Map Overlay Methods

Introduction

The ability to integrate data from two sources using map overlay is perhaps the key GIS analysis function. Using GIS it is possible to take two different thematic map layers of the same area and overlay them on one top of the others to form a new layer. The techniques of GIS map overlay may be linked to sieve mapping, the overlaying of tracings of paper maps on a light table. Map overlay has its origins in the work of McHarg (1969), who used large and complex manual map overlays in landscape architecture. Map overlay has many applications. At once it can be used for the visual comparison of data layers.

As with many other operations and analysis in GIS there are differences in the way map overlays are performed between the raster and vector worlds. In vector based systems map overlay is time-consuming, complex and computationally expensive. In raster based systems it is just the opposite – quick, straightforward and efficient.

Map overlay

Map overlay combines the geometry and attributes of two feature maps to create the output.(i.e.,) overlay where new spatial data sets are created involving the merging of data from two or more input data layers to create a new output data layer. The number of map features of the output map is not the sum of map features on the input and overlay maps but is usually much larger than the sum. Each map feature on the output contains a combination of attributes from the input and overlay maps. Two methods of overlay

      1) Vector overlay
      2) Raster overlay

Vector overlay

Vector map overlay relies heavily on the two associated disciplines of geometry and topology .The data layers being overlaid need to be topologically correct so that lines meet at nodes and all polygon boundaries are closed .To create topology for a new data layer produced as a result of the overlay process, intersections of lines and polygons from the input layers need to be calculated using geometry. In vector based systems overlay is time consuming, complex and computationally expensive. For complex data this is no small task and requires considerable computational power.
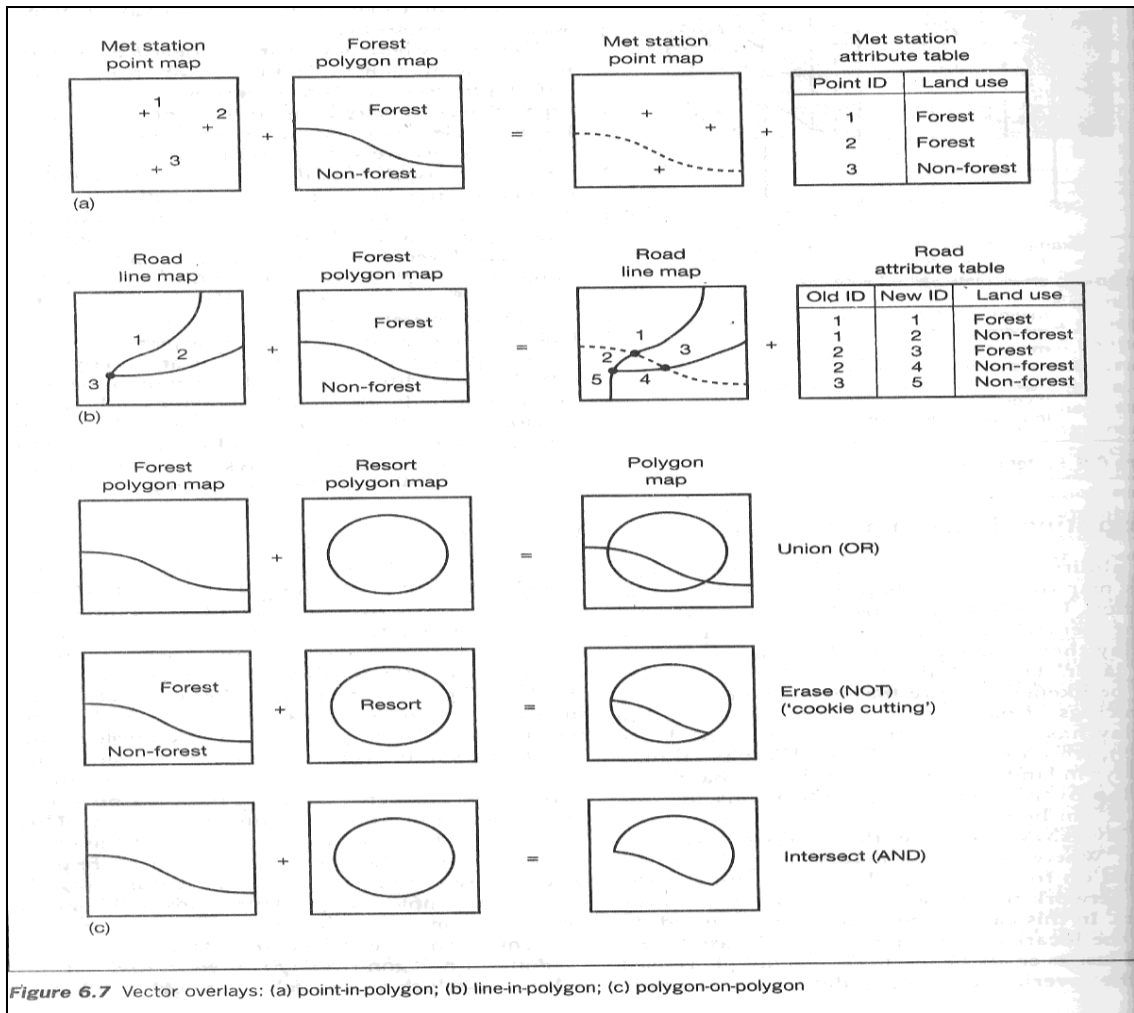
Types of vector overlay

