GIS Applications

LECTURE 2

Centrographic Statistics

- The most basic type of descriptors for the spatial distribution of crime incidents are *centrographic statistics. -basic parameters - They include:*
- 1. Mean center
- 2. Median center
- 3. Center of minimum distance
- 4. Standard deviation of X and Y coordinates
- 5. Standard distance deviation

Centrographic Statistics

- **Mean Center**
- The simplest descriptor of a distribution is the *mean center.*
- *This is merely the mean of the X and Y coordinates.*
- *It is sometimes called a center of gravity in that it represents the point in a distribution where all other points are balanced if they existed on a plane and the mean center was a fulcrum (Ebdon, 1988; Burt and Barber, 1996).*

Centrographic Statistics-Mean Center

$$\bar{X} = \sum_{i=1}^{N} \frac{X_i}{N} \qquad \bar{Y} = \sum_{i=1}^{N} \frac{Y_i}{N}$$

Centrographic Statistics- weighted mean center

- A weighted mean center can be produced by weighting each coordinate by another variable, Wi.
- For example, if the coordinates are the centroids of census tracts, then the weight of each centroid could be the population within the census tract.
- The formula above is extended slightly to

$$\bar{X} = \sum_{i=1}^{N} \frac{W_i X_i}{N} \qquad \bar{Y} = \sum_{i=1}^{N} \frac{W_i Y_i}{N}$$

Centrographic Statistics- Standard Deviation

- **Standard Deviation o f the X and Y Coordinates**
- In addition to the mean center and center of minimum distance, we can calculate various measures of spatial distribution, which describe
- the dispersion,
- orientation, and
- shape of the distribution of a variable (Hammond and McCullogh 1978; Ebdon 1988).

Centrographic Statistics- Standard Deviation

$$S_X = SQRT[\sum_{i=1}^{N} \frac{(X_i - \bar{X})^2}{N-1}]$$

$$S_Y = SQRT[\sum_{i=1}^{N} \frac{(Y_i - \bar{Y})^2}{N-1}]$$

Centrographic Statistics- Standard Deviation

- where Xi and Yi are the X (bar) and Y(bar) coordinates for individual points, X and Y are the mean X and mean Y, and N is the total number of points.

Centrographic Statistics- Standard Distance Deviation

- **Standard Distance Deviation**
- While the standard deviation of the X and Y coordinates provides some information about the dispersion of the incidents, there are two problems with it.
- First, it does not provide a single summary statistic of the dispersion in the incident locations and is actually two separate statistics (i.e., dispersion in X and dispersion in Y).

Centrographic Statistics- Standard Distance Deviation

- A measure which overcomes these problems is the *standard distance deviation or standard distance, for short.*
- *This is the standard deviation of the distance of each point from the mean center and is expressed in measurement units (feet, meters, and miles).*
- *It is the two-dimensional equivalent of a standard deviation.*

Centrographic Statistics

$$S_{XY} = \text{Sqrt}\left[\sum_{i=1}^{N} \left\{\frac{(d_{iMC})^2}{N-2}\right\}\right]$$

- where diMC is the distance between each point, i, and the mean center and
- N is the total number of points.
- Note that 2 is subtracted from the number of points to produce an unbiased estimate of standard distance since there are two constants from which this distance is measured (mean of X, mean of Y)

Network analysis-patterns of lines

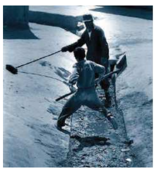
NETWORK ANALYSIS

- Networks are all around us.
- Roads,
- Railways,
- cables,
- pipelines,
- streams and
- even glaciers
- phenomena that frequently need to be represented and analysed as a network.

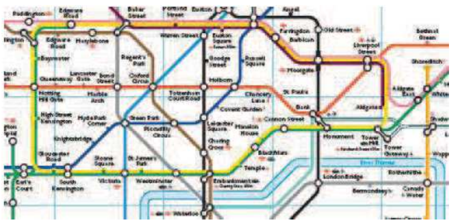
NETWORK ANALYSIS

- Networks are used to move people,
- transport goods,
- communicate information and
- control the flow of matter and energy.

