8.7 Network analysis

Line data that explicitly store topological information are called network data. Besides spatial operations, several methods of spatial analysis are applicable to network data.

8.7.1 Introduction

Some methods focus on the topological structure of the network, while others consider metric properties of the network.

The former is called 'topological analysis' while the latter is 'metric analysis'.

Topological analysis

In topology, a subfield in mathematics, spatial objects are regarded as equivalent if they can be transformed with each other by the rubber sheeting operation without changing their spatial structure. Equivalent objects are called 'isomorphic' objects.
Topological analysis focuses on the topological structure of a network; whether two nodes are directly connected, how many links a node is connected to, etc. Consequently, analysis of isomorphic networks yields the same result.

Topological analysis is often called 'graph-theoretic', because it regards the network as a 'graph'. The term 'graph' refers to a representation of the topological structure of a network, which neglects the length, shape, and other attributes of links.

Metric analysis, on the other hand, considers not only topological but also geometric properties of a network, say, the length, direction, and curvature of links.

Graph
A graph is a set of nodes and their connecting links.

Subgraph:
A subgraph is a part of a graph. It consists of a subset of nodes and links of the original graph.

Connected graph
A connected graph is a graph whose nodes are connected directly or indirectly with each other.

Disconnected graph
A disconnected graph is a graph in which some nodes are not connected either directly or indirectly with each other. A disconnected graph consists of a set of connected graphs, which are called connected elements.
Three connected elements

Figure: Connected and disconnected graphs

Planar graph
A planar graph is a graph in which links intersect only at nodes.

Non-planar graph
A non-planar graph is a graph in which some links intersect at points between nodes.

Figure: Planar and non-planar graphs

Note: A graph is a planar graph if it can be transformed by an isomorphic transformation into a planar graph.

Figure: Complete graphs

Complete graph
A complete graph is a graph in which every pair of nodes is connected directly by one link.
Circuit
A circuit is a set of links that starts from a node, visits several nodes, and returns to the starting node. If a circuit visits every node only once, the circuit is called a simple circuit.

Loop
A loop is a link whose ends are the same node.

Tree graph
A tree graph is a graph that does not contain a circuit.

In graph theory, it is permitted that two nodes are connected directly by more than one link. Loops can also exist in graph theory.
In network analysis, however, two nodes can be connected directly by only one link. Loops are not permitted.

One of the motivations of network analysis is to evaluate a network in terms of the connectivity among nodes, whether links are dense enough to provide a certain level of accessibility among nodes.

Network analysis is important in transportation planning, because the accessibility of residents to urban facilities is evaluated on a road network.

Dense networks are more convenient than sparse ones if they represent traffic networks.

Consequently, we evaluate the connectivity among nodes by measuring the density of links.

A network is well-connected if \( l \) is relatively larger than \( n \). Connectivity measures thus evaluate \( l \) in comparison with \( n \).

\[ \mu = l - n + c \]

A dense network has a large \( \mu \), which implies that nodes are well connected.

Among connected graphs (\( c=1 \)) a tree graph has the smallest \( \mu \). This indicates that, given a set of nodes, tree graphs are the most efficient graphs to connect all the nodes.
Given the number of nodes \( n \), we can calculate the maximum number of links. It is given by the complete graph, that is, \( n(n-1)/2 \).

If we consider only planar graphs, the maximum number of links is \( 3n-6 \), which gives the maximum \( \mu \), \( 2n-5 \). The domain of \( \mu \) is

\[
0 \leq \mu \leq 2n-5
\]

In network analysis we often implicitly assume planar graphs.

2) \( \alpha \) index

\( \alpha \) index is a standardized version of \( \mu \) index.

Dividing \( \mu \) by its maximum \( 2n-5 \), we obtain

\[
\alpha = \frac{\mu}{2n-5} = \frac{l-n+c}{2n-5}
\]

The domain of \( \alpha \) for planar graphs is

\[
0 \leq \alpha \leq 1
\]

3) \( \beta \) index

\( \beta \) index is defined by

\[
\beta = \frac{l}{n}
\]

A dense network has a large \( \beta \) as well as \( \mu \).

The domain of \( \beta \) for planar graphs is

\[
0 \leq \beta \leq \frac{2n-6}{n}
\]

4) \( \gamma \) index

\( \gamma \) index is a standardized version of \( \beta \) index:

\[
\gamma = \frac{\beta}{3n-6} = \frac{l}{3n-6}
\]

The domain of \( \gamma \) for planar graphs is

\[
0 \leq \gamma \leq 1
\]

Comparison of four measures

<table>
<thead>
<tr>
<th>( \mu )</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \gamma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.80</td>
<td>0.44</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.80</td>
<td>0.44</td>
</tr>
<tr>
<td>3.00</td>
<td>0.60</td>
<td>1.40</td>
<td>0.78</td>
</tr>
<tr>
<td>5.00</td>
<td>1.00</td>
<td>1.80</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Properties of connectivity measures

The connectivity measures show how densely nodes are connected by links. They provide a simple and efficient way of evaluating accessibility among nodes.

