**Network and Theory**

To understand the theory of networks, background knowledge about basic elements and various characteristics of networks is needed. In this unit, we will show how networks are composed and what kind of networks there are, respectively. The following concepts are introduced:

* Graphs
* Elements of graphs: nodes and edges
* Directed (asymmetric) or undirected (symmetric) edges
* Planar and nonplanar graphs
* Adjacency matrix

In graph theory, distinct terminology may apply. However, most terms have simple definitions. A graph consists of two types of elements, namely vertices (V) and edges (E) (e.g.: Graph A). Vertices represent objects that can have names and other attributes such as, for example, transfer points in a subway system. Every edge has two endpoints in the set of vertices, and is said to connect or join the two endpoints (they are adjacent), e.g. a flight connection between Baghdad and Basrah. Graphs are represented by drawing a dot or circle for every vertex, and drawing an arc between two vertices if they are connected by an edge (Graphs A-K). A graph is defined independently from its visualisation. Both figures A and B display the same graph.

A path from vertex s to vertex e in a graph is a list of adjacent vertices. A simple path is a path where no vertex is used more than once. In a connected graph there is an edge from each vertex to every other vertex. In a disconnected graph, there are disconnected parts of edges and nodes (Graph C).



A cycle is a simple path where the first and last vertex is the same (Graph E). A graph without cycles is called a tree (Graph K). Trees have a hierarchical structure and are often explicitly displayed.

The complete graph is a simple graph in which each vertex is adjacent to every other vertex (Graph F). Graphs are considered as sparse when they have a relatively small number of edges; graphs with a small number of possible edges missing are described as dense. In weighted graphs (Graph G), a number or a weight is assigned to each edge of the graph to represent distance (temporal or geometric) or cost. That way, more information is linked to the graph.

**Undirected Graphs:**



In directed graphs, edges are "one-way streets" (Graph H); an edge can lead from x to y but not from y to x. Directed graphs without directed cycles (directed cycles: all edges point in the same direction) are called directed acyclic graphs (Graph I). Directed and weighted graphs are called networks (Graph J). In colloquial language and in the field of geography the term net or network is often used for all kinds of graphs.

**Directed Graphs:**

****

There is another categorization: planar and non-planar graphs. A planar graph is one which can be drawn on the plane without any lines crossing (Graphs L and M).

**Planar vs. non-planar graphs:**

****

Another simple form of representation for graphs, which can also be processed by a digital computer, is an adjacency matrix. A matrix of size VxV is designed, where V is the number of nodes. The fields are set to 1 if an edge between the nodes exists, for example, a and b, and to 0 if no such edge exists. In the example we assume that an edge from each node to itself exists. Whether you set the size of the diagonal to 1 or 0, depends on the intended purpose. In some cases it is better to set the diagonal to 0. This matrix is called adjacency matrix because its structure indicates which nodes are neighbouring (i.e. are adjacent). Something similar is possible with weights: the weights (instead of the value 1) are written into the fields.



The notion of topology is often used in association with the description and analysis of networks. Topology and its concepts are discussed in detail in the module "Spatial Modeling". The networks shown in the figure below have topologically equivalent compounds, thus they are topologically equivalent graphs.



**Structural Properties of a Network**

After having discussed the basic building blocks of networks in detail, let us now deal with ways to capture and describe the structure of networks. The following measures are available for these tasks:

* Connectivity (Beta-Index)
* Diameter of a graph
* Accessibility of nodes and places
* Centrality / location in the network
* Hierarchies in trees

For most of these measures we will present one unweighted and one weighted (metric) case.

**Connectivity (Beta index)**

The simplest measure of the degree of connectivity of a graph is given by the Beta index (#). It measures the density of connections and is defined as:

$$β=\frac{E}{V}$$

where E is the total number of edges and V is the total number of vertices in the network.