There are mainly three types of vector overlay
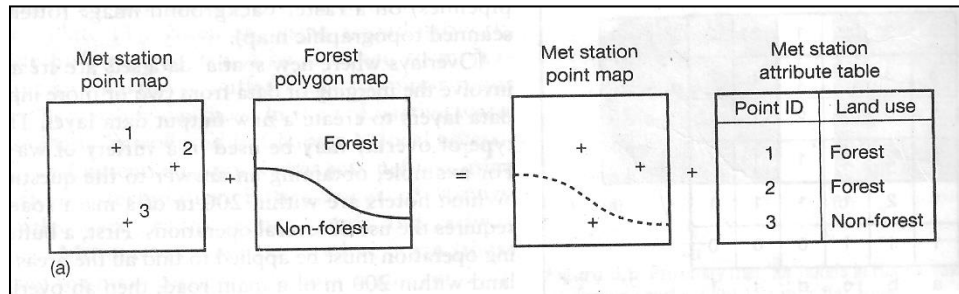Point-in-polygon
Line-in-polygon
Polygon-on-polygon

Figure 6.7 Vector overlays: (a) point-in-polygon; (b) line-in-polygon; (c) polygon-on-polygon

(Fig -1)

Point-in-polygon

In a point-in-polygon operation, the same point features in the input are included in the output but each point is assigned with attributes of the polygon with in which it falls. In the happy valley GIS meteorological stations are represented as points and land use as polygons. Using point-in-polygon overlay on these vector data layers it is possible to find out in which land use polygon each meteorological station is located. Fig 2 illustrates this overlay process. On the out put map a new set of rain gauge points is created with additional attributes describing land use.


(Fig-2)

Point-in-polygon analysis

1) The most common technique used in GIS is the '**half line**' or **Jordon method** (Laurini and Thompson, 1992).Consider a circular polygon with a point at its centre, and another point outside the circular polygon. If a line is extended from the central point to the edge of the data layer it will cross the polygon boundary once. A line extended from the point outside the circle will not cross the polygon boundary.
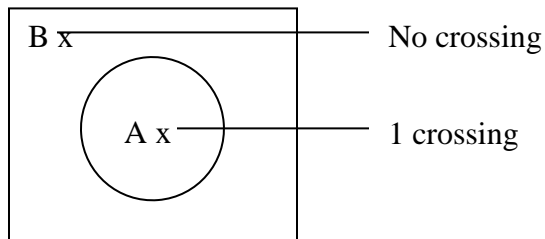
2) Now consider a more complex polygon with crenellated edges. A line extended from the central point will always cross the polygon boundary an odd number of times, while the line extended from point outside the polygon will cross the boundary an even number of times.
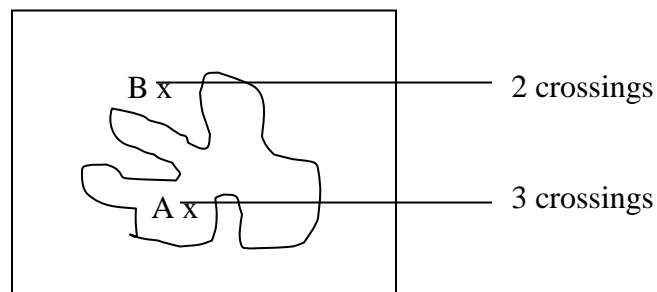
3) A difficulty with this method arises when points lie exactly on the polygon boundary. Are these considered in or out? Lauri and Thomson suggest that where the point has been produced using a mouse, it can be moved slightly to facilitate the in or out decision. Another problem situation is where the 'half line' is coinciding with the polygon boundary, making the counting of crossings impossible.
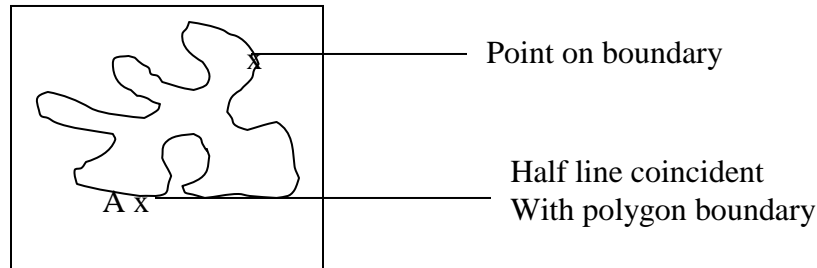
Line-in-polygon

In a line-in-polygon operation, the output contains the same line features as in the input but each of them is dissected by the polygon boundaries on the overlay map.

This overlay is more complicated. Imagine that we need to know where road pass through forest area to plan a scenic forest drive. To do this we need to overlay the road data on a data layer containing forest polygons. The output map will contain roads split into smaller segments representing 'roads in forest areas' and 'roads outside forest areas'.
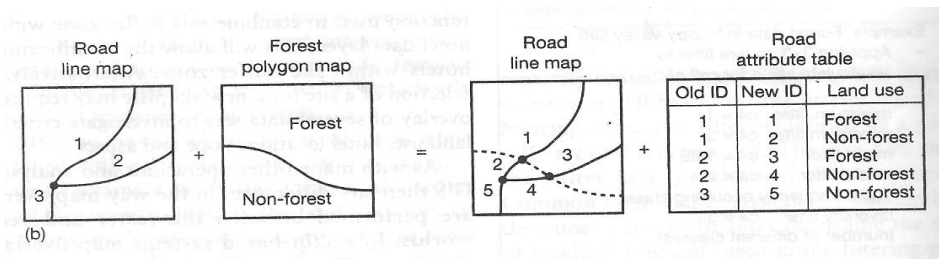
(Fig- 3)

Polygon-on-polygon

The most common overlay operation is polygon-on-polygon, involving two polygons maps. The output combines the polygon boundaries from the input and overlay maps to create a new set of polygons. This overlay could be used to examine the areas of forestry in the happy valley resort. Two input data layers are required – a forest data layer containing forest polygons, and resort boundary layer. Three different outputs could be obtained.

1) The output data layer could contain all the polygons from both of the input maps. This corresponds to the Boolean OR operation, or in mathematical terms, UNION.