The measures consider only the number of nodes, links, and connected elements. They drop detailed information about network connection, so they often cannot distinguish different graphs.
8.7.3 Topological analysis 2: accessibility measures

Connectivity measures describe the total (average) connectivity among nodes. In this sense they are global measures.

Accessibility measures, on the other hand, are local measures because they are defined for every node. Accessibility measures evaluate the accessibility from a node to the other nodes.

---

Terminology

Distance between two nodes
In graph theory, the distance between two nodes is defined as the minimum distance on the network.

Topological distance between two nodes
Topological distance between two nodes is defined as the minimum distance on the network, where the length of all the links is set to one. Consequently, topological distance between two nodes is the minimum number of links that connect the nodes.

1) König number

König number of a node is the topological distance to its farthest node. Any node is located within the topological distance given by the König number.
Small König number indicates that the node has high accessibility in the network, that is, it is located at the 'center' of the network. In transportation network, nodes of small König numbers are convenient locations in terms of accessibility to other nodes.

Using the König number we define the topological diameter of a network. The diameter of a network is the distance between the farthest pair of nodes, that is, the maximum König number.

If a graph has a round shape, its diameter is relatively small. A graph of an elongated shape has a large diameter.

2) Simbel number

Simbel number of a node is the sum of the topological distances to the other nodes.

As well as the König number, the Simbel number describes the accessibility of a node to other nodes, and consequently, evaluates the locational convenience of the node.

In metric analysis, on the other hand, we consider not only topological but also metric properties of a network, say, the length, curvature, and shape of links and the flow on the network.
### Terminology

**Length of a link**

In topological analysis, the length of a link is always set to one because it is the topological distance between adjacent two nodes. In metric analysis, on the other hand, the length of a link is defined by a metric measure such as the Euclidean distance, network distance, or time distance.

### Notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>The number of nodes</td>
</tr>
<tr>
<td>$l$</td>
<td>The number of links</td>
</tr>
<tr>
<td>$d_i$</td>
<td>The length of link $i$</td>
</tr>
<tr>
<td>$D_{ij}$</td>
<td>The distance between nodes $i$ and $j$</td>
</tr>
<tr>
<td>$D$</td>
<td>The diameter of a network ($D = \max D_{ij}$)</td>
</tr>
<tr>
<td>$f_i$</td>
<td>The flow on link $i$</td>
</tr>
</tbody>
</table>

### 1) $\pi$ index

$\pi$ index is the ratio of the total length of links to the diameter of the network. Mathematically it is defined as

$$\pi = \frac{\sum d_i}{D}$$

A dense network has a large $\pi$ which indicates that nodes are well-connected and that the network is conveniently structured.

The domain of $\pi$ is

$$1 \leq \pi$$

If a graph is disconnected, we calculate $\pi$ for each connected element separately and average the indices.

### 2) $\theta$ index

There are two types of $\theta$ index: $\theta_1$ and $\theta_2$. The former is a connectivity measure while the latter an accessibility measure.

$\theta_1$ index is the average flow per node:

$$\theta_1 = \frac{\sum f_i}{n}$$

In the real world, a large flow on a network implies that the nodes are closely related with each other. A large $\theta_1$, consequently, which reflects a large flow on the network, suggests that the nodes are well connected.
8.7.5 Metric analysis 2: accessibility measures

As well as topological analysis, metric analysis discusses the accessibility of nodes on a network.

1) $\eta$ index

$\eta$ index is the average length of links:

$$\eta = \frac{\sum d_i}{l}$$

If nodes are connected by short links, $\eta$ shows a small value, which implies high accessibility among nodes.

2) $\theta$ index

$\theta$ index is the average length of links per node:

$$\theta = \frac{\sum l_i}{n}$$

A small $\theta$ indicates that nodes are connected by short links, and consequently, high accessibility among nodes.

3) Degree of circuity

The degree of circuity is defined by

$$\frac{\sum (l_i - e_i)^2}{n}$$

where $e_i$ is the Euclidean distance between end nodes of link $i$. If links are close to straight lines, this index shows a small value, which indicates that high accessibility among nodes.

8.7.6 Application

Connectivity and accessibility measures are used

1. to evaluate existing transportation networks,
2. to analyze transportation networks in relation to landuse patterns, and
3. to analyze urban development process.

Figure: Urban development process
Homework Q.8.5 (10 pts)

1. Take (a part of) a subway network in a city, and represent it as a graph.
2. Examine whether the graph is a planar graph.
3. Calculate its connectivity measures, $\mu$, $\alpha$, $\beta$, and $\gamma$.

Homework Q.8.6 (30 pts)

Suppose a set of $n$ points denoted by $P$. You choose a point pair randomly from $P$, and link them. If you repeat this process $m$ times, what is the probability that two points are located within the distance $l$ measured in the topological distance? Discuss first the case of $l=1$, and then $l=2$, ...