*Beta index, calculated for different graphs*

In the figure above, the number of vertices remains constant in A, B, C and D, while the number of connecting edges is progressively increased from four to ten (until the graph is complete). As the number of edges increases, the connectivity between the vertices rises and the Beta index changes progressively from 0.8 to 2. Values for the index start at zero and are open-ended, with values below one indicating trees and disconnected graphs (A), and values of one indicating a network which has only one circuit (B). Thus, the larger the index, the

higher the density.

With the help of this index, regional disparities can be described, for example. In the figure below, the railway networks of selected countries are compared to general economic development (using the energy consumption index of the 1960s). Energy consumption is plotted on the *y*-axis and the Beta index on the *x*-axis. Where connectivity is high, the economic development is high as well.



Energy consumption compared to measure of connectivity (Haggett et al. 1977)

**Diameter of a graph**

Another measure for the structure of a graph is its diameter. Diameter # is an index measuring the topological length or extent of a graph by counting the number of edges in the shortest path between the most distant vertices. It is:

$$δ=max\_{ij}\left\{s\left(i,j\right)\right\}$$

where s(i, j) is the number of edges in the shortest path from vertex i to vertex j. With this formula, first, all the shortest paths between all the vertices are searched; then, the longest path is chosen. This measure therefore describes the longest shortest path between two random vertices of a graph.



The first two figures in graph A show possible paths but not the shortest paths. The third figure and figure B show the longest shortest path.



In addition to the purely topological application, actual track lengths or any other weight (e.g. travel time) can be assigned to the edges. This suggests a more complex measurement based on the metric of the network. The resulting index is *# = mT/m#*, where *mT* is the total mileage of the network and *m#* is the total mileage of the network's diameter. The higher *#* is, the denser the network.

**Accessibility of vertices and places**

A frequent type of analysis in transport networks is the investigation of the accessibility of certain traffic nodes and the developed areas around them. A measure of accessibility can be determined by the method shown. The accessibility of a vertex i is calculated by:

$$E\_{i}=\sum\_{j=1,j\ne i}^{v}n(i,j)$$

where v = the number of vertices in the network and *n* (*i*, *j*) = the shortest node distance (i.e. number of nodes along a path) between vertex *i* and vertex *j*. Therefore, for each node i the sum of all the shortest node distances *n*(*i*, *j*) are calculated, which can efficiently be done with a matrix. The node distance between two nodes *i* and *j* is the number of intermediate nodes. For every node the sum is formed. The higher the sum (node A), the lower the accessibility and the lower the sum (node C), the better the accessibility.

The importance of the node distance lies in the fact that nodes may also be transfer stations, transfer points for goods, or subway stations. Therefore, a large node distance hinders travel through the network.



*Calculation of the accessibility Ei*

As with the diameter of a network, a weighted edge distance can also be used along with the pure topological node distance. Examples of possible weighting factors are: distance in miles or travel time as well as transportation cost. For this weighted measure, however, the edge distance is used and not the node distance.

$$E\_{i}=\sum\_{j=1}^{e}s(i,j)$$

where e is the number of edges and s(i, j) the shortest weighted path between two nodes.

**Centrality / Location in the network**

The first measure of centrality was developed by König in 1936 and is called the König number *Ki*. Let *s*(*i*, *j*) denote the number of edges in the shortest path from vertex *i* to vertex *j*. Then the König number for vertex *i* is defined as:

$$K\_{i}=max\_{j\ne i}\left\{s\left(i,j\right)\right\}$$

where s(i, j) is the shortest edge distance between vertex i and vertex j. Therefore, Ki is the longest shortest path originating from vertex i. It is a measure of topological distance in terms of edges and suggests that vertices with a low König numbers occupy a central place in the network.



If you have determined the shortest edge distance between the nodes, then the largest value in a column in the König number (blue). In the example, the orange node is centrally located and the two green nodes are peripheral.

The method for determining the König number is also applicable to a distance matrix. The example of accessibility is shown again in the figure below. This time the matrix is used with the same values to calculate the König number.