2) The forest data layer could contain the whole of the resort area, and forest within this. The boundary of the resort would be used as the edge of the output map, and forest areas would be cut away if they fall outside it. This operation is referred to as 'cookie cutting' and it is equivalent to the mathematical IDENTITY operation.

3) The output data layer could contain the areas that meet both forest and within happy valley .This is the mathematical INTERSECT operation, and the out put map shows where the two input layers intersect. An output map would produced showing the whole of any forest polygons that are entirely with in the resort boundary , and 'cut 'away forest polygons which cross the resort boundary.
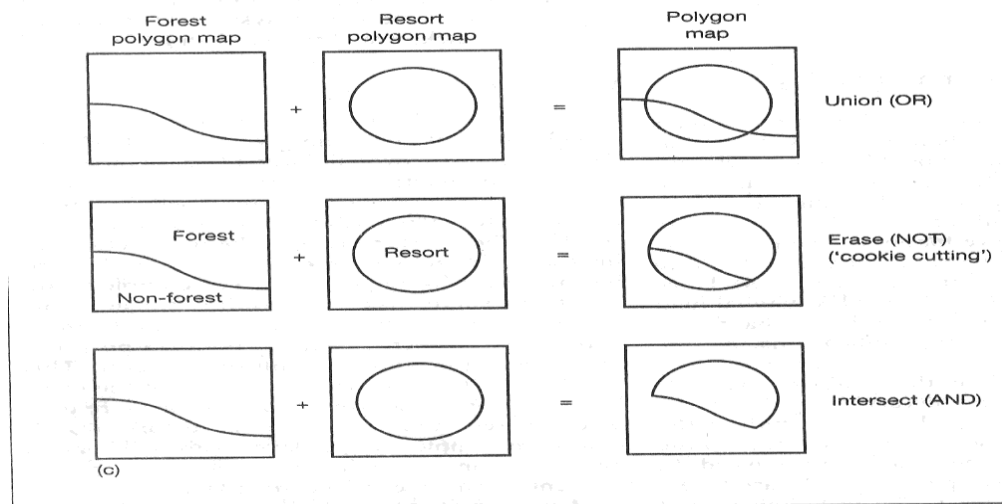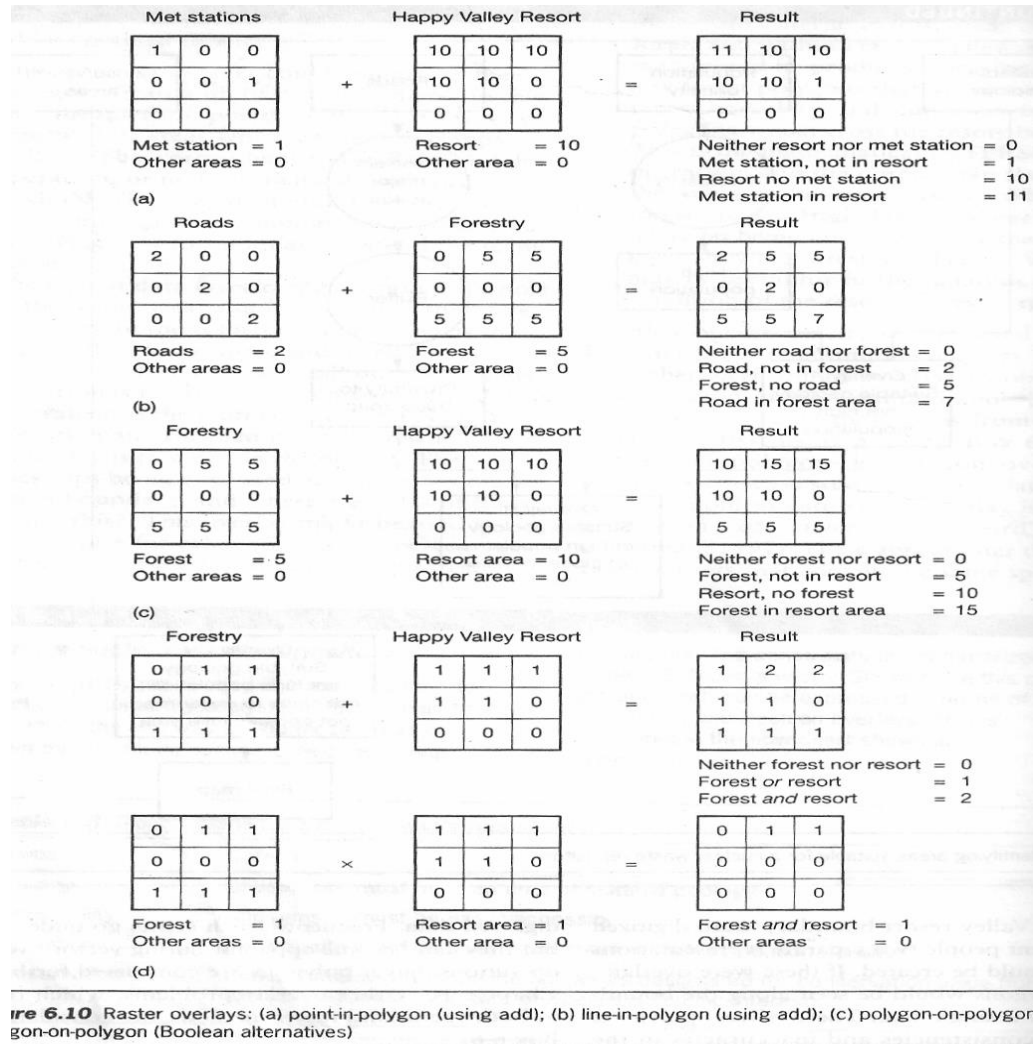

Figure 6.7 Vector overlays: (a) point-in-polygon; (b) line-in-polygon; (c) polygon-on-polygon

   One problem with vector overlay is the possible generation of **sliver polygons.**
These appear after the overlay of two data sets that contain the same spatial entities. If the happy valley resort boundary were digitized by two different people two separate representation of the area would be created. If these were overlaid, long thin polygons would be seen along the boundary instead of a single line. These sliver polygons arise from inconsistencies and inaccuracies in the digitized data.

