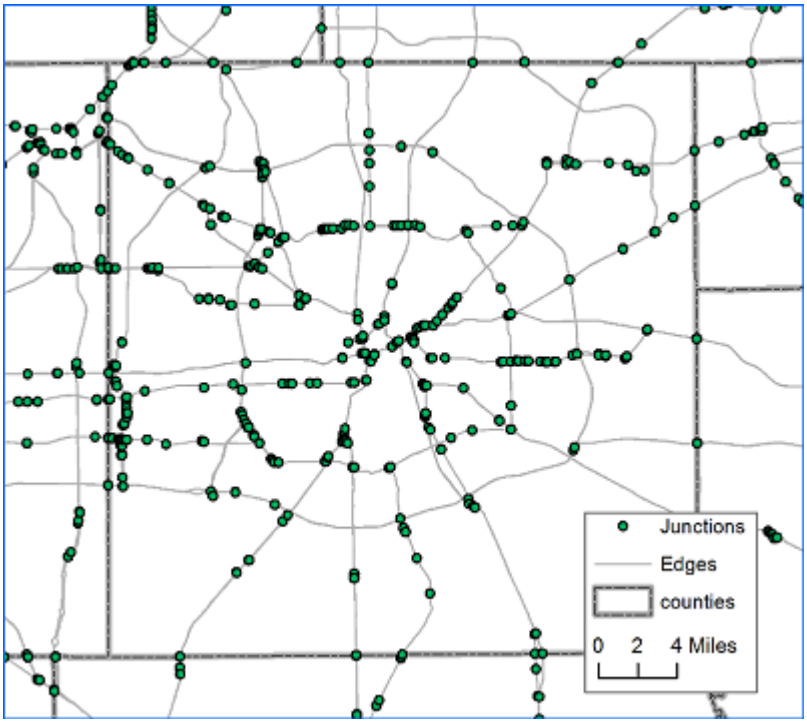
The size of a network is often measured using diameter D, which is defined as the number of links or steps needed to connect the two most remote nodes in the network. For example, given the network in Figure 3, the D is three which is identified in the pairs of nodes such as {v2, v5}, or {v1, v5}, or {v5, v6}. A spatial network may assign attributes to edges or vertices of a graph, often called a weighted graph, to represent properties of links and nodes in a network. For example, attributes may represent distance or cost of links (e.g., a speed limit or a driving time), and the size of cities or the capacity of facilities for nodes. Many realistic problems, referred to as network problems, are modeled and solved based on attributes of edges or nodes in the network. For details about the classes of network problems, refer to Drezner and Hamacher (2002) and Ahuja et al. (1993). A classical problem, the shortest path problem, is introduced as an example in the last section, and can also be reviewed at The Classic Transportation Problem here in the GIS&T Body of Knowledge.

4. Network Data Structure in GIS

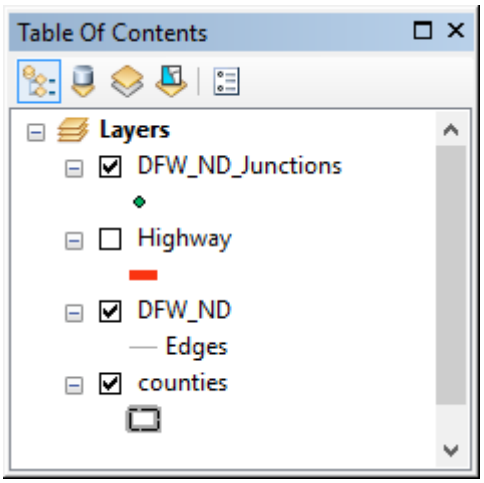
In GIS, a network is represented with features that represent geometries and properties. Often the geometries (or spatial locations) of links and nodes are organized with multiple separate layers (i.e., line and point layers). Their properties (such as traffic volume and the number of migrants that are often necessary for network analysis) can be stored in their attribute tables.

While such line and point layers can be utilized in cartographic representations of a network, network topology should be constructed for network analyses such as shortest path findings. For example, although Figure 1a displays the accurate locations of the highway segments, their logical connections are not ensued. In contrast, Figure 5a illustrates a network dataset constructed from the highway layer. Note that it is zoomed to Dallas County and corresponding layers are listed in Figure 5b. In the construction process, the junctions were created where highway segments are connected and their logical connections were collected and stored in the dataset. Figure 5c shows the network dataset (DFW_ND) in a geodatabase in Esri's ArcGIS: Highway is the source line layer, and two layers (i.e., DFW_ND_Junctions and Highway) are part of the network dataset.