NETWORK



NETWORK



NETWORK

NETWORK ANALYSIS

System requirements	Data requirements
Real time	Accurate
Compact	Up to date
Address conversion function	Topologically correct

System requirements	Data requirements
Output by visual display or synthetic voice	Attributes: • road conditions • classification • speed restrictions • one way streets • turning restrictions • width and height restrictions • junctions • roundabouts • reference landmarks

NETWORK ANALYSIS

A network is one of 5 basic entity types. A network can be defined as a set of linear features through which resources flow. Nodes (the end points of lines) are used as origins and destinations, and links (lines) travers from one node to the other. Nodes can have properties but in network analysis we are usually more concerned with the characteristics of the links (Laurini and Thompson, 1992). These include:

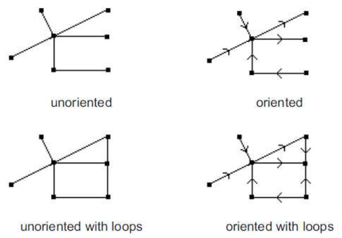
- length,
- direction,
- connectivity (lines must connect at least two points), and
- pattern.

NETWORK ANALYSIS

A classification of networks is discussed by Laurini and Thompson (1992). They suggest there are four main types:-

1. unoriented
2. oriented
3. unoriented with loops, and
4. oriented with loops.

NETWORK ANALYSIS



NETWORK ANALYSIS

- Rivers flow in one direction only- oriented
- Roads- oriented/not oriented- Depends!@
- ??
- One way streets are oriented
- Same with one rail track. Check

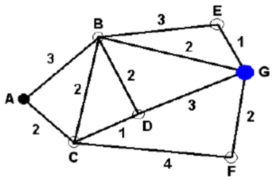
NETWORK ANALYSIS-**Algorithms for network operations**

- At the heart of a network analysis is the search procedure.
- One can, for example, select links that take you as far away from the start node as possible, never turning back.
- Alternatively one can search all the links that propagate from a node, moving out one link at a time and gradually accumulating a cost of travel in all directions from a start node.
- Or one can seek to find a route that passes through the fewest number of nodes, even if it is not necessarily the route of lowest cost.

The Dijkstra algorithm

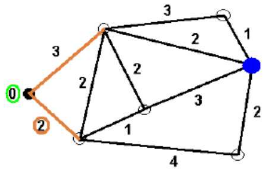
- A common question is "what is the shortest path?"
- for example, when planning the fastest route from one city to another.

The Dijkstra algorithm



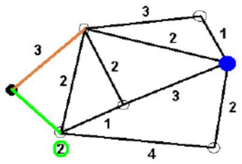
	Dist.	Parent	Incl.
A	0	-	No
B		-	No
C		-	No
D		-	No
E		-	No
F		-	No
G		-	No

The Dijkstra algorithm



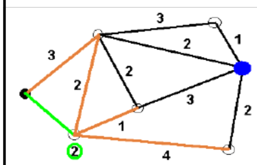
	Dist.	Parent	Incl.
A	**0**	**-**	**Yes**
B	3	A	No
C	2	A	No
D		-	No
E		-	No
F		-	No
G		-	No

The Dijkstra algorithm



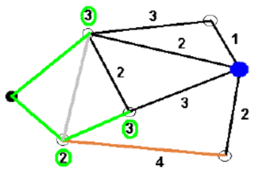
	Dist.	Parent	Incl.
A	**0**	**-**	**Yes**
B	3	A	No
C	**2**	**A**	**Yes**
D		-	No
E		-	No
F		-	No
G		-	No

The Dijkstra algorithm



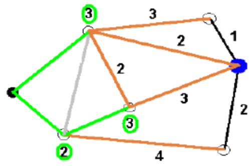
	Dist.	Parent	Incl.
A	**0**	**-**	**Yes**
B	3 or 4	A or C	No
C	**2**	**A**	**Yes**
D	3	C	No
E		-	No
F	6	C	No
G		-	No

The Dijkstra algorithm



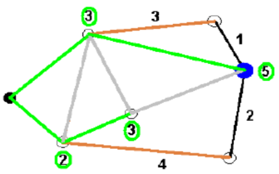
	Dist.	Parent	Incl.
A	**0**	**-**	**Yes**
B	**3**	**A**	**Yes**
C	**2**	**A**	**Yes**
D	**3**	**C**	**Yes**
E		-	No
F	6	C	No
G		-	No

The Dijkstra algorithm



	Dist.	Parent	Incl.
A	**0**	**-**	**Yes**
B	**3**	**A**	**Yes**
C	**2**	**A**	**Yes**
D	**3**	**C**	**Yes**
E	6	B	No
F	6	C	No
G	5 or 6	B or C	No