Raster overlay

   In the raster data structure everything is represented by cells- a point is represented by a single cell, a line by a string of cells and area by group of cells. The methods of performing overlays are different from those in vector GIS.Raster map overlay introduces the idea of map algebra or 'mapematics' (berry,1993) .A mathematical operations are performed on individual cell values from two or more input layers to produce an output value. Thus, the most important consideration in raster overlay is the appropriate coding of point, line and area features in the input data layers. Imagine that four of the happy valley data layers have been rasterized; the location of meteorological stations, the road network, the land use layer and the resort boundary.

Met stations

1	0	0
0	0	1
0	0	0

Happy Valley Resort

10	10	10
10	10	0
0	0	0

Result

11	10	10
10	10	1
0	0	0

+ =

Met station = 1
Other areas = 0

Resort = 10
Other area = 0

Neither resort nor met station = 0
Met station, not in resort = 1
Resort no met station = 10
Met station in resort = 11

(a)

Roads

2	0	0
0	2	0
0	0	2

Forestry

0	5	5
0	0	0
5	5	5

Result

2	5	5
0	2	0
5	5	7

+ =

Roads = 2
Other areas = 0

Forest = 5
Other area = 0

Neither road nor forest = 0
Road, not in forest = 2
Forest, no road = 5
Road in forest area = 7

(b)

Forestry

0	5	5
0	0	0
5	5	5

Happy Valley Resort

10	10	10
10	10	0
0	0	0

Result

10	15	15
10	10	0
5	5	5

+ =

Forest = 5
Other areas = 0

Resort area = 10
Other area = 0

Neither forest nor resort = 0
Forest, not in resort = 5
Resort, no forest = 10
Forest in resort area = 15

(c)

Forestry

0	1	1
0	0	0
1	1	1

Happy Valley Resort

1	1	1
1	1	0
0	0	0

Result

1	2	2
1	1	0
1	1	1

+ =

Neither forest nor resort = 0
Forest or resort = 1
Forest and resort = 2

Forestry

0	1	1
0	0	0
1	1	1

Happy Valley Resort

1	1	1
1	1	0
0	0	0

Result

0	1	1
0	0	0
0	0	0

× =

Forest = 1
Other areas = 0

Resort area = 1
Other area = 0

Forest and resort = 1
Other areas = 0

(d)

ire 6.10 Raster overlays: (a) point-in-polygon (using add); (b) line-in-polygon (using add); (c) polygon-on-polygon gon-on-polygon (Boolean alternatives)

To find out which meteorological stations are contained within the happy valley resort an equivalent to the vector **point-in-polygon overlay** is required. To do this one approach would be to add the two data layers.(figure 4a) The output map would contain cells with the following values:

➢ 0 for cells outside the resort boundary and without a meteorological station;
➢ 1 for cells containing meteorological stations but outside the resort boundary ;
➢ 10 for cells inside the resort boundary but without a meteorological station;
➢ 11 for cells inside resort boundary and containing a meteorological station.

In an operation to the vector **line-in-polygon method** (figure 4b), the sections of roads that pass through forest areas could be obtained. This would require the roads data layer, and a reclassified version of the land use map that contained only forest areas. Again the two maps would be added.

➢ 0 for cells with neither roads nor forest present ;
➢ 2 for cells with roads, but outside forest areas ;
➢ 5 for cells with forest present, but roads absent ;
➢ 7 for cells with both forest and roads present.

If the value '2' for a road were added to land use codes, the new value for a cell could be the same as that for another land use type (for example a road value 2 + land use value 2(water) = 4(which is the same as the value here for agricultural land)

Polygon-on-polygon analysis is conducted in just the same way. (figure 4c). Again the coding of input layers is the key to understanding the output produced by raster overlay. For example, adding the forest layer and resort boundary would produce the output layer with the following codes:

➢ 0 for cells outside the resort boundary and with forest absent ;
➢ 5 for cells outside the resort boundary and with forest present ;
➢ 10 for cells inside the resort boundary and with forest absent ;
➢ 15 for cells inside the resort boundary and with forest present

The coding of raster images used in overlay is very important, and frequently users employ Boolean images (using only codes 1 and 0)(figure   ).The algebraic manipulation of images in raster GIS is a powerful and flexible way combining data and organizing analysis.

There are two issues affecting raster overlay

➢ Resolution
➢ Scales of measurement

Resolution is determined by the size of cell used. (Ex) SPOT satellite data, are collected at a resolution of 10m. For some analysis, you may wish to overlay a SPOT image with 40m resolution. The result will be an output grid with a resolution of 10m, which is greater than the resolution at which the second data set was collected. Since you cannot disaggregate data with any degree of certainty. A better approach to the overlay of these two data sets would be to aggregate cells in the SPOT image to match the resolution of the second layer

The second issue is that of **scales of measurement**. Care is needed when overlaying raster data layers to determine whether the operation makes real sense according to the scales of measurement

Problems affecting Vector and Raster Overlay

1) The modifiable aerial unit problem (MAUP)

It occurs when arbitrarily defined boundaries are used for the measurement and reporting of spatial phenomena. The value of the entity recorded varies continuously over the space. (Ex) population data. The boundaries of these units are arbitrary and subject to change. Thus, changing the boundaries of units and desegregation and aggregation of units can affect the appearance of the data set. The overlay of two data sets, which both have modifiable units, should be undertaken with care, since overlay will almost lead to change in boundaries

2) Ecological fallacy

This occurs when it is inferred that data for areas under study can be applied to the individuals within those areas. For example, if analysis of geomorphic data identifies the average income level in a postal zone, it is an ecologically fallacy to assume that all the individuals within that postal zone have that level of income. Only when totally homogeneous data are aggregated into zonal data can the MAUP and the ecological fallacy be avoided.