(a)



(b)



(c)

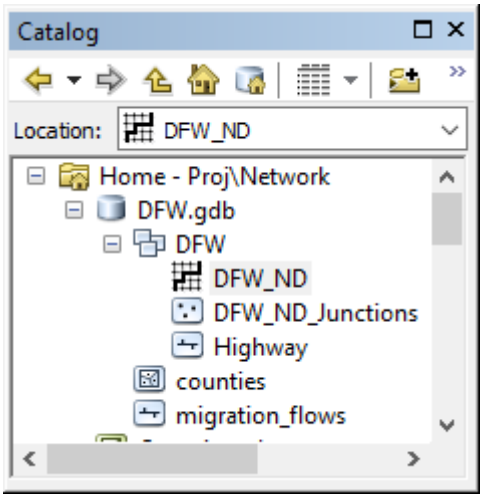


Figure 5. A network dataset for highways in Dallas County. Source: author.

In some applications, raster data structure can be utilized for a route finding problem that appears in continuous space. Here, each raster cell is interpreted as a node and links from each cell are connected to only its adjacent cells (e.g., 8 neighboring cells). That is, the connectivity is same at all cells. A route finding problem on raster, also called a least-cost path problem, utilizes this network structure and cell values for cost for the corresponding cells. This least-cost path problem has been utilized in identifying an optimal route for new utility facilities such as oil and gas pipeline and power lines (e.g., Feldman et al. 1995) and ecological conservation and environmental management (e.g., Dean 1997; Desrochers et al. 2011).

5. Application: The Shortest Path Problems

The shortest path problems are well studied. The standard form of the shortest path problems identifies the least length or cost path between two given nodes, and it can be expanded to find multi-shortest paths among a specific set (or all-pairs) of nodes. The objective function is to minimize the sum of the attributes, often called the network cost, on the edges which are the members of a path. The problems can be solved with different approaches including the matrix operation method, mathematical programming, or an algorithm-based approach.

- The matrix operation method uses matrix multiplication and matrix update procedures such that the attributes of links based upon their connectivity count the sum of attributes to possible paths. In this process, the diameter of the network determines the termination of operation (see O'Sullivan and Unwin 2010 for more detail).
- The mathematical approach interprets the problem as a network flow problem by setting an objective function that minimizes the total travel cost through selected decision variables of links between a source (s) and a destination node (t). A set of network constraints are formulated to ensure connectivity of paths and conservation of flow for the chosen path, and define integrality of decision variables (See Ahuja et al. 1993 for further detail).
- The algorithm-based approach effectively finds optimal solutions on large networks compared to mathematical programming. Two classical approaches by Dijkstra (1959) and Ford (1956) are well studied. Both algorithms adapt a systematic labeling procedure to find the optimal solution of the shortest paths. For example, Dijkstra's algorithm keeps track of specific set of links while building the least cost path from origin to destination node. At each iteration of the algorithm, the selection of nodes to build the least cost path is made by comparing the increased lengths when each of the candidate nodes becomes a member of the path. The procedure terminates when the path includes the destination node.

References

Ahuja, R. K., Magnanti, T. L., and Orlin, J. B. (1993). Network Flows: Theory, Algorithms, and Applications. Prentice Hall.

Dean, D. J. (1997). Finding optimal routes for networks of harvest site access roads using GIS-based techniques. Canadian Journal of Forest Research, 27(1), 11-22.

Desrochers, A., Bélisle, M., Morand-Ferron, J., and Bourque, J. (2011). Integrating GIS and homing experiments to study avian movement costs. Landscape Ecology, 26(1), 47-58.

Dijkstra, E. W. (1959). A note on two problems in connection with graphs. Numerische Mathematik, 1, 269-271.

Drezner, Z., and Hamacher, H. W. (eds.) (2002). Facility Location: Applications and Theory. Heidelberg: Springer Berlin.

Feldman, S. C., Pelletier, R. E., Walser, E., Smoot, J. C., and Ahl, D. (1995). A prototype for pipeline routing using remotely sensed data and geographic information system analysis. Remote Sensing of Environment, 53(2), 123-131.

Ford, L. R. (1956). Network flow theory. Technical Report P-932. Santa Monica, CA. Rand Corporation.

Golledge, R. G. (1987). Environmental cognition. In I. Altman and J. F. Wohlwill, Eds., Handbook of Environmental Psychology. New York: Wiley, vol. 1, pp. 131-174.

O'Sullivan, D. and Unwin, D. (2010) Geographic Information Analysis, 2nd Edition. John Wiley & Sons, Inc.