The Dijkstra algorithm



	Dist.	Parent	Incl.
A	**0**	**-**	**Yes**
B	**3**	**A**	**Yes**
C	**2**	**A**	**Yes**
D	**3**	**C**	**Yes**
E	6	B	No
F	6	C	No
G	**5**	**B**	**Yes**

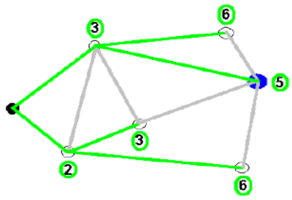
The Dijkstra algorithm

- The shortest path is 5 units. We extract its' route by 'reversing' though the table.
- Node G is reached from B which is reached from A i.e. shortest path = A B G.

The Dijkstra algorithm

- In this case none of the new links produced cumulative costs that were lower than other routes to those nodes so
- the links become redundant in the final analysis.
- Each node is now labeled with a number that represents the minimum cost.

The Dijkstra algorithm



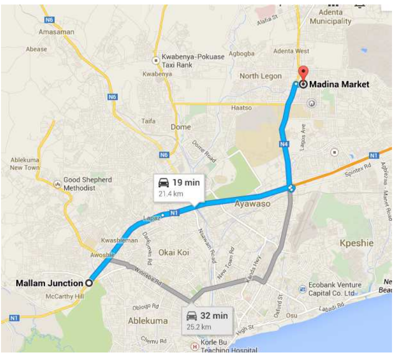
	Dist.	Parent	Incl.
A	**0**	**-**	**Yes**
B	**3**	**A**	**Yes**
C	**2**	**A**	**Yes**
D	**3**	**C**	**Yes**
E	**6**	**B**	**Yes**
F	**6**	**C**	**Yes**
G	**5**	**B**	**Yes**

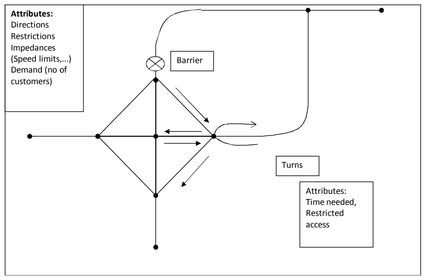
Imperfect solutions and heuristics

- It is necessary to use heuristics in many applications of network analysis to reduce the search space
- For example, in the example given a link might only be valid if it points towards the destination i.e. the angle between the link and a line from the node to the destination is less than 90 degrees.

Routing

- Finding shortest routes is probably the commonest routing problem to occupy GIS users.
- Finding the shortest route from A to B through a road network is crucial for
- emergency services,
- business journeys, or simply planning routes for holiday makers touring a region.



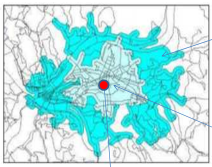


Travelling salesperson problem

- The Travelling Salesman Problem (TSP) is one of these. The travelling salesperson has to visit a number of customers then return to base. S/he needs to have details of the shortest tour possible for the sites to be visited.
- Tour
- Doctor

Isochrones

- Isochrones are lines joining points of equal time (just as isobars join points of equal pressure).
- Applications -for finding travel to work zones based on how long people will travel to a centre, or for the establishment of cost surfaces or zones for transportation of goods and services.
- For example, a furniture retailer may define zones A, B and C around the retail outlet. If you live in Zone A you can have goods delivered free of charge, if you live in Zone B, you will be charged a nominal fee



Network analysis design issues

- Many GIS will claim to contain network analysis operations.
- However, the majority at present only have capabilities for calculating shortest routes.
- data need to be up to date, accurate, topologically correct
- **Considerations** ->How much detail is needed? Who calculates the weightings for each link and how are they calculated?

NETWORK DATASET

- Network datasets are well suited to model transportation networks.
- They are created from source features, which can include simple features (lines and points)
- and turns, and store the connectivity of the source features. When you perform an analysis using ArcGIS Network Analyst, the analysis always happens on a network
- dataset.
- 41 steps in building a network dataset in ArcGIS!

NETWORK DATASET

- HierarchyMultiNet,
- Meters, Minutes,
- Oneway, RoadClass,
- TravelTime,
- WeekdayFallbackTravelTime, and WeekendFallbackTravelTime.

- Whenever the "Use Hierarchy" button is checked the solver goes simultaneously from origin and end points omitting lower hierarchy roads whenever it gets to a higher hierarchy road