3) Selecting Threshold criteria

In the radioactive waste disposal case study four different layers are combined in an overlay procedure to identify areas of the UK that simultaneously satisfy certain criteria. The problem is how to define the threshold criteria used for each constraint. Differences in the thresholds applied will produce differences in the output map. Fortunately, it is possible to investigate how much difference small changes in thresholds make to output maps. Thresholds can be re-applied to one or more of the input layers and overlay analysis repeated to view changes in the output.

4) Visual complexity

When two complex maps are overlain the output is likely to be more complex than either of the two input maps. The sheer complexity of some GIS outputs makes them very difficult to interpret. Generalization may be necessary, through with generalization comes a loss of details

NETWORK ANALYSIS

Network is a series of interconnecting lines along which there is a flow of data, objects or materials.
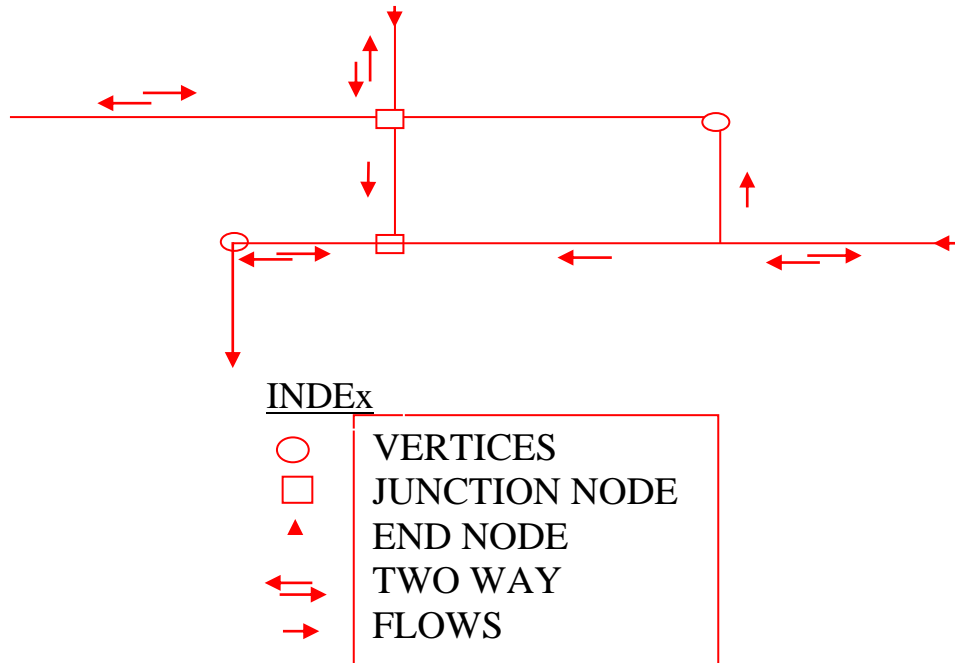
NETWORK MODELLING

- It is an abstract representation of the components and characteristic of their real world counterparts.
- It is defined as a line graph, which is composed of links representing linear channels of flow and nodes representing their connections.
- The network takes the form of edges connecting pairs of nodes can be junctions and edges can be segments of a road or a pipeline.
- In the network model the arc became network links representing the roads, railways, and air routes of transport network, the power line cable &pipe lines of the utilities networks and the rivers and streams of hydrological systems. the nodes in turn become network nodes, stops and centers.
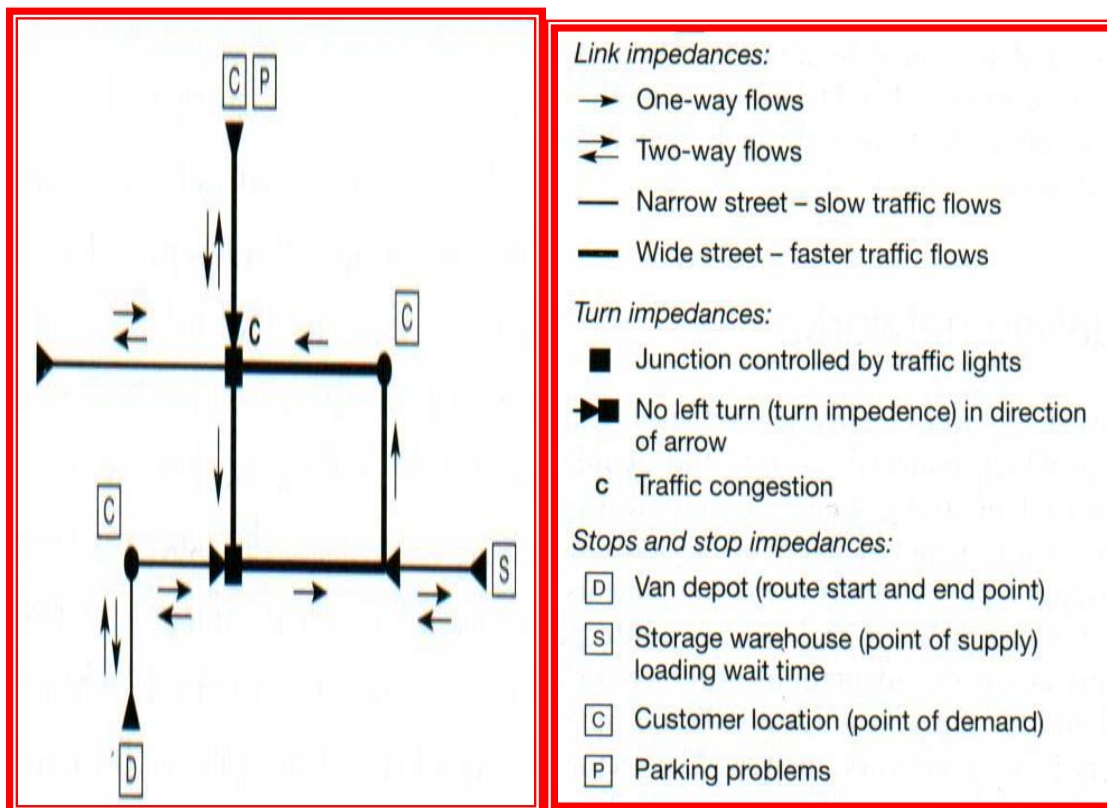
NETWORK DATA MODEL

Network data models becoming increasingly relevant to use a data model where a transportation network can be encoded, stored, retrived. modified, analyzed and displayed. Obviously, Gis have received a lot of attention over this issue since they are among the best tools to store and use network data models. Network data models are an implicit part of many GIS.

There are essentially adaptations of the vector data model and for this reason raster Gis are generally not very good at network analysis. the vector network model is made up of the same arc and node elements as any other vector data model but with addition of special attributed in the above figure.

NETWORK DATA MODEL

INDEx

○	VERTICES
□	JUNCTION NODE
▲	END NODE
⬌	TWO WAY
→	FLOWS

In the network model the arc became network lines representing the roads,railways,and air routs of transport networks, the power line cable & pipelines of transport networks, the power lines of the utilities networks and the rivers and streams of hydrological systems. The nodes in turn become network nodes, stops and centers.

TWO KEY CHARECTERISTICS OF NETWORKS ARE:

1   **Impedance**
2   **supply and demand**

Impedance

- Impedance is the cost associated with traversing a network   link, stopping, turning or visiting a center.
- Link impedance may be the time it takes to travel from one node to another along a network link.
- Turn impedance is also important and may be represented by the cost of making a particular turn.
- Impedance values are very important in determining the outcome of route finding, allocation and spatial interaction operation.

LINK, TURN, STOP IMPEDANCE AFFECTING THE JOURNEY OF A DELIVERY VAN


Link impedances:
- → One-way flows
- ⇄ Two-way flows
- — Narrow street – slow traffic flows
- ▬ Wide street – faster traffic flows

Turn impedances:
- ■ Junction controlled by traffic lights
- ▶■ No left turn (turn impedence) in direction of arrow
- c Traffic congestion

Stops and stop impedances:
- D Van depot (route start and end point)
- S Storage warehouse (point of supply) loading wait time
- C Customer location (point of demand)
- P Parking problems

Supply and demand

- Supply is the quantity of a resource available at a center that is   available to satisfy the demand associated with the links of a Network.
- Correct topology and connectivity are extremely important for   network analysis.

- Digital networks should be good topological representations of   the real world network.

NETWORK MODELLING IN RASTER GIS

In a raster GIS the grid cells only approximate the exact shapes of the lines in the network. The line and node attributes must be stored as a separate layer for each attribute. As a result, a network using a raster model normally consists of a vast number of layers. Since the grid has a given resolution, the cells will only approximate the exact length of the network.

Tracing a path from cell to cell in raster GIS generates a zigzag path instead of a straight line.
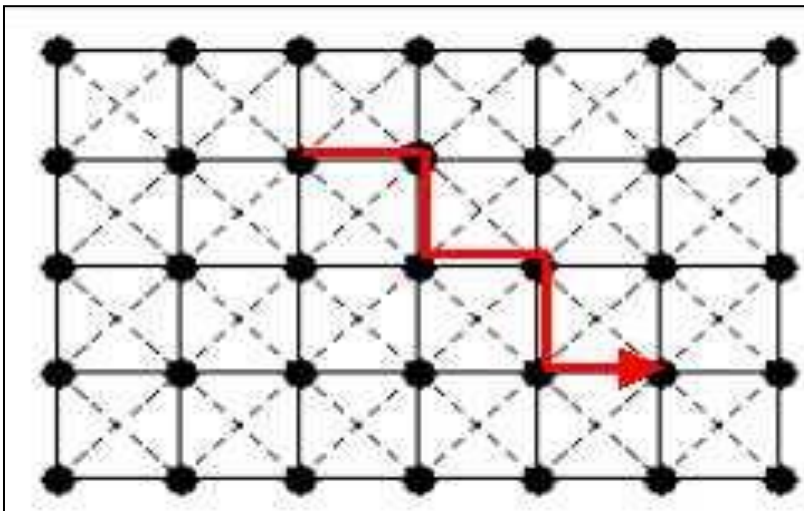
NETWORK MODELLING IN VECTOR GIS

A network in vector GIS takes the form of edges (or arcs) connecting pairs of nodes (or vertices). Nodes can be junctions and edges can be segments of a road. For a network to function as a real-world model, an edge will have to be associated with a direction and with a measure of impedance, determining the resistance or travel cost along the network.
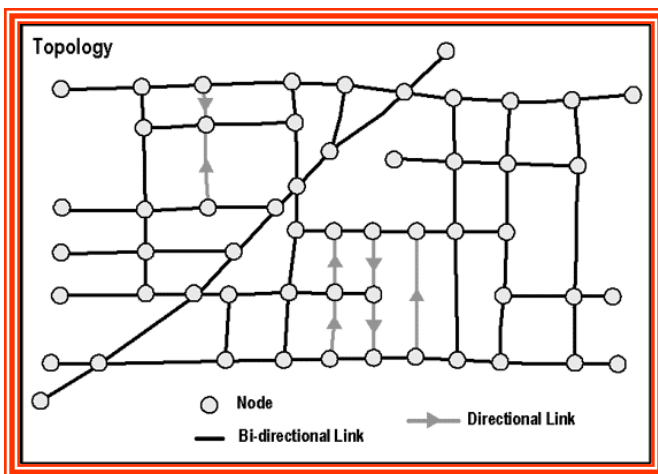
Network representation in vector Gis: nodes (numbers) and arcs (letters)
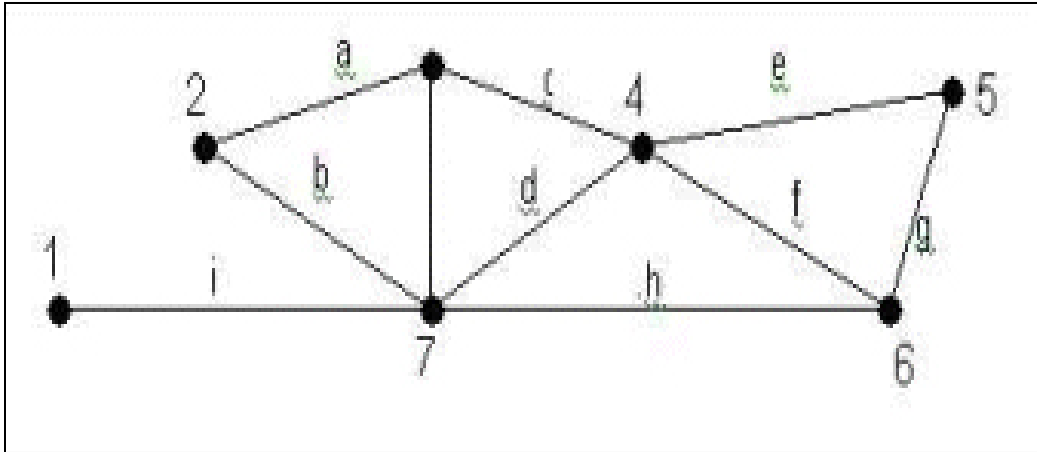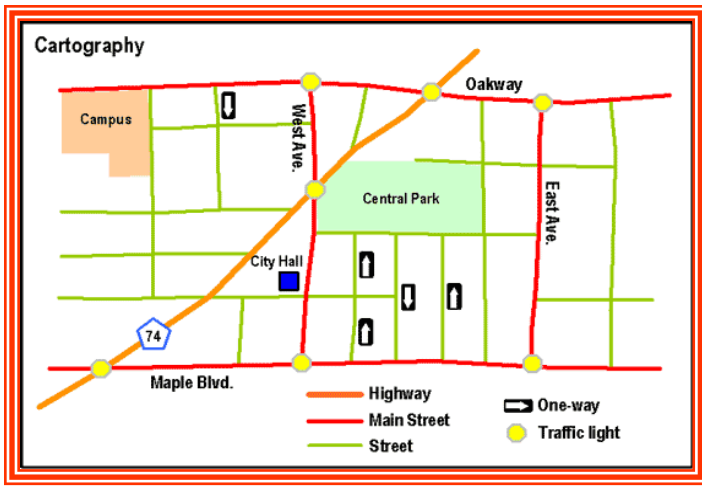
There are four basic application areas of network data models

1. Topology
2. Cartography
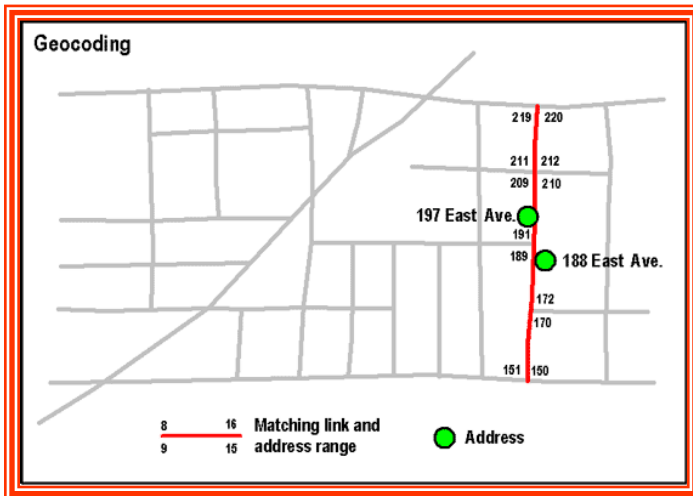3. Geocoding
4. Routing and assignment
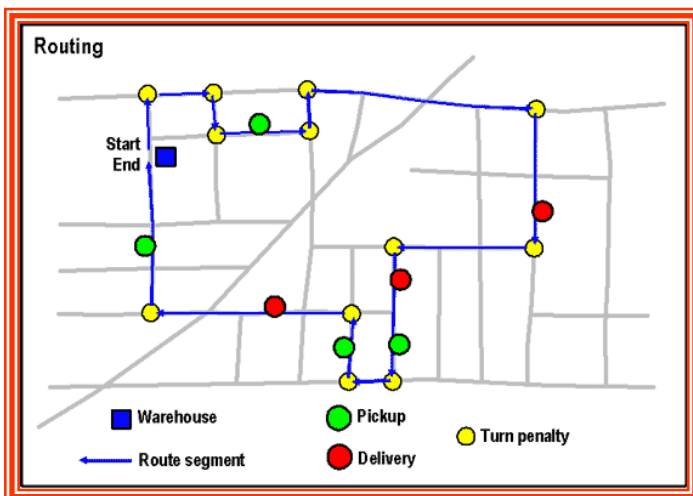
Topology of a Network Data Model

Cartography of a Network Data Model

Geocoding in a Network Data Model

Routing in a Network Data Model

NETWORK MODELING TECHNIQUES

It has been well developed in today's GIS.however to date its major applications have been in the areas of

- Transportation modelling
- Urban and facilities management
- Resource allocation
- Infrasture planning
- Environmental management
- Surface modelling
- Drainage networks

SPATIAL INTERPOLATION

Interpolation predicts values for cells in a raster from a limited number of sample data points. It can be used to predict unknown values for any geographic point data: elevation, rainfall, chemical concentrations, noise levels, and so on. Interpolation is based on the assumption that spatially distributed objects are spatially correlated; in other words, things that are close together tend to have similar characteristics. For instance, if it is raining on one side of the street, you can predict with a high level of confidence that it is also raining on the other side of the street. You would be less sure if it was raining across town and less confident still about the state of the weather in the neighboring province.

Visiting every location in a study area to measure the height, magnitude, or concentration of a phenomenon is usually difficult or expensive. Instead, dispersed sample input point locations can be selected and a predicted value can be assigned to all other locations. Input points can be either randomly, strategically, or regularly spaced points containing height, concentration, or magnitude measurements.

A typical use for point interpolation is to create an elevation surface from a set of sample measurements. Each point represents a location where the elevation has been measured. The values between these input points are predicted by interpolation.
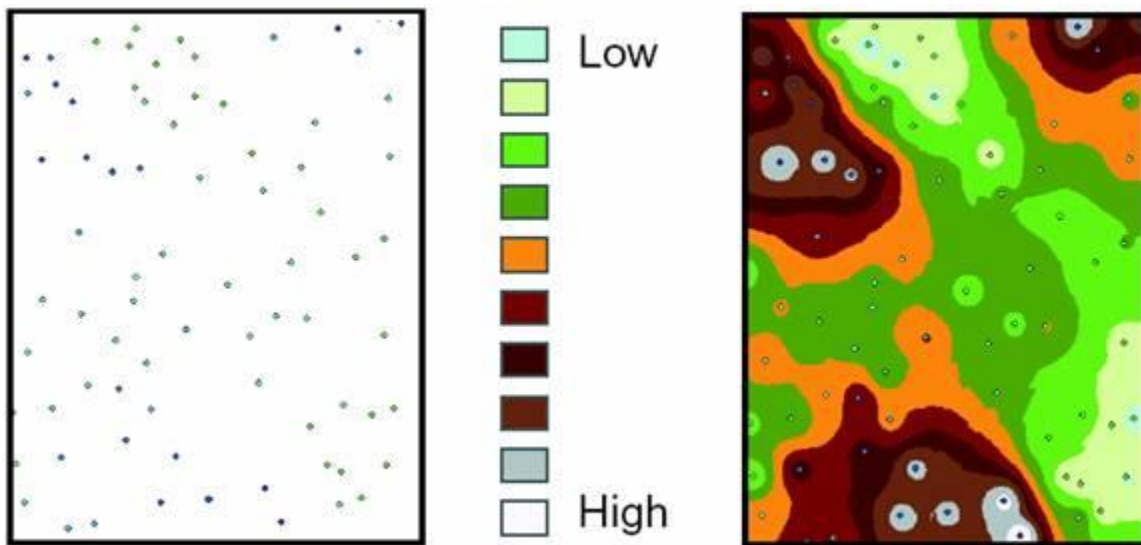
The resulting grid is a prediction of what the elevation is at any location on the actual surface.

Spatial interpolation methods

Inverse distance weighted

Inverse distance weighted is a deterministic estimation method where values at unmeasured points are determined by a linear combination of values at nearby measured points. Among available parameters, the power parameter can significantly affect the results. As the power parameter increases, IDW acts similarly to the nearest neighbor interpolation method in which the interpolated value is close to the value of the nearest measured value. The advantages of IDW are that it is simple, easy to understand, and efficient. Disadvantages are that it is sensitive to outliers and there is no indication of error.

Kriging

Kriging is a stochastic method similar to IDW in that it also uses a linear combination of weights at known locations to estimate the data value of an unknown location. Variogram is an important input in kriging interpolation. It is a measure of spatial correlation between two points. With known variograms, weights can change according to the spatial arrangement of the samples. A major advantage of kriging is that, in addition to the estimated surface, kriging also provides a measure of error or uncertainty of the estimated surface. A disadvantage is that it requires substantially more computing time and more input from users, compared to IDW and spline.

Spline

Spline is a deterministic method to represent two-dimensional curves on three-dimensional surfaces. It can be imagined as fitting a flexible surface through a set of known points using a mathematical function. A major advantage of spline is that it can create fairly accurate and visually appealing surfaces based on only a few sample points. Disadvantages of spline are that the resultant surface may have different minimum and maximum values from the input data set, it is sensitive to outliers, and there is no indication of errors.

SPATIAL MULTIMEDIA

Multimedia GIS is fully capable of providing services to both technical and non-technical users and users of multimedia GIS can retain more knowledge than users of traditional GIS. It can be applied to various fields such as tourism, environment, education etc. Some implementation issues such as data updating, information retrieval and data compression etc. are also addressed.

M M GIS has full capability to provide services to both technical and non-technical users.

M M GIS provides such an atmosphere in which user feels comfortable with the system environment before addressing the issues of data analysis. Basically MM GIS allow user to interact with datasets and help.

U sers of the MM GIS will retain more know ledge than users of traditional GIS. The comb-I nation of several media often results in powerful and richer presentation of information & ideas to stimulate and enhance information retention.
M M GIS allow traditional data in the electronic or dig ital formats and also allow updating the data easily and efficiently with low cost.

MMGIS is also capable of improving procedures for data collection, accuracy & management of temporal data, and understanding of the computerized tools used to handle uncertainty in digital spatial database.